

Professional paper / Stručni rad

Manuscript received: 2017-03-17

Revised: 2017-08-24

Accepted: 2017-09-12

Pages: 57 - 66

Automatska detekcija plagijata softverom PlagScan na Moodle platformi

Zoran Hercigonja

*Druga gimnazija
Varaždin, Hrvatska*

zoran.hercigonja@gmail.com

Sažetak: Plagiranje neprestano evoluirao i javlja se u gotovo svim jezičnim sredinama i područjima pritom zauzimajući nove forme i oblike. Stvaranju sofisticiranog plagijata, puno doprinose i jezične mogućnosti pojedinog jezika. Sve više sadržaja na Internetu, prilagođava se korisnicima i govornim područjima. Više na webu ne prevladava samo engleski jezik nego se pojavljuju i sadržaji na ostalim jezicima koji ne ulaze u skupinu svjetskih jezika. Plagijati sve više evoluiraju. Paralelno s njima, treba kreirati rješenja koja im mogu parirati. Jedno od takvih privremenih rješenja je i dodatak za platformu Moodle. Taj softver ima iznimne mogućnosti pronalaženja plagijata na hrvatskom jeziku. Dakle ne govori se o lokalnoj usporedbi i traženju plagijata, nego traženju podudarnosti plagiranog sadržaja s izvornim sadržajem web stranica na hrvatskom jeziku uz iznimno nisku cijenu.

