

# Otkrivanje plagijata upotrebom dodataka u sustavu Moodle

Zoran Hercigonja, univ.bacc.inf.

Dr.sc. Dijana Plantak Vukovac

Sveučilište u Zagrebu, Fakultet organizacije i informatike, Pavlinska 2, Varaždin

zohercigo@foi.hr, dijana.plantak@foi.hr

## **Sažetak**

*Internet kao jedna od aktualnih tehnologija današnjice čini lako dostupnim različite vrste podataka i sadržaja koji se mogu upotrijebiti u procesu obrazovanja i znanstvenog usavršavanja. S jedne strane olakšana je izrada dopunskih nastavnih materijala, učeničkih referata ili znanstvenih radova zbog korištenja dodatnih izvora s interneta, no s druge strane olakšano je neovlašteno preuzimanje sadržaja tj. plagiranje. Upotreba informacijsko-komunikacijskih tehnologija povećala je problem plagiranja i povrede akademske časti, ali istovremeno i omogućila razvoj računalnih programa i mrežnih usluga za otkrivanje plagijata. U ovom radu opisana su dva dodatka za detekciju plagijata u sustavu za e-učenje Moodle te su prikazani rezultati provedenih testova uspješnosti detektiranja plagiranog sadržaja.*

## **Uvod**

Većina autora definira pojam plagijarizma kao preuzimanje tuđih ideja, postupaka ili sadržaja u obliku teksta bez navođenja izvora, bez referenciranja ili parafraziranja i prezentiranje istih kao vlastitog rada. Jedna od definicija govori da je plagiranje neovlašteno korištenje ili vrlo bliska imitacija ideja, jezika, tema i predmeta [1]. Plagiranje kao hotimična ili nehotimična radnja, akt ili čin je vrlo zastupljena tema u sustavu školstva, a ponajviše visokog školstva. Okruženi frazeologijom i gotovim konstruktima, počinjemo određene sadržaje prisvajati, zaboravljajući na originalnost njihovih izvora. Deficit znanja u području pisanja radova u svim sustavima obrazovanja može u najveću ruku biti zaslužan za pojavu plagijata. Najčešći razlozi pribjegavanja plagiranju su: intelektualna nesigurnost, loše upravljanje vremenom, nedostatak jasnih argumenata, nedostatak kritičkih i analitičkih vještina, nedostatno istraživanje, loše preuzimanje bilježaka, loše vještine referenciranja, nerazvijene vještine pisanja te naposljetku varanje [2].

Razvoj i unapređenje informacijsko-komunikacijske tehnologije olakšao je neovlašteno preuzimanje sadržaja ili plagiranje. [3] Prvi ozbiljni pokušaji plagiranja javljaju se 1980. i 1981. godine razvojem tečajeva iz područja računalnih znanosti od strane IBM-a. Tečajevi su bili namijenjeni polaznicima za učenje osnovnih koncepata programiranja. Programska rješenja polaznika temeljila su se na preuzimanju gotovog programskog kôda od drugih polaznika. Tu je zabilježen prvi slučaj plagiranja u elektroničkom učenju.

Naime polaznici koji su redovito dobivali izvrsne rezultate na programskim zadacima, dosljedno su dobivali niske ocjene na kvizovima koji zahtijevaju pisanje kratkih segmenata programa, jer su preuzimali gotove programske konstrukte svojih "starijih kolega" [4] i metodom *copy/paste* ugrađivali u svoje programske zadatke. Tih godina posegnulo se za softverskim rješenjima u obliku programskih dodataka (eng. *Plugins*) za suzbijanje plagijata. Iako su to bili vrlo primitivni (ali uspješni) softveri za traženje sličnosti u uzorcima nad određenim slučajevima programskih kodova poput Jplag-a, koji osigurava predaju velikog broja sadržaja i MOSS-a koji je još bolji u tom pogledu [5], danas su prerasli u vrlo sofisticirana rješenja za detekciju plagijata poput raznih dodataka za sustave za upravljanje učenjem (eng. *Learning management systems, LMS*) ili sustave za elektroničko učenje općenito. Otkrivanje plagijata realizirano u okviru projekta *Detected Cheating On Programming Assignments* bilo je provedeno softverima na lokalnoj razini za pretraživanje radova koji se nalaze u bazi podataka institucije i globalne potrage za *online* tekstom [6]. MOSS softverom, koji je radio po metodi najmanjeg zajedničkog podniza, gdje se u uzorak uzimala jedna linija programskog koda, analizirani su radovi svih polaznika određenog tečaja. Rezultati detekcije plagijata izraženi su u postocima preuzetog sadržaja. MOSS se pokazao pouzdanim jer je već u testnoj verziji postigao da manji broj studenata plagira [7].

