

Ana Lavbič, I. gimnazija v Celju

INTERNETNI PLAGIARIZEM PRI POUKU SLOVENSKEGA JEZIKA

1. Kaj je plagiarizem?

Plagiarizem lahko opredelimo kot predstavitev oz. oddajo tujega dela kot lastno delo brez prisotnih citatov ali referenc. V Slovarju slovenskega knjižnega jezika je plagiat tisto, »kar je prepisano, prevzeto od drugod in objavljeno, prikazano kot lastno, navadno v književnosti«, plagiator pa tisti, ki »prepiše, prevzame tuje delo in ga objavi, prikaže kot lastno, navadno v književnosti« [SSKJ 2002]]. Glede na [Hinchliffe 1998] lahko oblike plagiarizma strnemo v naslednje skupine:

- uporaba dela druge osebe, ne da bi ta oseba to vedela,
- kopiranje ali uporaba drugih besed za predstavitev iste ideje brez navedbe virov,
- nakup dela od druge osebe oziroma organizacije ter oddaja le-tega,
- prenos besedila s spletnih mest, kjer so zbrana prosto dostopna gradiva.

Dijaki v srednjih šolah se plagiarizma poslužujejo bodisi namerno ali celo nezavedno, zlasti pa pri pisanju seminarskih in šolskih nalog ali spisov iz književnosti. Pri pripravi na pisanje takšne naloge iščejo različne vire in literaturo, ki bi jim bila v pomoč, tako v tiskani obliki (v knjigah, strokovnih revijah, časopisih) kot tudi v elektronski, pri tem pa besedilo pogostokrat bodisi dobesedno prepišejo bodisi le malenkostno spremenijo, ne da bi pri tem navedli vir informacije. Z uporabo računalnikov in dostopom do svetovnega spleta se je prepisovanje vsebine iz tujega dela zelo razširilo, saj lahko spletni uporabniki prosto kopirajo vsebino iz ogromne množice dokumentov z nekaj enostavnimi kliki kar od doma. Spletni uporabniki lahko z lahkoto prenesejo že pripravljene dokumente ali jih celo pridobijo proti plačilu. Najbolj priljubljeni spletni mesti, kjer lahko dijaki prosto pridobijo gradivo za potrebe opravljanja obveznosti pri pouku slovenskega jezika, sta:





- **Dijaški.net** (<http://dijaski.net>),
- **ARCO** (<http://www.stjost.si/arco>).

O plagiarizmu v Sloveniji in tujini je bilo v zadnjem času veliko govora. Nekaj odmevnih primerov plagiarizma v znanstvenih in strokovnih delih je sprožilo razprave o tem, kaj plagiarizem je, kako ga je mogoče odkriti pa tudi o kazenskih sankcijah zoper plagiatorja [Habjanič 2012: 27–28]. Ker je plagiarizem danes aktualna tematika, nas je zanimalo, ali se ga poslužujejo tudi dijaki pri pouku slovenskega jezika, kako se ta kaže in kakšni so načini ugotavljanja le-tega.

2. Na kaj morajo biti pozorni učitelji?

Za lažje odkrivanje plagiarizma med dijaki morajo biti učitelji pozorni predvsem na naslednje indikatorje:

