

Ikaros.cz 26. 6. 2009

Kontrola plagiátů v seminárních pracích prostřednictvím Odevzdej.cz

Autoři:

- Stančík, Martin, Pekárková, Lucie, Mikulášová, Zuzana, Krhutová, Růžena, Brandejsová, Jitka, Brandejs, Michal

Klíčová slova:

- seminární práce, plagiátorství, Masarykova univerzita

Rubriky:

- Informační zdroje, služby, aplikace

English title:

Searching for Traces of Plagiarism in Seminar Papers via Odevzdej.cz

Abstract:

Plagiarism and its prevention have recently become topical not only in connection with theses, but also seminar and other papers that students are supposed to write during their studies. The development team of IT specialists working for the Faculty of Informatics, Masaryk University, headed by the associate professor Michal Brandejs, met multiple requests on the part of teachers working for various universities for a system capable of searching for traces of plagiarism in seminar papers. The Odevzdej.cz project (<http://odevzdej.cz/>) has been joined by ten public universities. It represents a follow-up to a similar joint project of National Registry of Theses named Theses.cz (<http://theses.cz/>) (successfully completed in 2008), which now involves 23 universities from Czech Republic and Slovakia.

Cíle a možnosti systému Odevzdej.cz

Na mnohých školách, například v ekonomických nebo humanitních oborech, jsou seminární práce důležitou součástí studia. Současně na internetu roste počet serverů, které shromažďují seminární a školní práce a umožňují jejich opisování. Proto si projekt Odevzdej.cz klade za cíl nejen snižovat opisování mezi seminárními, závěrečnými pracemi a jinými dokumenty, ale také co nejvíce eliminovat smysl uvedených serverů.

Obr. 1: Titulní stránka projektu Odevzdej.cz

Odevzdej.cz umožňuje učitelům zřizovat pro účely odevzdání seminárních prací tzv. odevzdáárny, kam mohou studenti učitelům vkládat práce podle pokynů pedagoga. Jde o e-learningový model pro odevzdávání prací, který přináší pedagogům i mnohé jiné výhody než jen vyhledávání podobných souborů (podezřelých plagiátů). Například učitel neztrácí čas osobním sbíráním prací, nemusí si nechávat práce zasílat e-mailem jako přílohy, neboť

stahování velkých souborů, je-li nízká přenosová rychlost připojení nebo omezená kapacita poštovní schránky, může způsobovat problémy. Učitel není nucen vytvářet si vlastní systém adresářových složek na svém lokálním počítači tak, aby se v narůstajícím počtu souborů vyznal. Systém nahradí učiteli práci s archivací či zálohováním pro případ, že by se poškodil lokální počítač. Učitelé odpadá sledování v kalendáři, do kdy kterým studentům zadal termín odevzdání prací, a nastaví přesné datum pro období odevzdání práce (od – do). Práce lze také učitelem ohodnotit a studenti si mohou hodnocení přečíst. Každý učitel si může založit hned několik odevzdávacích podle počtu svých předmětů nebo skupin studentů a sdílet je podle potřeby s dalšími svými pedagogickými kolegy. Po zpracování vložené práce může vyučující nechat práci kdykoliv zkontrolovat systémem na odhalování podobností, který ji porovná s ostatními vloženými pracemi včetně závěrečných prací uložených v Theses.cz (<http://theses.cz/>) a s dalšími evidovanými zdroji dat (např. Wikipedie).

Mezi deset veřejných vysokých škol, které se v letošním roce zúčastňují rozvojového projektu, patří Česká zemědělská univerzita v Praze, Jihočeská univerzita v Českých Budějovicích, Masarykova univerzita, Ostravská univerzita v Ostravě, Univerzita Jana Evangelisty Purkyně v Ústí nad Labem, Univerzita Karlova v Praze, Univerzita Palackého v Olomouci, Vysoká škola ekonomická v Praze, Vysoká škola polytechnická Jihlava a Vysoká škola technická a ekonomická v Českých Budějovicích. Kromě těchto projevily zájem o využívání systému i další střední a základní školy. Školy mají dvě možnosti, jak využívat tento systém. Buď mohou ponechat vkládání prací na učitelé a ten si je vloží sám (například dokumenty sbalené zipem), nebo zadá studentům, kam mají do systému práce vkládat. Anebo mohou školy přenášet do systému Odevzdej.cz seminární práce ze svých lokálních úložišť automatizovaně. Sběr prací může probíhat centralizovaně na celouniverzitní úrovni nebo jen na úrovni jednotlivých pracovišť.

Obr. 2: Neautentizované vložení souboru do odevzdávacího systému

Obr. 3: Odevzdávací systém se slovně ohodnocenými pracemi

Princip volného přístupu a kontroly dokumentu

K tomu, aby byl nakonec nástroj pro odhalování plagiátů dán k dispozici i veřejnosti, inspirovaly vývojový tým zkušenosti s více než jednoletým provozem systému Theses.cz. Na pracovníky týmu z Masarykovy univerzity se obracelo stále více a více lidí z řad veřejnosti, kteří žádali porovnání své práce systémem, a nyní tak mohou učinit sami.

Jedinečnost takového volného přístupu k systému spočívá v tom, že kdokoli má možnost zdarma zjistit, zda je vložený dokument podobný jinému dokumentu v systému, a přispět tak k vyšší ochraně autorských práv. Stačí, když do systému Odevzdej.cz vloží soubor nebo odkaz z internetu. Systém data analyzuje a po určité době pošle výsledky na uživatelem zadanou e-mailovou adresu. Vkládaná práce může být ve formátu Word, Open Office, TeX atd. Formát se vždy převede do PDF a holého textu. Součástí systému je i automatická antivirová kontrola souborů.

