

CENTRÁLNY REGISTER PUBLIKAČNEJ ČINNOSTI CREPČ

Mgr. Ján Grman
SVOP spol. s r.o., Bratislava
grman@svop.sk

INFORUM 2008: 14. konferencie o
profesionálných informačných zdrojích
Praha, 28. - 30.5. 2008

MOTIVÁCIA

- ◉ Systém VŠ SR a jeho úlohy
 - Výskum a vývoj
 - Pedagogická činnosť
- ◉ Evidencia publikačnej činnosti ako dôležitý parameter hodnotenia VŠ
 - Akreditácia odborov, Register garantov
 - Grantové systémy a posudzovanie
 - Vplyv na výšku finančnej dotácie (rastie)

VÝCHODISKOVÉ CIELE

- ⊙ **Automatizácia procesov získavania a vyhodnocovania údajov EPČ od VŠ**
 - Výmena údajov na priereze používaných IS
 - Implementácia kontrolných mechanizmov
- ⊙ **Zverejňovanie informácií o EPČ - portál**
 - Štandardné prehľadávanie unikátnej bázy dát na priereze pracovísk, rokov a autorov
 - Prezentácia štatistík
 - Autorizované sekcie pre potreby vkladania údajov ale aj manuálne vykonávané kontroly dodaných dát nezávislými hodnotiteľmi

ANALÝZA - STAV

- ◉ **Dotazníkové zisťovanie stavu na VŠ**
 - celkový objem EPČ a ročný prírastok
 - počty pracovníkov EPČ, spôsob zberu dát (centrálne v knižnici, formuláre pre autorov, ...)
 - používaný informačný systém pre evidenciu a jeho schopnosti / možnosti v oblasti exportu dát
- ◉ **Závery**
 - Väčšina VŠ používa profesionálne IS - typicky modul knižnično-informačného systému univ. knižnice
 - Niektoré VŠ majú vlastné riešenia evidencie a časť realizovala dodávku štatistík pre účely MŠ SR ručnou kompiláciou v neštruktúrovanej podobe

ANALÝZA - VÝMENNÝ FORMÁT

- **Vnútrotnými formátmi** evidencie väčšiny VŠ sú MARC21, UNIMARC, pri časti to sú iné vlastné typy uloženia
- **Voľba padla na formát XML**
 - profesionálne KIS v prieskume deklarovali schopnosť produkovať XML
 - pri riešení, ktoré si VŠ realizovali samé, išlo zvyčajne o riešenia nepracujúce priamo z formátom typu MARC a voľba niektorého formátu MARC by viedla k nutnosti zložitejších konverzií medzi MARC formátmi a dotkla by sa viac producentov
 - formát XML je ľahko implementovateľný aj do systémov nie priamo zameraných na spracovanie EPČ, navyše plne rešpektuje legislatívu a odporúčania pre výstavbu štátnych informačných systémov, ich súčastí a protokolov pre výmenu dát

