Výstupní datové formáty a uchovávací strategie v kontextu proměn správy dokumentů

Mgr. Zbyšek Stodůlka

Konference ISSS 2021, Hradec Králové

Výstupní datové formáty a eSSL

- Určení původci zajistí příjem dokumentů … alespoň v datových formátech stanovených jako výstupní datové formáty nebo formáty dokumentů, které jsou výstupem z autorizované konverze dokumentů obsažených v datové zprávě. Pokud veřejnoprávní původci umožní příjem dokumentů v digitální podobě také v dalších datových formátech … zveřejní jejich přehled na úřední desce nebo, nezřizují-li úřední desku, na svých internetových stránkách. § 64 odst 1 zák. č. 499/2004 Sb.

- Po vyřízení věci se spis uzavře. Uzavřením spisu se rozumí kompletace všech dokumentů patřících do spisu, kontrola a doplnění údajů podle § 66 odst. 3 před uložením do spisovny a převedení dokumentů v digitální podobě do výstupního datového formátu a jejich opatření metadaty podle národního standardu. § 65 odst 5 zák. č. 499/2004 Sb.

- Veřejnoprávní původce před uložením kontroluje uzavřený spis a vyřízený dokument, jsou-li úplné, zda jsou v evidenční pomůcce správně zpracovány všechny povinné údaje a zda jsou dodrženy podmínky uzavření spisu. Předmětem kontroly je zejména … převedení dokumentu v digitální podobě do výstupního datového formátu. § 19 odst. 2 písm. f) vyhl. č. 259/2012

- Výstupním datovým formátem dokumentů v digitální podobě se rozumí **a)** datový formát výstupu z elektronického systému spisové služby, **b)** datový formát dokumentu ukládaného ve spisovně, která je součástí elektronického systému spisové služby, **c)** datový formát pro předávání do digitálního archivu… § 23 odst. 1 vyhl. č. 259/2012

Výstupní datové formáty a eSSL

- eSSL umožňuje uživateli, který přijímá e-mailovou zprávu v proprietárním formátu, uložit ji ve více datových formátech, včetně otevřeného a výstupního. NSESSS VMV čá. 57/2017 požadavek 2.2.8
- Skenovací funkce eSSL ukládá snímky ve výstupních datových formátech uvedených v právním předpisu upravujícím podrobnosti výkonu spisové služby. NSESSS VMV čá. 57/2017 požadavek 2.3.2
- eSSL identifikuje datový formát komponenty a v případě, že se nejedná o výstupní datový formát podle prováděcího právního předpisu upravujícího podrobnosti výkonu spisové služby a) změní datový formát na výstupní v případě statických textových dokumentů, statických kombinovaných textových a obrazových dokumentů, statických obrazových dokumentů, dynamických obrazových dokumentů, zvukových dokumentů a databází… VMV čá. 57/2017 požadavek 2.3.9
- eSSL zajistí, aby komponenty, které nejsou ve výstupních datových formátech podle prováděcího právního předpisu upravujícího podrobnosti výkonu spisové služby, byly do těchto formátů ztvárněny postupem podle zákona a) při příjmu po provedení požadavku 2.1.15, b) při uzavření spisu, c) při uzavření typového spisu, součásti nebo dílu, d) při vyřízení dokumentu vloženého přímo do věcné skupiny. VMV čá. 57/2017 požadavek 5.3.1
- eSSL umožní uživateli zvolit, zda komponenty dokumentů ztvárňovaného spisu budou a) ponechány ve stávajícím datovém formátu, nebo b) ztvárněny do výstupního datového formátu podle prováděcího právního předpisu upravujícího podrobnosti výkonu spisové služby. VMV čá. 57/2017 požadavek 5.3.8
- eSSL při ztvárnění do výstupního datového formátu PDF/A použije verzi PDF/A-2 nebo vyšší v případě podepisování ztvárnění kvalifikovaným elektronickým podpisem. eSSL při ztvárnění do výstupního datového formátu PDF/A použije verzi PDF/A-2 nebo vyšší v případě podepisování ztvárnění kvalifikovaným elektronickým podpisem. VMV čá. 57/2017 požadavek 5.3.10
- Jestliže je pro ztvárnění do výstupního datového formátu použita verze PDF/A-3, eSSL neumožní do PDF vložit komponentu v datovém formátu, který není výstupním datovým formátem, pokud není současně vloženo ztvárnění těchto komponent ve výstupním datovém formátu. VMV čá. 57/2017 požadavek 5.3.11
- eSSL u datového formátu, který nemá definován prováděcím právním předpisem upravujícím podrobnosti výkonu spisové služby výstupní datový formát, datový formát nepřevádí. VMV čá. 57/2017 požadavek 5.3.12

