

# SIGNIFIKANTNÍ VLASTNOSTI: PŘÍSPĚVEK KE KOLEKTIVNÍMU NEVĚDOMÍ

## SIGNIFICANT PROPERTIES: CONTRIBUTION TO COLLECTIVE UNKNOWN

**Miroslava Beňačková, Pavlína Kočišová, Vojtěch Kopský,  
Frank Richard Mally, Natalie Ostráková**

*Národní knihovna ČR*

### **Abstrakt**

*Účel* – Článek pojednává o problematice signifikantních vlastností digitálních objektů a stanovuje si za cíl tento koncept aplikovat v rámci Národní digitální knihovny. V textu jsou představeny jak různé přístupy k otázce signifikantních vlastností a jejich určování, tak i jejich různé druhy. Cílem je ale pokusit se určit data-centrické technické vlastnosti zvukových a obrazových souborů v rámci NDK, které je vhodné uchovat a představit návrh jejich zápisu. Zařazením se jedná o výzkumnou studii.

*Design/metodologie/přístup* – Metodologicky se jedná o analýzu. Shrnuje rozličné přístupy k signifikantním vlastnostem v mezinárodních institucích a postupy při jejich stanovení. Blíže pojednává o možnostech jejich zápisu do standardu PREMIS a věnuje se také duplicitě informací. V návaznosti na to předkládá výčet technických vlastností, které jsou námi považovány za signifikantní.

*Výsledky* – Analýza technických vlastností v rámci standardů NDK a signifikantních vlastností doporučených jinými institucemi umožnila určení signifikantních vlastností pro zvukové a obrazové soubory jednotlivých objektů v rámci NDK. Tento výčet může být v budoucnu použit jako podklad při případné implementaci signifikantních vlastností do standardů NDK.

*Originalita/hodnota* – Přínosem článku je ucelený pohled na problematiku signifikantních vlastností a vytvoření podkladů k dalšímu vývoji standardů NDK. Zmiňuje také další typy signifikantních vlastností, kterým je potřeba se v budoucnu věnovat.

*Klíčová slova*: signifikantní vlastnosti, významné vlastnosti, PREMIS, Národní digitální knihovna, digitální archivace, technická metadata, zvukové soubory, obrazové soubory

### **Abstract**

*Purpose* – This article deals with the matter of significant properties of digital objects and it aims to apply this concept within the NDL. In the text, we present various approaches to the question of significant properties and how to identify them as well as their different kinds. The goal is to try to identify data-centric technical properties of audio and image files within NDL, which need to be preserved and to present a way how to record them. It can be classified as a research study.

*Design/Methodology/Approach* – Regarding used methodology, this article is an analysis. It sums up different approaches of international institutions to the significant properties and the ways of identifying them. It deals more closely with possibilities of recording them in PREMIS and it addresses duplicity of information, too. Building on that, it presents a list of technical properties which we consider significant.

*Results* – The analysis of the technical properties within NDL standards and the significant properties recommended by other institutions allowed identification of the significant properties for audio and image files for single objects within NDL. This list can be used in the future as a groundwork for potential implementation of significant properties to the NDL standards.

*Originality/Value* – The benefit of the article is a comprehensive overview of the matter of significant properties and creation of groundwork for future development of NDL standards. It also mentions other types of significant properties, which need to be addressed in the future.

*Keywords*: significant properties, significant characteristics, PREMIS, National Digital Library, digital archiving, technical metadata, audio files, image files

## Úvod

Signifikantní vlastnosti, též významné či klíčové vlastnosti, digitálních objektů jsou důležitým nástrojem v oblasti dlouhodobého uchovávání kulturního dědictví. Jedním z cílů digitální archivace je uchovat významné vlastnosti archivovaných digitálních objektů pro další generace, a to v průběhu času a napříč ochrannými opatřeními (migrace). Stanovování signifikantních vlastností tak představuje určení vlastností objektů, které je třeba v co největší míře při migracích zachovat. Identifikace signifikantních vlastností je ovšem poměrně komplexní aktivita, která zahrnuje i reflexi potřeb cílové komunity. V naší studii se pokusíme pro tuto analýzu shrnout současná teoretická východiska a přestavit naše dosavadní výsledky, zejména na poli určování signifikantních vlastností pro rastrové obrazy a zvukové dokumenty.

## 1 Související terminologie

V českém prostředí se pro označení vlastností sledovaných v této studii setkáváme s několika termíny, resp. s několika překlady a anglických termínů „significant properties“ a „significant characteristics“. Cubr (2017) termín překládá jako „signifikantní vlastnosti“, Konečný (2018, 2019) pracuje s termínem „významné vlastnosti“. Oba překlady považujeme za vhodné a zaměnitelné, nicméně pro potřeby tohoto textu budeme užívat termín „signifikantní vlastnosti“, který je pro tuto práci, zaměřenou v první polovině na čistě teoretická východiska celého konceptu, vhodnější i z hlediska nezaměnitelnosti s podobnými výrazy.

V kontextu dlouhodobého uchovávání se pojem významné vlastnosti (v originále jako *essential characteristics*) objevil poprvé na konci 90. let 20. století, a to konkrétně například v textech autorů Rothenberga a Biksona (1999) a u Clifforda Lynche (1999). Rothenberg a Bikson tehdy uvedli, že pokud je záměrem, aby si objekt (v originále záznam) během celé své existence v dlouhodobé archivaci podržel svoji funkčnost (ve všech ohledech) a základní charakteristiky, potom musí být tyto základní charakteristiky identifikovány a záměrně uchovávány.

Koncept signifikantních vlastností byl ústředním tématem v několika výzkumných projektech na poli digitální archivace především v první dekádě 21. století. Jako příklad uvedme CEDARS projekt (1998-2001), InSPECT (2007-2009) a PLANETS (2006-2010).

K označení vlastností, jež dnes nazýváme signifikantní, významné či klíčové bylo v průběhu času používáno několik označení, různých termínů. Knight (2010a) a metadatový standard PREMIS pracují s termínem “significant properties”, Angela Dappertová a další používají termín “significant characteristics”, dále je možné setkat s “essential properties” (e.g. Thibodeau 2002; viz též Vermaaten, Lavoie, and Caplan 2012, kde je tentýž termín použitý v odlišném významu) a “essential characteristics” (e.g. Rothenberg and Bikson 1999, Lynch 1999, Heslop, Davis, and Wilson 2002). V normě OAIS jsou tyto vlastnosti dle Giaretta a kol (2009) označeny pojmem transformační vlastnost informací (*transformational information property*), jež jsou zde definovány jako vlastnosti, u kterých „je uchování jejich hodnoty pokládáno za nutné, ale ne dostačující, aby bylo možné ověřit, zda při jakékoliv nevratné transformaci byl zachován informační obsah“ (ISO 14721:2012 in Cubr, 2017, s. 90).

Faniel a Yankel (2011) používají termín „contextual metadata“ k označení signifikantních vlastností (Faniel, Yankel, 2011), tento termín se ale spíše používá pro označení metadat osvětlujících okolnosti vzniku dokumentu (kulturní, technické apod.) a plně se nepřekrývá s obsahem pojmu signifikantních vlastností, tak jak je dnes obvykle chápán.

V anglicky psané odborné literatuře se nejčastěji setkáváme s termíny „significant properties“ (viz projekty PLANETS, CEDARS, InSPECT či metadatový standard PREMIS) a „significant characteristics“ (viz Dappert a Farquhar, 2009). Dappertová a Farquhar (2009) však jasně rozlišují mezi označením „characteristics“ a „properties“, kde „characteristics“ používají pro označení vztahu mezi vlastností (property, např. bitová hloubka, počet otáček za minutu) a konkrétní hodnotou této vlastnosti (např. 16 bitů nebo 76 rpm).

Stejně jako v případě pojmenování existuje také více různých definic, které signifikantní vlastnosti vysvětlují. Nejčastěji citovanou definicí je dle Konečného (2019) definice Andrewa Wilsona, která je použita v projektu InSPECT:

“vlastnosti digitálních objektů, které musí být v průběhu času času udržovány, aby byla zabezpečena trvalá dostupnost, použitelnost a význam těchto objektů” (“the characteristics of digital objects that must be preserved over time in order to ensure the continued accessibility, usability, and meaning of the objects”) (Wilson 2007).

