

Národní repozitář: Účel a vývoj

Schůzka koordinační skupiny k tématům Open Science

16. 9. 2020

Miroslav Šimek, Petra Černohlávková

miroslav.simek@techlib.cz, petra.cernohlavkova@techlib.cz

- NTK provozuje od roku 2009 Národní úložiště šedé literatury, které obsahuje
 - VŠKP
 - zprávy (grantové, závěrečné, ...)
 - metodiky
 - firemní literaturu
 - ... a další
- Některé certifikované metodiky jsou ukládány do NUŠLu ze zákona
- NUŠL se postupem času stal i sirotčím repozitářem a obsahuje i osobní archivy

Proč změna?

- NUŠL je provozován na starší platformě Invenio (CERN), kterou je nutno zaktualizovat
- Invenio 3+
 - obecnější v datovém modelu než DSPACE
 - rychlejší a méně náročné na prostředky než Fedora Commons
 - principiálně odlišné od předchozí verze
 - MARC is dead, let JSON live
 - napsané v pythonu, se kterým má NTK zkušenost
- Platforma Invenio umožňuje vytvořit jakýkoliv hierarchický metadatový popis

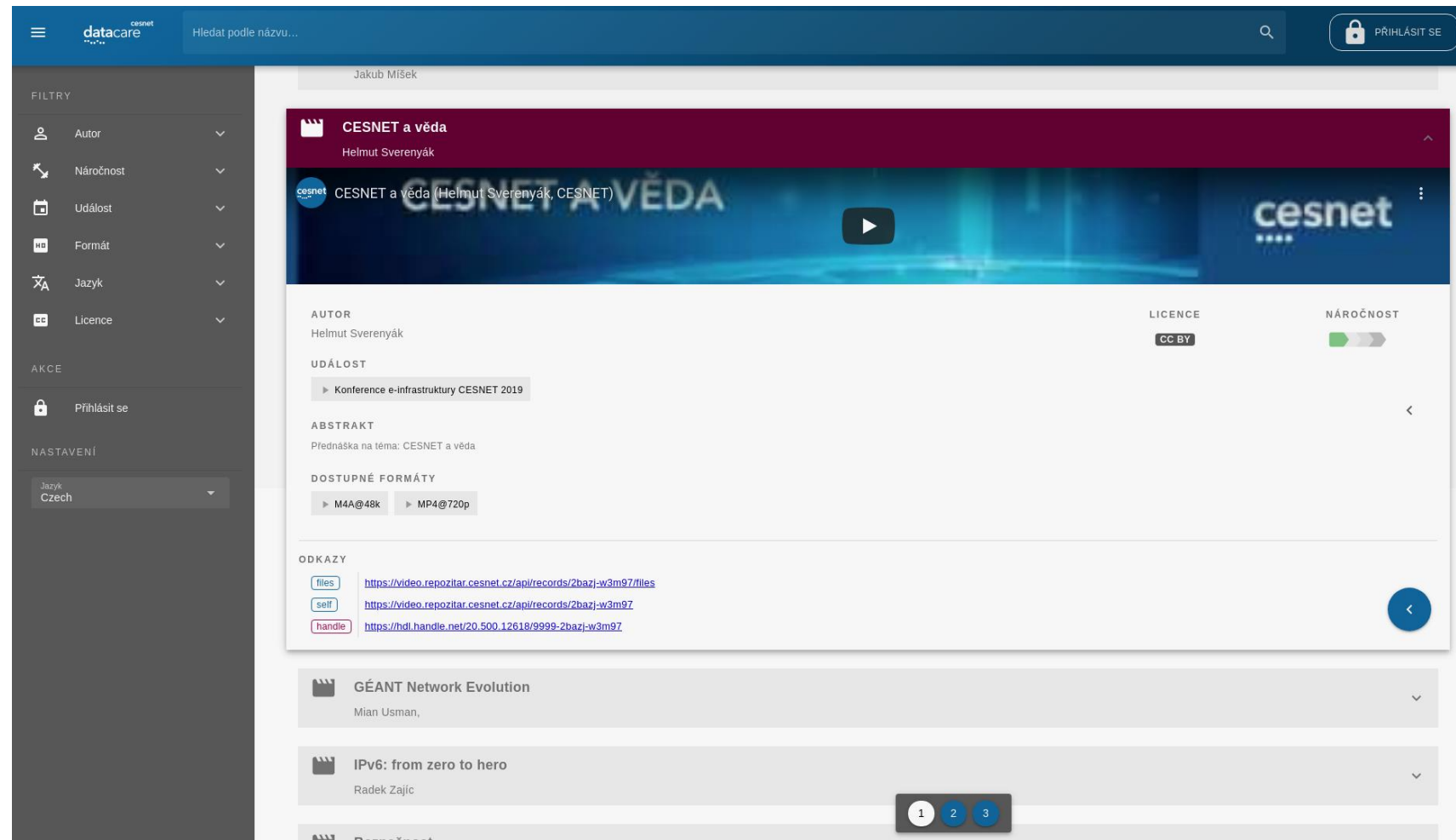
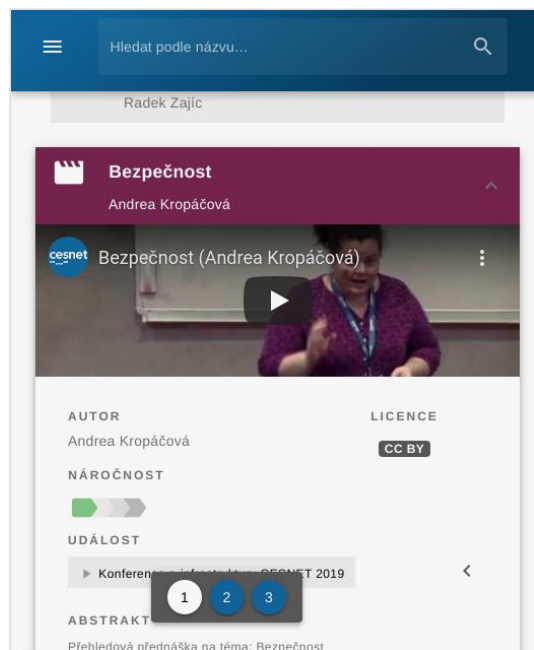
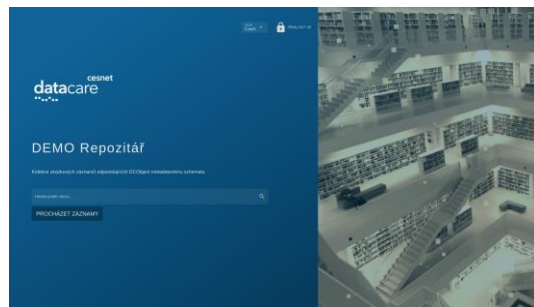
NUŠL → Národní repozitář

