

PREMIS – JAK MU POROZUMĚT

Library of Congress Network Development and MARC Standards Office

Autor: Priscilla Caplan

Revize: Ediční komice PREMIS

Původní vydání 2009, revidované 2017, aktualizované v červnu 2021

Copyright © 2009 Knihovna Kongresu, platí pro území mimo USA

Translation © 2018 Univerzita Karlova

Z anglického originálu Understanding PREMIS přeložily Bc. Michaela Málková, Bc. Marie Odehnalová a Bc. Hana Světelská.

Odborná revize Mgr. Jan Hutař, Ph.D. a Mgr. Marek Melichar

Na překladu se dále podílela PhDr. Radka Římanová, Ph.D.

V případě užití jakékoliv části tohoto díla musí být dílo citováno.

Obsah

PREMIS A JAK MU POROZUMĚT	1
1. PREMIS V KONTEXTU	1
1.1. Co jsou archivační metadata?	1
1.2. Co je PREMIS?.....	2
1.3. Co je Datový slovník PREMIS?	2
1.4. Jak používat PREMIS?	4
1.5. Měli byste používat PREMIS?	4
2. KLÍČOVÉ KONCEPTY PREMIS	5
2.1. Sémantické jednotky	5
2.2. Kontejnerové jednotky a podjednotky	6
2.3. Kontejnerové jednotky rozšíření	6
3. DATOVÝ MODEL PREMIS	7
3.1. Objekty	8
3.2. Události	10
3.3. Agenti	11
3.4. Práva.....	11
4. DATOVÝ SLOVNÍK.....	12
4.1. Ukázkové heslo Datového slovníku pro jednoduchou sémantickou jednotku.....	12
4.2. Ukázka hesla z Datového slovníku pro kontejnerovou jednotku.....	14
5. PREMIS V PRAXI	15
5.1. PREMIS v XML.....	15
5.2. Soulad s PREMIS	16
6. DOPLŇKOVÉ INFORMACE O PREMIS	18
Příloha A: Příklady	19
A.1. Příklad objektu	19
A.2. Příklad události.....	22
A.3 Příklad agenta	23
Příloha B: Slovníček termínů	25

PREMIS A JAK MU POROZUMĚT

Tento průvodce je relativně stručným náhledem na PREMIS, standard archivačních metadat. Neposkytne vám dostatek informací, jak PREMIS implementovat, ale získáte představu, o čem PREMIS vlastně je. Mnoha čtenářům to bude stačit. Pro ty, kteří se budou zabývat Datovým slovníkem PREMIS, který má více než 250 stran, může tento průvodce sloužit jako lehký úvod, který Vám slovník PREMIS přiblíží.

1. PREMIS V KONTEXTU

1.1. Co jsou archivační metadata?

Pokud pracujete v knihovně, archivu nebo muzeu, pravděpodobně jste seznámeni s metadaty a popisem zdrojů. Nejspíše víte, že metadata jsou rozdělena podle cíle, který mají plnit: popisná metadata usnadňují nalezení a identifikaci zdrojů, administrativní metadata zajišťují jejich správu a sledování, a strukturální metadata naznačují, jak jsou komplexní digitální objekty složeny, aby mohly být náležitě interpretovány. Stejně tak archivační metadata podporují aktivity určené k zajištění dlouhodobé použitelnosti digitálního zdroje.

Datový slovník PREMIS definuje archivační metadata jako „informace, které repozitář užívá k podpoře procesu digitální archivace“. Zde jsou příklady archivačních aktivit a jak jsou zajišťovány pomocí metadat:

- Soubor musí být uložen bezpečně, aby ho nikdo nemohl neúmyslně (nebo se zlým úmyslem) modifikovat. Pomocí kontrolního součtu uloženého jako metadata může být určeno, zda došlo ke změně uloženého souboru mezi dvěma časovými body.
- Soubory musí být uloženy na médiu, které současné počítače umí přečíst. Pokud je médium zničené nebo zastaralé (jako osmipalcové diskety používané v 70. letech 20. století), může být obtížné nebo nemožné data obnovit. Metadata mohou podporovat správu médií tím, že zaznamenají typ a stáří úložného média a data, kdy byly soubory naposledy obnoveny.
- Během dlouhých časových období i oblíbené formáty mohou zestárnout, tzn. žádná současná aplikace je nemusí umět interpretovat. Správci archivu musí využít archivační strategie k zajištění trvalé použitelnosti zdrojů. To může znamenat migraci starých formátů na novější ekvivalenty nebo emulaci starého interpretačního prostředí na novějším hardwaru a softwaru. Obě strategie, migrace i emulace, vyžadují metadata o originálním formátu souboru a o hardware i software prostředí.

- Archivační strategie mohou mít za následek změnu originálních souborů (migrace) nebo změnu způsobu jejich interpretace (emulace). To může zpochybnit autenticitu zdroje. Metadata ji mohou pomoci podpořit tím, že zdokumentují digitální původ zdroje – jeho historii převodu odpovědnosti za správu a autorizovanou historii změn.

1.2. Co je PREMIS?

PREMIS je zkratka pro „PREservation Metadata: Implementation Strategies“¹, což je název mezinárodní pracovní skupiny, která byla mezi roky 2003 a 2005 sponzorována OCLC² a RLG³. Tato pracovní skupina vytvořila zprávu nazvanou Datový slovník PREMIS, která zahrnuje datový slovník, a i poměrně mnoho podrobností o archivačních metadatech. Knihovna Kongresu pak publikovala sadu schémat PREMIS pro reprezentaci elementů metadat v Datovém slovníku v XML. V březnu 2008 bylo zveřejněno druhé, aktualizované vydání Datového slovníku a nové podpůrné schéma, následováno menšími revizemi, verzemi 2.1 a 2.2. V červnu 2015 byla zveřejněna zásadní revize, verze 3.0.

Aktivní rozvoj PREMIS je sponzorován Knihovnou Kongresu. Zahrnuje odkazy na webové stránky různých druhů oficiálních i neoficiálních informací o PREMIS, diskusní skupinu a „wiki“ pro uživatele PREMIS. Existuje také Ediční komise PREMIS zodpovědná za revize Datového slovníku a schématu. Kongresová knihovna se také snaží šířit povědomí o PREMIS, sponzoruje tutoriály a implementační schůzky o používání PREMIS a zadává studie a publikace se vztahem k PREMIS, jako je tento průvodce.

Pokud se mluví o PREMIS, obvykle mají na mysli Datový slovník. Občas mohou také odkazovat na XML schéma, na pracovní skupinu nebo na celé úsilí.

Datový slovník PREMIS: www.loc.gov/premis/v3/premis-3-0-final.pdf

webová stránka PREMIS: www.loc.gov/premis

diskusní skupina Implementační skupiny PREMIS: pig@listserv.loc.gov

pokud chcete skupinu odebírat, pošlete e-mail na adresu listserv@listserv.loc.gov
se zprávou: subscribe pig [vaše jméno]

1.3. Co je Datový slovník PREMIS?

Datový slovník PREMIS definuje základní sadu elementů metadat (obvykle „sémantické jednotky“, ale o tom budeme mluvit později), které by repozitáře měly obsahovat, aby byly schopny plnit archivační funkci. Archivační funkce se mezi repozitáři mohou lišit, ale obecně budou zahrnovat akce k zajištění toho, že digitální objekty zůstanou čitelné (tzn. že mohou být z daného média přečteny) a interpretovatelné (tzn. že mohou být zobrazeny, přehrány nebo jinak interpretovány aplikačním softwarem). Stejně tak by měly zajistit, aby digitální