Další vybrané definice:

OCLC/RLG 2002:

“vlastnosti zobrazovaného datového objektu s obsahem, jenž musí být uchovány a spravovány napříč cykly konzervačních procesů.” (“properties of the Content Data Object’s rendered content which must be preserved or maintained during successive cycles of the preservation process”)

Grace, Knight, and Montague 2009 (překlad dle Konečný, 2019):

„vlastnosti informačního objektu, které musí být v průběhu času zachovány, aby byla zajištěna trvalá přístupnost, použití a význam těchto objektů, a jejich způsobilost vypovídat o původním účelu“. (“the characteristics of an Information Object that must be maintained over time to ensure its continued access, use, and meaning, and its capacity to be accepted as evidence of what it purports to record”)

PREMIS 2015:

“vlastnosti konkrétního objektu subjektivně určené za natolik důležité, aby byly udržovány napříč archivačními opatřeními.” “characteristics of a particular object subjectively determined to be important to maintain through preservation actions”

Signifikantními vlastnosti a jejich stanovením se zabývalo mnoho projektů. Některé jiné se jimi přímo nezabývaly, ale přesto svými poznatky k jejich výzkumu nepřímo přispěly; například CURL Exemplars in Digital Archives (CEDARS), působící mezi lety 1998 a 2002. Z výstupu tohoto projektu si teorie signifikantních vlastností bere zejména poznatek o potřebě výběru vhodných formátů dat, které zajistí při migracích zachování a úplnost vlastností, jež si instituce stanoví jako signifikantní. Projekt CAMELEON (1999-2003) se zabýval emulacemi, migracemi dat a užitím formátu XML jakožto prezervační strategie.

Pro výzkum signifikantních vlastností v kontextu strategie digitální ochrany byl pak důležitý zejména projekt DELOS a (2004-2007) a projekt PLANETS (2006-2010). Poslední jmenovaný byl přímo zaměřen na vývoj řešení k automatizaci procesů identifikace, validace a extrakce základních vlastností digitálních objektů a metodologie, jak toho dosáhnout. Kromě toho byly v rámci projektu vyvíjeny nástroje a služby, které by toto umožnily u širokého spektra objektů a registr PRONOM, který by tyto charakteristiky, reprezentativní vlastnosti a informace, uchovával. Tento registr v současnosti užívá většina významných institucí v oboru; např. Britská národní knihovna, Národní archivy Spojeného království, Národní knihovna Nizozemí, Microsoft Research Limited, a mnoho dalších.

Přímo výzkumem signifikantních vlastností se pak zabývaly projekty Investigating the Significant Properties of Electronic Content over Time (InSPECT, 2007-2009), o kterém se v této práci zmiňujeme na různých místech, a projekt InterPARES 1 a 2 (International Research on Permanent Authentic Records in electronic Systems, 1999-2001, 2002-2007).

## 2 Stanovení signifikantních vlastností

Metodické přístupy, jež se používají při stanovení signifikantních vlastností, je možné rozdělit do dvou základních skupin přístupů: na přístupy zaměřené na data (data-centrické) a přístupy zaměřené na člověka. Přehled přístupů s tím, jak se v průběhu času objevovaly v odborné literatuře je znázorněn v následující tabulce:

Zaměřený na data	Zaměřený na člověka
= analýza dat užito např. v InterPARES 1 & 2	= analýza u zúčastněných stran (tzv. stakeholderů) - průkopníky jsou Rothenberg a Bikson 1999
→ používá se v odborné literatuře k digitálnímu kurátorství	→ používá se v literatuře o opakovaném užití dat (data reuse)
zaměřuje se primárně na vlastnosti digitálního objektu	zaměřuje se primárně na producenty a určené komunity; - zdůrazňuje vlastnosti, které jsou důležité pro udržování významu obsahu v čase a umožňující určené komunitě rekonstruovat význam vlastností
věnuje se hlavně technickým vlastnostem pro zpřístupňování, zobrazování digitálních objektů	věnuje se převážně vlastnostem důležitým pro uchování významu v čase
<b>Signifikantní vlastnosti, zmiňované v literatuře:</b>	
<ul style="list-style-type: none"> <li>• funkcionalita, vztahy mezi daty, vzhled (Coyne at al. 2007);</li> <li>• vzhled a funkce/atmosféra (look and feel) (Hedstrom et al. 2006; Matthews et al. 2008)</li> <li>• informace o právech (proprietární algoritmy nebo copyright) (Ashley, Davis, Pinsent 2008; Matthews et al. 2008)</li> <li>• účel a užití (Ashley, Davis, and Pinsent 2008)</li> <li>• výpočetní prostředí (Morrissey 2010)</li> <li>• záměr tvůrce (Coyne a Stapleton 2008)</li> <li>• užití (Morrissey 2010)</li> </ul>	<ul style="list-style-type: none"> <li>• čištění dat a manipulační procesy (Carlson and Anderson 2007)</li> <li>• souvislosti vzniku a úplnost metod pro sběr dat (Faniel and Jacobsen 2010)</li> <li>• uvedení informací o kontrole kvality dat (Carlson and Anderson 2007)</li> <li>• úroveň dokumentace (Corti 2007)</li> <li>• standardy pro dokumentaci (Faniel 2009)</li> <li>• identita tvůrce dat/sběratele dat (Knorr-Cetina 1999; Van House 2002; Van House, Butler, and Schiff 1998; Zimmerman 2008)</li> <li>• práva subjektů a lidí, spojených s objekty (Carlson a Anderson 2007)</li> </ul>

Tab. 1 Přístupy ke stanovení signifikantních vlastností

S ohledem na dlouhodobou archivaci a na to, jak je zřejmě použití funkcionality signifikantních vlastností v této oblasti zamýšleno, tj. pro uchování a zpřístupnění digitálních objektů v podobě co nejpodobnější jejich původní podobě a s respektem k jejich původnímu účelu, je optimální kombinace obou přístupů, tj. analýza archivovaných dat i analýza uživatelů, tzv. stakeholderů (zájmová skupina, též cílová skupina). Analýzu objektů a stakeholderů doporučuje i framework InSPECT, jež Konečný (2019, s. 11) považuje nejucelenější a nejpokročilejší aktivitu věnující se signifikantním vlastnostem a jejich stanovení. Jedná se o britský projekt probíhající v letech 2007-2009, v jeho rámci bylo definováno pět kategorií vlastností a to:

- Obsah (Content)
- Kontext (Context)
- Struktura (Structure)
- Reprodukce (Rendering)
- Chování (Behaviour)

S touto kategorizací signifikantních vlastností pracuje též PREMIS 3 (s. 51), který ji využívá pro ilustraci možného plnění elementů popisujících signifikantní vlastnosti. Pro kategorii Obsah tak například uvádí jako příklad zápis „all textual content and images“, tj., že se jedná o textový dokument s obrázky. Pro kategorii Chování uvádí jako příklad hodnotu “editable”, tj., že je objekt editovatelný, má funkci editace.

Pro identifikaci a klasifikaci signifikantních vlastností pak Framework Inspect uvádí následující základní kroky (podle Stephena Grece /InSPECT/, v Hockx-Yu a Knight 2008):

- 1) Definice intelektuálních složek dokumentu
- 2) Identifikace technických vlastností (požadovaných k opětovnému vyvolání obsahu) každé definované složky z prvního kroku
- 3) objasnění funkce každé vlastnosti z druhého kroku a posouzení její relativní hodnotu/významu
- 4) zhodnocení (měření) každé vlastnosti za použití předem definované hodnotící metody

Knight (2009) pak jednotlivé kroky stanovení signifikantních vlastností v rámci projektu InSPECT dále rozpracovává. Výsledkem je podrobnější postup zahrnující jak analýzu datových objektů, tak i analýzu stakeholderů.

Při analýze objektů hodnotitel studuje samotný objekt (musí mít k dispozici reprezentativní vzorek), technickou specifikaci nebo standard, jež popisují strukturu a složení objektu a používá vhodné nástroje pro analýzu objektu. Analýze se podrobuje buď obecný typ objektu (audio soubor, video soubor, obraz) nebo specifitější typ objektu např. mluvené slovo vs. hudební zvukový záznam. U hodnocených objektů zjišťuje obvykle technické vlastnosti, jako je např. formát souboru, velikost v bytech, použité kodeky, použitý kompresní algoritmus. Tato analýza probíhá zpravidla s použitím nástrojů pro extrakci metadat.

