

# Mezi mlýnskými kameny aneb validace digitálního obsahu

PAVLÍNA NIMRICHTROVÁ  
NÁRODNÍ ARCHIV

---

ISSS

HRADEC KRÁLOVÉ, 13. 5. 2024



- **Validace digitálního obsahu**
- Datové formáty
- Výsledky nástrojů v praxi veřejných archivů
- Novinky ve validaci SIP balíčků z eSSL



# Validace digitálního obsahu

## Požadavky na primární využití informací - původce

Autenticita

Čitelnost

Důvěryhodnost

Správnost, úplnost, přesnost

Aktuálnost

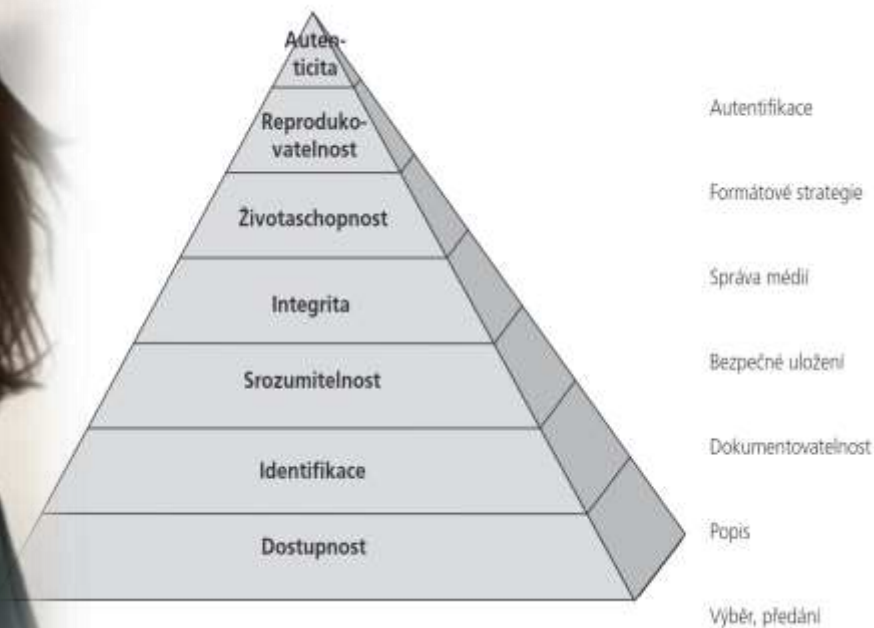
Dostupnost

Otevřenost

Bezpečnost



## Cíle dlouhodobého uchování



# Validace digitálního obsahu

- Ověření struktury, úplnosti
- Metadata/formáty/dokumentace
- Nyní existují informační balíčky jen pro eSSL, geodata, tzv. volné soubory
- Je pravděpodobné rozšíření eHEALTH, školy, justice ...

- Validace digitálního obsahu
- **Datové formáty**
- Výsledky nástrojů v praxi veřejných archivů
- Novinky ve validaci SIP balíčků z eSSL



# Datové formáty

- Identifikace pomocí formátového registru PRONOM  
<https://www.nationalarchives.gov.uk/PRONOM/Default.aspx> (nástroj DROID)

- Formátová pravidla NA

<https://portal.nacr.cz/cro/uvodni-stranka/formatova-pravidla-narodniho-archivu/>

- Národní standard formátů pro archivaci

[https://archi.gov.cz/znalostni\\_baze:archivni\\_formaty](https://archi.gov.cz/znalostni_baze:archivni_formaty)



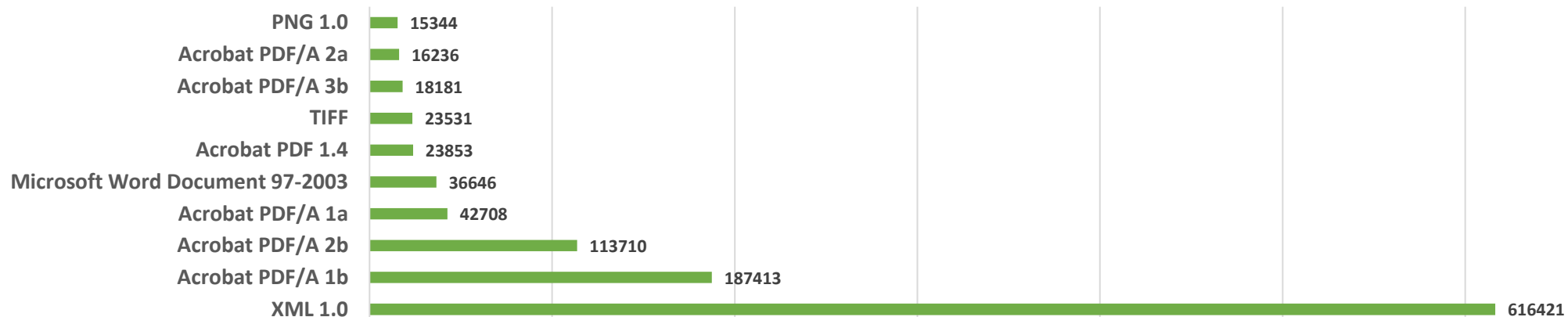
[Library of Congress](#)

