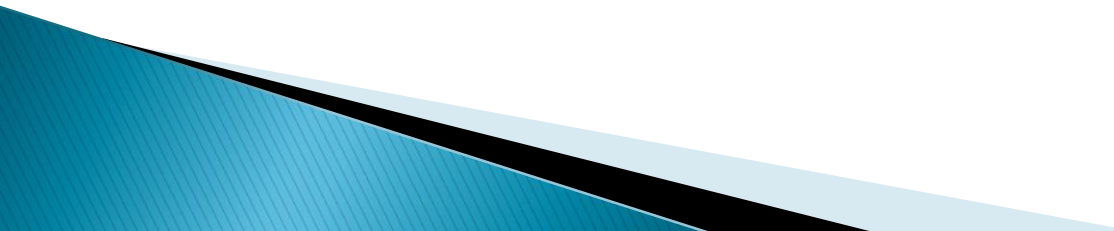


# Článková bibliografia v procese digitalizácie textových dokumentov

# Článková bibliografia a digitalizácia v SNK

- ▶ Objekty
  - ▶ Aktuálny stav
  - ▶ Postupy
  - ▶ Budúcnosť
- 

# Digitalizácia v SNK

- ▶ **Externá digitalizácia – mimo DC SNK**
- ▶ **Externý spracovateľ – Tender Media Group**

# Objekty digitalizácie

- ▶ **zviazaný ročník novín a seriálu**
- ▶ **číslo tlačенých novín**
- ▶ **číslo tlačенého časopisu**
- ▶ **jednotlivý článok z tlačенých novín alebo časopisu**
- ▶ **jednotlivý článok z tlačенého vedeckého alebo odborného časopisu a monografie**
- ▶ **jednotlivý článok z tlačенého vedeckého alebo odborného zborníka**

# Digitalizácia článkov – aktuálny stav

- ▶ **Počet titulov periodík: 720**
- ▶ **Počet čísel: 62 000**
- ▶ **Počet článkov: 879 000**

# Výstup digitalizácie článkov



**PSP balíček** - producer submission package

Balíček dát a metadát, ktorý prichádza od  
producenta dát

Štandardy:

**Definícia metadátových formátov pre  
digitalizáciu periodík (SNK, 2013)**

**Definícia metadátových formátov pre  
digitalizáciu monografií (SNK, 2013, 2015)**

# Vstupy digitalizácie – SNK

**Kontrola fyzického stavu a komplet  
periodika/zborníka**

**Materský BIB  
záznam periodika/zborníka**

**Skeny periodika,  
zborníka**



**BIB záznamy CL  
KIS Virtua  
NB SNK  
Reeditácia  
Reindexácia  
Deduplikácia**

# Externý spracovateľ

1. Spracuje skeny

2. Spáruje:

DIGI obrázky CL  
+ BIB záznamy  
+ materský záznam  
periodika

3. Tvorí BIB záznam  
ak nie je BIB zo SNK



Vytvorí PSP balíček



Balíček dát a metadát, ktorý  
prichádza od producenta



Výstup digitalizácie



# PSP balíček

- ▶ **archívna kópia (1 MC – archívnu kópiu pre každú stranu)**
- ▶ **používateľská kópia (1 UC pre každú vzniknutú MC, teda stranu)**
- ▶ **CR - ALTO XML súbor pre každú stranu**
- ▶ **OCR - TXT súbor - možnosť stiahnuť si len text dokumentu, vyhľadávanie / indexácia**
- ▶ **metadáta pre MC**
- ▶ **bibliografické metadáta - MODS a DC**
- ▶ **štrukturálne metadáta – METS**
- ▶ **technické metadáta - MIX, premis**
- ▶ **administratívne metadáta - premis, METS**
- ▶ **kontrolné metadátové súbory (s kontrolnými súčtami a údajmi o vzniku dát a pod.)**