Ključne riječi: Moodle, PlagScan, Detekcija, Plagijat, Dodatak

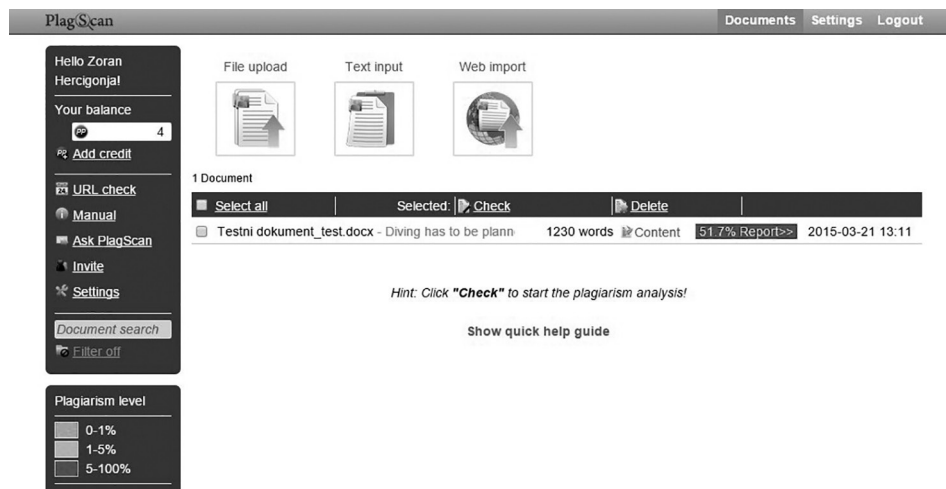
UVOD

Okruženi frazeologijom i gotovim konstruktima, počinjemo određene sadržaje prisvajati, zaboravljajući na originalnost njihovih izvora, zaboravljajući da ih je netko napisao, osmislio i sastavio. S druge strane nedosljednost, nedostatak vještine u izražavanju, kritičkom razmišljanju ili rezoniranju, put je prema plagiranju i prisvajanju tuđih originalnih misli, definicija i rezultata. Prisvajanje i kopiranje tuđeg sadržaja bez referenciranja ili parafraziranja, predstavlja čin krađe kojim se narušava etički kodeks i povređuju etička prava pojedinca, izvornog autora sadržaja određenih znanstveni, istraživačkih ili stručnih radova. Nove tehnologije kao što je na primjer Internet, osim što osiguravaju dostupnost sadržaja, osiguravaju i prepisivanje u obliku direktnog kopiranja sadržaja. Sam razvoj i unapređenje informacijske tehnologije olakšava neovlašteno preuzimanje teksta [1]. Razvoj tehnologije olakšava pristup sadržajima ali i prisvajanje frazeologije i odlomaka jednostavnom naredbom „copy/paste“ („kopiraj/zalijepi). Dakle 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 [1]. Prvi pokušaji doslovnog preuzimanja sadržaja zabilježeni su negdje u doba „primitivnog“ interneta. Sustavi za elektroničko učenje danas su najzastupljeniji oblik programskih rješenja koji podupiru učenje na daljinu. Jedan od najpoznatijih sustava za izradu e-tečajeva temeljen na socio-konstruktivističkoj paradigmi učenja je sustav za upravljanje učenjem Moodle [2]. U posljednje vrijeme, svjedoci smo sve učestalijeg korištenja platformi za elektroničko učenje. Neke od poznatijih platformi su Moodle, Claroline i Webinar. Takve platforme predstavljaju i repozitorij sadržaja gotovih materijala za učenje i studentskih radova. Mjesta na kojima se nalazi pozamašan broj radova, vrlo su pogodna i za plagiranje u raznim oblicima, a najčešće metodom kopiraj/zalijepi. Iako ima mnogo alternativnih softvera za detekciju plagijata izvedenih u obliku desktop verzija, prioritet njihovog korištenja je ekonomičnost. Budući da se u visokim učilištima vrlo intenzivno primjenjuju sustavi za elektroničko učenje, potrebno je koristiti rješenja u obliku nadogradnje koja bi omogućila trenutnu provjeru sadržaja i traženje plagijata. Na taj način nema rasipanja ni vremena ni resursa. Desktop verzije softvera zahtijevaju određene performanse računala, dok nadogradnje na sustav za elektroničko učenje nisu toliko zahtjevne za računalo korisnika. Do sada je već izvedena nekolicina softvera za otkrivanje plagijata u obliku dodataka za pojedine platforme za e-učenje. Primjerice korišteni su i testirani dodaci VeriCite i Crot pro na platformi Moodle. Pokazali su iznimne rezultate pretraživanjem web izvora [3]. Napravljeno je testiranje nad traženjem podudarnosti s web izvorima. Drugo istraživanje iste vrste namjerava se napraviti na Riječkom Sveučilištu korištenjem dodatka Turnitin to jest automatskog lovca na plagijate, koji će skenirati sve diplomske i ostale radove [4]. Svi ti dodaci daju veliku nadu u smanjenje stope plagiranja, no problem je najčešće cijena i moguća ograničenja korištenja. Primjerice cijena licence Turnitin softvera godišnje iznosi 12 tisuća dolara [4]. Rješenju koje se nudi ovim radom, ne može parirati dosad ni jedan dodatak-softver za otkrivanje plagijata na hrvatskom jeziku u cjenovnoj domeni i domeni uspješnosti pronalaska plagijata.

PROGRAMSKI DODATAK PLAGSCAN

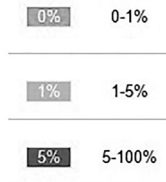
Komercijalni programski dodatak PlagScan na platformi Moodle, dostupan je u repozitoriju dodataka na web stranici Moodle-a. Njegova najveća osobitost je mogućnost kvalitetnog pronalaženja plagijata na hrvatskom jeziku. PlagScan je dodatak specijaliziran za detekciju Copy&Paste ili kopiraj/zalijepi pojavnog oblika plagijata. Radi na principu traženja podudarnosti sadržaja s internetskim stranicama. Dulje korištenje dodatka se plaća, no probna verzija je besplatna. Dodatak je komercijalan. [5]. Probna verzija podrazumijeva korištenje u roku od trideset dana s dodatnim ograničenjima. Primjerice jedno od očiglednih ograničenja je ograničenje u broju riječi pojedinog dokumenta. Probna verzija traži određen broj riječi, 2000 riječi po jednom dokumentu. Prati se i broj pokušaja testiranja na mjesečnoj razini. Moguće je u trideset dana napraviti samo dvadeset pokušaja. Probna verzija je izvedena kao web servis koji ne omogućava detekciju na lokalnoj razini između dva ili više predanih dokumenata na jedno mjesto. Sama kvaliteta traženja uzoraka, utvrđivanja razine plagiranog sadržaja kao i iscrpnost izvještaja, jednaka je kvaliteti licenciranih verzija PlagScana. Podržava više Moodle verzija: 2.0, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8. Podržava nekoliko različitih tipova dokumenata: pdf, doc i docx [5]. PlagScan ima veliku osjetljivost uzoraka. Provodi detekciju na razini riječi. Dakle kao uzorak se uzima riječ. U izvještaju koji je vrlo iscrpan, svakom detektiranom plagijatu PlagScan pridružuje po nekoliko web izvora s kojima je utvrđena podudarnost.

