

Antiplagiátorský systém (APS) kontroluje originalitu prác uložené v Centrálnom registri záverečných a kvalifikačných prác (CRZP). APS indexuje práce uložené v CRZP a navyše zdroje z voľne dostupného internetu s dôrazom na zdroje v slovenskom jazyku.

Samotný algoritmus nie je závislý na jazyku. To sa prejavuje aj aktuálne na výsledkoch testovania prác v anglickom, maďarskom, ale aj ukrajinskom jazyku. Pre vysokú spoľahlivosť a presnosť detekcie podobnosti v konkrétnom jazyku je potrebné urobiť aj množstvo nastavení a vytvoriť špecifické, jazykovo závislé dátové štruktúry.

Aktuálne hlavným zámerom APS je skvalitňovanie procesu vzdelávania. V APS je však skrytý aj potenciál nástroja podporujúceho skvalitňovanie procesov vo vede a výskume. Tento potenciál je potrebné využiť a rozšíriť záber CRZP a APS aj na dokumenty neohraničené dikciou vysokoškolského zákona (záverečné a kvalifikačné práce) a pojať tam dokumenty vypracované celou vedeckou komunitou na Slovensku.

Stanovený cieľ je možné dosiahnuť v týchto hlavných krokoch:

- Vypracovanie štúdie
 - ❖ o aktuálnom stave spracovania záverečných a kvalifikačných prác na SAV s cieľmi
 1. analýza aktuálneho stavu úložísk
 2. návrh na unifikáciu úložísk
 3. návrh na zapojenie úložísk do CRZP
 - ❖ o aktuálnom stave publikačnej a edičnej činnosti SAV a akademických pracovísk v SR s cieľmi
 1. analýza možností poskytovať služby APS pre účely kontrolu publikačnej činnosti (s garanciou nezverejňovania)
 2. realizovateľnosť rozšírenia CRZP a APS o práce z publikačnej a edičnej činnosti SAV a akademických pracovísk v SR
 - ❖ o legislatívnych aspektoch predkladaného riešenia
 - ❖ o potrebe analytických a programátorských prác
 - ❖ o potrebe finančných zdrojov
 - ❖ o potrebe ľudských zdrojov
 - ❖ o časovej náročnosti
- Implementácia modifikovaného CRZP a APS a jeho pomocných systémov
- Vypracovanie dokumentácie
- Pilotná prevádzka modifikovaného CRZP a APS a jeho pomocných systémov

- Rutinná prevádzka modifikovaného CRZP a APS a jeho pomocných systémov

Hlbším ponorom do problematiky detekcie plagiátov sme zistili, že kvalitu procesu detekcie dokážu pozitívne ovplyvniť:

- Implementácia slovníka synonym slov slovenského jazyka do tezauru APS
- Implementácia koreňových morfém slovenského jazyka do tezauru APS
- Implementácia „stop slov“ slovenského jazyka do tezauru APS
- Implementácia „čistenia“ korpusu slov slovenského jazyka (vyraďovanie nesprávnych, skomolených slov)
- Implementácia slovenského WordNetu
- Budovanie korpusu anglického jazyka

Rutinná prevádzka APS s rozšírenými detekčnými schopnosťami nad záverečnými a kvalifikačnými prácami a nad prácami publikačnej a edičnej činnosti SAV a akademických pracovísk v SR by mohla byť cieľovým stavom v období pre najbližšie roky.

V úlohách rozšírenia rozsahu testovaných dokumentov je možné postupovať dvomi smermi. Prvým smerom je rozširovania záberu CRZP o dokumenty VaV komunity, ktoré v CRZP zatiaľ nie sú podchytené. Druhým smerom je rozširovanie indexovania internetu aj o špecifické alebo bežne nedostupné elektronické zdroje.

V oblasti rozšírenia CRZP je možné uvažovať o týchto okruhoch dokumentov:

- Zapojenie ústavov SAV do procesu bežného testovania záverečných a kvalifikačných prác ktoré sú produkované na pracoviskách SAV
 - vyžaduje analýzu existujúceho stavu s evidenciou prác na SAV
 - môže vyžadovať vybudovanie úložiska EZP na SAV
- Umožniť VŠ a SAV dobrovoľné testovanie elektronických dokumentov (iných ako záverečné a kvalifikačné práce) ktoré vznikajú na ich pôde, s cieľom zvýšiť ich ochranu a kvalitu

Ak hovoríme o dokumentoch ktoré vznikajú ako výsledky vedeckej činnosti, máme na mysli predovšetkým tieto typy:

- dokumenty ktoré sú výstupom publikačnej činnosti (články v zborníkoch a časopisoch, ...)
- dokumenty ktoré sú výstupom edičnej činnosti (učebnice, skriptá, ...)

Špecificky zaujímavou oblasťou použitia APS je jeho zapojenie do procesu recenzovania publikácií.