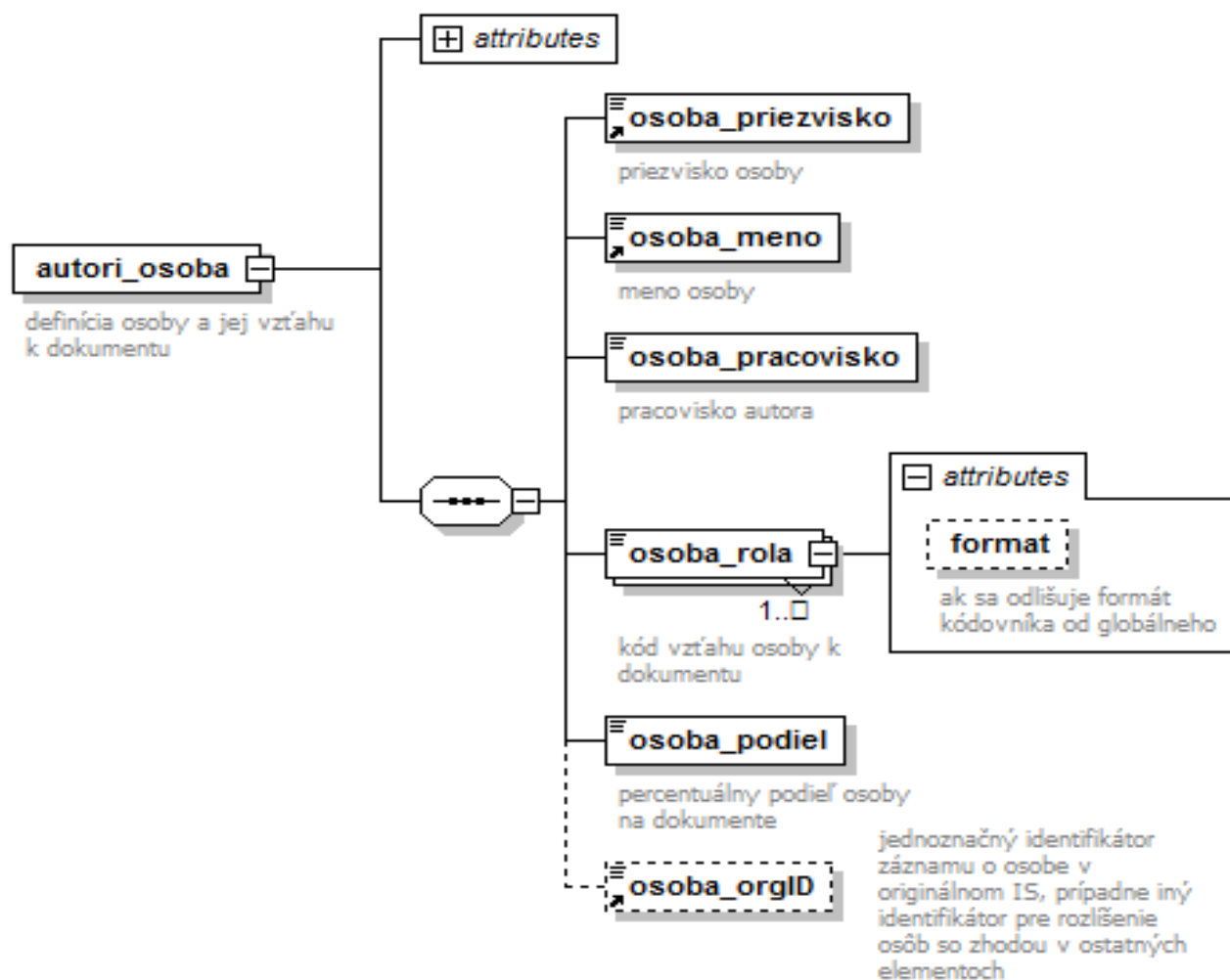
CREPČ-XML

- striktne definovať **minimálny záznam** pre potreby CREPČ a to najmä:
 - identifikáciu zdroja katalogizácie,
 - druh dokumentu,
 - kód kategórie dokumentu a časové zaradenie,
 - príslušnosť záznamu k pracovisku / pracoviskám, príslušnosť záznamu k autorom (informácie o roli autora a jeho podiele na zázname)
- využiť súbor údajov pre potreby bibliografického spracovania a prezentácie záznamu,
- definovať **povinné i voliteľné elementy** a striktne definovať ich vzájomné vzťahy,
- minimalizovať **prácnosť prevodu** z používaných formátov UNIMARC a MARC21,
- **zabezpečiť** maximálnu **unifikáciu** záznamov pochádzajúcich z rôznych zdrojov a interných formátov,

CREPČ-XML

- ◉ zabezpečiť **prevod** všetkých **dát** nevyhnutných pre funkcie registra (povinné polia) a ďalších dát, ktoré prinášajú dôležitú informačnú hodnotu pre potenciálnych používateľov CREPČ (vedecké agentúry, akreditačné komisie, verejnosť a pod.),
- ◉ **minimalizovať chyby** vyplývajúce z rozdielov implementácie noriem rôznymi informačnými systémami, či dokonca rôznymi pracoviskami používajúcimi systém rovnakého výrobcu,
- ◉ zabezpečiť **formát ľahko implementovateľný** aj v informačných systémoch nepoužívajúcich normu ISO 2709 a niektorý z MARC formátov,
- ◉ použiť **formát, ktorý bude ľahko rozšíriteľný a modifikovateľný** aj pre budúce (a dosiaľ neznáme) funkcie CREPČ alebo jeho systémové rozšírenia

CREPC-XSD



REALIZAČNÁ FÁZA

- Vybudovanie **serverovej** kapacity
 - Windows 2003 Server
 - MS SQL 2005
- Implementácia **centrálnej databázy** a komunikačného portálu pre import dát a ich prezentáciu
- Implementácia **exportných funkcií** na strane IS jednotlivých producentov dát
- **Pilotná dávka dát** za obdobie 1.11.2006 až 30.10.2007 (rok vykazovania 2007)

PROCES KONTROLY

- ◉ **Automatizovaná kontrola úplnosti položiek záznamov** (realizovanej už vlastnosťami výmenného formátu),
- ◉ **Automatizovaná kontrola duplicít** (duplicita publikácie v rámci odovzdanej dávky, duplicita publikácie autora v rámci viacerých VVS s kontrolou rozdelenia
- ◉ **Poloautomatizovaná kontrola údajov vo vnútri CREPC** prostredníctvom špecializovaného rozhrania pre hodnotiteľov so zámerom umožniť:
 - kontrolu zaradenia publikácie do kategórie na základe rozsahu dokumentu,
 - kontrolu zaradenia publikácie do kategórie na základe obsahu dokumentu,
 - potlačiť falošné duplicity, vyznačiť naopak skryté duplicity
 - evidovať poznámky ku kvalite popisu