# PSP – našou rečou povedané

- ▶ **1 PSP = 1 číslo periodika/zborníka + jeho články**
- ▶ **Obsahuje:**
  - ▶ **Digitálny obraz 1 čísla periodika/zborníka**
  - ▶ **Digitálne obrazy jeho článkov**
  - ▶ **Bibliografické metadáta:**
    - ▶ **Materský BIB záznam periodika/zborníka**
    - ▶ **Záznam o čísle periodika/zborníka**
    - ▶ **BIB záznamy článkov z daného čísla periodika/zborníka**

# BIB metadátá v PSP balíčku

Materský záznam  
periodika



Metadátový záznam 1  
čísla periodika



Články z 1 čísla  
periodika



# Konkrétne

## ► Slovenská archeológia

- Bratislava : Slovenská akadémia vied, 1953-
- Nitra : Archeologický ústav Slovenskej akadémie vied, 1995-

Slovenská archeológia.  
Roč. LIV, č. 1 (2006)

Články



# PSP balíčky v SNK

## Kontrola PSP balíkov



# Kontrola PSP balíčkov

- ▶ kontrola úplnosti (obrazové súbory, logická a fyzická štruktúra)
- ▶ kontrola bibliografických metadát a popisných údajov
- ▶ kontrola článkov - bibliografických metadát, segmentácia a úplnosť textu
- ▶ kontrola zhody metadát s katalógom Slovenskej národnej bibliografie v systéme Virtua
- ▶ kontrola novovytvorených bibliografických metadát článkov

# Archivácia PSP a SIP balíčkov

- ▶ Uloženie SIP do digitálneho archívno-informačného systému Rosetta (SNK)
- ▶ Uloženie SIP do CDA (*Centrálny dátový archív*) - UKB
- ▶ Uloženie SIP do CAIR (*Centrálna aplikačná infraštruktúra a registratúra*)
- ▶ **CAIR** = projekt Národného osvetového centra - vytvára systém na centrálnej úrovni pre všetky inštitúcie v rezorte kultúry
- ▶ **Výstup CAIR-u – prezentačný portál Slovakiana**
- ▶ on-line prístup k výstupom digitalizácie inštitúcií v rezorte kultúry

**SIP balíček – balík dát a metadát pre aplikáciu sprístupnenia**

# Import PSP → Katalóg SSNB Články

- ▶ **identifikuje všetky digitalizované články (uuid, urn:nbn)**
- ▶ **obohatí sa o obrovské množstvo nových článkových záznamov, čím sa eliminuje ich absencia z obdobia pred rokom 1976 a po roku 2010 (cca 450 000 BIB záznamov)**
- ▶ **následným pridelením URL adresy zabezpečí prepojenie katalógu s digitálnym obrazom dokumentu**
- ▶ **umožní budúcu editáciu dát digitalizovaného objektu (napr. zmeny vo všetkých typoch autorít a pod.)**



# Čo dostane používateľ

- ▶ **Celé číslo periodika** → listing a fulltext
- ▶ vyhľadávanie nad všetkými periodikami
- ▶ **Jednotlivé články** → cielené vyhľadávanie
- ▶ vyhľadávanie iba vo vedeckých a odborných periodikách
- ▶ **digitálna edícia** vedeckých a odborných článkov v digitálnej knižnici

# Kde to dostane používateľ

- ▶ **Digitálna knižnica SNK**
- ▶ **Prezentačný portál Slovakiana**
  - ▶ portál, prostredníctvom ktorého sa verejnosť – laická i odborná – dostane on-line k výstupom digitalizácie
- ▶ **Katalóg Slovenská knižnica/KIS3G (URL)**

# Formáty zobrazenia

- ▶ **JPEG s OCR vrstvou**
  - ▶ **PDF**
  - ▶ **EPUB**
- 
- ▶ **pre jednoduchú publikáciu napr. pre smartfóny alebo elektronické čítačky (CAIR Slovakiana)**