- Prisotnost nenavadno oblikovanega besedila oziroma oblike, ki ne ustrezajo zahtevam dane naloge (Hinchliffe 1998). (Harris 2012) je predlagal, da je treba biti še posebej pozoren na nenavadne robove, raztegnjene tabele, pomešane stile poglavij ipd. Pomembno je preveriti tudi posebno uporabo majhne oz. velike začetnice, pri neposredni uporabi podatkov s spletnih strani pa smo pozorni tudi na številke strani, datume idr. (Hinchliffe 1998).
- Uporabo različnih oblikovalskih stilov oz. različnih načinov pisanja v okviru odstavka. To je namreč jasen znak, da je bila vsebina kopirana iz različnih virov (Harris 2012). Indikatorji, na katere je treba biti pozoren, so tudi uporabljen žargon in napredna uporaba besedoslovja, tj. raba izrazja, ki je glede na učenčevo znanje prezahtevno.
- Preverjanje uporabljenih referenc v prispevku predvsem z vidika časovne dimenzije. Če so reference npr. starejše od 10 let ali podajajo trditve, ki ne držijo več (npr. da je predsednik Republike Slovenije Janez Drnovšek).
- Skladnost vsebine prispevka z zahtevano nalogo, kjer se preverja predvsem, ali vsebuje vse zahtevane informacije. V primeru, da ne, Harris (2000) opozarja, da obstaja velika verjetnost, da je avtor do vsebine v zadnjem trenutku prišel s pomočjo kopiranja iz drugih virov. Ravno tako je negativen signal prisotnost dodatnih informacij, ki jih naloga sploh ni zahtevala.

3. Napotki za preprečevanje plagiarizma

Strokovnjaki so prepričani, da obstajajo pristopi, ki so pri preprečevanju plagiarizma učinkoviti. Nekaj smernic podajam v naslednjem seznamu:

- Učitelj naj dijakom ponudi možnost izbire alternativnih tematik, ki jih dijaki predlagajo sami in uskladijo z učiteljem. Vnaprej pripravljen seznam tematik dijakom predstavi osnovne smernice, medtem ko jim daje tudi dovolj svobode pri izbiri teme, ki je njim najbolj zanimiva. Vsako leto je smiselno seznam posodobiti, predvsem zaradi tega, da se prepreči ponovna uporaba nalog iz prejšnjih let.

- Prispevek naj bo vnaprej opredeljen z okvirnim številom referenc iz različnih virov. Zahteva je lahko npr. uporaba reference iz dveh knjig, dveh revij in dveh spletnih virov. Dijakom je potrebno predhodno predstaviti ustrezne metode citiranja. Dobra praksa pa je tudi ta, da dijaki poleg svojega dela oddajo tudi fotokopije uporabljenih virov (Harris 2012).
- Učitelj naj sproti preverja napredek skozi vse leto. Za vsako zahtevano nalogo naj opredeli točne datume oddaje osnutkov, vmesnih in končnih izdelkov. Pri tem naj dijakom ne dopušča, da posamezne faze preskočijo, ampak zahteva vse vmesne rezultate. S tem je dijakom otežena uporaba dela nekoga drugega in tudi zmanjšana skušnjava po prepisovanju.
- Dijakom naj predstavi pojem plagiarizma in kako se mu izognemo. Pove jim naj, da je nujno dati v narekovaje vse, kar uporabijo neposredno iz besedila. Besedilo lahko tudi parafrazirajo, vendar morajo pri tem paziti, da ne zamenjajo le tu in tam kakšne besede. Po zaključenem besedilu zapišejo vire in literaturo, ki jih navajajo v besedilu, slikah in preglednicah, in sicer po abecednem vrstnem redu.
- Prav tako jim je smiselno razkriti, da tudi učitelj pozna spletna mesta za potrebe opravljanja obveznosti pri pouku slovenskega jezika, ki jih sami pogosto uporabljajo.

Na spletu obstajajo tudi številne informacijske rešitve, ki omogočajo razkritje plagiarizma. Po navadi so takšne rešitve v obliki storitev, na katere se je treba naročiti. Slaba stran takšnih pristopov je seveda, da v večini niso brezplačni, medtem ko nam pogosto prihranijo veliko truda in nam dokaj samodejno poiščejo potencialne oblike plagiarizma. V okviru našega eksperimenta smo si takšno rešitev zamislili sami in jo tudi podrobneje predstavljamo v nadaljevanju.