VÝSLEDKY KONTROLY

- ◉ Fáza 1.1 - reporty z podozreniami na chyby (nie všetky duplicity nimi naozaj boli)
- ◉ Fáza 1.2 - opravy dát na strane producentov
- ◉ Fáza 1.3 - dodávka opravených dát a slovného komentára k výsledkom kontroly - opakovaný import dát
- ◉ Fáza 2.0 - nová komplexná a očistená databáza
- ◉ Základné dôsledky
 - Opravy dát resp. odmazanie duplicitných záznamov
 - Štart konštruktívnej diskusie o samotnom procese evidencie
 - Získanie skúseností a námety na ďalšie zlepšenia

VÝSTUPY PROJEKTU

- ◉ **bezpečné uloženie**, archivácia a jednotná prezentácia informácií,
- ◉ **vyhľadávanie** záznamov podľa definovaných kritérií (autor, kľúčové slová, pracovisko autora, rok vydania atď.),
- ◉ **možnosť kontroly dát** z viacerých pohľadov a súvislostí,
 - pohľad na priereze rokov sa prejaví až importom dát za rok 2008
 - rok 2008 sa bude dodávať na viac častí
- ◉ **priame výstupy a štatistiky** pre potreby manažmentu finančných prostriedkov,
- ◉ **značná redukcia času** od získania údajov, cez ich kontrolu, až po získanie požadovaných analytických podkladov

ĎALŠIE DÔSLEDKY

- ◉ vznik relevantného zdroja informácií aj pre iné zložky participujúce na procesoch súvisiacich s problematikou výskumu, vývoja, ich evidencie a hodnotenia
- ◉ zjednotením metodiky hodnotenia EPČ sa umožní používať dáta CREPČ aj **grantovým agentúram** (VEGA, KEGA, ...) a znížiť tak administratívne nároky na VŠ redukciou duplicitnosti vykazovania
- ◉ cenný zdroj informácií pre verejnosť ale i samotné VŠ a ich pracovníkov
- ◉ tlak na **zlepšovanie kvality** a presnosti evidencie ale prenesene i na samotných autorov (vedeckých pracovníkov) - pod drobnohľadom verejnosti

NÁMETY

- ◉ **problém štandardizácie číselníkov pracovísk** v rámci interných KIS univerzít a CREPC a propagáciu a údržbu zmien (ktoré sú pomerne dynamické),
- ◉ **problematika evidencie ohlasov**, ich aktualizácie s ohľadom na prax ich zápisu v MARC formátoch,
- ◉ **problematika pravidiel pre kontrolu rozsahu publikácií** a percentuálne podiely spoluautorov v rámci viacerých VVŠ, **problematika evidencie autorských hárkov**,
- ◉ **problematika automatizácie overovania karentovaných časopisov** voči aktuálnej databáze
- ◉ **rozširovanie funkcionality portálu**

ĎALŠIE CIELE A MIESTO CREPČ

- sprístupniť a zapojiť do CREPČ aj **d'alšie subjekty** realizujúce vzdelávaciu a vedeckú činnosť (napríklad štátne a súkromné vysoké školy, SAV a podobne).
 - pozitívne je, že samotné VŠ mimo CREPČ majú záujem participovať a samé oslovujú riešiteľov s otázkou na možnosti a podmienky zapojenia sa - aktuálne bola zapojená Akadémia ozbrojených síl gen. M.R.Štefánika v Liptovskom Mikuláši
- CREPČ plne zapadá do koncepcie budovania **štátneho informačného systému**
 - automatizácia procesov a tvorba centrálnych registrov v rezorte školstva (register študentov, register pracovníkov a garantov, ...)
- riešiteľom sa podarilo pilotnú fázu projektu realizovať naozaj v **krátkom čase** (menej ako 6 mesiacov vrátane prípravných fáz) a úspešne uviesť do praxe.

ZÁVER

◉ Projekt

- Centrálny projekt evidencie publikačnej činnosti (www.crepc.sk)

◉ Prevádzkovateľ

- Ministerstvo školstva SR (www.minedu.sk)

◉ Riešiteľský kolektív

- Univerzita Konštantína Filozofa v Nitre (www.ukf.sk)
- Žilinská univerzita v Žiline (www.uniza.sk)
- SVOP spol. s r.o. Bratislava (www.svop.sk)

◉ Zapojené VŠ

- Všetky verejné vysoké školy (celkovo 20 VVŠ)
- Jedná štátna VŠ (AOS LM)

ĎAKUJEM ZA POZORNOST

Diskusia a otázky