# Prínos digitalizácie článkov

- ▶ **Sociálny**
- ▶ **Vplyv → veda, výskum, vzdelávanie**
- ▶ **Výberová digitálna edícia pre:**
- ▶ **prístup, šírenie a uchovávanie vedeckých a odborných informácií**
- ▶ **Dôležitý zdroj informácií pre:**
- ▶ **vedeckú a odbornú verejnosť**
- ▶ **univerzity**
- ▶ **vysoké školy**
- ▶ **študentov**

# Digitalizácia a bibliografia (nielen článková)

- ▶ Bibliografia prechádza do novej vývojovej fázy
- ▶ **Pradávný sen bibliografov sa naplňa**
- ▶ **Bibliografia privádza používateľa ku konkrétnemu dokumentu**
- ▶ **Sprístupňuje :**
  - ▶ BIB záznamy o dokumentoch
  - ▶ Kompletné dokumenty - digitálne obrazy (plné texty) dokumentov
- ▶ **Usmerňuje používateľa v digitálnej knižnici**

# **Po skončení NP DIKDA**

- ▶ **Pokračovanie digitalizácie článkov**
- ▶ **Digitalizačný plán na obdobie udržiateľnosti NP DIKDA (2016-2020)**
- ▶ **Oddelenie spracovania článkov NB SNK**
- ▶ **1. 10. 2015**

# Nová článková bibliografia v SNK

- ▶ **Efektívna symbióza**
  - ▶ Práca s digitálnym obsahom
  - ▶ Práca v digitalizačnom prostredí (WF)
  - ▶ Práca v Súbornom katalógu Články
- 
1. *klasifikácia,*
  2. *indexovanie*
  3. *deskriptívne bibliografické metadáta*

# Ciele 1

- ▶ digitalizácia a bibliografia článkov v reálnom čase – súbežne
- ▶ tvorba bibliografických metadát priamo v procese digitalizácie a katalogizácie
- ▶ bibliografické metadáta pre SK Články a zároveň pre digitálny objekt
- ▶ proces prebieha na 1 mieste bez zbytočných časových a logistických strát



# Ciele 2

- ▶ manažment kvality metaúdajov v SK Články a DK
- ▶ manažment sémantickej interoperability v SK Články a DK
- ▶ navigácia používateľa v digitálnej knižnici

# Činnosť OSČ NB SNK

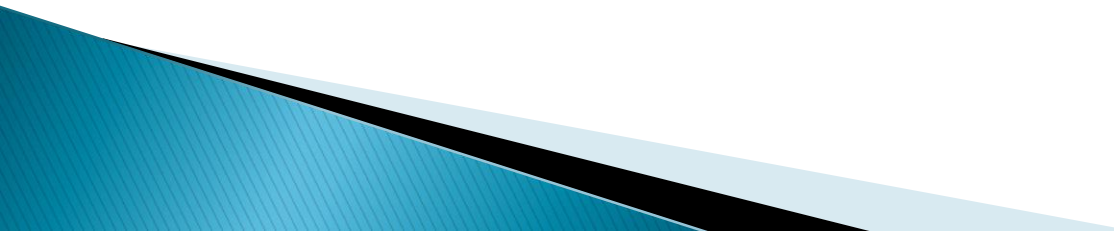
- ▶ Súbežné spracovanie článkov z vedeckých a odborných periodík a zborníkov
- ▶ Kompletizácia krátkych záznamov článkov vytvorených externým dodávateľom v NP DIKDA (deskriptívne metadáta, predmetové heslá)
- ▶ Manažment kvality metaúdajov v SK Články a DK

# Sme spokojní ?

**Jediný spôsob, ako byť skutočne spokojný, je veriť, že to, čo robíš, je skvelá práca.**

**A jediný spôsob, ako robiť skvelú prácu, je milovať, to čo robíš.**

**Steve Jobs**



Ďakujem za pozornosť

[ludmila.rohonova@snk.sk](mailto:ludmila.rohonova@snk.sk)