Následně se zjištěné technické vlastnosti přiřadí dle své role do jedné z následujících kategorií:

- Obsahové vlastnosti „vztahují se k informačnímu obsahu (text, statické obrazy, pohyblivé obrazy)“ (Konečný, 2019), příkladem vlastnosti je například počet slov, počet znaků, ilustrace.
- Kontextové vlastnosti popisují souvislosti vzniku informačního obsahu, příkladem vlastnosti je například jméno autora, vydavatel.
- Poskytovací (rendering) vlastnosti zahrnují informace, jež jsou nutné pro zpřístupnění, poskytnutí obsahu v podobě, v jaké byl zamýšlen. Jedná se například o použité písmo (font, barva, velikost).
- Strukturální vlastnosti popisují vztahy mezi částmi uvnitř objektu, které jsou nezbytné pro zobrazení objektu, tak jak byl zamýšlen. Příkladem je email a jeho přílohy, titulky k filmu apod.

(Knight in Konečný, 2019)

V dalším kroku se hodnotitel zamýšlí nad různými způsoby, jakými může uživatel objekt užít a vytváří seznam těchto aktivit, jež zároveň determinují očekávané chování, odpovědi objektu. Jako příklad uvádí Konečný (2009) následující chování uživatele:

- chce si zobrazit historický text v původní podobě
- chce text statisticky analyzovat
- chce použít ilustrace ve své vědecké práci

Zjištěné chování, jež může být od objektu vyžadováno, se následně rozpracuje do potřebných funkcí a funkcionalit objektu.

Jak bylo výše uvedeno, vedle analýzy dat by měly být pro určení signifikantních vlastností zvažovány i potřeby uživatelské komunity. V procesu identifikace signifikantních vlastností dle projektu InSPECT tedy následuje analýza stakeholderů. Hodnotitel by tedy měl identifikovat stakeholdery, kteří mají nějaký vztah k informačnímu objektu, a měl by zjistit a určit funkce/chování, jež uživatelé od objektu očekávají. Oslovit vybrané stakeholdery je možné pomocí běžných dotazovacích metod (dotazníky, rozhovory, pozorování). Zvolí se digitální objekty, které stakeholderi obvykle užívají, a hodnotitel s nimi nad těmito objekty provádí např. rozhovor, který se věnuje především funkcím objektů. V dalším kroku je cílem určit skutečné chování objektu, tj. co pravděpodobně budou stakeholderi s objektem dělat. Může se jednat o podmnožinu předpokládaného chování a užití objektu. V tomto kroku mohou být stakeholderi požádáni o demonstraci, jak objekty používají.

Chování je dále klasifikováno do funkcí, jimiž by měl objekt disponovat. Tyto funkce se porovnají s funkcemi, zjištěnými analýzou objektu, a z výsledku porovnání vznikne seznam technických vlastností, které jsou nezbytné, aby objekt podporoval uvedené funkce.

Pro jednotlivé vlastnosti se pak určí hodnoty (rozsah, minimum, maximum apod.), jichž může nabývat, a které jsou pro stakeholdera přijatelné. Tyto hodnoty pak mohou sloužit k zjišťování, případně měření,

zda má po transformaci objekt stále signifikantní vlastnosti, a do jaké míry a zda je tedy akceptovatelný pro stakeholdera. (Konečný, 2019)

### 3 Zaznamenání signifikantních vlastností v metadatech

Aby bylo možné posoudit, zda jsou v průběhu času a napříč ochrannými opatřeními zachovány signifikantní vlastnosti digitálních objektů, je potřeba je zaznamenat do metadat, jež tyto digitální objekty doprovázejí. Přímou s pojmem signifikantní vlastnosti (significant properties) pracuje metadatový standard PREMIS již od své první verze. V aktuální verzi standardu PREMIS 3.0 je možné signifikantní vlastnosti popsat u všech úrovní objektů, tj. u intelektuální entity, reprezentace, souboru a bitového toku pomocí sémantické jednotky nazvané significantProperties, zahrnující tyto elementy:

- 1) significantPropertiesType: aspekt, fasety nebo atribut objektu, jehož je záměrem uchovat. Příkladem hodnot jsou: obsah, struktura, chování, počet stránek, šířka strany apod. Tj., jedná se o název vlastnosti.
- 2) significantPropertiesValue: v tomto elementu se zaznamená podoba této vlastnosti u konkrétního popisovaného objektu, tj. například číselně se zapíše počet stran, pro vlastnost “chování” se jako hodnota může uvést “editovatelný” apod.
- 3) significantPropertiesExtension: umožňuje vložení externího metadatového schématu je-li potřeba.

Metadatový standard PREMIS tak nabízí velkou flexibilitu při zápisu signifikantních vlastností:

- Umožňuje posuzovat objekty dle různých aspektů a jejich různé atributy a je tak široce použitelný.
- Umožňuje popsat vlastnosti měřitelné (počet stran, šířka strany) i subjektivně považované za důležité.
- Umožňuje popsat vlastnosti na různém stupni podrobnosti (např. “textový obsah s obrázky” vs. počet stran textu, počet obrázků).

Na druhou stranu ale standard PREMIS není pro uživatele v popisu signifikantních vlastností příliš návadný. Ve srovnání s plněním ostatních sémantických jednotek, které mají obvykle přesně definovaný obsah (což znamená, že napříč institucemi a aplikacemi standardu PREMIS je malá variabilita), je popisování signifikantních vlastností výrazně variabilní a odvislé od lokální implementace a přístupu k signifikantním vlastnostem – a to i v případě paměťových institucí stejného typu. V praxi (možná proto) není tato sekce standardu PREMIS příliš užívána a definování signifikantních vlastností se považuje za obtížné. Objeví-li se v této oblasti nějaký postup, který bude komunitou označen za příklady dobré praxe, pak je možné očekávat, že se tímto budou paměťové instituce stejného typu inspirovat.

Signifikantními vlastnostmi, které konkrétní archiv u svých objektů zaznamenává, indikuje, že má v úmyslu tyto vlastnosti sledovat a dlouhodobě udržovat spíše než ty, které explicitně neuvede.

## 4 Užití standardu PREMIS pro zápis signifikantních vlastností ve vybraných zahraničních institucích

Výzkum zahraniční praxe naznačuje, že užití výše představených elementů standardu PREMIS pro popis signifikantních vlastností není příliš časté.

Digitální knihovna HathiTrust používá pro zápis archivačních metadat standard PREMIS ve verzi 1.0 a 2.0 (plánují migraci z verze 1.0 do verze 2.0.) S verzí 2.0 implementovali i elementy pro popis signifikantních vlastností, které používají pro popis úrovně reprezentace, tj. SIP balíčku a zaznamenávají počet souborů v balíčku a počet stránek (viz. obr. 1). (HathiTrust, 2015)

```
<PREMIS:object xsi:type="PREMIS:representation">
  <PREMIS:objectIdentifier>
    <PREMIS:objectIdentifierType>identifier</PREMIS:objectIdentifierType>
    <PREMIS:objectIdentifierValue>dul1.ark:/13960/t13n2vj0t</PREMIS:objectIdentifierValue>
  </PREMIS:objectIdentifier>
  <PREMIS:significantProperties>
    <PREMIS:significantPropertiesType>file count</PREMIS:significantPropertiesType>
    <PREMIS:significantPropertiesValue>960</PREMIS:significantPropertiesValue>
  </PREMIS:significantProperties>
  <PREMIS:significantProperties>
    <PREMIS:significantPropertiesType>page count</PREMIS:significantPropertiesType>
    <PREMIS:significantPropertiesValue>320</PREMIS:significantPropertiesValue>
  </PREMIS:significantProperties>
</PREMIS:object>
```