Aktivita PA4 NCIP VaVal - Národní centrum pro informační podporu výzkumu, vývoje a inovací

- upgrade SW
- ukládání publikačních výsledků výzkumu zejména pro potřeby organizací, které repozitář nevlastní
- rozšíření spektra typů vkládaných dokumentů
- výhled: dostupnost – za podmínek stanovených vlastníkem výsledku – všechny výsledky VaVal, přinejmenším jejich metadata s odkazem na soubor

Kdo jsme

- **NTK**
 - vývoj
 - metadatový popis
 - kurátorství
- **CESNET**
 - vývoj
 - provozování repozitáře v k8s, S3, databáze, indexace
 - zálohování a LTP, návaznost na národní infrastrukturu úložišť
- **VŠCHT**
 - vývoj
- **... další zváni**



Vlastnosti NR

- **Verifikace metadat**

- do depositu je možno umístit jakkoliv (ne)validní metadata
- automatické kontroly a přiřazení chyb
- publikovány jsou pouze validní záznamy
- hromadné opravy

- **Přikládání souborů**

- indexace plných textů pro vyhledávání

- **Schvalovací proces a publikování záznamů**

- podporován vícestupňový schvalovací proces
- hlavní kurátor jako role metadatového administrátora



Vlastnosti NR

- **Archivace**
 - po publikování záznamu vzniká archivní informační balíček (AIP)
 - ten je uložen do CESNET LTP a může putovat zpět do instituce
- **Spolehlivost**
 - Kontrolní součty a periodická verifikace
- **Centrální řízení přístupu**
 - eduID.cz / Perun s možností zřízení účtů pro ty, co nemají identitu v eduID.cz
- **Přidělování HANDLE**

Kde jsme? *Smlouvy*

- Oznámení partnerským institucím o změně
- Dokončuje se návrh
- Jedna smlouva více typů spolupráce
- Specifika v přílohách
- Probíhá zpětná vazba od ČNB, UK a NPÚ
- K partnerům do konce roku

Kde jsme? *Typologie*

- Průzkum typologie a funkcionality (červen - červenec)
- Návrh typologie
 - Vychází z NUŠL, RIV, OpenAIRE a COAR
 - Struktura – definice, mapování RIV a OpenAIRE a určení ŠL
- následuje získání zpětné vazby

Kde jsme? *Schémata*

- Vypracování nových metadatových schémat
 - Vychází z NUŠL, RIV, OpenAIRE (for repositories + for CRIS), UK a COAR Vocabularies
 - JSON
 - 4 schémata: common, theses, Nresults, eventRelatedDocs
- Testovací nastavení validací a nasazení řízených slovníků
- Migrace VŠKP z NUŠL

Kde jsme? *Slovníky*

- COAR Access Rights
- Instituce (provider, publisher, degreeGrantor)
- Práva (copyright, CC)
- Role přispěvatele (s mapováním MARC a DataCite)
- Výsledky typu N
- Věcný popis – PSH, MeSH-CZ, Medvik

Kam směřujeme?

2021 → **Národní
repozitář
Repo,**

Děkujeme za pozornost

Otázky?

miroslav.simek@techlib.cz, petra.cernohlavkova@techlib.cz