Uchovávací strategie

- **bit-level ochrana** - nejjednodušší strategie uchovávání digitálních objektů, **vhodná pro krátkodobé až střednědobé uchovávání**, založená na **fyzické ochraně před ztrátou, změnou nebo havárií nosiče nebo souboru**. **Nemění se integrita digitálního objektu, který je možné migrovat z datového nosiče na jiný datový nosič nebo do nového technologického prostředí**. Nevýhodou strategie je možná absence možnosti reprodukce objektu v současném či budoucím prostředí.

- **logická ochrana digitálních dat** - spočívá v činnostech prováděných během životního cyklu digitálního objektu tak, aby byla **zajištěna trvalá použitelnost informačního obsahu.** Použitelnost je obecný pojem pro vyhledatelnost, zobrazitelnost, pochopitelnost a autenticitu obsahu. K jejímu zajištění **musí provozovatel digitálního repozitáře (např. eSSL/elektronické spisovny)/digitálního archivu sledovat technologický vývoj a přizpůsobit jeho změnám své procesy a postupy**. Spolu s datovými objekty musí být uchovávána také **odpovídající vysvětlující informace v podobě metadat o původu objektu, kontextu, integritě atd. která jsou v průběhu životního cyklu neustále doplňována**.

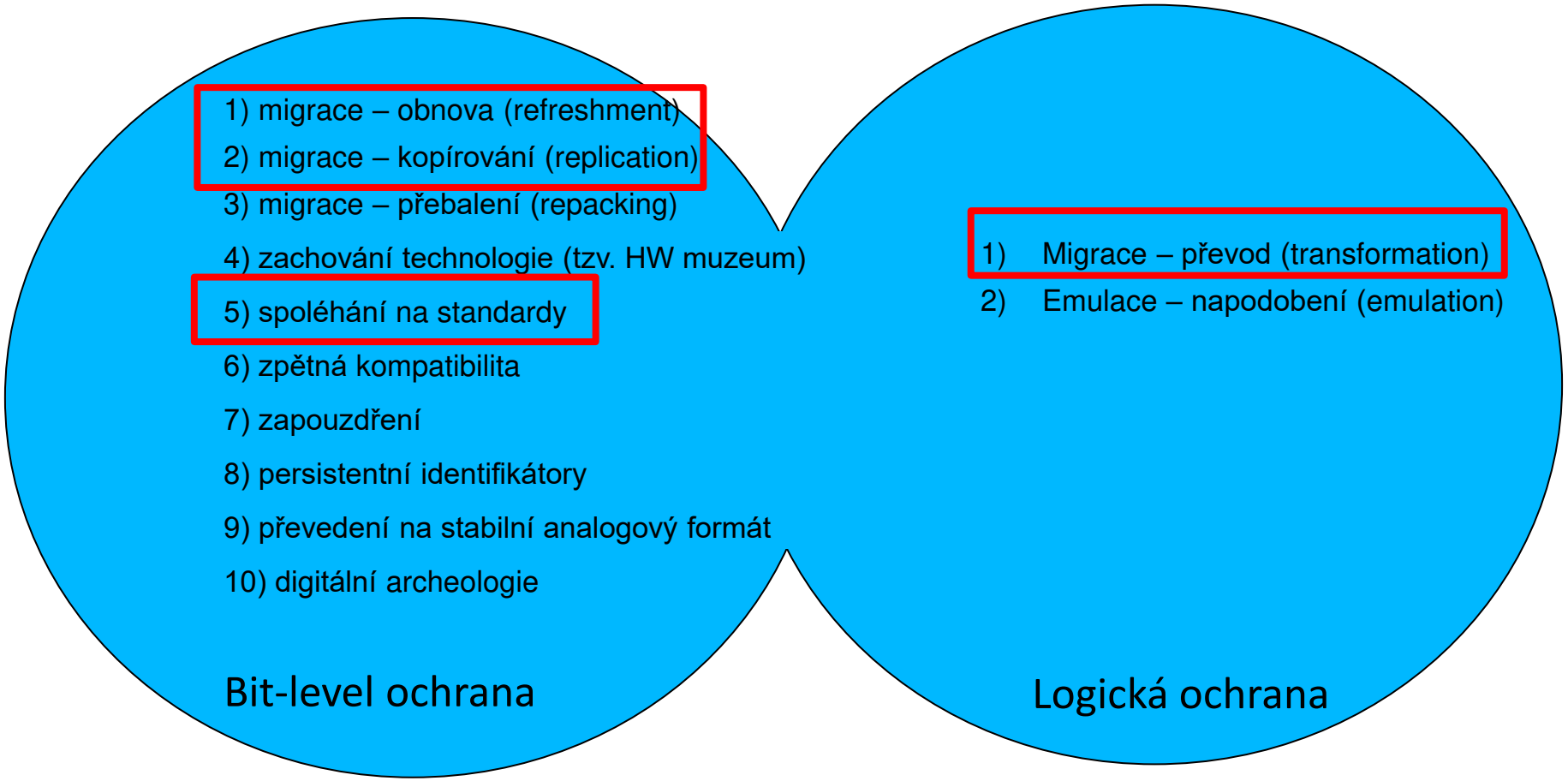
Uchovávací strategie

- Volba vhodné strategie uchovávání může být ovlivněna následujícími faktory:

- 1) proveditelnost (existuje software /hardware k provedení);
- 2) udržitelnost (jak dlouho dané řešení vydrží);
- 3) praktičnost (lze jej využít s rozumnými náklady);
- 4) vhodnost (jaké typy objektů uchovat a za jakým účelem)

- Podle délky uchovávání:
- 1) krátkodobé (3-5 let)
- 2) střednědobé (5-10 let)
- 3) dlouhodobé (10 let+)