Obr. 1 Použití elementů v HathiTrust

V roce 2015 prezentoval Nizozemský institut pro zvuk a obraz (The Netherlands Institute for Sound and Vision, NISV) na workshopu k implementaci standardu PREMIS při konferenci iPress svojí implementaci standardu PREMIS pro repozitář počítačových her. Signifikantní vlastnosti zjišťují a zaznamenávají, aby mohli kontrolovat autenticitu digitálních objektů po migraci. Implementovali jednak odpovídající části standardu PREMIS i vlastní externí schéma, jež umisťují do elementu significantPropertiesExtension. Externí schéma umožňuje zaznamenat typ informace (zvuk, obraz, video, software apod.), roli objektu v archivu (archivní kopie, náhled apod.), popisná metadata, délka záznamu, počet pixelů, barva objektu, rozměry objektu, zda je obsažena zvuková stopa apod. (De Vos, 2015; Steeman, 2018)

Z Registru implementací standardu PREMIS (<http://www.loc.gov/standards/premis/registry/index.php>) je možné zjistit, že v minulosti sledované elementy používaly i další instituce. Například Národní knihovna ve Walesu ve vzorovém záznamu z roku 2008, který do registru umístila jeden z elementů následujícím způsobem (viz obr. 2).

```

▼<premis:formatRegistry>
  <premis:formatRegistryName>PRONOM</premis:formatRegistryName>
  <premis:formatRegistryKey>fmt/10</premis:formatRegistryKey>
  <premis:formatRegistryRole>specification</premis:formatRegistryRole>
</premis:formatRegistry>
</premis:format>
<premis:significantProperties>internal content only</premis:significantProperties>
</premis:objectCharacteristics>
▼<premis:relationship>

```

Obr. 2 Použití elementů v Národní knihovně Walesu

V roce 2008 pracovali s verzí 1.0 standardu PREMIS, který pro zápis signifikantní vlastnosti zatím obsahoval pouze jediný dále nestrukturovaný metadatový element, který použili zřejmě k zápisu informace, že popisovaný objekt nemá žádné externí vazby. (National Library of Wales, 2008)

Švédský národní archiv, jehož vzorový záznam je možné v Registru implementací standardu PREMIS také nalézt, v roce 2015 používal sledované elementy zřejmě pro zápis názvu souboru (jméno stránky) uvnitř nějaké složky, datového balíčku.

```

▼<premis:object xsi:type="premis:file">
  ▼<premis:objectIdentifier>
    <premis:objectIdentifierType>SE/RA</premis:objectIdentifierType>
    <premis:objectIdentifierValue>R1004768/c/047681RA.tif</premis:objectIdentifierValue>
  </premis:objectIdentifier>
  ▼<premis:significantProperties>
    <premis:significantPropertiesType>PageName</premis:significantPropertiesType>
    <premis:significantPropertiesValue>SE/RA/83004/RA/2013/MPO/R1004768/047681RA.tif</premis:significantPropertiesValue>
  </premis:significantProperties>
  ▼<premis:objectCharacteristics>

```

Obr. 3 Použití elementů v Národním archivu Švédska

(Riksarkivet, 2015?)

LTP systém Rosetta, resp. jeho interní metadatové schéma DNX, které využívá některých částí/sémantických jednotek standardu PREMIS také pracuje se signifikantními vlastnostmi dle standardu PREMIS. Aktuálně se zde zaznamenávají extrahovatelné technické vlastnosti jednotlivých souborů. V dokumentaci k systému se dále uvádí, že jako signifikantní vlastnosti mohou být popsány i vlastnosti dalších úrovní objektů (intelektuální entita, reprezentace, bitová sekvence) a že se nemusí jednat jen o údaje, které jsou extrahovány metadatovými extraktory. (Ex Libris, 2020)

Správci lokálních instalací systému Rosetta mohou využít defaultního nastavení systému, který má předdefinované mapování širokého spektra signifikantních vlastností nebo mohou tuto část systému konfigurovat a definovat a následně namapovat vlastnosti, které jsou signifikantní z jejich pohledu. (Německá národní vědecká a technická knihovna /TIB/, 2020).

Ze vzorového záznamu, jež do registru implementací uložily Archivy Nového Zélandu, které používají systém Rosetta, lze vydedukovat, že elementy jsou plněny informacemi extrahovanými nástrojem Jhove (viz obr. 4 a 5).

```

▼<section id="significantProperties">
  ▼<record>
    <key id="significantPropertiesType">AES.analogDigitalFlag</key>
    <key id="significantPropertiesValue">FILE_DIGITAL</key>
    <key id="significantPropertiesExtension"/>
  </record>
  ▼<record>
    <key id="significantPropertiesType">AES.encoding</key>
    <key id="significantPropertiesValue">PCM audio in integer format</key>
    <key id="significantPropertiesExtension"/>
  </record>
  ▼<record>
    <key id="significantPropertiesType">AES.byteOrder</key>
    <key id="significantPropertiesValue">1</key>
    <key id="significantPropertiesExtension"/>
  </record>

```

Obr. 4 Použití sémantické jednotky signifikantní vlastnosti v Archivech Nového Zélandu

```

▼<record>
  <key id="significantPropertiesType">AES.schemaVersion</key>
  <key id="significantPropertiesValue">1.02b</key>
  <key id="significantPropertiesExtension"/>
</record>
▼<record>
  <key id="significantPropertiesType">AES.useOtherType</key>
  <key id="significantPropertiesValue">[OTHER, JHOVE_validation]</key>
  <key id="significantPropertiesExtension"/>

```

Obr. 5 Použití sémantické jednotky signifikantní vlastnosti v Archivech Nového Zélandu s informací o použitém nástroji

(Archives New Zealand, 2009).

S konceptem signifikantních vlastností explicitně pracuje i systém Archivemata. Ten pro několik typů dat stanovuje významné vlastnosti, které jsou v systému evidovány. Signifikantní vlastnosti sleduje u následujících typů dat: zvukové soubory, rastrové a vektorové obrazy, tabulkové procesory, textové dokumenty, emailové zprávy, prezentace, videa a webové stránky. U rastrových obrazů se jako signifikantní vlastnosti zaznamenává: šířka a výška obrazu, sekvence obrazů v případě vícestránkového souboru TIFF, rozlišení, počet bitů na kanál a počet barevných komponent. U prezentací se například může zaznamenat počet stránek, slov, znaků, odstavců, řádek, tabulek, grafických znázornění, použité fonty. (Significant properties, 2020)

## 5 Signifikantní vlastnosti v aplikaci NDK

Na základě výše uvedeného průzkumu se zamýšlíme nad užitím sémantických jednotek pro signifikantní vlastnosti ze standardu PREMIS i v digitalizaci Národní knihovny. Užití těchto elementů/sémantických jednotek je spíše výjimečným jevem a je tedy málo příkladů, kterými se lze inspirovat. První návrhy pro aplikaci v NDK pracují se signifikantními vlastnostmi pro úroveň souboru, tak jak je v zahraničních aplikacích, jež jsme prezentovali dříve, běžnější. Popis na úrovni souborů je relativně také snadno implementovatelný do současného Standardu NDK, jež aktuálně pracuje s popisem technických vlastností jednotlivých souborů. Výhledem do budoucna je ale prozkoumat aplikaci i na vyšší úrovni, tj. na úrovni reprezentace (tj. celého balíčku) a na úrovni intelektuální entity. Předpokládáme, že stanovení těchto signifikantních vlastností bude ve spolupráci s uživatelskou komunitou, kterou v České republice tvoří zejména instituce, digitalizující z projektu VISK7 a některé soukromé společnosti, které poskytují digitalizaci službou na zakázku. Zástupci této odborné komunity se při Národní knihovně sdružují do Formátového výboru a k němu přidružených pracovních skupin. Tyto pracovní skupiny ve spolupráci s pracovníky Odboru digitalizace pak vzájemnou diskusí přenášejí potřeby praxe do metodologie digitalizace a naopak.

Předpokládáme, že signifikantní vlastnosti pro Intelektuální entitu budou korespondovat se současným způsobem výběru základních popisných metadat, tj., podle zásad tzv. Minimálního záznamu, převzatých do digitalizační praxe z katalogizačních pravidel. Ve Standardu NDK je tento Minimální záznam obsažen v metodikách pro všechny typy dokumentů, byť se liší skladba povinných údajů podle jejich specifik a specifik jejich nosičů. Signifikantní vlastnosti jednotlivých druhů dokumentů budou tato specifika pochopitelně následovat.