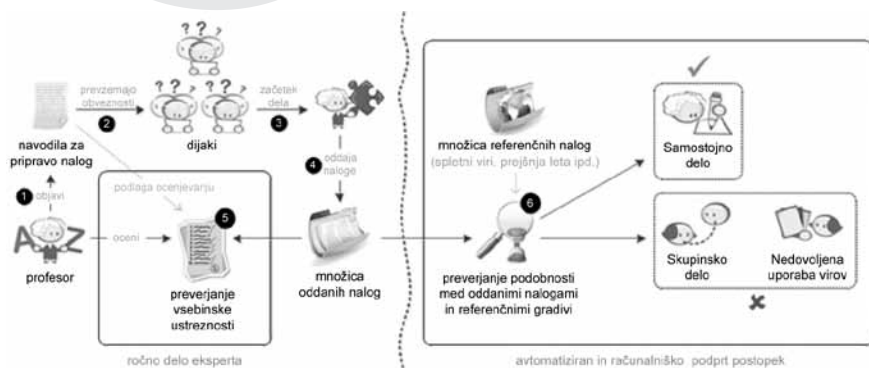
4. Eksperiment

Za potrebe ugotavljanja prisotnosti plagiarizma med dijaki slovenskih srednjih šol je bil izveden eksperiment, katerega namen je bil potrditi in analizirati različne načine plagiarizma, ki se pojavljajo pri pouku slovenskega jezika. Osnovna ideja je bila zelo preprosta (glej sliko 1), tj. pomagati profesorju pri vrednotenju samostojnega dela dijakov iz književnosti pri pouku slovenskega jezika. Izbira problemske domene je bila takšna, da je bilo dijakom na voljo več spletnih virov, kjer je bilo moč pridobiti dodatno gradivo v elektronski obliki. V nadaljevanju bosta na kratko predstavljeni metoda dela in analiza rezultatov.

4.1 Metoda

Pri pouku slovenskega jezika so dijaki treh oddelkov 1. letnika gimnazije dobili nalogo napisati esej z naslovom »Človek – ime ti je nasprotje!« na podlagi Shakespearovega dela Hamlet. Zahteve so bile podane v obliki že prej omenjenega naslova in dolžine vsaj ene strani besedila (glej korak 1 na sliki 1). Dijaki so imeli en teden časa, nalogo pa so lahko opravili doma

(glej korake 2, 3 in 4 na sliki 1). Izdelek je moral biti pripravljen v elektronski obliki in tudi oddan je bil na takšen način v spletni učilnici sistema Moodle.



▼ Slika 1: Postopek ocenjevanja in preverjanja ustreznosti

Za potrebe preverjanja plagiatov med omenjenimi deli smo vsa oddana dela najprej prečistili (odstranili razna oblikovanja in izluščili zgolj tekst) in za analizo uporabili prosto dostopen program WCopyfind [1]. Omenjena programska rešitev (glej korak 6 na sliki 1) nam omogoča primerjavo dokumentov (eseji vseh dijakov, ki so oddani ter besedila, ki so prosto dostopna na spletu) in prikaz podobnosti glede na besede in fraze.

Proces odkrivanja plagiatov zahteva opredelitev dveh množic dokumentov:

- referenčna gradiva, za katere obstaja večja verjetnost, da se bodo uporabili za prepisovanje (npr. dokumenti iz spleta, knjižnice itd.) in
- gradiva, ki jih dijaki oddajo v okviru predpisane obveznosti.

Za referenčna gradiva smo uporabili spletni mesti Dijaški.net in ARCO. S spletnega mesta Dijaški.net smo identificirali 34 potencialnih dokumentov in pri ARCO 12 potencialnih dokumentov, ki bi se lahko uporabili pri prepisovanju. Svojo nalogo je oddalo 86 dijakov. Dejansko oddanih nalog je bilo na koncu 75.

Pri odkrivanju prepisovanj in plagiatov smo temeljili na odkrivanju večbesednih fraz, katerih ujemanje smo iskali v več virih. Za vsako oddano nalogo se je preverilo ujemanje z vsemi viri v:

- množici referenčnih nalog in
- množici preostalih oddanih nalog.