To daje veće izgleda u otkrivanju plagijata. Osjetljivost uzoraka i pronalaženje podudarnosti s više izvora daju veće izgleda u detektiranju plagijata i to rezultira zadovoljstvom korisnika. Omogućava prijenos dokumenata i direktno unošenje teksta u za to predviđeno mjesto.



Slika 1: Početno sučelje

PlagScan prilikom automatske aktivacije odmah nakon predaje dokumenta, generira izvještaj s postotkom kopiranog sadržaja. Ima procjenu razine plagiranog sadržaja u tri boje: zelena, žut i crvena.



Slika 2: Razine plagiranog teksta

Originalnim tekstom se ovdje smatra postojanje dopuštenog kopiranog sadržaja u omjeru do 1%, dok se sve ostalo interpretira kao plagijat. Plagijatom u omjeru 1-5% smatra se modificirani sadržaj što podrazumijeva dodavanje ili ispuštanje riječi, segmenata i rečenica. Plagijatom iznad 5% smatra se doslovno preuzeti sadržaj.

- ✓ [0] (24 Matches, 25.9%/28.1%) from bucearte.com/article/enfermedad-descompresiva?locale=en
- ✓ [1] (23 Matches, 26.1%/28.4%) from www.diversalernetnetwork.org/medical/artic..._ess_What_Is_It_and_What
- ✓ [2] (15 Matches, 16.0%/18.3%) from www.scubaboard.com/forums/diving-medicin..._ot-her-ow-shes-been-di
- ✓ [3] (15 Matches, 17.7%) from docs4ever.blogspot.com/2013_05_17_archive.html
- ✓ [4] (15 Matches, 17.7%) from www.appletreomedicalgroup.com/travel_medicine/travel_listing.php?id=13
- ✓ [5] (15 Matches, 17.5%) from www.netdoctor.co.uk/travel/diseases/decompression_sickness.htm
- ✓ [6] (15 Matches, 17.7%) from www.traveldoctor.co.uk/dcs.htm
- ✓ [7] (14 Matches, 16.3%) from dragondiverscyprus.com/diving-in-cyprus-dragons-blog/
- ✓ [8] (12 Matches, 14.2%) from draomd.com/newblog/decompression-sickness/
- ✓ [9] (6 Matches, 4.3%/6.5%) from colonialcatering.com.au/images/afm3/0189273.php
- ✓ [10] (6 Matches, 5.4%) from www.answers.com/Q/What_are_the_effects_o...Decompression_Sickness_is_
- ✓ [11] (5 Matches, 3.8%/4.9%) from www.londondivingchamber.co.uk/index.php?id=dci&page=11
- ✓ [12] (5 Matches, 5.8%) from static.ukund.com/manuals/URKUND_View6_in_brief_EN.pdf
- ✓ [13] (5 Matches, 4.7%) from static.ukund.com/manuals/URKUND_Master_Class_-_Duplicates_EN.pdf
- ✓ [14] (4 Matches, 7.8%) from www.pbcr.org/pdfs/TrainingAndSafety/COPD/OCSProjects/RescueLts/Barotra
- ✓ [15] (4 Matches, 7.3%) from ahmedredsea.blogspot.com/2012_11_01_archive.html
- ✓ [16] (4 Matches, 7.6%) from ahmedredsea.blogspot.com/2012/11/decompression-sickness.html
- ✓ [17] (4 Matches, 6.8%) from www.adurbac.co.uk/index.php?option=com_content&view=article&id=61&Item
- ✓ [18] (4 Matches, 4.6%) from scubadiverlife.com/2011/04/24/decompression-sickness-scuba-diving/
- ✓ [19] (2 Matches, 2.5%) from www.scubaboard.com/forums/archive/index.php?t-482181.html
- ✓ [20] (2 Matches, 2.5%) from www.scubaboard.com/forums/diving-medicine/482181-so-worry-about-dcs.htm
- ✓ [21] (1 Matches, 4.4%) from www.diveisac.org.uk/goby27.pdf
- ✓ [22] (1 Matches, 4.5%) from www.powershow.com/view1/c45ca-ZDc1Z/Diving_Safety_Myths_powerpoint_pj
- ✓ [23] (1 Matches, 4.3%) from yosh.katbox.net/comic/wake-up-fall/
- ✓ [24] (1 Matches, 3.3%) from www.admfoundation.org/emergencyprocedures.html
- ✓ [25] (1 Matches, 1.2%) from www.preservearticles.com/2012022423687/11-disorders-of-respiratory-system
- ✓ [26] (1 Matches, 0.7%) from myrome.club/portaldelayed-vertigo-after-flying

