



## **Využití technických metadat pro dlouhodobou ochranu digitálních objektů v systému Rosetta**

Mgr. Eliška Pavlásková

# Technická metadata

---

- Vlastnosti digitálních dat – souborů
  - Obecné vlastnosti – velikost, formát
  - Specifické vlastnosti – rozlišení, typ kodeku
  - Informace o prostředí - SW, ve kterém byl soubor vytvořen či upraven
- Součást reprezentační informace dle OAIS
  - Informace, které umožní cílovým uživatelům objektu jej **smysluplně použít**

# Datový model - PREMIS

Intelektuální entita  
Digitalizovaný rukopis

Reprezentace  
Archivní kopie

Soubor  
TIFF

Reprezentace  
Uživatelská kopie

JP2000

# Technická metadata - zdroje

---

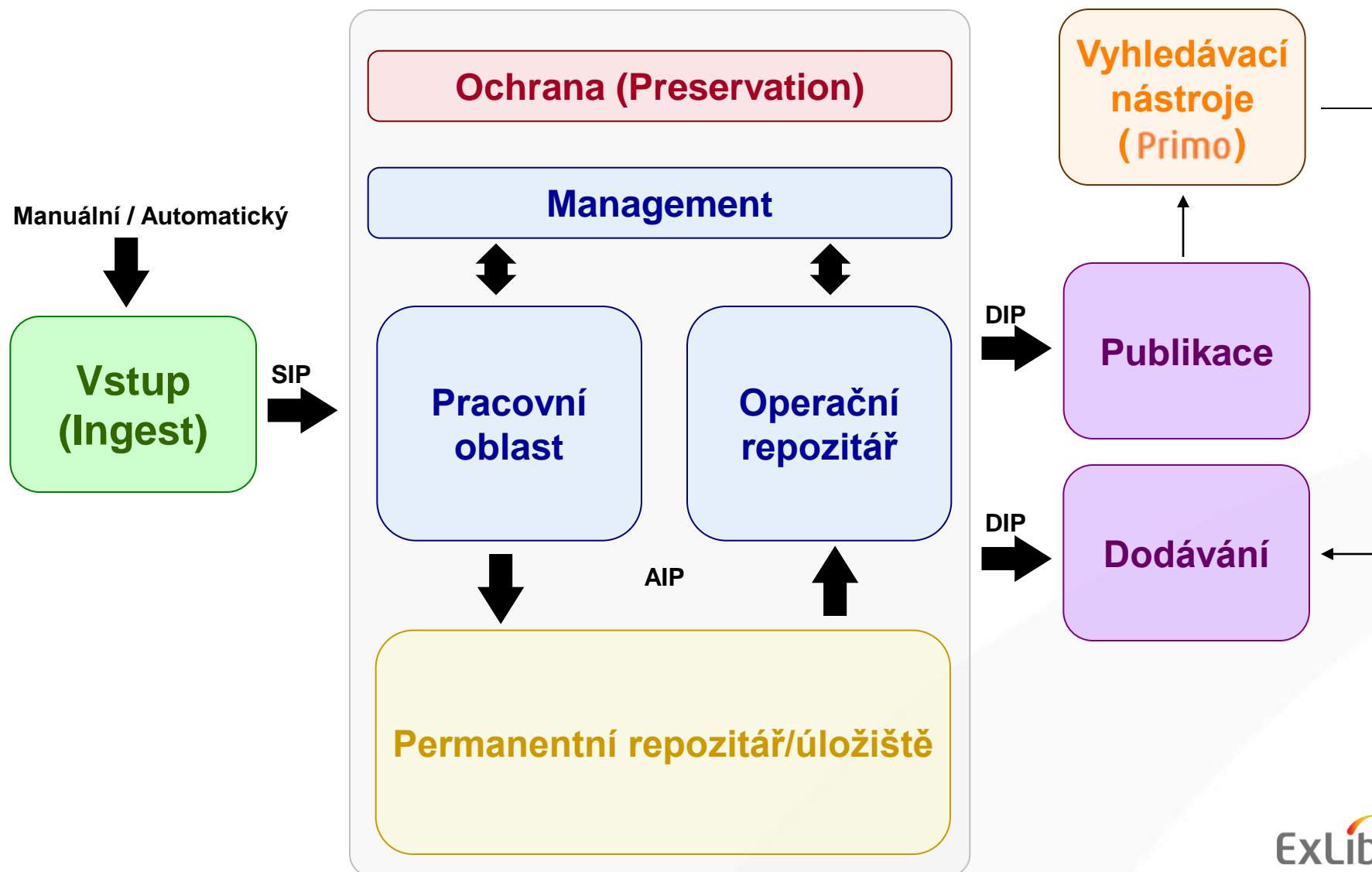
- Manuální tvorba
  - Obvykle při vzniku objektu
- **Automatická extrakce**
  - S využitím existujících nástrojů

# Využití

---

- **Vstup do systému**
  - Validace dat – jsou soubory tím, za co se vydávají?
  - Filtrace – chceme daný soubor?
  - Transformace - migrace, tvorba odvozených kopií
- **Výstup** – šíření objektů
  - Zobrazení
- **Uložení**
  - Identifikace zastaralých formátů
  - Hodnocení výsledků migrace
  - ...