Prvi pokušaji doslovnog preuzimanja sadržaja zabilježeni su 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 [8] je sustav za upravljanje učenjem Moodle. Moodle omogućuje spremanje i predaju sadržaja i ujedno služi i kao repozitorij radova studenata, učenika i nastavnika. U ovom radu opisana su dva dodatka za otkrivanje plagijata u sustavu Moodle te je prikazana uspješnost detektiranja plagiranog sadržaja temeljem provedenog testiranja dodataka.

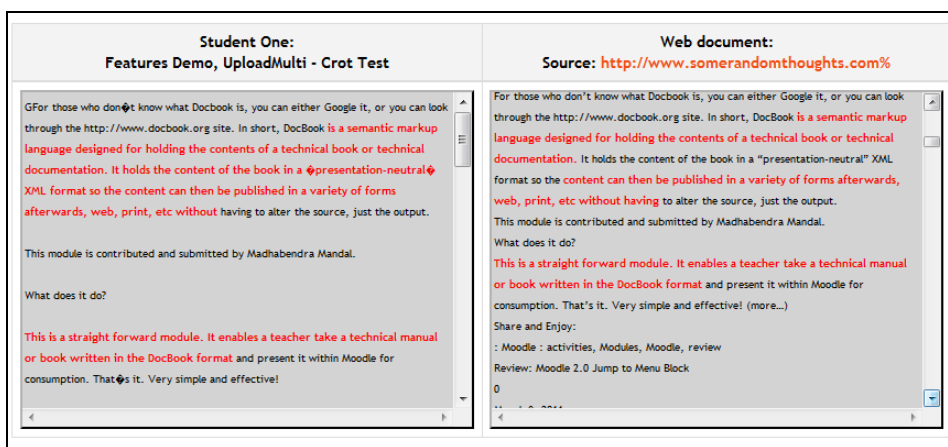
### **Dodaci za Moodle za detekciju plagijata**

Programski dodaci su softverske komponente koje poboljšavaju i proširuju funkcionalnosti druge softverske aplikacije ili sustava i obično se mogu izvoditi samostalno. Dodaci su najčešće dostupni kao besplatni *download* softver s interneta. Većina aplikacija ima mogućnost nadogradnje određenom kategorijom dodataka uključujući i sustav Moodle. Dodaci za Moodle za detekciju plagijata su automatizirani alati kojima je cilj otkriti plagirane sadržaje i u repozitoriju dodataka na web stranici Moodla (*Plugins>Plagiarism*, <https://moodle.org/plugins/browse.php?list=category&id=35>) na raspolaganju je 12 dodataka.

U nastavku ovog poglavlja opisana su dva dodatka za otkrivanje plagijata: besplatan dodatak *Crot Pro* i komercijalan dodatak *VeriCite*.

### ***Crot Pro plagiarism***

*Crot Pro* je besplatan dodatak za detekciju plagijata i nasljednik je dodatka *Crot*. Podržava više formata dokumenata: PDF, doc, docx i txt dokumente. Namijenjen je pretežito za starije verzije Moodle sustava (2.0, 2.1, 2.2, 2.3) [9], ali je kompatibilan i s novijim verzijama, pogotovo s verzijom 2.6. Ukupan broj preuzetih dodataka za 2014. i 2015. godinu je 763 što govori o popularnosti dodatka s obzirom na neke druge dodatke za detektiranje plagijata. *Crot Pro* ne detektira različite razine plagiranog sadržaja, tj. postotak izmjene plagiranog sadržaja u odnosu na original, kao komercijalni dodaci. *Crot Pro* provodi samo uspoređivanje sadržaja koji se nalazi u modulu "Zadaci" (eng. *Assignment*) s web izvorima. Nakon što je provedena detekcija plagijata u navedenom modulu, automatski se generira izvještaj s postotkom ukupno plagiranog sadržaja. Izvještaj se sastoji od dva dijela. Paralelno se uspoređuje plagirani tekst i originalni tekst. Preuzeti sadržaj je označen crvenom bojom koja naglašava uočljivost preuzetog izvornog sadržaja (vidi Sliku 1).