Slika 3: Iscrpna analiza

PlagScan ima odvojene dijelove iscrpne analize. U izvještaju su posebno odvojeni detektirani izvori preuzetog sadržaja s postocima, a posebno označeni dijelovi teksta s primjerenom bojom kao razinom plagiranja. Svakom detektiranom odlomku ili rečenici, pridružene su brojke u uglatim zagradama što je prikazano slikom 5.

Analyzed document

Diving has to be planned and carried out in a responsible manner, making sure that first aid equipment and relevant phone numbers are at hand if an accident should take place [3] [4] [5] [6] ... Finally, it is important to know the signs of decompression sickness and to be able to give first aid to the diver who has been affected [3] [4] [6] [5] ... Beware of diving bells though; they have become more and more popular at tropical beach resorts lately as a method of experiencing nautical adventures without proper training or brains [12] [13] Diving bells is an obsolete way of diving and should only be utilized by professionals in salvaging operations at great depths, like the Swedish Regal ship Vasa or the German WWII battle cruiser Bismarck [12] ... Not for looking at coral at a depth of a measly 9m in some lagoon. It can be dangerous. Don't do it. What is decompression sickness? [3] [4] [5] [6] ... Decompression sickness, by many referred to as "the bends" or "diving sickness", is triggered when bubbles of nitrogen forms in the blood and tissue of a body. These bubbles follow when a person rises from deep water to shallow water, where the pressure is lower, too quickly. The symptoms of the bends occur almost immediately after a completed dive and can lead to catalepsy or even death in serious cases [13] ... If you or a dive-mate exhibits symptoms of decompression sickness, immediately stop the dive, start administering first

Slika 4: Analiza dokumenta

U probnoj verziji, iscrpna analiza dodatka PlagScan šalje se na e-mail korisnika i može se preuzeti u PDF ili docx formatu dokumenta.

MATERIJALI I METODE

U svrhu testiranja, i provjere kvalitete ovog dodatka, napravljena su tri testa. Sva tri testa, provedena su na razini traženja podudarnosti testiranog sadržaja s web izvorima. Cilj je bio provjeriti postotak uspješnosti detektiranog sadržaja na webu, trajanje procesa detekcije plagijata te broj otkrivenih izvora.

PRVI TESTNI DOKUMENT

Prvi test bio je osmišljen kroz kreirani dokument metodom kopiraj/zalijepi. Dakle prvi dokument sadržava ukupno 844 riječi. Broj riječi po dokumentu je namjerno ograničen zbog ograničenja koje proizlazi iz korištenja probne verzije PlagScana. Naime PlagScan dozvoljava maksimalno 2000 riječi po dokumentu u probnoj inačici. Sadržaji korišteni u sastavljanju prvog testnog dokumenta, kopirani su sa stranica Hrčka. Točnije, radi se o sadržajima radova časopisa iz kategorije Računalstvo objavljenima u protekle tri godine. Sadržaj je nasumično odabran. Odabrano je ukupno devet nasumičnih radova na hrvatskom jeziku. Dakle u testni dokument ulaze sadržaji s ukupno devet izvora. Konačni dokument od 844 riječi, pospremljen je u docx format. Cilj je bio koristiti sadržaje isključivo na hrvatskom jeziku metodom kopiraj/zalijepi.

DRUGI TESTNI DOKUMENT

Za izradu drugog testnog dokumenta, odabrano je pedeset (50) izvora na hrvatskom jeziku. Testni dokument sadržavao je informacije o brodogradnji. Rečenice i fragmenti su doslovno preuzimani s web stranica i chat grupa metodom kopiraj/zalijepi. No da bi provjerili kvalitetu dodatka, rečenice i segmente u dokumentu se ispremiješalo do mjere da nije bilo moguće utvrditi smislen i suvisao logički redoslijed nizanja rečenica. Primjerice rečenicu iz jednog izvora, umetnulo se u preuzeti segment nekog drugog izvora s kojim ta rečenica uopće nije bila povezana. Takav način plagiranja je podosta kompleksan i kompliciran za softvere za detekciju plagijata. Za sastavljanje dokumenta korišteno je 50 različitih izvora isključivo na hrvatskom jeziku. Konačni testni dokument s tri stranice sadržaja na hrvatskom jeziku s ukupno 1337 riječi, spremljen je u docx format.