Eventuální analýza signifikantních vlastností metadatových záznamů bude pro digitalizaci Národní knihovny zřejmě výhledově nutností zejména z důvodu ztráty dat v transformacích mezi formáty, která je již nyní realitou v případě transformací mezi metadaty ve formátu MODS a Dublin Core. Zápis signifikantních vlastností se v tomto případě jeví nejen jako odůvodněný, ale i jako nevyhnutelný.

Podíváme-li se do již zmiňovaného Minimálního záznamu, nalezneme základní vodítko k jejich určení. Nejsnadnější to zřejmě bude pro nejširší skupinu dokumentů, které NDK digitalizuje (a ukládá institucím z programu VISK7) - tedy pro digitalizované tištěné dokumenty (tj. ve výsledku obrazové dokumenty). Naopak problematickými se pak zdají zvukové dokumenty, v nichž je variabilita popisných údajů obrovská. Z tohoto důvodu je snadnější vykonat analýzu "technických" vlastností dokumentů, které jsou snadněji postihnutele a měřitelné. V Národní knihovně byly prozatím analýze podrobeny technické vlastnosti zvukových a obrazových dokumentů, které jsou aktuálně přijímány do úložiště LTP, nebo se to v blízké budoucnosti plánuje. Část analýzy je diskutována dále.

## 5.1 Signifikantní vlastnosti zvukových dokumentů

Pro zvukové dokumenty byly definovány signifikantní vlastnosti na základě obsahu a potřeb *Definice metadatových formátů pro digitalizaci zvukových dokumentů, část 1 gramofonové desky, verze 0.3* a též na základě výstupu projektu InSPECT (Knight, 2008). Některé z těchto vlastností jsou již v metadatech balíčku SIP zaznamenávány, některé vlastnosti aktuálně do metadat zaznamenat nelze.

V následující tabulce jsou vyjmenovány vlastnosti, které tato analýza považuje za signifikantní, zároveň je v ní uveden návrh zápisu do metadat. Aktuálně jsou některé z níže uvedených signifikantních vlastností již zapsány ve stávající Definicí pro metadatové formáty pro digitalizaci zvukových formátů (převážně ve formátu AES57). V tabulce je uvádíme tučným písmem, a zároveň uvádíme i příslušný element technických metadat, užívaný k zápisu. Pro zbývající vlastnosti se zvažuje, zda budou umístěny do elementů pro signifikantní vlastnosti (`significantPropertiesType`), nebo se zapíše na jiné místo (možná i za použití dalšího metadatového schématu) do metadatových souborů tak, aby byly zařazeny mezi podobné informace, tj., v logické návaznosti na informace, ke kterým patří.

<b>signifikantní vlastnost</b>	<b>poznámka</b>	<b>hodnota significantProperties Type</b>	<b>hodnota significantProperties Value</b>	<b>přítomnost v DMF</b>
<b>délka souboru</b>	délka zvukového souboru ve frames	duration	časový údaj 11223344	duration v AES57
<b>velikost souboru</b>	velikost zvukového souboru v bytech	file size	číselný údaj 12345	size v PREMIS Object
<b>bitová hloubka</b>	bitová hloubka digitálního zvukového souboru	bit depth	číselný údaj hodnota nejméně 24	bitDepth v AES57
<b>vzorkovací frekvence</b>	vzorkovací frekvence zvukového souboru v herzích	sample rate	číselný údaj 96000	sampleRate v AES57
<b>počet kanálů</b>	počet kanálů zvukového souboru	number of channels	číselný údaj 1	numChannels v AES57

kvalita zvuku	možnost blíže popsat kvalitu zvuku a co jí mohlo ovlivnit	sound quality	volný text Nahrávka je kvůli svému stáří hůře slyšitelná.	
historie souboru	výčet procesů a událostí, vykonaných s popisovaným zvukovým souborem (migrace apod.)	history of the file	volný text Nahrávka byla 28.11.2019 migrována do formátu mp3. Ztráta obsahu nebyla zjištěna.	
<b>formát souboru</b>	formát souboru	format	slovní hodnota, stačí uvést příponu souboru wav	format, resp. formatName v PREMIS Object
bezztrátovost souboru	indikace, jestli je formát souboru bezztrátový	losslessness	slovní hodnota, možné hodnoty: <ul style="list-style-type: none"> <li>• ano (bezztrátový formát)</li> <li>• ne (ztrátový formát)</li> </ul>	
<b>typ nosiče</b>	typ nosiče, ze kterého zvukový soubor pochází	carrier	slovní hodnota, v současnosti může být použita hodnota: <ul style="list-style-type: none"> <li>• gramofonová deska</li> <li>• fonografický váleček</li> </ul>	v MODS na úrovni sound recording v elementu form; kontrolovaný slovník z 338
frekvenční rozsah	frekvenční rozsah dané nahrávky	frequency range	slovní hodnota 27-200 Hz	

Tab. 2 Signifikantní vlastnosti zvukových dokumentů, analýza NDK

Kromě výše uvedených vlastností se nabízí také několik dalších, které budou ještě předmětem další diskuze; jsou pro popis digitalizovaného zvukového dokumentu relevantní, ale popisují daný objekt obecněji. Za prvé, se jedná o informaci, zda je daný zvukový dokument zdrojovým audiem, archivní kopií nebo uživatelskou kopií – z nichž se každá liší formátem, velikostí a určením. Tento rozdíl by měl být patrný z názvu souboru a technických metadat, ale v případě, že dojde k přejmenování souboru nebo neúplnosti technických metadat, by tato informace byla přínosná.

V souvislosti s možným přejmenováním se naskytá otázka, zda v nich pro budoucí dohledatelnost nezaznamenat i původní název dokumentu, ze kterého byl zvukový záznam pořízen. Nakonec, pro správné posouzení stavu objektu by bylo vhodné zaznamenat i fixity, tj. stálost objektu, jejímž prostřednictvím lze poznat, že je dokument kompletní.

Níže jsou uvedeny možnosti zápisu zmíněných vlastností:

signifikantní vlastnost	vysvětlivka	hodnota significantProperties Type	hodnota significantPropertiesValue	přítomnost v DMF
<b>originální název</b>	původní jméno souboru vč. přípony	original name	volný text 123456789_01.wav	originalName v PREMIS Object
typ souboru	indikace, jakým výstupem digitalizace soubor je	file type	slovní hodnota, možné hodnoty: <ul style="list-style-type: none"> <li>• source audio</li> <li>• master copy audio</li> <li>• user copy audio</li> </ul>	
<b>úroveň ochrany</b>	Úroveň ochrany daného souboru vzhledem na jeho formát	preservation level	<ul style="list-style-type: none"> <li>• slovní hodnota, možné hodnoty:</li> <li>• logická ochrana</li> <li>• bitová ochrana</li> </ul>	preservationLevelValue v PREMIS Object
<b>úplnost souboru</b>	stálost souboru, jeho kompletnost	fixity	číselný údaj, hodnota kontrolního součtu 3B408677FB465F162069D6330F1553D0	fixity v PREMIS Object

Tab. 3 Signifikantní vlastnosti zvukových dokumentů, analýza NDK

Z tabulky je zřejmé, že velká část údajů, které aktuálně považujeme za signifikantní vlastnosti je již ve Standardu NDK zaznamenávána. I když duplicita údajů není vysloveně na škodu, je na místě se ptát, zda je pouze za účelem zápisu určených signifikantních vlastností nutné informace, které již zaznamenané jsou (a nadále se zapisovat do technických metadat budou) zapisovat ještě jednou do

speciálních elementů ve standardu PREMIS. Na druhou stranu může být přínosné tyto vlastnosti shromáždit na jedno dedikované místo, čímž se usnadní jejich pozdější extrakce ze systému LTP, a následně možná i kontrola výsledků migrace. Pro pozdější uživatele a správce těchto dat může být navíc zajímavé, jaké vlastnosti byly považovány za významné v době uložení objektů k dlouhodobému uchování. Zároveň je pro lidské čtení příhodnější záznam v elementech, které jsou jako významné vlastnosti nazvané a jim přímo dedikované.