¹ česky Archivační metadata: strategie implementace, poznámka překladatelů

² Online Computer Library Center

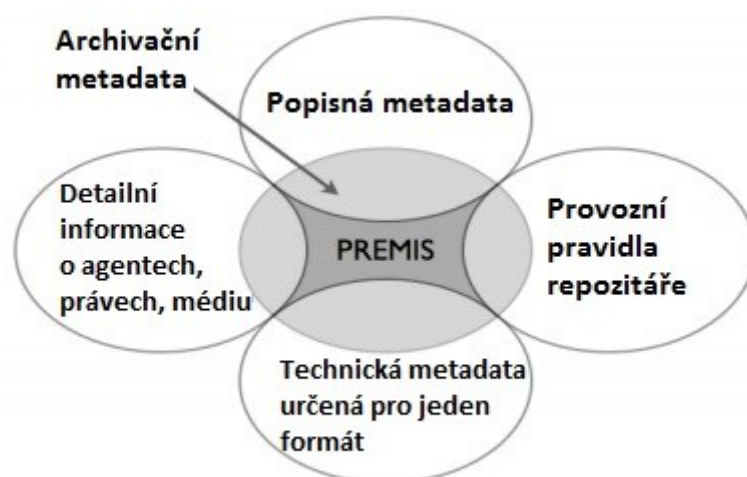
³ Research Library Group

objekty v repozitáři nebyly neúmyslně modifikovány, a že legitimní změny jsou zdokumentovány.

Záměrem Datového slovníku není definovat všechny možné elementy archivačních metadat, jen ty, které budou většinou v repozitářích potřebné. Několik kategorií metadat je vyloučeno, včetně těchto:

- Metadata určená pro jeden formát, tj. metadata, která se týkají pouze jednoho formátu nebo skupiny formátů jako je audio, video nebo vektorová grafika.
- Metadata zaměřená na implementaci a provozní pravidla, tj. metadata, která popisují politiku nebo pokyny individuálního repozitáře, např. způsob poskytování přístupu k materiálům.
- Popisná metadata. Přestože popis zdroje je pro archivaci zjevně relevantní, může být použito mnoho jiných nezávislých standardů, včetně MARC, MODS a Dublin Core.
- Detailní informace o médiu nebo hardwaru. Opět, i když zjevně relevantní pro archivaci, tato metadata jsou ponechána k definici jiným komunitám.
- Detailní informace o **agentech** (lidé, organizace nebo softwaru), nad rámec základní informace nezbytné k jejich identifikaci.
- Obsáhlé informace o **právech** a oprávněních; pozornost je zaměřena na ty, které ovlivňují archivační funkce.

Pokud zvážíme všechna metadata potřebná pro organizaci, která spravuje archivní repozitář, můžeme si PREMIS představit jako podmnožinu uprostřed. Netýká se na jedné straně vyhledávání a zpřístupnění a nesnaží se definovat detailní metadata zaměřená na formát na straně druhé. Definuje pouze metadata, která jsou běžně potřebná k vykonávání archivačních funkcí nad všemi materiály.



Obrázek 1: PREMIS jako podmnožina všech archivačních metadat

Obrázek 1 zobrazuje všechna archivační metadata jako kruh uprostřed diagramu. Zahrnují některá popisná metadata, některá provozní pravidla, některá detailní technická metadata a některé detailní informace o **agentech**, **právech**, médiu a hardware. PREMIS je malé jádro v srdci archivačních metadat, které vylučuje všechny tyto typy.

1.4. Jak používat PREMIS?

Datový slovník PREMIS definuje, co má archivní repozitář obsahovat. Soustředí se na systém a správu repozitáře, ne na autory digitálního obsahu, lidi, kteří skenují nebo jinak konvertují analogový obsah na digitální, nebo personál, který hodnotí a licencuje komerční elektronické zdroje. PREMIS lze především použít při návrhu a hodnocení repozitáře a pro výměnu archivovaných informačních balíčků.

Ti, kdo navrhují a vytvářejí softwarové komponenty archivního repozitáře, by měli používat PREMIS jako návod pro to, které informace by měla aplikace získávat a zaznamenávat nebo které by měly být jinak známy správcům repozitáře.