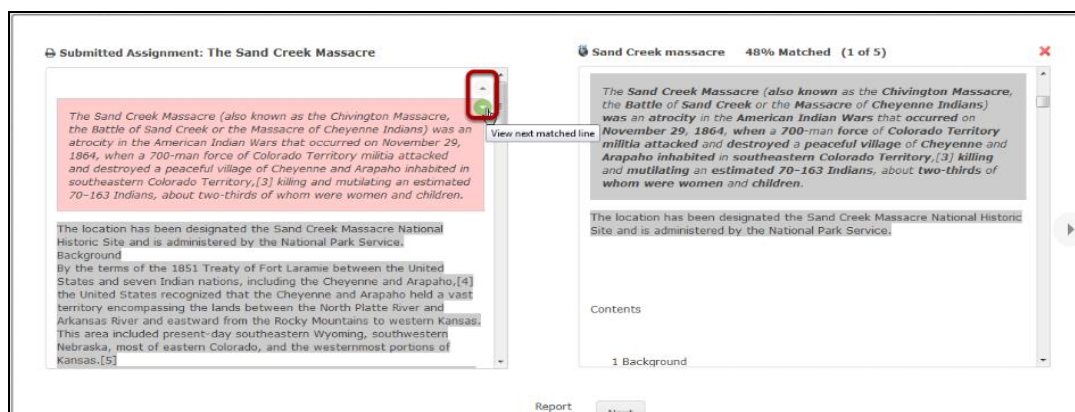
**Slika 1: Izvještaj Crot Pro plagiarsim dodatka**

Kao besplatna inačica dodatka za detekciju plagijata, ovaj dodatak nema iscrpnih izvještaja ni utvrđivanje razina plagiranog sadržaja. Doduše, omogućava označavanje sadržaja za koji je utvrđeno da je potencijalan plagijat. Kao i komercijalni dodaci, ovaj dodatak omogućava poveznice na web izvore s kojih je preuzet sadržaj.

### ***VeriCite plagiarism***

*VeriCite* je dodatak za detekciju plagijata dostupan kao dodatak za sustav Moodle ili sustav Sakai, ali i kao samostalna inačica IMS alata za učenje. Uz spomenutu

interoperabilnost i prilagodljivost platformama za učenje, omogućava detektiranje plagijata nad dokumentima različitih formata (pptx, html, docx, PDF, TXT, RTF, ppt, doc, ODF). Podržava više verzija Moodla: 2.5, 2.6, 2.7, 2.8. [10]. Rad mu se svodi na upotrebu metode traženja podudarnosti u podnizu i podslijedu između predanih dokumenata na sustav Moodle i izvornih dokumenata. Prilikom detekcije plagijata, omogućava generiranje izvještaja koji su dostavljeni odmah nakon što je provedena aktivnost detektiranja plagijata. Izvješća su dinamička što znači da će se rezultati ažurirati pri svakoj novoj predaji dokumenta na sustav Moodle. Svi dostavljeni radovi su indeksirani u šifriranom obliku i pospremljeni u repozitorij. Ovaj dodatak se može koristiti kao probna verzija trideset dana nakon instalacije, a za dugotrajnije korištenje, potrebno je platiti licencu u mjesečnom iznosu od 93.00 \$. U probnoj verziji te licenciranoj verziji *VeriCite* dodatka u proces detekcije plagijata, uključeni su lokalni repozitoriji i baze podataka kojima *VeriCite* ima pristup. Isto tako detaljni i iscrpni izvještaji jednake su kvalitete i kod probne verzije i kod licencirane verzije. Ukupan broj preuzetih dodataka za 2014. i 2015. godinu iznosio je 360 preuzimanja. *VeriCite* zbog osjetljivosti uzoraka prepoznaje i razine plagiranog sadržaja koje označuje bojama. Boje se protežu od tamno plave koja označava originalni tekst u postocima od 0-10% do jarko crvene koja označava visoku stopu preuzetog sadržaja 91-100%. Označavanje teksta bojama pomaže pojedincu u utvrđivanju razine originalnosti pojedinog sadržaja (vidi Sliku 2).



Slika 2: Izvještaj VeriCite plagiarsim dodatka

*VeriCite* je kvalitetan dodatak za detekciju plagijata koji se ubraja u najbolje komercijalne dodatke te sadrži velik broj skladištenog sadržaja u repozitorijima raznih domena.