Uchovávací strategie



Výhody a nevýhody VDF

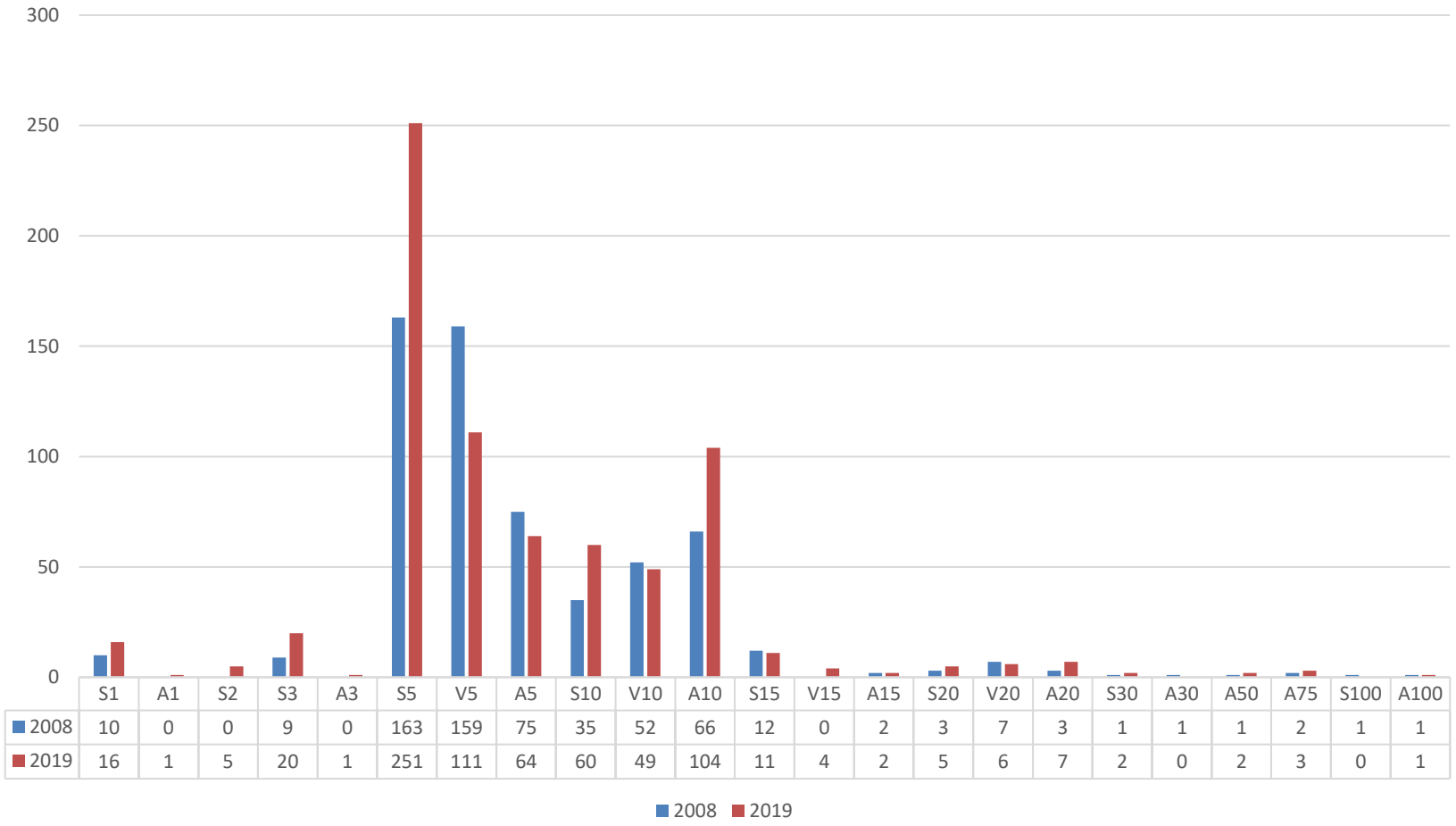
+	-
reprodukovatelnost (zobrazení) dokumentu občanem/původcem s využitím běžných softwarových nástrojů	původce vytvářející dokument musí disponovat dostatečně kvalitními softwarovými nástroji v eSSL (možnost sdílení služby)
podpora udržitelnosti uchovávání (i dlouhodobého) u původce, ev. po výběru za archiválie v archivu	
vizuální kontrola původcem při migraci (změně datového formátu)	
původce může omezit příjem formátů, které dokáže zpracovat	
možnost řešit problémové ztvárnění současnými technologickými prostředky (např. formulářové PDF s XFA není podporováno v PDF/A-4/PDF 2.0)	

Kritéria pro dlouhodobého uložení formátů (výběr)

Otevřenost	specifikace formátu je úplná a veřejně přístupná
Licenční omezení	použití formátu není licenčně omezeno
Rozšíření	formát je široce rozšířen, lze očekávat dlouhé migrační cykly
Funkčnost	formát reprezentuje podstatné vlastnosti ve své kategorii
Implementace	existuje vícero implementací daného formátu
Množství záznamu informace	formát umožňuje vysokou úroveň záznamu informace
Ověřitelnost	existují metody a nástroje pro rozpoznání formátu a jeho validaci
Složitost	formát je co možná nejméně složitý
Vlastní dokumentovatelnost	formát zahrnuje metadata a umožňuje jejich využití
Příklady dobré praxe	formát se již využívá pro dlouhodobé uchovávání (např. v archivech)
Perspektiva	formát je z archivního hlediska perspektivní i v budoucnosti a má potenciál

