

## **O letošním ročníku semináře *Systemy pro zpřístupňování vysokoškolských prací 2007***

( zpřístupnění | šedá literatura | plagiátorství | informační systémy | elektronické vysokoškolské kvalifikační práce | digitální knihovny | Česká republika | Asociace knihoven vysokých škol České republiky | Zprávy, reportáže a glosy )

### Abstrakt:

Druhý ročník semináře *Systemy pro zpřístupňování vysokoškolských kvalifikačních prací 2007*, který uspořádala Odborná komise pro otázky elektronického zpřístupňování vysokoškolských kvalifikačních prací, se konal se pod záštitou Asociace knihoven vysokých škol ČR dne 16. 10. 2007. Organizace se ujala Ústřední knihovna VUT. Program semináře byl rozdělen do čtyř bloků: projekty, zkušenosti provozovatelů systémů, technologie a řešení dílčích technických úkolů. Nejprve byly představeny dva velké projekty: společný projekt sedmnácti vysokých škol „**Národní registr VŠKP a systém na odhalování plagiátů**“ a projekt Státní technické knihovny „**Národní úložiště šedé literatury**“. Dále prezentovaly krátké vstupy o konkrétních zkušenostech zástupci Univerzity v Pardubicích, Univerzity Karlovy a Akademie múzických umění v Praze. Odpoledne byly představeny produkty na odhalování plagiátů firmy SEFIRA s.r.o a Mendelovy lesnické a zemědělské univerzity v Brně. Poté zazněly referáty a diskuze na téma převody dat do formátu PDF, technická metadata a identifikátory. Prezentace jsou vystavené na adrese <http://www.evskp.cz/> v rubrice Dokumenty / Prezentace.

### Obsah:

Seminář „***Systemy pro zpřístupňování vysokoškolských kvalifikačních prací 2007***“ se konal 15. 10. 2007 již jako druhý ročník. Jeho organizační zabezpečení na sebe opět vzala Ústřední knihovna VUT v Brně, finančně jej podpořila Asociace knihoven vysokých škol ČR (dále AKVŠ), o program se postarala Odborná komise pro otázky elektronického zpřístupňování vysokoškolských kvalifikačních prací (dále Komise). Program semináře byl rozdělen do čtyř bloků: projekty, zkušenosti provozovatelů systémů, technologie a řešení dílčích technických úkolů. Jedním z velmi diskutovaných témat byla otázka plagiátorství včetně zamýšlení se nad jeho důsledky a možnostmi postihu. Záměrem pořadatelů bylo opět poskytnout především prostor k výměně zkušeností a k diskusím, což se bohatě naplnilo. Seminář navštívilo 40 osob, ačkoliv registrujících se zájemců bylo více. Bohužel se ne všichni – když se s nimi již organizačně

počítalo – nakonec dostavili. O to více času se tak získalo na to, aby se všichni zájemci mohli vyjádřit ke všem předneseným příspěvkům. Diskuse dokonce obvykle časově přesahovaly vlastní prezentace a bylo patrné, že zvolená témata účastníky zaujala a že k nim také mají co říci.