## Testiranje dodataka za Moodle za otkrivanje plagijata

U ovom poglavlju prikazani su rezultati testiranja dodataka *Crot Pro* i *VeriCite*. Prije testiranja izrađen je testni dokument prema dva kriterija testiranja: tehnička ograničenja

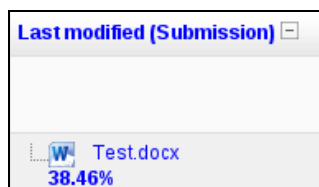
dodataka te odabir teme s web poveznicama na stranice koje su globalno rasprostranjene na internetu. Komercijalan dodatak *VeriCite* uvjetovao je određen broj riječi po stranici dokumenta (najmanje 500, a najviše 2000 riječi po dokumentu). Stoga je kreirani testni dokument sadržavao tri stranice s točno 1410 riječi. Dodatno ograničenje prilikom testiranja, odnosilo se na nemogućnost detekcije plagijata na lokalnoj razini što je bio slučaj s besplatanim dodatkom *Crot Pro*. To znači da ne postoji mogućnost usporedbe i traženja podudarnosti između sadržaja tj. datoteka koje su predane pod aktivnost za predaju "Zadaće" u Moodleu. Stoga je napravljena detekcija plagijata traženjem podudarnosti sa sadržajem iz isključivo web izvora.

Za temu testnog dokumenta odabrana je vrlo pristupačna i rasprostranjena tema na internetu o ronjenju. Prilikom izrade testnog dokumenta trebalo je obratiti pozornost i na korištenje jezika. Naime, izvori koji su konzultirani za sastavljanje testnog dokumenta na hrvatskom jeziku nisu dali nikakvog rezultata prilikom inicijalnog testiranja dodataka. Preuzeti sadržaji nisu bili prepoznati. Stoga je testiranje provedeno na tekstu napisanom na engleskom jeziku.

Testni dokument sadržavao je osnovne informacije o ronjenju, informacije o profesionalnom ronjenju, vrsti i kvaliteti ronilačke opreme, informacijama o rekompresiji vode, liječničkim savjetima vezanim za razna oštećenja i moguće profesionalne deformacije. Rečenice i fragmenti su doslovno preuzimani s web stranica, internacionalnih blogova i chat grupa metodom kopiraj/zalijepi. Sadržaj je odabran nasumično koristeći se poveznicama na sljedeću web lokaciju. Dijelovi sadržaja su uzeti i iz .pdf dokumenata na koje je postojala direktna poveznica s početne web stranice. Za sastavljanje dokumenta korišteno je 50 različitih izvora na engleskom jeziku. Konačni testni dokument s tri stranice sadržaja na engleskom jeziku s ukupno 1410 riječi, spremljen je u .docx format kojeg podržavaju oba Moodle dodatka za detekciju plagijata, kako komercijalni tako i besplatni.

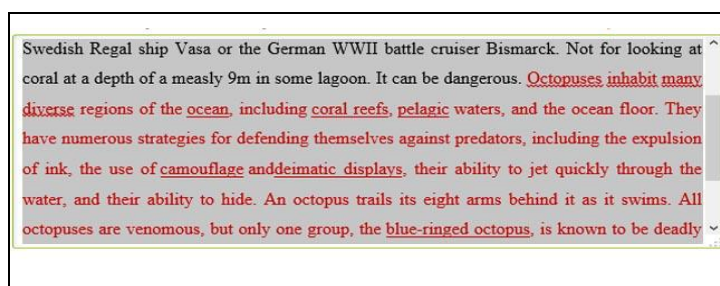
### **Rezultati provedenih testova**

*Crot Pro* kao besplatan dodatak je pokazao iznimno niske rezultate. Postotak plagiranog sadržaja za ovaj testni dokument iznosio je 38.46% s pronađenih 9 izvora preuzetog sadržaja. S obzirom na utrošeno vrijeme procesa detekcije plagijata koje je trajalo čak 24 sata, ovo je razmjerno mali postotak otkrivenih izvornih sadržaja. Isto tako detektirano je samo devet potencijalnih izvora preuzetih sadržaja od korištenih pedeset (vidi Sliku 3).