formát	verze	PUID	kódování	standard	výstup eSSL	intermediát (mezanin)	archivace
Broadcast wave format (WAV/BWF)	2.00	fmt/703	modulace PCM	EBU-Tech 3285, DIN SPEC 15587	výstupní formát	ano	ano (*vhodné pro digitalizaci)
Free Lossless Audio Codec (FLAC)	1.2.1 a vyšší	fmt/279		FLAC Format Specification version 1.21. <a href="http://flac.sourceforge.net/format.html">http://flac.sourceforge.net/format.html</a>	§ 23 odst. 5	ano	ano
Waveform audio format (WAV)		fmt/6	modulace PCM	Multimedia Programming Interface and Data Specifications 1.0. IBM Corporation and Microsoft Corporation, August 1991. <a href="http://www.tactilemedia.com/info/MCI_Control_Info.html">http://www.tactilemedia.com/info/MCI_Control_Info.html</a> , Multimedia Data Standards Update April 15, 1994. <a href="https://www-nrmlsp.ecn.mcgill.ca/Documents/AudioFormats/WAVE/Docs/RIFFNEW.pdf">https://www-nrmlsp.ecn.mcgill.ca/Documents/AudioFormats/WAVE/Docs/RIFFNEW.pdf</a>	výstupní formát	ne	ano
MPEG-1 Audio Layer II nebo MPEG-2 Audio Layer II (MP2)		fmt/198		ISO/IEC-11172-3, ISO/IEC 13618-3	výstupní formát	ne	individuální posouzení digitálním archivem/NDÁ (*nevhodné pro digitalizaci)

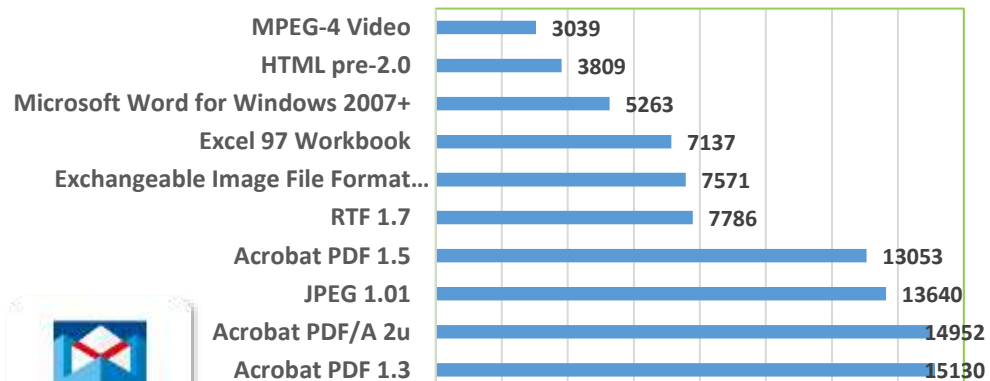
# Datové formáty

- Trvale uloženo 191267 AIP/originálních souborů 1217557 - 133 formátů dle PRONOM

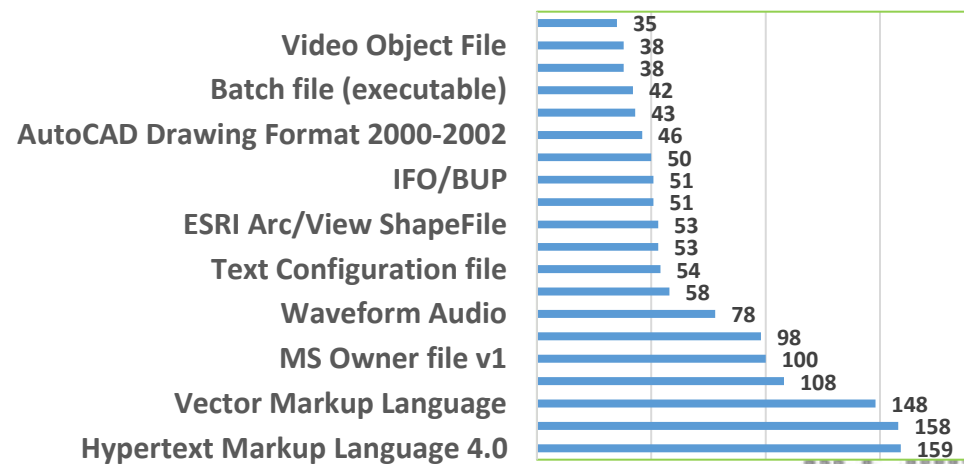
## 10 nejčastějších formátů



## Druhá desítka



## Pátá desítka



- Validace digitálního obsahu
- Datové formáty
- **Praxe veřejných archivů**
- Novinky ve validaci SIP balíčků z eSSL





# Praxe veřejných archivů

Aktuální problémy souvisejí s:

- dodáním komponent – ve výběru i při trvalém předání (nemožnost napárovat či ztotožnit dodané sady)
- diakritikou v názvech souborů
- použitím pevných křížových odkazů
- využíváním věcných skupin a spisových plánů a jejich aktualizace
- otázkou velkých balíků a souborů (např. audiovizuální)



- Validace digitálního obsahu
- Datové formáty
- Praxe veřejných archivů
- **Novinky ve validaci SIP  
2024**



# Novinky ve validaci SIP 2024

- Validace dle NSESSS 2017, 2012 se nemění; příp. pouze opravy chyb
- <https://validatorsip.nacr.cz/test/>
- Připomínky - Github, issues [https://github.com/NACR-CZ/Validator\\_SIP/wiki](https://github.com/NACR-CZ/Validator_SIP/wiki)
- WSDL služba
- Umístění
- Nekontroluje se množství
- RokSkartacniOperace
- Digitalizát, kontejner
- Problém se spisy v dílu
- Nyní přechodné období -> tolerují se obě varianty (sk. znak V, křížové odkazy do dílu, dokumentu, dokumenty samostatně)



# Děkuji za pozornost

PAVLÍNA NIMRICHTROVÁ

---

ISSS

PAVLINA.NIMRICHTROVA@NACR.CZ