Po zahájení semináře předsedkyní AKVŠ a Komise pro eVŠKP následoval jeden z hlavních příspěvků ing. Jitky Brandejsové, který připravila ve spolupráci s ing. Janem Machem z VŠE. Měl za úkol seznámit se záměry projektu „**Národní registr VŠKP a systém na odhalování plagiátů**“, který v letošním roce nově vznikl a kterého se účastní sedmnáct vysokých škol. Účastníci projektu chtějí přispět ke zvýšení kvality závěrečných prací a vysokoškolského vzdělávání a k rychlé dostupnosti informací vytvořením celonárodního registru vysokoškolských kvalifikačních prací (VŠKP) spojeného s odhalováním plagiátů (tj. podobných textů) v digitálním úložišti kvalifikačních prací. Proto podaly tento společný projekt v rámci centralizovaných rozvojových projektů na Ministerstvo školství, mládeže a tělovýchovy. Za koordinaci projektu zodpovídá Masarykova univerzita (MU) a těžiště její práce bude spočívat zejména v technologické realizaci projektu. MU navazuje na své vlastní dlouholeté zkušenosti s vytvořením archivu závěrečných prací a se službou pro odhalování plagiátů, kterou provozuje v rámci Informačního systému Masarykovy univerzity (IS MU). Jedním z významných partnerů je Vysoká škola ekonomická v Praze, která bude koordinovat spolupráci na návrzích struktur a funkcí národního registru VŠKP. Tato část projektu bude vycházet z doporučení a dosavadních výsledků Odborné komise pro otázky elektronického zpřístupňování vysokoškolských kvalifikačních prací. Komise bude využívána jako konzultační a poradenský orgán. Součástí návrhu národního registru bude také řešení autorskoprávní problematiky, metodických doporučení pro oblast provozu a budování lokálních registrů, citování zdrojů apod. Do projektu se zapojily kromě Masarykovy univerzity a Vysoké školy ekonomické v Praze také Akademie múzických umění v Praze, Česká zemědělská univerzita v Praze, Janáčkova akademie múzických umění v Brně, Jihočeská univerzita v Českých Budějovicích, Ostravská univerzita v Ostravě, Slezská univerzita v Opavě, Univerzita Hradec Králové, Univerzita Jana Evangelisty Purkyně v Ústí nad Labem, Univerzita Palackého v Olomouci, Univerzita Tomáše Bati ve Zlíně, Vysoká škola báňská - Technická univerzita Ostrava, Vysoká škola polytechnická Jihlava, Vysoká škola technická a ekonomická v Českých Budějovicích, Vysoká škola uměleckoprůmyslová v Praze a Západočeská univerzita v Plzni. Další podrobnosti o projektu lze sledovat na <http://www.theses.cz/>.

Podle očekávání se nejbohatší diskuse rozvinula právě po této přednášce. Byly zvažovány např. otázky potřeby a možností porovnávat plné texty prací také s

internetem, možnost budoucí účasti dalších škol stojících nyní mimo projekt (počítá se s bezplatným zapojením zájemců), vysvětlovány byly některé detaily praxe a provozu systému na MU (např. upřednostňuje se sběr plných textů v nativním formátu a systém je automaticky převádí do formátu PDF), hovořilo se o předpokládaných nákladech aj.

Ve druhém příspěvku dopoledního bloku nás Petr Novák z Státní technické knihovny v Praze seznámil se záměry a koncepcí projektu Národní úložiště šedé literatury, který byl letos podán do veřejné soutěže ve výzkumu a vývoji na rok 2008 vypsané Ministerstvem kultury. STK tak i v roce 2007 pokračuje ve svých aktivitách zacílených mj. na vybudování informačního systému registrujícího a zpřístupňujícího českou produkci šedé literatury. Projekt počítá také se spoluprací s NR VŠKP. Jedním z jeho cílů je stanovit pravidla trvalého uchovávání digitálních objektů, zásady ochrany dat a povinná technická metadata. Součástí prezentace bylo také připravované dotazníkové šetření, jehož výsledkem bude harmonizace představ a konkretizace potřeb producentů šedé literatury v ČR. V diskuzi se zvažovala otázka jednoznačné kategorizace šedé literatury.

Po přestávce následoval blok věnovaný konkrétním zkušenostem provozovatelů systémů pro zpřístupňování. Lucie Vyčítalová z Univerzity v Pardubicích seznámila účastníky s praktickými zkušenostmi se zaváděním digitálního repozitáře DSpace na Univerzitě Pardubice. Nasazení systému bylo řešeno v rámci projektu podpořeného z FR VŠ. Směrnice školy č. 13/2007 ukládá studentům odevzdávat tištěnou a elektronickou verzi své bakalářské, diplomové, rigorózní nebo disertační práce. Podmínky přístupu k plným textům prací jsou určeny licenční smlouvou, která je uzavřena mezi studentem a fakultou, právo omezit přístup k práci má též fakulta. Ke sběru dat a plných textů prací je využíván studijní systém STAG, z něhož budou záznamy a plné texty pravidelně exportovány do celouniverzitního digitálního repozitáře **DSpace**. V souvislosti se zaváděním systému byl proveden jednorázový export záznamů VŠKP z knihovního systému Daimon do DSpace, byla specifikována metadatová pole dle Metadatového standardu Komise pro eVŠKP, byla provedena lokalizace DSpace verze 1.4.2 do češtiny, v neposlední řadě byl aktivován trvalý identifikátor handle. Část příspěvku byla též věnována konkrétním problémům, které bylo nutné vyřešit (duplicity záznamů přebíraných ze studijního systému, problémy popisu prací studenty apod.).