Skartační režimy (vzorový SSP ORP)

Počet věcných skupin dle skartačních režimů ve vzorovém SSP ORP 2008 a 2019



	S1	A1	S2	S3	A3	S5	V5	A5	S10	V10	A10	S15	V15	A15	S20	V20	A20	S30	A30	A50	A75	S100	A100
2008	10	0	0	9	0	163	159	75	35	52	66	12	0	2	3	7	3	1	1	1	2	1	1
2019	16	1	5	20	1	251	111	64	60	49	104	11	4	2	5	6	7	2	0	2	3	0	1

Skartační režimy (vzorový SSP ORP)

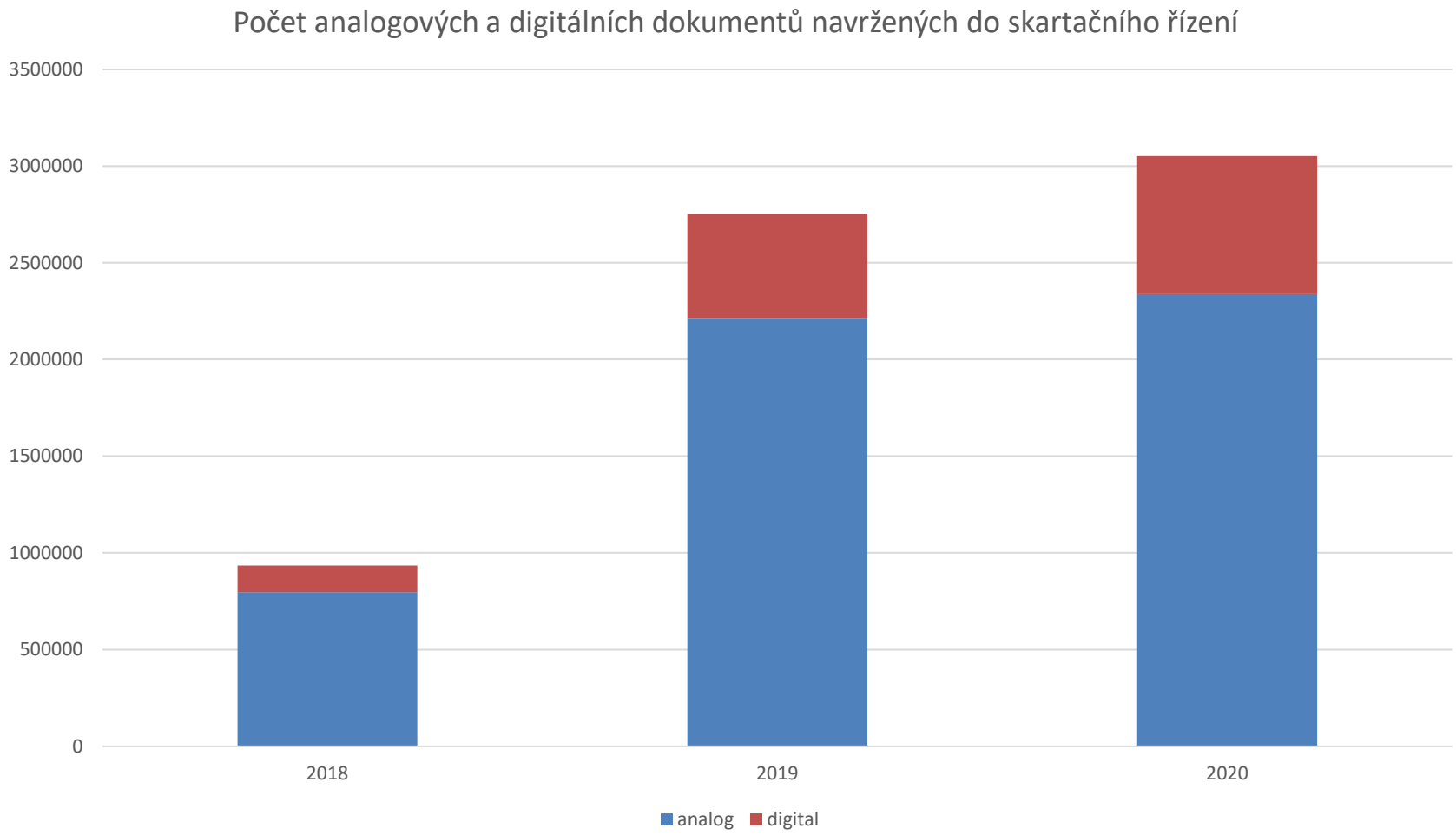
- u vzorového SSP 2019 dochází k precizaci věcných skupin pro výběr (+ 124 věcných skupin)