TREĆI TESTNI DOKUMENT

Cilj je bio usporediti kvalitetu rada plugina provodeći proces detekcije sadržaja na engleskom i hrvatskom jeziku. Pratilo se vrijeme procesa detekcije te postotak otkrivenih plagijata. Izrađena su dva testna dokumenta: jedan na hrvatskom jeziku, a drugi na engleskom jeziku na istu temu. Svaki dokument je sadržavao jednak broj riječi: 873

riječi. Broj riječi je sveden na 873 zbog ograničenja probne verzije PlagScana. Sadržaji su preuzeti iz istog izvora Wikipedije na temu "Motori s unutarnjim izgaranjem". Korištena je Wikipedija zbog mogućnosti višejezičnog prikazivanja sadržaja, ali i učestalosti posjećivanja. Poznato je da su na Wikipediji objavljeni sadržaji na većem broju jezika. To je bio razlog korištenja tih sadržaja. Sadržaji su preuzeti direktno metodom "Copy/Paste".

REZULTATI PROVEDENIH TESTOVA

U sva tri provedena testa, željela se utvrditi razina kvalitete dodatka. Zato su određeni sljedeći pokazatelji kvalitete: 1. vrijeme trajanja procesa detekcije, 2. postotak plagiranog sadržaja te 3. broj otkrivenih web izvora.

REZULTATI PRVOG TESTA

Rezultati prvog testa su i više nego zadovoljavajući. Postupak detekcije plagijata, trajao je svega 2 sekunde. Pri tome je detektirana visoka razina plagiranog sadržaja označena crvenom bojom od 79,6%. Od korištenih devet izvora u sastavljanju prvog testnog dokumenta, pronađeno ih je devet. To znači da je PlagScan uspio pronaći podudarnost s baš svim izvorima iz kojih je sadržaj i bio preuzet.

PlagScan Filename: Copy_Paste_Testni_dokument.docx Date: 2016-01-06 15:22
 Results of plagiarism analysis from 2016-01-06 15:23

15 matches from 5 sources, of which 5 are online sources.

Print | Download as PDF | Show comparison | View as list | Word document (docx) with annotations | Share

- [0] (8 matches, 62.4%) from <https://pogledkrozprozor.wordpress.com/feed/> [Highlight matches in source](#)
- [1] (2 matches, 10.4%) from <https://pogledkrozprozor.wordpress.com/category/primjeri-dobre-prakse/> [Highlight matches in source](#)
(+ 2 documents with identical matches)
- [4] (2 matches, 6.3%) from bib.irb.hr/datoteka/478147.Bognar_2009_Uitelj_i_uenici_akcijski_istraivali.pdf [Highlight matches in source](#)

9 matches Citation | Remove marking | Save changes

ostvaruje, budući da se odrađuju „na licu mjesta“ ([6],[8]). Većina suvremenih zanimanja, a posebno učiteljsko, susreće se s kompleksnim zahtjevima prakse u kojoj probleme nije moguće unaprijed uočiti, a još više ponuditi jasna rješenja i rješenja koja se mogu primjeniti. U današnjoj školi učitelji postaju aktivni sudionici u procesu istraživanja. Akcijsko istraživanje predstavlja sistematičan proces promatranja, opisivanja, planiranja, djelovanja, refleksije, evaluacije, modifikiranja, a kao kombinacija akcije i istraživanja može dati zanimljive rezultate za unapređenje kvalitete odgojno – obrazovne prakse ([1],[2],[7]).

