

## Ako čítať a interpretovať „Protokol o kontrole originality“

Protokol o kontrole originality (ďalej len protokol) je výstupom zo „Systému na odhaľovanie plagiátov pre potreby slovenských akademických a výskumných inštitúcií“ (ďalej ANTIPLAG), ktorý beží nad Centrálnym registrom záverečných a kvalifikačných prác (ďalej CRZP). Protokol je elektronický dokument generovaný algoritmom na odhaľovanie plagiátov. Úlohou tohto dokumentu je vysvetliť jednotlivé časti protokolu.

Systém ANTIPLAG je systém pre podporu rozhodovania. Protokol, ktorý je jeho výstupom:

- nie je potvrdením, že testovaná práca je ORIGINAL (nemôže odhaliť kopírovanie myšlienky a ani zdrojov ktoré nemá k dispozícii a v indexe)
- nie je dôkazom, že testovaná práca je PLAGIÁTOM (napriek vysokému percentuálnemu prekryvu textu testovanej práce môže ísť o text správne citovaný alebo preberaný v súlade so zámerom práce)

Protokol má byť pomôckou. Upozorniť na dokumenty ktoré mohli uniknúť pozornosti školiteľa alebo oponenta.

### Všeobecné údaje protokolu

Strany protokolu sú očíslované. Protokol je chránený heslom proti manipulácii. V päte dokumentu je jednoznačný identifikátor protokolu (PID) vo forme textu aj čiarového kódu (Code39), dátum vytvorenia protokolu a identifikáciu portálu.

### Záhlavie protokolu

V záhlaví protokolu je identifikácia testovanej práce. Vychádza z dát dodaných lokálnym úložiskom Záverečných a kvalifikačných prác (ZP) v procese výmeny dát. Zároveň obsahuje údaj o globálnom percente prekryvu textu (označené ako GP). Globálne percento podobnosti nie je jednoduchým súčtom podobností testovaného dokumentu s jednotlivými dokumentami, ktoré boli vyhodnotené ako zaujímavé z pohľadu množstva podobností. Dôvodom je fakt, že časť textov môže byť podobná viacerým dokumentom (rovnaký odsek cituje alebo kopíruje viac prác). Celkové percento podobnosti je teda menšie, alebo rovné súčtu parciálnych podobností v jednotlivých dokumentoch ( $P^1, P^2, \dots, P^N$ ).

### Zoznam dokumentov

Každá práca je porovnávaná voči indexu systému ANTIPLAG. Zo všetkých dokumentov sú vybrané tie, v ktorých sa nachádza nadprahové množstvo podobného textu. Tabuľka „Práce s nadprahovou hodnotou podobnosti“ obsahuje identifikáciu nájdených dokumentov spolu s údajom o percente textového pokrytia. Tento údaj je podielom súčtu dĺžok nájdených odsekov textu a celkovej dĺžky textu testovaného dokumentu (v príklade označené ako P1 a P2).

### Podobný text – podobný odsek

Systém je schopný hľadať podobnosti už pre sekvencie slov od dĺžky troch plnovýznamových slov bez ohľadu na ich poradie. V praxi sú samozrejme zaujímavé dlhšie sekvencie slov a viet. V tomto smere je systém parametrizovateľný a tieto parametre sa môžu upresňovať pre rôzne typy dokumentov.

System neoznačuje len texty, ktoré sa vyznačujú absolútnou zhodou. Je schopný riešiť situácie, keď sú slová vo vetách zamenené v poradí, vety sú doplnené o doplnkové slová a podobne.

### **Detaily – zistené podobnosti**

Z praktických dôvodov (aby protokoly nemali 100 strán), ale aj ideových dôvodov (je potrebné identifikovať a prezentovať ako problém jedinú izolovanú vetu ?) sa definuje minimálny počet plnovýznamových slov, ktoré sa musia nachádzať v odseku označenom ako podozrivý. Navyše tieto slová musia na „ploche“ tohto odseku tvoriť nadprahové percento textu odseku. Hovoríme o percente spoľahlivosti odseku a toto číslo je uvádzané pod poradovým číslom odseku.

Text citovaný v protokole je text, ktorý sa nachádza v testovanom dokumente (po prevode na plain text). V nájdenom dokumente by sa mal nachádzať text ktorý je tomuto odseku podobný (slová môžu mať iné poradie, niektoré môžu chýbať, niektoré môžu byť naopak pridané, atď...).