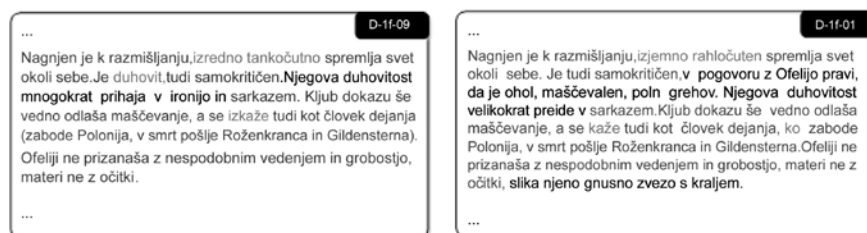
Na takšen način smo lahko preverili, ali je dijak prepisoval iz obstoječih gradiv s svetovnega spleta ali od svojih sošolcev oziroma dijakov, ki so imeli isto obveznost.

Naša analiza je temeljila na naslednjih predpostavkah:

- Pri iskanju ujemanja se upoštevajo fraze dolžine vsaj 6 besed.
- Oddana naloga je bila sumljiva takrat, ko se je pojavila vsaj 1 fraza iz izvirnega dokumenta v vsaj 1 dokumentu referenčnih nalog.
- Pri iskanju ujemanja fraz se je dopuščala nepopolnost stopnje 2, kar pomeni, da sta lahko največ 2 besedi pri ujemanju zapisani tudi nekoliko drugače (npr. „včeraj je prišel domov zelo pozno“ in „včeraj je prišel domov res pozno“).

- Delež v ujemanju fraz je bil nastavljen na 80 %, kar je dopuščalo tudi nekoliko šuma pri preverjanju, v primeru, da so dijaki nekoliko spremenili prepisano besedilo.
- Končna ločila pri ujemanju ne igrajo vloge in se ne upoštevajo.
- Velika začetnica ne vpliva na rezultate ujemanja (npr. beseda Zajec je enaka besedi zajec).

Primer majhnega izseka primerjave dveh podobnih besedil je prikazan na sliki 2. Gre za primer prepisovanja med sošolcema, v katerem je bilo prepisano kar 2/3 celotne obveznosti (glej sliko 3).

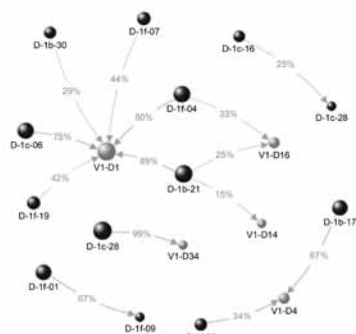


↘ Slika 2: Primer ujemanja besedila

Z rdečo barvo so označeni deli besedila, ki so bili prepisani v celoti, z zeleno barvo pa deli besedila, ki se med seboj nekoliko razlikujejo, vendar še vedno obstaja sum, da so bila prepisana in naknadno malo spremenjena z namenom prikrivanja prepisovanja.

4.2 Rezultati

Izmed vseh oddanih (75) nalog se je pri 11 dijakih izkazalo, da je njihovo delo plagiat. To predstavlja kar 15 % delež vseh oddanih nalog. Pri teh nalogah je bilo moč opaziti od 15 % do 99 % prepisanega besedila iz predhodno omenjenega referenčnega gradiva (glej sliko 3). Na sliki so kot točke grafa prikazani dijaki (zapis v obliki D-**-**) in spletni viri (V*-D**), ki so bili uporabljeni pri prepisovanju. Medsebojna povezava med dvema točkama in smer predstavljata, da je začetna točka prepisovala od končne točke. Oznaka povezave pa predstavlja delež prepisane vsebine. Velikost točke grafa odraža pomembnost (npr. spletni vir V1-D1 je bil uporabljen pri veliko dijakih in dijak D-1c-26 je prepisal skoraj celotno vsebino).