Digitalni materijali učiteljima služe za olakšavanje procesa poučavanja dok učenicima pomažu u procesu učenja. Međutim, postavlja se pitanje da li digitalni materijali sami po sebi mogu utjecati na kvalitetu nastave fizike. Rezultati brojnih edukacijskih istraživanja ukazuju na to da moderne tehnologije samo djelomično ispunjavaju očekivanja te se ujedno ističe važnost njihove kompetentne primjene radi osiguravanja plodovitih nastavnih situacija koje olakšavaju učenje [4]. Iako su metodčki aspekti prepoznati kao veoma bitni, oni u najvećem broju izvještaja o empirijskim studijama primjene modernih medija ostaju u sjeni opisivanja same tehnologije i učenličkih stavova u odnosu na moderne tehnologije [5]. Digitalni materijali su efikasni samo u onoj mjeri koliko nastavnici uspijevaju iskoristiti njihove didaktičke potencijale.

Pravi interaktivni eLearning materijali imaju moć preobraziti sadržaj i učenje u smisleno iskustvo za učenika. Takvi materijali omogućavaju učenicima da aktivno obrađuju sadržaj, primjenjuju ga u interaktivnim okruženjima i na taj način duže zadrže usvojeno znanje. Učenje pomoću interaktivnih materijala često je i brže jer korisnik koristi više razine misaonih

Slika 5: Iscrpni izvještaj

Isto tako PlagScan je unutar sadržaja napravio dodatnu selekciju. Dio sadržaja je obojao u crvenu boju i označio je kao *Exact match* ili potpuna podudarnost te *Possibly altered text* to jest mogućnost izmjena nad tekстом. To se može protumačiti kao prepoznavanje teksta koji nije bio potpuno preuzet metodom kopiraj/zalijepi nego je u preuzetim dijelovima bilo umetnuto ponešto vlastitih riječi autora. Znači npravljen su omanje modifikacije izvornog teksta.

REZULTATI DRUGOG TESTA

Uzimajući u obzir broj korištenih izvora za izradu testnog dokumenta te nasumično i spajanje rečenica i segmenata testnog dokumenta, može se reći da su rezultati iznad polovice. Postotak otkrivenog plagiranog sadržaja je 55,5%. Oglašena je visoka razina plagiranog sadržaja označena crvenom bojom. Vrijeme trajanja samog proces detekcije je 6,4 sekunde. U tom vremenu otkriveno je 47 web lokacija s kojima je utvrđena podudarnost sadržaja.

PlagScan Filename: Shake_PsteTestni dokument_test.docx Date: 2016-01-06 15:09
Results of plagiarism analysis from 2016-01-06 15:10

218 matches from 23 sources, of which 23 are online sources.

[Print](#) [Download as PDF](#) [Show comparison](#) [View as list](#) [Word document \(docx\) with annotations](#) [Share](#)

- [0] (23 matches, 24.8%/27.0%) from www.diversalernetnetwork.org/medical/artic...ess_What_Is_It_and_What_Is_The_Treatment [Highlight matches in source](#)
- [1] (24 matches, 23.7%/25.8%) from bucearte.com/article/enfermedad-descompresiva?locale=en [Highlight matches in source](#)
- [2] (15 matches, 15.2%/17.4%) from www.scubaboard.com/forums/archive/index.php/t-221547.html [Highlight matches in source](#)
- [3] (15 matches, 16.8%) from www.traveldoctor.co.uk/dcs.htm [Highlight matches in source](#)

(+ 1 documents with identical matches)

46 matches [Citation](#) [Remove marking](#) [Save changes](#)

Analyzed document

=====1/3=====

Diving has to be planned and carried out in a responsible manner, making sure that first aid equipment and relevant phone numbers are at hand if an accident should take place. Finally, it is important to know the signs of decompression sickness and to be able to give first aid to the diver who has been affected. Beware of diving bells though; they have become more and more popular at tropical beach resorts lately as a method of experiencing nautical adventures without proper training or brains. Diving bells is an obsolete way of diving and should only be utilized by professionals in salvaging operations at great depths, like the Swedish Regal ship Vasa or the German WWII battle cruiser Bismarck. Not for looking at coral at a depth of a measly 9m in some lagoon. It can be dangerous. Don't do it. What is decompression sickness? Decompression sickness, by many referred to as "the bends" or "diving sickness", is triggered when bubbles of nitrogen forms in the blood and tissue of a body. These bubbles follow when a person rises from deep water to shallow water, where the pressure is lower, too quickly. The symptoms of the bends occur almost immediately after a completed dive and can lead to catalepsy or even death in serious cases. If you or a dive-mate exhibits symptoms of decompression sickness, immediately stop the dive, start administering first aid, and call a specialist in divers' medicine for help. The only treatment is to breathe 100 % oxygen on site as well as in transport to treatment in a