### **Obmedzenia**

- systém je schopný nájsť sekvencie slov, ktoré sú kratšie ako nastavená prahová úroveň, ale neprezentuje ich
- v niektorých prácach sa ako podozrivé vyhodnotia podobné dlhé sekvencie patriace do hlavičky práce alebo do literatúry – tento jav by bolo možné redukovať odsekávaním začiatkov a koncov dokumentov – v tejto fáze projektu sme k tomuto kroku zatiaľ nepristúpili
- slovník plnovýznamových slov bude postupne upravovaný tak aby vylučoval

### **Vysvetlivky technických termínov používaných v texte**

- **Index** je dátová štruktúra ktorá vzniká spracovaním dodaných dokumentov a vzniká za účelom rýchleho porovnania dokumentov.
- **Plain text** vyniká prevodom originálneho dokumentu na text očistený od značiek, formátovania a ďalších atribútov textu pre porovnanie nevhodných

### **Spätná väzba**

Budeme vďační za pripomienky, návrhy vylepšení, prípadne ohlásenie problémov s protokolom. Tieto informácie prosíme doručovať na adresu [grman@svop.sk](mailto:grman@svop.sk).

Za tvorivú skupinu ANTIPLAG

Mgr. Ján Grman, PhD.  
SVOP spol. s r.o.

26.4.2010, verzia 1.0

## Protokol o kontrole originality

### Kontrolovaná práca

Citácia	Percento*	
BIBLIOGRAFICKÁ IDENTIFIKÁCIA TESTOVANÉHO DOKUMENTU plagID: 766	<b>13,61%</b>	<< GP

\*Číslo vyjadruje percentuálny podiel textu, ktorý má prekryv s indexom prác korpusu CRZP.

### Práce s nadprahovou hodnotou podobnosti

Dok.	Citácia	Percento*	
1	BIBLIO.IDENTIFIKÁCIA NÁJDENÉHO DOKUMENTU č.1	2,44%	<< P1
2	BIBLIO.IDENTIFIKÁCIA NÁJDENÉHO DOKUMENTU č.2	13,08%	<< P2

\*Číslo vyjadruje percentuálny prekryv testovaného dokumentu len s dokumentom uvedenom v príslušnom riadku.

Dátum vygenerovania protokolu



25.04.2010

Identifikátor protokolu a dokumentu v CRZP  
(aj vo forme čiarového kódu)