Slika 3: Crot Pro izvještaj trenutnog stanja

U iscrpnom izvještaju (Slika 4) crvenom su bojom označeni odlomci tekstova u kojima je otkriven izvorni sadržaj. Uspoređujući osjetljivost pronalaženja identičnih uzoraka u odnosu na kasnije testiran *VeriCite* dodatak, može se reći da *Crot Pro* ima malu osjetljivost na uzroke.



Slika 4: Crot Pro iscrpni izvještaj

Kod testiranja dodatka *VeriCite*, zapanjujuća je bila činjenica da je postupak detekcije plagijata trajao vrlo kratko: svega 3 sekunde. U testnom dokumentu *VeriCite* je detektirao 95% plagiranog sadržaja (vidi Sliku 5), te je u procesu detekcije utvrđeno 43 web izvora od ukupno 50.

Overview

Manage Assignments

Filter by: Assignment

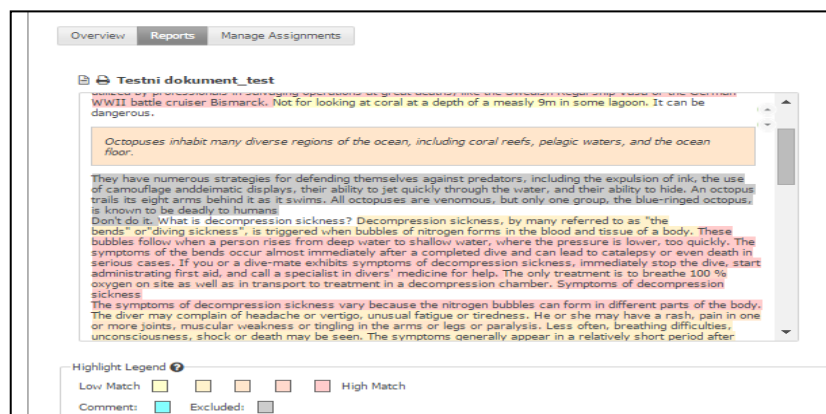
User

Status

| Assignment | User | Status    | Submission           | Date Submitted     | Report Score | Grade                                    |
|------------|------|-----------|----------------------|--------------------|--------------|------------------------------------------|
| None       |      | Submitted | Testni dokument test | 03/21/2015 08:33AM | 95           | <div>View Report</div> <div>Delete</div> |

Slika 5: VeriCite izvještaj trenutnog stanja

Iz detaljnog izvještaja na Slici 6, vidljivo je gradiranje plagiranog sadržaja po razinama. Za primijetiti je da ima dosta žute i crvene boje što podrazumijeva vrlo visoku razinu plagiranja. Crvena je najviša razina s visokim postotkom kopiranog izvornog sadržaja, a žuta podrazumijeva omanje izmjene dodavanjem ili brisanjem pojedinih dijelova izvornog sadržaja. Klikom na pojedini sadržaj označen crvenom ili žutom bojom, s desne strane se otvara izvorna adresa plagiranog sadržaja te postotak izvornosti po rečenici ili odlomku.



Slika 6: VeriCite iscrpni izvještaj

U tablici 1. prikazana je usporedba rezultata testiranja dodataka za detekciju plagijata.

| Rezultati testiranja dodataka za detekciju plagijata |                        |                         |
|------------------------------------------------------|------------------------|-------------------------|
| Dodaci/Značajke                                      | VeriCite               | Crot Pro                |
| Vrijeme potrebno za detekciju plagijata              | 3 sekunde              | 24 sata                 |
| Postotak plagiranog sadržaja                         | 95%                    | 38.46%                  |
| Broj detektiranih web izvora od ukupno 50            | 43                     | 9                       |
| Osjetljivost prepoznavanja uzoraka                   | Visoka na razini retka | Niska na razini odlomka |

Tablica 1: Usporedba rezultata provedenih testova

Prema ovim rezultatima možemo zaključiti da je *VeriCite* daleko kvalitetniji alat za identificiranje plagijata u odnosu na besplatan *Crot Pro* te je postigao puni potencijal svojih funkcionalnosti. Omjer kvalitete izražen cijenom od 93.00\$ potpuno odgovara kvaliteti njegovih funkcionalnosti. S druge strane *Crot Pro* i njegove funkcionalnosti nisu u rang *VeriCite*-a. No, uzimajući u obzir da je *Crot Pro* besplatna inačica, s oko 39% detektiranog plagiranog sadržaja, može se reći da je postigao prihvatljive rezultate.