Slika 6: Iscrpni izvještaj

Rezultat je podosta dobar ako se uzme u obzir broj pronađenih web lokacija. Znači od 50 korištenih web lokacija, pronađeno je 47. To je i više nego zadovoljavajuće. No postotak detektiranog sadržaja je samo 55.5%. Ali i taj postotak je još vrlo visok ako se uzme u obzir sam sadržaj koji je bio podosta kompleksan. S obzirom na kompleksnost sadržaja stvorenog metodom plagiranja Shake & Paste (protresi/zalijepi) koja predstavlja varijaciju kopiraj/zalijepi plagijata gdje su odlomci organizirani bez nekog logičnog reda i ne postoji jasan prijelaz između odjeljaka, odlomaka i cjelina [6], sve ispod 50% se ne može smatrati plagijatom. Softver nije u stanju sa sigurnošću utvrdi što je plagijat. U međuostalom kao što je gore navedeno PlagScan je dodatak specijaliziran za detekciju Copy&Paste ili kopiraj/zalijepi pojavnog oblika plagijata, a ne Shake & Paste pojavnog oblika plagijata. Unutar teksta, PlagScan je napravio dodatnu analizu. Sadržaj je obojao crvenom, plavom i zelenom bojom. U tekstu je uspio čak prepoznati i dijelove sadržaja koji su bili citirani. Ta kategorija se zove *Marked as quotation* ili označeno kao citat.

Tablica 1: Prikaz rezultata provedenih testova

Rezultati provedenih testiranja za prva dva testa		
Testovi/Značajke	Prvi test	Drugi test
Vrijeme potrebno za detekciju plagijata	2 sekunde	6,4 sekunde
Postotak plagiranog sadržaja	79,6%	55,5%
Broj detektiranih web izvora	9 od 9	47 od 50

Dakle PlagScan je s obzirom na ukupne rezultate, dokazao svoju kvalitetu i izvrsnost detaljnom i iscrpnom analizom, bogatim izvještajem i brzinom provođenja procesa detekcije plagijata i to nad isključivo izvorima na hrvatskom jeziku. Može se reći da je opravdao svoju kvalitetu i reputaciju kvalitetnog softvera.

REZULTATI TREĆEG TESTA

Treći test se sastojao od usporedbe kvalitete uspješnosti rada plugina provodeći proces detekcije sadržaja na engleskom i hrvatskom jeziku.

TESTNI DOKUMENT NA ENGLISKOM JEZIKU

U trajanju procesa detekcije od 6.4 sekunde, otkrivena je visoka razina podudarnosti sadržaja od 94.0%. Oglašena je visoka razina plagiranog sadržaja označena crvenom bojom. Vrijeme trajanja samog proces detekcije je 6,4 sekunde. Utvrđene su dvije razine plagiranja: potpuni plagijat i djelomični plagijat. Dio sadržaja je obojen u crvenu boju i označen kao *Exact match* ili potpuna podudarnost te u zelenu boju kao *Possibly altered text* to jest mogućnost izmjena nad tekstem. To se može protumačiti kao prepoznavanje teksta koji nije bio potpuno preuzet metodom "Copy/Paste" nego je u preuzetim dijelovima bilo umetnuto ponešto vlastitih riječi autora. Unutar teksta, PlagScan je napravio dodatnu analizu. Sadržaj je obojao crvenom, plavom i zelenom bojom. U tekstu je uspio čak prepoznati i dijelove sadržaja koji su bili citirani. Ta kategorija se zove *Marked as quotation* ili označeno kao citat. Broj lociranih podudarnosti internetskim izvorima je 29.

TESTNI DOKUMENT NA HRVATSKOM JEZIKU

U trajanju procesa detekcije od 2 sekunde, otkrivena je visoka razina podudarnosti sadržaja od 94.3%. Broj lociranih podudarnosti internetskim izvorima je 11. PlagScan je unutar sadržaja napravio dodatnu selekciju. Dio sadržaja je obojao u crvenu boju i označio je kao *Exact match* ili potpuna podudarnost te zelenom bojom kao *Possibly altered text* to jest mogućnost izmjena nad tekstem. U tekstu nisu prepoznati dijelovi poput citiranog sadržaja, iako je takvog sadržaja u testnom dokumentu bilo. Dakle nije prepoznata kategorija *Marked as quotation* ili označeno kao citat. To je osnovni nedostatak u usporedbi s prethodnim testom sa sadržajima na engleskom jeziku.