- vyšší počet věcných skupin, u kterých je definována spouštěcí událost = fakticky delší doba uchovávání dokumentu (např. po pořízení nové evidence; po dovršení plnoletosti dítěte; po skončení projektu; po odstranění stavby atd.)

- pokles využití skartačního znaku V; vyšší počet věcných skupin u nichž je skartační lhůta 10 a více let

- nutné zohlednit počátek běhu skartační lhůty od 1. ledna následujícího kalendářního roku, kdy nastala spouštěcí událost a návrh do skartačního řízení v roce následujícím po uplynutí lhůty – pokud dokument vyřízen v roce 2009 mající skartační lhůtu 10 let – do skartačního řízení možné navrhnout v roce 2021

Vývoj počtu navržených dokumentů dle formy



PDF/A

- formát vytvořený pro podporu dlouhodobého uchovávání
- oproti PDF omezena možnost změnit zobrazení (zakázány skripty, šifrování atd.)
- 3 úrovně požadavků (do PDF/A-3, PDF/A-4 dle standardu PDF/A-UA)
  - úroveň A: součástí je vždy textová vrstva, struktura, popis obrázků aj.
  - úroveň B: bez textové vrstvy, pouze ztvárnění
  - úroveň U (Unicode): od PDF/A-2; textová vrstva v kódování Unicode

Verze	Odvozeno z	Změny
PDF/A-1 (ISO 19005:2005)	PDF 1.4	výchozí verze
PDF/A-2 (ISO 19005:2011)	PDF 1.7	podpora JPEG 2000; fonty OpenType; podpora elektronického podpisu PAdES
PDF/A-3 (ISO 19005:2012)	PDF 1.7	vložení jiného souboru (XML, DOC, XLS aj.)
PDF/A-4 ISO (19005-4:2020)	PDF 2.0	vložení komponent bez vizuální reprezentace jen ve specifickém profilu (f); nový profil (e) pro 3D objekty (nahrazuje PDF/E)

Závěr

- formátová normalizace je jedním z efektivních nástrojů elektronizace správy dokumentů v ČR

- v rámci eSSL nutné pravidelně kontrolovat a vyhodnocovat plnění legislativních požadavků i efektivitu/nákladovost výkonu spisové služby (např. původce „tiskne“ desetitisíce formulářů ve formátu FO, neboť jej eSSL neumí převést do VDF)

- sladění formátů pro příjem/odeslání/archivaci není bez obtíží – např. pro archivaci relačních databází využívá Národní archiv formát SIARD a pro ukládání GIS dat připravuje GeoSIP (export u obojího nutné kontrolovat)

- jako pomůcku pro původce zveřejňuje Národní archiv Formátová pravidla, kde i doporučení pro alternativní využití jiného formátu (https://validatorsip.nacr.cz)

- při další úpravě legislativy snaha o precizaci podmínek vkládání příloh u formátu PDF/A-3 (dnes v některých eSSL nutné manuální vytěžování; riziko obcházení množiny přijímaných formátů; kyberbezpečnost – některé antivirové programy nekontrolují přílohy atd.); doplnění audiovizuálního formátu MPEG-4, strojové čitelnost doplněním textové vrstvy u vyhotovovaných dokumentů

Díky za Vaši pozornost!

zbysek.stodulka@nacr.cz