V tejto súvislosti by bolo možné APS ponúknuť na použitie vydavateľom časopisov a zborníkov a výmenou získavať plné texty. Hlavné prínosy spočívajú:

- v rozšírení možností recenzenta pre napísanie objektívnej recenzie
- v ochrane autorov ktorých príspevky sú po indexovaní chránené nielen právne ale aj fakticky, pretože sa dostávajú do porovnávania so všetkými dokumentmi, ktoré prídu do APS po vytlačení daného článku
- systémy APS, najmä jeho prezentačné časti je tiež možné použiť aj pre zverejňovanie publikácií autorov ktorí by prejavili záujem a publikácie nie sú nijak licenčne viazané a tým zabezpečiť ich vyššiu dostupnosť a citovanosť vo vedeckej komunite.

Analýza a implementácia slovníkov slovenského jazyka je dôležitá pre účely zvyšovania spoľahlivosti detekcie podobných textov. Sťažuje mechanické preberanie cudzích myšlienok do textu. Budovanie systémov špeciálnych slov musí byť systematické aby pokrylo čo najväčšiu slovnú zásobu jazyka. Z korpusu dokumentov CRZP boli vypracované zoznamy stop slov pre SJ a AJ. Dokumenty, ktoré vstupujú do APS obsahujú rôzne slová, mnohé preklepy alebo sekvencie písmen bez významu. Strom slov je z týchto dôvodov veľmi dôležitý čistiť. Zároveň je možné využiť významové závislosti slov. Okrem krátkych a často používaných slov ktoré nemajú reálny vplyv na výsledky zisťovania podobností (tzv. stop slová – alebo slovník neindexovaných slov) je možné využiť slovníky morfém a synonym. V tejto súvislosti je možné definovať tieto čiastkové úlohy:

- Analýza a implementácia konverzného algoritmu pre prevod slovníka morfém do APS
- Analýza a implementácia konverzného algoritmu pre prevod slovníka synonym do APS
- Analýza a implementácia konverzného algoritmu pre prevod slovníka WordNet pre AJ do APS
- Práce ručného dočistenia stromu slov a získanie relevantného WordNETu pre SJ
- Implementácia notifikačného mechanizmu pre čistenie „nových slov“ ktoré sú indexované už po procese prvotného očistenia slovníka

Prípady plagiátorstva sa nevyhýbajú žiadnej vrstve spoločnosti a dnes ešte stále nie sú dostatočne pranierované a postihované. Ani prípady plagiátorstva medzi politikmi či poslancami nie sú výnimkou. Potrebujeme implementáciu etických štandardov dostať na podstatne vyššiu úroveň. Veľmi dôležitou oblasťou je aj naďalej zvyšovanie povedomia o etike, plagiátorstve, o autorských právach, o duševnom vlastníctve. Významnú úlohu pri zvyšovaní povedomia musí zohrať rodina, škola, pracovisko a všetky druhy médií. Nie sme úplne na začiatku, deti, mládež a dospelí si plne uvedomujú, čo je krádež (hmotná). Potrebujeme im však dostať do povedomia aj krádež iného druhu – krádež ideí, textov, obrázkov, ... a oboznámiť ich so základnými etickými princípmi. Nebude to ľahká úloha, pretože v dnešnom internetovom svete prevláda u detí a mládeže a časti dospelaj populácie názor, že čo je na internete to je aj moje. Na tomto poli je ešte potrebné vykonať veľa práce.

K zvýšeniu povedomia o plagiátorstve prispelo už samotné oznámenie o tom, že všetky vysoké školy na Slovensku budú povinne využívať systém na odhaľovanie plagiátov. Až po tomto oznámení začal vysoko významne stúpať počet vyhľadávaní tohto pojmu na internete a tým sa verejnosť – najmä študentská – bližšie oboznamovala s významom tohto pojmu. Okrem internetu stúpol aj počet článkov v médiách (aj printových) o tomto pojme, ale nie až tak výrazne.

Existujú však skupiny ľudí, ktorí nehľadajú poctivé cesty k získaniu vysokoškolských titulov a aj skupiny, ktoré poskytujú služby týmto skupinám. Nie je zložité za niekoľko minút nájsť niekoľko internetových stránok poskytujúcich vypracovanie podkladov k rôznym druhom prác.

Každá vysoká škola má pripravenú vlastnú smernicu týkajúcu sa záverečných a kvalifikačných prác, ktorú si pripravila na základe všeobecnej smernice, ktorú vydalo MŠVVaŠ SR. Dodržiavanie smerníc týkajúcich sa postupov pri príprave, odovzdávaní a vyhodnocovaní protokolov o kontrole originality prispieva k zvýšeniu kvality odovzdávaných prác.

Od 1.9.2011 sú práce prístupné na portáli CRZP (tie, ktoré došli do CRZP po 31.8.2011) a sú teda nielen pod kontrolou vysokých škôl, ale aj pod kontrolou verejnosti.