## 5.2 Signifikantní vlastnosti rastrových obrazů

Zatímco většina charakteristik ve výčtu z výstupu projektu InSPECT – konkrétně šířka obrázku, výška obrázku, počet bitů na vzorek a počet vzorků na pixel – jsou rozhodně vlastnostmi, jejichž znalost je nezbytná pro vykreslení obrázku, jsou zároveň vlastnostmi, aktuálně zaznamenávanými ve Standardu NDK (metadata MIX). Analogicky se situací popsanou výše pro zvukové soubory, je tedy třeba se ptát, zda je uvádět i do metadat PREMIS – a mít je tudíž v SIP balíčku duplicitně. Redundance je sice základem bezpečnosti, ale duplicita uvnitř jednoho souboru je v tomto smyslu asi nejméně účinnou formou tohoto konceptu. Kontejner <PREMIS:significantProperties> tedy považujeme za vhodné využít pouze pro ty vlastnosti, které v MIX nejsou.

Většina technických metadat je samozřejmě zároveň vložená (embedded) do obrazového souboru samotného. Vzhledem k tomu, že většina z nich popisuje vlastnosti, které jsou souboru vlastní (*intrinsic*) může mít i toto nádech duplicity, přesto je to důležitou součástí ochrany, protože neshoda těchto metadat se skutečnými vlastnostmi souboru by indikovala poškození, nebo manipulaci se souborem. Vlastnosti, které takto podchyceny nejsou, jsou spíše vnější (*extrinsic*) vůči souboru. Přesto by jejich uvedení ve vložených metadatech mohlo přinést přidaný kontrolní prvek.

Vzorkovací frekvence je jediná ze seznamu vlastností InSPECT, která je vnější vůči obrazovému souboru, a v podstatě popisuje vztah digitálního obrázku k analogovému originálu/předloze. Protože není nezbytná pro vykreslení, je nezdědčena z metadat souboru. Jedná se nicméně o vlastnost zcela zásadní pro chápání souboru a zachování jeho autenticity. Ve Standardu NDK se již vyskytuje v metadatech MIX, ale ve vložených metadatech souborů JPEG2000 často chybí, případně je uvedena v *display resolution box*, což neodpovídá významu, který sledujeme. Proto v rámci implementace významných vlastností uvažujeme pro zanesení tohoto údaje o možnosti využívání *capture resolution box*.

Přestože InSPECT tvrdí opak (Montague, 2010), barevný prostor je významná vlastnost, protože umožňuje vykreslení barev. Zde musíme důsledně varovat před tvrzením InSpect, že barevný prostor není třeba zachovat mezi migracemi. Naopak, jakmile je obrázek jednou naskenován ve zvoleném barevném prostoru, je nutné tento prostor přísně zachovávat; změna barevného prostoru může vést pouze ke zhoršení – zkreslení barev a ztrátě informace.

Digitální obrázky naskenovaných tištěných stran jsou v našem archivačním procesu běžně podrobeny oříznutí. Velikost ořezu je vlastnost, která by měla zůstat zaznamenána pro zachování autenticity dokumentu. Oříznutí by mělo být představováno čtyřmi hodnotami pro každý okraj stánky. Protože jako jediné nejsou dosud uvedeny v mets, navrhujeme pro ně využít kontejner <PREMIS:significantProperties>. Bylo by samozřejmě výhodné, kdyby bylo možné tento údaj vložit i do metadat souboru, ale to je problematické proto, že se jedná o údaj na pomezí technického a provenienčního a ve struktuře metadat JPEG2000 pro něj není vhodný prostor.

Příklad možného zápisu:

```
—<PREMIS:significantProperties>
—<PREMIS:significantPropertiesType>cropTop</PREMIS:significantPropertiesType>
—<PREMIS:significantPropertiesValue>25 mm</PREMIS:significantPropertiesValue>
—</PREMIS:significantProperties>

—<PREMIS:significantProperties>
—<PREMIS:significantPropertiesType>cropBottom</PREMIS:significantPropertiesType>
—<PREMIS:significantPropertiesValue>15 mm</PREMIS:significantPropertiesValue>
—</PREMIS:significantProperties>

—<PREMIS:significantProperties>
—<PREMIS:significantPropertiesType>cropLeft</PREMIS:significantPropertiesType>
—<PREMIS:significantPropertiesValue>20 mm</PREMIS:significantPropertiesValue>
—</PREMIS:significantProperties>

—<PREMIS:significantProperties>
—<PREMIS:significantPropertiesType>cropRight</PREMIS:significantPropertiesType>
—<PREMIS:significantPropertiesValue>15 mm</PREMIS:significantPropertiesValue>
—</PREMIS:significantProperties>
```

## Shrnutí a diskuze

Teoretický výzkum signifikantních vlastností ve formátu PREMIS probíhal zejména v prvním desetiletí nového století, kdy byly ustanoveny jeho základní teze a metodologie analýzy k jejich stanovení. Na svém významu nabývá s nutností konverzí dat do nových technických i metadatových formátů, aby byla zajištěna kontinuita přenášených informací a co nejdřívejší přenesení objektů pro další archivaci.

Z průzkumu praxe v zahraničních institucích a diskusí v rámci konferencí se zdá, že plnění speciálních elementů pro signifikantní vlastnosti ve standardu PREMIS probíhá extrahováním vlastností archivovaných souborů, nikoliv intelektuálních entit a jejich metadatových záznamů. Při hledání vlastní cesty se tedy nelze dost dobře opřít o zkušenosti jiných. I z tohoto důvodu náš příspěvek v této oblasti prozatím nenachází konkrétní odpovědi, a zůstává proto pouze příspěvkem ke “kolektivnímu nevědomí”. V případě, že k analýze metadatových záznamů přistoupíme, je ale jisté, že to bude za účasti určené komunity, kterou v našem případě tvoří členové pracovních skupin při Formátovém výboru NDK.

V NDK momentálně zpracované signifikantní vlastnosti pro zvukové dokumenty a rastrové obrazy lze poměrně lehce zpracovat, jelikož už se buď v nějaké formě v metadatech vyskytují, nebo je možné je snadno doplnit zejména do souborů AMD\_METS, které již nyní pro zápis vlastností objektů a eventů, které s nimi souvisejí, používáme. V případě specifického zápisu bychom pak spíše řešili případné duplicity v uvádění hodnot. Pokud bychom do budoucna uvažovali o provedení analýzy signifikantních vlastností i pro intelektuální entitu, reprezentaci a metadatový záznam – a jejich následný zápis do archivačního balíčku – bylo by tedy nejspíš nutné zvážit také změnu způsobu, jakým momentálně metadata zaznamenáváme (např. v případě vlastností, jež nelze extrahovat automatizovaně), jelikož v současné době pro zápis popisných metadat formát PREMIS neuvádíme.

## Závěr

Koncept signifikantních vlastností představuje pro komunitu zabývající se digitální ochranou kulturního dědictví příležitost pro reflexi, zda údaje, které doprovázejí archivované digitální objekty, jsou dostatečné pro správné pochopení, zobrazení, funkci a chování digitálního objektu tak, aby pokud možno zprostředkoval uživateli všechny informace nesené původně originálem, které uživatel potřebuje, či zkušenost, která je s originálem srovnatelná. Pro určení signifikantních vlastností existují návody jak pro interní analýzu objektů, tak pro analýzu potřeb uživatelů (označovaných v této souvislosti pojmem *stakeholder*). V NDK se v současnosti věnujeme analýze signifikantních vlastností pro zvukové soubory a rastrové obrazy. Naše dosavadní zkušenosti ukazují, že tento přístup přináší další zpřesnění informací, předávaných uživateli, a posílení autenticity námi uchovávaných záznamů.

Při dlouhodobém uchovávání digitálních objektů je nutné počítat s tím, že archivované objekty nebude možné zachovat v takové podobě, v jaké byly do dlouhodobého úložiště přijaty. Již nyní se tedy musíme připravovat na jejich migraci. Stanovení signifikantních vlastností a zejména jejich zaznamenání pak může být pomyslným vodítkem při budoucích transformacích a nástrojem, jak transformace kontrolovat. V uvedeném textu jsme se zabývali výhradně datovými soubory, nicméně, do budoucna budeme nutně muset zvážit stejnou analýzu pro metadatové záznamy, jelikož v jejich případě riziko ztráty není pouze potenciální, ale reálné. Již nyní ztráty mezi transformacemi mezi formáty zaznamenáváme v případě MARC-to-MODS, respektive, MODS-to-DC. Analýza a následné zapracování signifikantních vlastností minimálně metadatového záznamu (a nejlépe i intelektuální entity a reprezentace) by tedy jistě bylo přínosem k budoucímu uchovávání dokumentů.