2391A9FE66FD4429A227FA21BF0B6BC2

- 1 -

www.crzp.sk

ukážka textu (môže byť krátky)  
s vyznačením zhodných  
pínavýznamových slov

**Details - zistené podobnosti**

Odsek	Citácia
<b>BIBLIOGRAFICKÁ IDENTIFIKÁCIA NÁJDENÉHO DOKUMENTU</b>	
plagiD: 290	
1 80%	činu podvodu sa vyžaduje, aby bolo preukázané, že páchateľ v čase, keď sa zmluvný záväzok uzatváral, konal s úmyslom, že záväzok v dohodnutom čase nespĺní, alebo ho nebude môcť splniť a tým veriteľa uvádza do omylu, aby sa obohatil na jeho škodu. Majetkom sa rozumie všetky majetkové hodnoty, ktoré tvoria riaden vecí, ale aj pohľadávky, iné práva, peniaze a iné oceňiteľné hodnoty. Časťim majetkom sa rozumie majetok, ktorý nepatrí páchateľovi alebo nepatrí výlučne len jemu. Škodou na cudzom majetku sa rozumie len ujma majetkovej povahy. Obohatením seba alebo iného sa rozumie neoprávnené rozumnoženie majetku páchateľa alebo inej osoby, či už rozšírením majetku, alebo ušetrením nákladov, ktoré by inak boli z majetku páchateľa alebo inej osoby vynaložené. Ide o neoprávnené obohatenie. Obohatenie sa nemusí rovnáť spôsobená škoda. Uvedením do omylu je konanie, ktorým páchateľ predstiera okolnosti, ktoré nie sú v súlade so skutočným stavom vecí. Uvedenie do omylu môže byť spôsobené nielen konaním, ale aj opomenutím konania, alebo aj konštudentným činom. Vyššie omylu spočíva v tom, že páchateľ sám k vyvolaniu omylu nepripel, ale po zistení omylu iného a v príčinom vzťahu k nemu koná tak, aby na škodu cudzieho majetku seba alebo iného obohatil. Subjektívnu
2 76%	konanie, ktorým páchateľ predstiera okolnosti, ktoré nie sú v súlade so skutočným stavom vecí. Uvedenie do omylu sa môže spáchať nielen konaním ale aj opomenutím alebo aj konštudentným činom. - Pri využití omylu iného páchateľa sám na vyvolanie omylu ne pripel, ale po zistení omylu niekoho iného a v príčinom vzťahu k nemu koná tak, aby na škodu cudzieho majetku seba alebo niekoho iného obohatil. - Osobou, ktorú páchateľ pri podvode
3 84%	činu podvodu vo vzťahu špeciálny (lex specialis). Úmyslom zákonodarcu bolo umožniť šírku diferenciácie podvodných konaní a zabezpečiť ich preventívny účinok. Trestný čin úverového podvodu pozostáva z dvoch samostatných skutkových podstát: 1. páchateľ vyhlásil od iného úver alebo zabezpečenie úveru tým, že ho uvedie do omylu, pokiaľ ide o splnenie podmienok na poskytnutie úveru alebo na spĺnanie úveru, a tak mu spôsobí malú škodu (§ 222 ods. 1 Trestného zákona). 2. páchateľ ako zamestnanec, člen, štatutárny orgán, zástupca alebo iná osoba oprávnená konať za toho, kto úver poskytuje, umožní získať úver tomu, o kom vie, že nespĺňa podmienky určené na jeho poskytnutie, v tomto prípade sa vyžaduje špeciálny subjekt (§ 222 ods. 2 Trestného zákona). Vyhlásenie úveru sa týka predovšetkým bankových úverov. Úver sa poskytuje na základe zmluvy o úvere, ktorou sa veriteľ zaväzuje, že na požiadanie dlžníka poskytne v jeho prospech peňažné prostriedky do určitej sumy, dlžník sa zaväzuje peňažné prostriedky vrátiť a zaplatiť úroky. Jednotlivé komerčné banky majú podmienky na poskytovanie úverov kontrolované pre svoje potreby v obchodných podmienkach. Na trestnosť páchateľa podľa odseku 1 sa vyžaduje, aby páchateľ 1. iného uviedol do omylu v otázke splnenia podmienok na poskytnutie úveru a 2. spôsobil tým malú škodu. Pri úverovom podvode škoda nevzniká už vyhlásením úveru, ak sa úver riadne spláca, ale až nesplácaním splátok lebo uvedením kritérií musia byť splnené súčasne, t. j. vyhlásenie úveru, a tým spôsobenie malej škody. Trestný
4 87%	špeciálne k trestnému činu podvodu. Pozostáva z dvoch samostatných skutkových podstát: 1. páchateľ vyhlásil od iného poisťné plnenie tým, že ho uvedie do omylu v otázke splnenia podmienok na jeho poskytnutie, a tak mu spôsobí malú škodu (§ 223 ods. 1 Trestného zákona). 2. páchateľ ako zamestnanec, člen, zástupca alebo iná osoba oprávnená konať za toho, kto poisťné plnenie poskytuje, napomáha získať poisťné plnenie, tomu, o kom vie, že nespĺňa podmienky na jeho poskytnutie t. j. špeciálny subjekt (§ 223 ods. 2 Trestného zákona). Poisťovací podvod sa vzťahuje na nároky vyplývajúce z poisťenia, ku ktorým dochádza: - uzavretím poisťnej zmluvy, napríklad o úrazovom poisťení, o havarijnom poisťení motorového vo zidla, o poisťení zodpovednosti za škody spôsobené výkonom povolenia a pod., - na základe zákonného poisťenia. Osobitný prívrny predpis môže fyzickej osobe a právnickej osobe uložiť povinnosť uzavrieť poisťnú zmluvu. Napríklad povinnosť uzavrieť poisťnú zmluvu má pri tuzemskom motorovom vozidle len, kto je ako držiteľ motorového vozidla zapísaný v dokladoch
5 86%	tom, že páchateľ v súvislosti s ponukou, predajom alebo rozširovaním cenných papierov alebo iných listín, ktoré sľubujú účasť na majetkových výnosoch podniku, alebo v súvislosti s ponukou zvýšiť výnosy takého investovania v prospektoch alebo v iných propagačných materiáloch alebo prehľadoch týkajúcich sa majetkových pomerov alebo výnosov podniku vo vzťahu k väčšiemu počtu osôb uvádza nepravdivé údaje alebo nereálne údaje o výnosoch investovania alebo o majetkových pomeroch podniku, do ktorého sa má investovať, alebo kto nevýhody takého investovania zamlčí". Predpisuje,



2391A9FE66FD4429A227FA21BF0B68C2