↘ Slika 3: Interakcija dijakov med seboj in z referenčnimi viri



Pri dijakih, ki so prepisovali, lahko identificiramo naslednje skupine (glej sliko 3):

- a) prepisovanje od sošolcev (2 primera)
 - dijaka D-1f-01 in D-1c-16 sta prepisovala od
 - dijakov D-1f-09 in D-1c-28

- b) prepisovanje iz enega izbranega spletnega vira (7 primerov) in
 - dijaka D-1f-08, D-1b-17 sta prepisovala iz vira V1-D4,
 - dijak D-1c-26 je prepisoval iz vira V1-D34 in
 - dijaki D-1f-07, D-1b-30, D-1c-06 in D-1f-19 so prepisovali iz vira V1-D1

- c) prepisovanje iz več izbranih spletnih virov (2 primera).
 - dijak D-1f-04 je prepisoval iz virov V1-D1 in V1-D16,
 - dijak D-1b-21 je prepisoval iz virov V1-D1, V1-D14 in V1-D16.

5. Sklep

Iz predstavljene analize lahko sklepamo, da je lažje prepisovati od doma s spleta kot od sošolcev. To kaže tudi na to, da gre za novodobni plagiarizem, saj se mlajše generacije najraje in najlažje poslužujejo besedil s spleta, ki so lahko dostopna. Ena izmed najbolj pogostih oblik plagiarizma je tako prepisovanje iz enega spletnega vira (točka b zgoraj). Kot prikazuje slika 1 predlagamo, da si učitelj poleg ročnega preverjanja vsebinske ustreznosti pomaga še z najnovejšimi, računalniško podprtimi postopki, ki mu pomagajo pri odkrivanju nepravilnosti pri preverjanju in ocenjevanju določenih obveznosti pri pouku. Predstavljen primer je pokazal, da je pojavov plagiarizma v šoli kar nekaj (v našem primeru okrog 15 %). To težavo lahko identificiramo s poznavanjem stanja na svetovnem spletu in vključitvijo obstoječih, sorodnih virov (oddane naloge dijakov iz prejšnjih let, spletna mesta, kamor dijaki oddajajo svoje delo v skupno rabo ipd.) v analizo ugotavljanja plagiarizma.

POVZETEK

V prispevku smo predstavili vedno večji problem plagiarizma v srednjih šolah in podali nekaj smernic, ki lahko učitelju pomagajo pri zmanjšanju rastočega trenda uporabe tujega dela kot lastnega dela. Predstavili smo učinkovit postopek avtomatiziranega preverjanja prisotnosti prepisovanja iz različnih virov.

↘ Viri in literatura

- Bloomfield, L., WCopyfind. 2012. [uporabljeno 24. 10. 2012]. Dostopno na URL: <http://plagiarism.bloomfieldmedia.com/z-wordpress/software/wcopyfind/>.
- Harris, Robert, 2012. Anti-Plagiarism Strategies for Research Papers. [uporabljeno 2012-10-24]. Dostopno na URL: <http://www.virtualsalt.com/antiplag.htm>.
- Hinchliffe, Lisa. 1998. Cut-and-paste plagiarism: preventing, detecting and tracking online plagiarism. [uporabljeno 24. 10. 2012]. Dostopno na URL: <http://alexia.lis.uiuc.edu/~janicke/plagiary.htm>.
- Inštitut za slovenski jezik Frana Ramovša, Slovar slovenskega knjižnega jezika. 2002. [uporabljeno 24. 10. 2012]. Dostopno na URL: <http://bos.zrc-sazu.si/sskj.html>.
- Habjanič, Sandra, 2012: Plagiarizem v diplomskih delih iz področja družboslovja v Sloveniji: diplomsko delo. Maribor: Univerza v Mariboru. Filozofska fakulteta.