---

## Dedikace

Článek vznikl na základě institucionální podpory dlouhodobého koncepčního rozvoje výzkumné organizace poskytované Ministerstvem kultury ČR.

## Seznam literatury

Archives New Zealand. *04\_NewZealandNationalLibrary.xml* [online]. 2009 [cit. 2020-07-01]. [http://www.loc.gov/standards/premis/registry/examples/04\\_NewZealandNationalLibrary.xml](http://www.loc.gov/standards/premis/registry/examples/04_NewZealandNationalLibrary.xml)

Ashley, K., Davis, R. & Pinsent, E. (2008). *Significant Properties of E-learning Objects (SPeLOs)*. Tech. rep. Draft version 0.91. JISC Information Environment Committee. [http://pubs.ulcc.ac.uk/54/1/spelos\\_report.pdf](http://pubs.ulcc.ac.uk/54/1/spelos_report.pdf).

Carlson, S. & Anderson, B. (2007). What are data? The many kinds of data and their implications for data re-use. *Journal of Computer-Mediated Communication* 12 (2): 635–651. doi:10.1111/j.1083-6101.2007.00342.x. <https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.1083-6101.2007.00342.x>.

THE CEDARS PROJECT, *Cedars Guide to Digital Collection Management*, (2002). University of Leeds. Dostupné také z: <https://www.webarchive.org.uk/wayback/archive/20050410120000mp/http://www.leeds.ac.uk/cedars/guideto/collmanagement/guidetocolman.pdf>

Corti, L. (2007). “Re-using archived qualitative data: Where, how, why?” *Archival Science* 7 (1): 37–54. doi:10.1007/s10502-006-9038-y. <https://link.springer.com/content/pdf/10.1007%2Fs10502-006-9038-y.pdf>

Coyne, M., Duce, D., Hopgod, B., Mallen, G. & Stapleton M. (2007). *The Significant Properties of Vector Images*. Tech. rep. Version 4.3. JISC Digital Preservation Programme. [https://www.webarchive.org.uk/wayback/archive/20140613220103/http://www.jisc.ac.uk/media/documents/programmes/preservation/vector\\_images.pdf](https://www.webarchive.org.uk/wayback/archive/20140613220103/http://www.jisc.ac.uk/media/documents/programmes/preservation/vector_images.pdf)

Coyne, M. & Stapleton, M. (2008). *The significant properties of moving images*. Tech. rep. JISC Digital Preservation Programme.

Cubr, L. (2017) *Autenticita a digitální informace*. (Disertační práce). Dostupné z: <https://is.cuni.cz/webapps/zzp/detail/105596>.

Dappert, A. & Farquhar, A. (2009). “Significance is in the Eye of the Stakeholder”. In *13th European Conference, ECDL 2009, Corfu, Greece, September 27-October 2, 2009. Proceedings*, ed. by Maristella Agosti et al., 5714:297–308. Lecture Notes in Computer Science. European Conference on Digital Libraries, Springer Berlin, Heidelberg. isbn: 978-3-642-04346-8. doi:10.1007/978-3-642-04346-8. [https://link.springer.com/content/pdf/10.1007%2F978-3-642-04346-8\\_29.pdf](https://link.springer.com/content/pdf/10.1007%2F978-3-642-04346-8_29.pdf)

De Vos, J.(2015). *Domain specific implementation of PREMIS*. Prezentace prezentovaná na/v: iPRES 2015 workshop: PREMIS implementation fair – Nov. 6th 2015. Dostupné na WWW: <https://www.loc.gov/standards/premis/pif/2015/index.html>

EX LIBRIS. *Rosetta AIP Data Model* [online]. Ex Libris Limited, (June 2020). [cit. 2020-07-01]. Dostupné z:

[https://files.mtstatic.com/site\\_11811/39700/4?Expires=1593609950&Signature=njx57zG0HvBvcCUaHAQYjKM2B9BiLYA6QQRtYMDE2DDFNQ8C0GU-L9mPUtK0rjUfla5ob93WxZC9H~aVJBaD47xIfAdyihAw8yStcZ2s-67J3vPjNcrjKCAMet8wD4g2DLcbIBeci2~qHIY2epTlFIH70khXkLERm56Ym3HN8tk\\_&Key-Pair-Id=APKAJ5Y6AV4GI7A555NA](https://files.mtstatic.com/site_11811/39700/4?Expires=1593609950&Signature=njx57zG0HvBvcCUaHAQYjKM2B9BiLYA6QQRtYMDE2DDFNQ8C0GU-L9mPUtK0rjUfla5ob93WxZC9H~aVJBaD47xIfAdyihAw8yStcZ2s-67J3vPjNcrjKCAMet8wD4g2DLcbIBeci2~qHIY2epTlFIH70khXkLERm56Ym3HN8tk_&Key-Pair-Id=APKAJ5Y6AV4GI7A555NA)

Faniel, I. M. (2009). “Unrealized Potential: The Socio-Technical Challenges of a Large Scale Cyberinfrastructure Initiative”.

[https://pdfs.semanticscholar.org/e6f3/23e6106b9c6d89f7419f489be72eb921df10.pdf?\\_ga=2.155260862.409180184.1571031725-715198704.1570454413](https://pdfs.semanticscholar.org/e6f3/23e6106b9c6d89f7419f489be72eb921df10.pdf?_ga=2.155260862.409180184.1571031725-715198704.1570454413)

Faniel, I. M., and Jacobsen, T. E. (2010). “Reusing Scientific Data: How Earthquake Engineering Researchers Assess the Reusability of Colleagues’ Data”. *Computer Supported Cooperative Work* 19:355–375. doi:10.1007/s10606-010-9117-8. <https://link.springer.com/content/pdf/10.1007/2Fs10606-010-9117-8.pdf>

Faniel, I. M., & Yakel, E. (2011). Significant Properties as Contextual Metadata. *Journal of Library Metadata* 11:155–165. issn: 1937-5034. doi:10.1080/19386389.2011.629959

German National Library of Science and Technology (TIB). *Core Trust Seal: Assessment Information: Repository: TIB Digital Long-Term Archive* [online]. (2020). [cit. 2020-07-01]. <https://www.coretrustseal.org/wp-content/uploads/2020/06/TIB-Digital-Long-Term-Archive.pdf>

Giaretta, D., Matthews, B., Bicarregui, J., Lambert, S., Guercio, M., Michetti, G., & Sawyer, D. (2009). Significant Properties, Authenticity, Provenance, Representation Information and OAIIS Information. *UC Office of the President: California Digital Library* [online]. 2009-05-10 [cit. 2018-11-14]. <https://escholarship.org/uc/item/0wf3j9cw>

Grace, S., Knight, G. & Montague, L. (2009). InSPECT Project Final Report. Tech. rep. Version 1.0. King’s College London. <https://ndownloader.figshare.com/files/13131539>.

HathiTrust. *HathiTrust PREMIS Implementation - Version 2.0* [online]. updated April 20, 2015 [cit. 2020-07-01].

<https://docs.google.com/document/d/1UTZNIzRfelVixIYJ9nZ12tnieIe6FNZSMi9Fdhh2z5c/edit#heading=h.k2fhyw155gi>

Hedstrom, M. L., Lee, C. A., Olson, J.S. & Lampe, C. A., (2006). The Old Version Flickers More: Digital Preservation from the User’s Perspective. *The American Archivist*: 159–187. <https://americanarchivist.org/doi/pdf/10.17723/aarc.69.1.1765364485n41800>

Heslop, H., Davis, S. & Wilson, A. (2002). *An approach to the preservation of digital records*. Tech. rep. National Archives of Australia (NAA).