# Rosetta - moduly



# Metadatové extraktory

---

- **JHOVE** - JSTOR/Harvard Object Validation Environment
- **JHOVE2**
- **Novozélandský metadatový extraktor**
  
- **DROID** – identifikace formátu

# Registry formátů

---

- **PRONOM**
  - Online registr formátů a SW nástrojů schopných tyto formáty zpracovat
  - Dobré technické zázemí
  - PUID PRONOM Unique Identifier
- **Unified Digital Format Registry**
  - Kooperativní tvorba dat
  - Leden 2012



Home > Preservation > Manage Local Libraries > Formats List > List of Formats

List of Formats

Find:  in: All

1 - 10 of 827 Formats 1 2 3 >>

	<u>Associated/ Disassociated</u>	<u>Name</u> ▲	<u>Description</u>	<u>Version</u>	<u>Classification</u>	<u>Registry Type</u>	<u>Registry ID</u>		
1	Disassociated	<a href="#">ExL-Fmt-21</a>	Finale Notation Form...	-	Generic	EX Global	ExL-Fmt-21	<a href="#">View</a>	<a href="#">Edit</a>
2	Disassociated	<a href="#">ExL-Fmt-22</a>	CDX File Format	-	Text	EX Global	ExL-Fmt-22	<a href="#">View</a>	<a href="#">Edit</a>
3	Disassociated	<a href="#">ExL-Fmt-23</a>	ENIGMA Transportable...	-	(	EX Global	ExL-Fmt-23		<a href="#">Edit</a>
4	Disassociated	<a href="#">ExL-Fmt-41</a>	Open Publication Str...	2.0 v1.0	€	EX Global	ExL-Fmt-41		<a href="#">Edit</a>
5	Associated	<a href="#">ExL-Fmt-61</a>	MPEG-4 Media File	-	(	PRONOM	<a href="#">fmt/199</a>		<a href="#">Edit</a>
6	Associated	<a href="#">ExL-Fmt-62</a>	Microsoft Office Ope...	2007	(	PRONOM	<a href="#">fmt/189</a>		<a href="#">Edit</a>
7	Disassociated	<a href="#">ExL-Fmt-241</a>	Genealogy GEDCOM Fil...	5.5	(ed)	EX Global	ExL-Fmt-241		<a href="#">Edit</a>
8	Disassociated	<a href="#">ExL-Fmt-261</a>	ARTIST 2 Interchange...	2.1.2	eBook	EX Global	ExL-Fmt-261	<a href="#">View</a>	<a href="#">Edit</a>
9	Associated	<a href="#">ExL-Fmt-24417</a>	Free Lossless Audio ...	-	Audio (NLNZ)	PRONOM	<a href="#">fmt/279</a>	<a href="#">View</a>	<a href="#">Edit</a>
10	Associated	<a href="#">fmt/1</a>	Broadcast WAVE	0	Audio (AES)	PRONOM	<a href="#">fmt/1</a>	<a href="#">View</a>	<a href="#">Edit</a>



**Back**

**Switch to Global**

# Rizika vážící se k formátům

---

- Zastarání formátu
  - Nejsou dostupné aplikace potřebné pro **zobrazení**
  - Nejsou dostupné aplikace potřebné pro **migraci**
- Formát má rizikovou **vlastnost**
- **Metadatový extraktor** nesprávně zpracovává daný formát

# Prevence rizik

---

- Využití **registrů formátů**
  - Získání informace o riziku
- **Kooperativní** tvorba znalostních bází
  - Šíření informací o riziku
- Včasné odhalení rizika
  - Průběžná **analýza** obsahu repozitáře
- Včasná reakce na riziko
  - Tvorba **plánu ochrany**

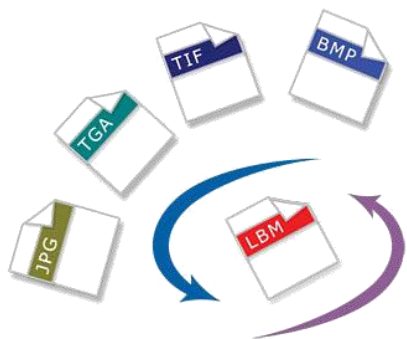
# Práce s riziky

---

- **identifikace** ohrožených objektů
- příprava **variant** plánu ochrany
- **testování** variant
- volba **optimální** varianty
- **provedení** plánu

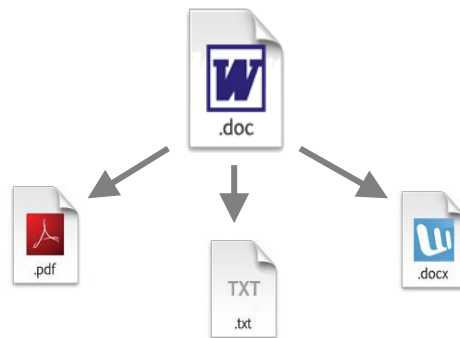
# Ochranné akce

Komplexní systém pro aktivní ochranu



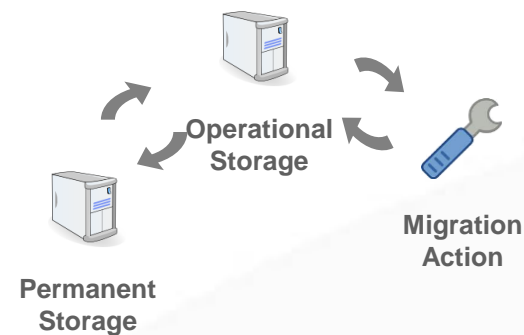
## Plánování ochrany

(formát + risk = plán ochrany)



## Hodnocení alternativ

(technická i obchodní kritéria)



## Provedení plánu

(vytvoření a uložení nové verze)

# Problémy (a řešení)

---

- Nedostatek informací o rizicích
  - Nutnost kooperace
- Technické nedostatky extrakčních nástrojů
  - Vývoj a implementace nových nástrojů či jejich verzí
- Nutnost automatizace procesu práce s riziky
  - Optimální software
- Podcenění významu prevence
  - Vzdělávání

---

[eliska.pavlaskova@multidata.cz](mailto:eliska.pavlaskova@multidata.cz)