## Zaključak

Plagiranje je vrlo česta pojava koja se sve više potencira upotrebom informacijsko-komunikacijskih tehnologija. Zbog toga je detekcija plagijata prijeko potrebna i nužna, naročito u sustavu obrazovanja i znanosti. S obzirom na činjenicu da je proces detekcije plagijata kompleksan i iscrpljujući posao, osmišljeni su softveri u obliku programskih dodataka, uključujući dodatke za sustave za elektroničko učenje.

U ovom radu prikazana su i testirana dva dodatka za otkrivanje plagijata u sustavu Moodle, *Crot Pro* i *VeriCite*. Komercijalan dodatak *VeriCite* identificirao je daleko više

plagiranog sadržaja i preuzetih izvora koji su se koristili u testnom dokumentu u odnosu na besplatan dodatak *Crot Pro*. Uzimajući u obzir prethodno opisane tehničke prednosti i opremljenost ovog softvera, *VeriCite* je postigao izvrsne rezultate u otkrivanju plagijata. S druge strane, *Crot Pro* je bio manje uspješan. Iako dodatak *Crot Pro* sadrži nekoliko ograničenja (nepostojanje vlastitih repozitorija i baza podataka te nepostojanje razina razlikovanja plagiranog sadržaja), zbog činjenice da je besplatan, može se reći da je postotak otkrivenog plagiranog sadržaja zadovoljavajuć i opravdavajuć, naročito u situaciji kada ne raspolažemo resursima za kupovinu komercijalnog rješenja.

Ovo istraživanje usmjerava pažnju na delikatan problem akademskog svijeta, plagiranje, te prikazuje mogućnosti korištenja dodataka za sustave za e-učenje u otkrivanju plagijata. Zbog učestale pojave sve sofisticiranijih oblika plagiranja, softvere treba stalno nadograđivati i težiti dubljim razinama detekcije semantičke i sintaktičke razine plagiranog sadržaja.

### Popis literature

1. Hrannabuss S. *Contested texts: issues of plagiarism*, Library Management MCB University Press: 2001 [citirano 05.06.2015. ]
2. The Learning Centre at University of New South Wales. 2015. [citirano 05.06.2015.] Dostupno na < <https://student.unsw.edu.au/how-does-plagiarism-happen>>
3. Baždarić, K., Pupovac, V., Zulle, L., Petrovečki, M. *Plagiranje kao povreda znanstvene i akademske čestitosti*. Zagreb: 2009 [citirano: 05.06.2015.] Dostupno na <<http://hrcak.srce.hr/38691>>
4. Bowyer, K. W., Lawrence, O. H. *Experience Using „MOSS“ to Detected Cheating On Programming Assignments*. Florida: 1999. [citirano 05.06.2015.] Dostupno na <<http://www3.nd.edu/~kwb/nsf-ufe/1110.pdf>>
5. Prechelt, L., Malpohl, G., Phippsen, M. *JPlag: Finding plagiarism among a set of programs*. Njemačka: 2000. [citirano 05.06.2015.] Dostupno na <<http://page.mi.fu-berlin.de/prechelt/Biblio/jplagTR.pdf>>
6. Butakov, S. *Crowdclimbing Approach for Community Based Plagiarism Detection Service*. Kanada: 2014. [citirano 06.06.2015.] Dostupno na <<http://www.conference.org/proceedings/www2014/companion/p1093.pdf>>
7. Zhigang, S., Xiaohong S. *Moodle Plugins for Highly Efficient Programming Course*. Kina: 2012. [citirano 06.06.2015.] Dostupno na <<http://www.slideshare.net/mrc12/moodle-plugins-for-highly-efficient-programming-courses>>
8. Jadrić, M., Ćukušić, M., Lenković, M. (2012) *e-učenje, Moodle u praksi*, Split: 2012 [citirano: 14.08.2015.]
9. Moodle Plugins Library. *Plagiarism: Crot Plagiarism Checker*. 2015. [citirano 05.06.2015.] Dostupno na <[https://moodle.org/plugins/view/plagiarism\\_crot](https://moodle.org/plugins/view/plagiarism_crot)>
10. VeriCite Longsight manual. 2015. [citirano 05.06.2015.] Dostupno na <<https://www.longsight.com/vericite/help/docs/VeriCite-full-guide.pdf>>