<http://www.imaginar.org/taller/dppd/DPPD/40%20pp%20Approach.pdf>

Hockx-Yu, H. & Knight, G. (2008). What to Preserve?: Significant Properties of Digital Objects. *International Journal of Digital Curation*. Bath: University of Bath, 3(1), 141-153. DOI: <https://doi.org/10.2218/ijdc.v3i1.49> ISSN 1746-8256.

Knight, G. (2008). *Significant Properties Testing Report: Audio Recordings* [online]. CeRch, [cit. 2019-11-27]. Dostupné z: [https://figshare.com/articles/InSPECT\\_Significant\\_Properties\\_Testing\\_Report\\_Audio\\_Recordings/7137818](https://figshare.com/articles/InSPECT_Significant_Properties_Testing_Report_Audio_Recordings/7137818)

Knight, G. (2010a). *Significant Properties Data Dictionary*. Tech. rep. Version 1.1. InSPECT Project. <https://ndownloader.figshare.com/files/13131551>.

Knight, G. (2009). *InSPECT Framework Report*. Londýn. Dostupné také z: <http://significantproperties.kdl.kcl.ac.uk/inspect-framework.pdf>

Knorr-Cetina, K. (1999). *Epistemic cultures: how the sciences make knowledge*. Harvard University Press, Cambridge, Mass. isbn: 0674258932.

Konečný, M. (2018). *Významné vlastnosti digitálních objektů*. Presentace prezentovaná na/v: Národní technická knihovna na semináři 15. listopadu 2018

Konečný, M. (2019) *Významné vlastnosti digitálních objektů a jejich určování*. Interní studie pro NK.

Matthews, B., McIlwrath, B., Giaretta D. & Conway, E. (2008). *The significant properties of software: A study*. Tech. rep. Version 1.0. JISC Digital Preservation Programme. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.456.122&rep=rep1&type=pdf>

Montague, L., Brown, A., Knight, G. & Stephen G. (2010). InSPECT Significant Properties Testing Report: Raster Images. *InSPECT Project Document*. Dostupné také z: [https://figshare.com/articles/InSPECT\\_Significant\\_Properties\\_Testing\\_Report\\_Raster\\_Images/7137803](https://figshare.com/articles/InSPECT_Significant_Properties_Testing_Report_Raster_Images/7137803)

Morrissey, S. (2010). The economy of free and open source software in the preservation of digital artefacts. *Library Hi Tech* 28 (2). doi:10.1108/07378831011047622. <https://www.emerald.com/insight/content/doi/10.1108/07378831011047622/full/pdf?title=the-economy-of-free-and-open-source-software-in-the-preservation-of-digital-artefacts>

Lynch, C. (1999). Canonicalization: a fundamental tool to facilitate preservation and management of digital information. *D-Lib Magazine*, no. 9. <http://www.dlib.org/dlib/september99/09lynch.html>

National Library of Wales. (2008). *11\_Wales\_NLW\_wills.xml* [online]. [cit. 2020-07-01]. [http://www.loc.gov/standards/premis/registry/examples/11\\_Wales\\_NLW\\_wills.xml](http://www.loc.gov/standards/premis/registry/examples/11_Wales_NLW_wills.xml)

OCLC/RLG, Working Group on Preservation Metadata. (2002). *Preservation Metadata and the OAIS Information Model – A Metadata Framework to Support the Preservation of Digital Objects*. Tech. rep. [https://www.oclc.org/content/dam/research/activities/pmwg/pm\\_framework.pdf](https://www.oclc.org/content/dam/research/activities/pmwg/pm_framework.pdf)

PREMIS Editorial Committee, (2015). *PREMIS Data Dictionary for Preservation Metadata: version 3.0*. Washington, D.C.: The Library of Congress. <http://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf>

Riksarkivet. *Premis\_45\_SNA\_2015.xml* [online]. 2015? [cit. 2020-07-01].  
[http://www.loc.gov/standards/premis/registry/examples/premis\\_45\\_SNA\\_2015.xml](http://www.loc.gov/standards/premis/registry/examples/premis_45_SNA_2015.xml)

Rothenberg, J. & Bikson, T. (1999). *Carrying Authentic, Understandable and Usable Digital Records Through Time*. Report to the Dutch National Archives and Ministry of the Interior. Tech. rep. RAND-Europe. [https://www.rand.org/content/dam/rand/pubs/rand\\_europe/RE99/RE99-016/RE99-016.pdf](https://www.rand.org/content/dam/rand/pubs/rand_europe/RE99/RE99-016/RE99-016.pdf).

Significant characteristics. *Archivematica* [online]. last modified on 11 February 2020, at 18:33 [cit. 2020-07-01]. [https://wiki.archivematica.org/Significant\\_characteristics](https://wiki.archivematica.org/Significant_characteristics)

Steeman, M., ed. (2018) *Preservation Metadata Dictionary 2.0*. Tech. rep. Netherlands Institute for Sound and Vision. <https://publications.beeldengeluid.nl/pub/61>

Thibodeau, K. (2002). Overview of technological approaches to digital preservation and challenges in coming years. *In The State of Digital Preservation: An International Perspective, Conference Proceedings*, 107:4–31. Council on Library / Information Resources Washington, D.C. isbn: 1-887334-92-0. <https://clir.wordpress.clir.org/wp-content/uploads/sites/6/pub107.pdf>.

Van House, N. A. (2002). Digital libraries and practices of trust: Networked biodiversity information. *Social Epistemology* 16 (1): 99–114. doi:10.1080/02691720210132833. <https://www.tandfonline.com/doi/abs/10.1080/02691720210132833>

Van House, Nancy A., Mark H. Butler, and Lisa R. Schiff. 1998. Cooperative knowledge work and practices of trust: Sharing environmental planning data sets. *In Proceedings of The 1998 ACM Conference on Computer Supported Cooperative Work*, 335–343. <http://people.ischool.berkeley.edu/~vanhouse/cscwdist.pdf>.

Vermaaten, S., Lavoie, B. & Caplan, P. (2012). Identifying threats to successful digital preservation: the SPOT model for risk assessment. *D-Lib Magazine* 18 (9/10). doi:10.1045/september2012-vermaaten. <http://www.dlib.org/dlib/september12/vermaaten/09vermaaten.html>.

Wilson, A. (2007). *Significant Properties Report, InSPECT Work Package 2.2*. Tech. rep. Version 2. InSPECT Project. [https://figshare.com/articles/InSPECT\\_Significant\\_Properties\\_Report\\_WP2\\_2\\_/7137812](https://figshare.com/articles/InSPECT_Significant_Properties_Report_WP2_2_/7137812).

Zimmerman, A. S. (2008). New knowledge from old data: The role of standards in the sharing and reuse of ecological data. *Science, Technology, & Human Values* 33 (5): 631–652. doi:10.1177/2F0162243907306704.

[https://www.researchgate.net/publication/239566188\\_New\\_Knowledge\\_from\\_Old\\_Data\\_The\\_Role\\_of\\_Standards\\_in\\_the\\_Sharing\\_and\\_Reuse\\_of\\_Ecological\\_Data](https://www.researchgate.net/publication/239566188_New_Knowledge_from_Old_Data_The_Role_of_Standards_in_the_Sharing_and_Reuse_of_Ecological_Data)

---

## Poznámka o autorech

*Autory této analýzy jsou současní a bývalí metadatoví a formátoví specialisté Oddělení pro standardy Národní knihovny ČR, které analyzuje, zavádí a rozšiřuje standardy potřebné pro dlouhodobou ochranu digitálních dokumentů. Toto oddělení také zastrešuje standardizaci v rámci projektu Národní digitální knihovny (NDK), a poskytuje konzultace ostatním digitalizujícím knihovnám a institucím.*

Natalie Ostráková, e-mail: [natalie.ostrakova@nkp.cz](mailto:natalie.ostrakova@nkp.cz)

Pavčina Kočišová, e-mail: [pavlina.kocisova@nkp.cz](mailto:pavlina.kocisova@nkp.cz)

Miroslava Beňáčková, e-mail: [miroslava.benackova@nkp.cz](mailto:miroslava.benackova@nkp.cz)

Vojtěch Kopský, e-mail: [vojtech.kopsky@nkp.cz](mailto:vojtech.kopsky@nkp.cz)

Frank Richard Mally, e-mail: [richardmally@seznam.cz](mailto:richardmally@seznam.cz)