Tablica 2: Usporedba rezultata trećeg testa

Testovi/Značajke	Test na engleskom jeziku	Test na hrvatskom jeziku
Vrijeme potrebno za detekciju plagijata	6,4 sekunde	2 sekunde
Postotak plagiranog sadržaja	94,00%	94,03%
Broj detektiranih web izvora	29	11

ZAKLJUČAK

Novoosmišljeni softver za suzbijanje plagijata na hrvatskom govornom području konačno će dovesti do smanjenja broja plagijata na sveukupnoj domeni obrazovanja. PlagScan je pokazao vrlo visoku kvalitetu korištenja kroz svoje funkcionalnosti koje uistinu mogu biti od velike pomoći školstvu. On bi trebao biti budućnost i kamen temeljac za izgradnju sličnih softvera. No u izradi visoko sofisticiranog softvera za detekciju plagijata postavlja se pitanje plaćanje licence s jedne strane i isplativost korištenja s druge strane. Ako je korisnik nastavnik i ima potrebu detektiranja plagijata zadaća ili seminarskih radova nekoliko puta na godinu, potencijal korištenja dodatka je neiskorišten. Cijena licence dodatka od nekoliko tisuća Dolara na godinu ne podmiruje potencijal korištenja kroz godinu. Taj problem bi trebao postati tema za daljnja razmatranja i planove za izgradnju novih softvera.

POPIS LITERATURE

- [1] Baždarić, K., Pupovac, V., Zulle, L., Petrovečki, M. *Plagiranje kao povreda znanstvene i akademske čestitosti*. Zagreb: 2009 [citirano: 18.01.2017.] Dostupno na <<http://hrcak.srce.hr/38691>>
- [2] Jadrić, M., Ćukušić, M., Lenković, M. *e-učenje, Moodle u praksi*, Split: 2012 [citirano: 18.01.2017.]
- [3] Hercigonja, Z. Plantak Vukovac, D. Detekcija plagijata upotrebom dodatka za Moodle [citirano 18.01.2017.] Dostupno na < https://radovi.cuc.carnet.hr/modules/request.php?module=oc_program&action=summary.php&id=50 >
- [4] Srednja. hr, 2017 [citirano 18.01.2017.] Dostupno na < <http://m.srednja.hr/Studenti/Vijesti/U-Hrvatsku-stigao-lovac-na-plagijate-Nema-vise-prepisivanja-seminarskih-diplomskih-i-doktorskih-radova> >
- [5] PlagScan Plagiarism Checker (2016) [citirano 18.01.2017.] Dostupno na <http://www.plagscan.com/api/PlagScan_Moodle_v1_Manual.pdf>
- [6] Weber-Wulff D, Möller C, Touras J, Zincke E (2013) Plagiarism Detection Software Test 2013, Berlin Dostupno na <http://plagiat.htw-berlin.de/softwareen/test2013/report-2013/> [citirano: 22.08.2017]
- [7] Moodle repozitorij <https://moodle.org/plugins/browse.php?list=category&id=35>

Automatic plagiarism detection with PlagScan Moodle plugin

Abstract: Plagiarism is constantly evolving and it occurs in almost all language areas, taking different kind of forms and shapes. To create a sophisticated plagiarism, the linguistic abilities of a particular language are needed. Today's Internet content, adapts more and more to its users and language areas. English is not any more so prevalent on the web because contents in other languages occurs. As plagiarism increasingly evolves development of new solutions that they can compete with is crucial. One of these temporary solutions is the PlagScan plugin for the learning management system Moodle. PlagScan has among other things an exceptional ability to find plagiarism in the Croatian language. So we're not talking about local search and comparison of plagiarism, but rather finding matches of plagiarized content with the original web content in the Croatian language with extremely low cost.

Keywords: Moodle, PlagScan, Detection, Plagiarism, Plugin

List of figures:

Figure 1: Initial Interface

Figure 2: Plagiarized text levels

Figure 3: Comprehensive analysis

Figure 4: Document Analysis

Figure 5: An exhaustive report

Figure 6: An exhaustive report

List of tables:

Table 1: View the results of the tests performed

Table 2: Comparison of the results of the third test