

IS.MUNI.CZ a služba pro odhalování plagiátů

Michal Brandejs, Jitka Brandejsová, Jan Kasprzak, Miroslav Křipač, Martin Stančík

Abstrakt

V roce 2009 oslaví projekt Informačního systému MU (pod vedením stejných tvůrců jako Theses.cz) 10 let svého vzniku. V době svého 7. roku existence kromě mnoha jiných vlastností nabídnul svým uživatelům také službu pro vyhledávání podobných souborů (odhalování potenciálních plagiátů). Jde o vlastní implementaci algoritmu vyhledávání podobných souborů ve studijním informačním systému (přesahující rámec studijních informačních systémů). Nejprve byla dána možnost vyhledávání podobných souborů pro běžné uživatele, ale po několika měsících zkušenostech z provozu byla přidána další speciální možnost vyhledávání podobných souborů pro manažery. Trvalo určitý čas, než se vykrytalizovaly všechny organizační postupy na univerzitě, byť díky centrálnímu sběru závěrečných prací do Archivu závěrečných prací IS MU od roku 2004 a zveřejňování závěrečných prací světu od roku 2006, se na MU nastavila pravidla v krátkém čase. Jak vyhledávají uživatelé IS MU podobné práce vůči pracím vloženým do Theses.cz (a později do Odevzdej.cz) bude součástí příspěvku včetně dalších zkušeností se závěrečnými pracemi na MU.

Historie služby na odhalování plagiátu v rámci Informačního systému Masarykovy univerzity (IS MU) sahá do roku 2004, kdy byla spuštěna funkce na elektronický sběr závěrečných prací pomocí jednotného systému (tedy ne pouze papírově nebo na datových nosičích). V té době se také univerzita rozhodla, že zveřejní všechny takto odevzdané práce všem studentům a zaměstnancům školy v plném rozsahu. Vedle jasného přínosu pro akademickou obec, kdy se každý uživatel jednoduše dostane k výsledkům práce ostatních, na které může navázat nebo ze kterých může čerpat, však již v té době řada uživatelů pozorovala na nebyvalé usnadnění práce studentům, kteří se rozhodnou opisovat. Proto začaly první úvahy na vytvoření počítačového systému, který pokud na jedné straně usnadňuje elektronické zpřístupnění textů, na druhé straně pomůže snížit negativní vlivy, které toto zveřejnění přináší.

V roce 2006, kdy se podařilo zveřejňování prací prosadit také na legislativní úrovni pro celou širokou veřejnost, tak Masarykova univerzita nejen plně zpřístupnila všechny své texty závěrečných prací od daného data, ale poskytla svým uživatelům možnost prohledávat tyto texty na podobnosti a získat tím nástroj, který usnadní odhalení skutečných podvodníků.

Postupně se pak systém na odhalování plagiátů rozšířil z původní funkce na srovnání konkrétního dokumentu o vlastnosti, které umožňují vyhledání těch souborů, ke kterým systém našel automaticky po vložení podobnost ve větším rozsahu. Toto tzv. globální vyhledávání umožňuje kontrolu nad rámcem konkrétních předmětů či studentů na vyšší úrovni. Pověřený pracovník na dané katedře, fakultě či celé univerzitě tak může periodicky kontrolovat úplně všechny dokumenty.

S touto funkcí souvisí také nutná organizační opatření, které postupně musela univerzita zavést. Na celouniverzitní úrovni je nejprve pověřen pracovník, který kontroluje napříč celou školou opakovaně. V době největšího sběru prací (obvykle v jarních měsících až do června) je pak interval kontroly kratší (systém vždy vypisuje nové dokumenty resp. nechá označit ty podobné dokumenty, které byly vyhodnoceny jako korektní tak, aby nebyly dále uvažovány). Případná podobnost, která není přímo vyhodnocena jako oprávněná, je pak předána na fakultu. Na této úrovni už práce povinně kontroluje vedoucí závěrečné práce, který jednak prověřuje podezření předané z centrální úrovně a jednak dle nařízení fakulty (neboť různé fakulty MU mají různá zaměření a z nich plynoucí nároky na kontrolu podobností) předkládá výsledky kontroly jím vedených prací buďto pro každou práci nebo pouze v případě pozitivního nálezu. Řešení případného nálezu je v kompetenci fakult, tedy obvykle v řešení vyučujícího a případně děkanem ustanovené disciplinární komise.

Na základě těchto opatření a pozitivních zkušeností s provozem systému už od roku 2004 se postupně některé fakulty začaly rozhodovat, že zruší povinnost odevzdávat práce v papírové formě (zatím dvě fakulty) a většina z nich již práce v papírové formě nearchivuje, čímž ušetří pro univerzitu v současné době velmi cenné prostory.

Navíc se poslední době rozbíhají projekty, které systematicky digitalizují starší závěrečné práce, které pak systém IS MU automaticky po vložení převádí do prostého textu a tím umožní rovněž kontrolu podobností a tím vyhledání možných plagiátů zpětně.

Ze zkušeností z ostrého provozu plyne, že preventivní účinek systému je velmi silný. Studenti, kteří vědí, že jejich práce jsou systematicky kontrolovány raději investují do řádného vypracování svých úkolů, než by riskovali odhalení nebo se snažili složitě systém odhalování plagiátů obejít. Zajímavé také je, že nejvíce disciplinárních řízení se studenty se týká právě přestupků souvisejících s opisováním. Jako významný přínos, který není vůbec běžný, pak bereme fakt, že celý systém funguje naprosto nezávisle na správcích systému, kteří skutečně pouze spravují software a hardware sloužící přímo koncovým uživatelům.