Následoval referát Evy Březinové z Ústavu výpočetní techniky UK, v němž informovala o aktuálním stavu zavádění systému DigiTool na UK i o problémech, který tento proces provázejí. V současnosti se zde podle plánu pracuje na programování migrační procedury metadat ze studijního systému do repozitáře, kam se dosud vkládají data ručně. Vznikají také související interní normativní předpisy, avšak až na úrovni jednotlivých fakult, což vytváří odlišné podmínky a situaci to komplikuje. Univerzita

Karlova se bohužel také nepřipojila k projektu sedmnácti škol, ačkoliv to bylo původně přislíbeno.

Poslední dopolední krátký vstup byl od Ivy Horové z Akademie múzických umění v Praze na téma kvalifikační práce netextové povahy. Seznámila s prvními zkušenostmi se sběrem těchto prací na AMU, z nichž především vyplývá značná nejasnost této problematiky jak obecně, tak ale zatím i v rámci AMU. Bude třeba se systematicky pustit do kategorizace těchto prací, pokusit se stanovit pravidla pro popis, datové formáty a struktury apod. Problémem je však např. také otázka názvů těchto prací – např. u koncertů, kde obecné označení je duplicitní (např. bakalářský koncert) a přitom z názvů jednotlivých interpretovaných skladeb jej vytvořit také nelze.

Odpolední program byl ve dvou blocích zaměřen na technologie a řešení dílčích technických záležitostí. Společnost SEFIRA představila prostřednictvím Ivy Celbové vlastní připravované řešení na odhalování plagiátů vysokoškolských kvalifikačních prací, které využívá produkt ConText CZ/SK umožňující fulltextové vyhledávání v českých, resp. slovenských textech. Po představení produktu ConText CZ/SK na loňském semináři byli účastníci seznámeni s jeho novými funkcemi.

Milan Šorm nazval svůj příspěvek Boj s plagiáty. Seznámil v něm s praxí a zkušenostmi s pilotním řešením systému na Mendelově lesnické a zemědělské univerzitě v Brně (MZLU), který vznikl na základě požadavku vedení školy. Popsal v něm proces porovnávání a vyhodnocování textů. Tento systém vznikl jako alternativa a do jisté míry i konkurence antiplagiátorského systému MU a budoucnost ukáže, která a jak se prosadí. MZLU má ambice i na poli tvorby Národního registru VŠKP, prozatím se do společného projektu nepřipojila, ale je připravena nabídnout své služby.

Druhý příspěvek Jana Macha se zabýval aktuální problematikou převodu textů do formátu PDF, následnou extrakcí textu z PDF pro potřeby fulltextového prohledávání a nastavení restrikcí pro uživatele na užití dokumentu: zamezení kopírování či tisku textu a jeho změnám. Tato problematika se stává aktuální v souvislosti s budováním elektronických registrů VŠKP. Jedním z doporučených nástrojů pro konverzi do PDF je zdarma dostupný PDF Creator, který může být nasazen jako virtuální tiskárna, a to i ve verzi síťového nastavení s předdefinovanými metadatovými poli PDF souboru či plně automatickým převodem dokumentů. Pro extrakci textu či HTML z PDF jsou prezentovány dva volně dostupné programy, a to program XPDF a PDFBox. Otázka zabezpečení PDF souborů je velmi složitá a bohužel současné nástroje nemají problém odstranit restrikce a Master Password, pokud je uživatel schopen dokument zobrazit.

Otázkou výběru a implementace konkrétního systému trvalých (persistentních) identifikátorů v případě Národního registru VŠKP i dalších digitálních knihoven šedé

literatury se ve své druhé prezentaci zabýval Petr Novák z STK. Tato otázka dosud čeká na své vyřešení. Cílem příspěvku bylo základní nastínění existujících technologií, sestavení žebříčku dle kritérií vhodnosti pro daný účel a zejména otevření diskuse.

Akce byla kladně hodnocena účastníky již na místě a je zřejmé, že pořadatelé i organizátoři již počítají s jejím dalším ročníkem v roce 2008.

**Horová, Iva.** O letošním ročníku semináře Systémy pro zpřístupňování vysokoškolských prací 2007. *Ikaros* [online]. 2007, roč. 11, č. 12 [cit. 2008-01-29]. Dostupný na World Wide Web: <<http://www.ikaros.cz/?q=node/4435>>. URN-NBN:cz-ik4435. ISSN 1212