Institucionální přístup

Systém je kromě vysokých škol k dispozici také dalším vzdělávacím institucím, nejčastěji středním nebo základním školám. Ty mohou využívat například výše uvedeného volného

přístupu, který je založen na tom, že jakýkoli učitel si může založit až tři odevzdávárny a do každé vložit max. 33 prací. Přístup mu však po šesti měsících vyprší. Pokud si však škola přeje využívat systém dlouhodobě, se všemi možnostmi a řádnou uživatelskou podporou, může se obrátit na kontaktní adresu uvedenou na stránkách systému Odevzdej.cz.

Obr. 4: Neautentizovaná kontrola souboru na plagiáty (výsledek přijde na uvedený e-mail)

Obr. 5: Výpis podobných souborů

Připravuje se

V současné době pracuje vývojový tým z Fakulty informatiky MU spolu s dalšími školami na propojení s lokálními systémy škol. Školy budou moci práce nejen automatizovaně importovat do centrálního systému Odevzdej.cz, ale také automatizovaně přebírat výsledky o kontrole plagiátů. Mezi nejvýznamnější novinky patří integrace algoritmu vyhledávání podezřelých souborů na internetu do systému a již se realizují zkušební testy. Vyvíjí se nástroje pro vkládání velkého množství souborů (zipem) a další aplikace. Systém se připravuje také pro překlad do angličtiny.

Ocenění

Systém na odhalování plagiátů v závěrečných pracích zavedla Masarykova univerzita jako první v České republice v roce 2006 ve svém Informačním systému MU pro podporu administrativy studia a e-learningu. V loňském roce se k využívání systému pro kontrolu závěrečných prací Theses.cz (Národní registr vysokoškolských kvalifikačních prací a systém na odhalování plagiátů) připojilo více než 20 škol. Počet souborů, s nimiž se práce porovnávají, vzrostl na téměř 1,5 milionu dokumentů, z toho 73 tisíc tvoří závěrečné práce. Systém byl oceněn cenou Inforum 2009 za významný počin v oblasti elektronických zdrojů. Ocenění v oblasti IT získal tým z Masarykovy univerzity již podruhé. Poprvé to bylo v roce 2007 za Archiv závěrečných prací, který je součástí celouniverzitního Informačního systému (IS MU) pro podporu studia. Detaily jsou k dispozici v článku na adrese http://is.muni.cz/clanky/2009_Inforum_cena.pl.

Pracovníci vývojového týmu se zúčastnili také mezinárodní soutěže ve vyhledávání plagiátů (1st International Competition on Plagiarism Detection), která se konala v červnu 2009. Software odhalil více než 69 % ze všech opsaných úseků textu, které zadavatel do soutěžních dokumentů vygeneroval. Tento výkon spolu s dalšími kritérii hodnocení zajistil týmu informatiků z Fakulty informatiky MU druhé místo v detekci externího plagiátorství i celkové druhé místo v prvním ročníku mezinárodní soutěže. Podrobnější informace najdete v článku na adrese http://is.muni.cz/clanky/2009_soutez_plagiaty.pl.

Literatura:

1. Masarykova univerzita. Vysoké školy zahajují kontrolu plagiátů v seminárních pracích pomocí softwaru (Tisková zpráva). *IS MU: Informační systém Masarykovy univerzity* [online]. 2009 [cit. 2009-07-21]. Dostupný z [www: <http://is.muni.cz/clanky/2009_odevzdej.pl>](http://is.muni.cz/clanky/2009_odevzdej.pl).
2. Masarykova univerzita. Systém Masarykovy univerzity odhalil v mezinárodní soutěži nejvíce plagiátů (Tisková zpráva). *IS MU: Informační systém Masarykovy univerzity*

- [online]. 2009 [cit. 2009-07-21]. Dostupný z www:
<http://is.muni.cz/clanky/2009_soutez_plagiaty.pl>.
3. Masarykova univerzita. Systém na odhalování plagiátů získal cenu Inforum 2009 (Tisková zpráva). *IS MU: Informační systém Masarykovy univerzity* [online]. 2009 [cit. 2009-07-21]. Dostupný z www:
<http://is.muni.cz/clanky/2009_Inforum_cena.pl>.
 4. BRANDEJSOVÁ, J.; BRANDEJS, M.; LUNTER, P.; KŘIPAČ, M.; KASPRZAK, J. Odhalování plagiátů v seminárních pracích. *IS MU: Informační systém Masarykovy univerzity* [online]. 2009 [cit. 2009-07-21]. Dostupný z www:
<http://is.muni.cz/clanky/2009_euba_odevzdej.pl>.

Brandejs, Michal; Brandejsová, Jitka; Krhutová, Růžena; Mikulášová, Zuzana; Pekárková, Lucie; Stančík, Martin. Kontrola plagiátů v seminárních pracích prostřednictvím Odevzdej.cz. *Ikaros* [online]. 2009, roč. 13, č. 8. [cit. 2009-08-11]. Dostupný na World Wide Web:
<<http://ikaros.cz/node/5641>>. URN-NBN:cz-ik5641. ISSN 1212-5075.

Zdroj: <<http://ikaros.cz/node/5641>>