Ti, kdo plánují implementovat archivní repozitář, by měli používat PREMIS jako kontrolní seznam k hodnocení navrhovaných softwarů. Systémy, které podporují Datový slovník PREMIS, budou lépe schopny uchovat informační zdroje dlouhodobě.

Občas může být potřeba převést uložené informační balíčky z fungujícího repozitáře do repozitáře jiného. Například správcovská organizace může migrovat obsah z jednoho systému repozitáře do druhého, zákazník může chtít změnit jednu službu třetí strany na jinou nebo instituce může použít archivační služby jiné instituce v důvěryhodném digitálním repozitáři. PREMIS poskytuje obecnou sadu elementů dat, které používá exportující i importující repozitář, obzvláště pokud je použito PREMIS XML schéma.

1.5. Měli byste používat PREMIS?

To záleží na více faktorech. Většiny personálu v knihovnách, archivech, muzeích a jiných organizacích kulturního dědictví se digitální archivace přímo netýká. V takovém případě je dostačující, že víte, co to PREMIS je: datový slovník pro archivační metadata. Jestliže vaše pracovní pozice zahrnuje určitou zodpovědnost za některý z aspektů digitální archivace, pravděpodobně pro vás bude užitečné znát PREMIS blíže. Týká-li se vás hodnocení nebo implementace institucionálního repozitáře nebo archivačního systému, měli byste PREMIS znát detailně. Zvažte, zda by nebylo vhodné použít tréninkové materiály na webové stránce PREMIS nebo se zúčastnit školení PREMIS, pokud vám bude nabídnuto.

Pokud pracujete na digitalizačních projektech, možná přemýšlíte, jestli byste měli vytvářet PREMIS metadata pro pozdější užití. Většina elementů PREMIS je navržena tak, aby je automaticky dodávala aplikace archivního repozitáře. (To samozřejmě neznamená, že je

dodávají všechny aktuálně dostupné aplikace.) Některé informace byste však měli zaznamenat, pokud je to možné:

Inhibitory. Inhibitory jsou definovány jako jakékoliv vlastnosti objektu určené k zamezení přístupu, použití nebo migrace. Příklady inhibitorů zahrnují ochranu heslem a šifrování. Software repozitáře nemusí být schopný inhibitor identifikovat, protože inhibitor může softwaru zabránit analyzovat objekt, proto je důležité inhibitory zaznamenat, pokud víte, že je soubor má. PREMIS definuje sémantické jednotky pro typ inhibitoru, cíl (akce, které jsou blokovány) a klíč (heslo nebo jiný mechanismus, jak inhibitor obejít).

Původ. Digitální původ je záznam provozního řetězce a historie změn digitálního objektu. Pokud vaše instituce objekt vytvořila, okolnosti vzniku jsou zřejmě důležitou částí jeho původu. Název a verze aplikace, pomocí které byl objekt vytvořen, a datum vzniku mohou být často extrahovány z hlavičky souboru, ale ne vždy, proto je doporučeno tyto informace zaznamenávat. PREMIS umožňuje zaznamenat historii změn jako informace o **událostech**, která je popsána níže. Řízený slovník typů **událostí** je dostupný na adrese: id.loc.gov/vocabulary/preservation/eventType.

Mnoho typů **událostí** PREMIS je navrženo tak, aby popisovaly akce, které nastanou poté, co je něco odesláno na příjem do repozitáře, přestože jiné jsou použity pro **události**, které se staly před přijetím, jako zachycení nebo vstup.

Významné vlastnosti. Významné vlastnosti jsou charakteristiky objektu, které by měly být zachovány během archivačních opatření. Například, zda jsou v dokumentu zásadní jen slova a obrázky, nebo zda jsou stejnou měrou důležité i fonty, pozadí, formátování a jiné formální znaky. Myšlenka významných vlastností je jedním z nejdůležitějších konceptů digitální archivace a také jedním z nejméně pochopených. Nicméně, jakákoli instituce, která vytváří nebo získává materiály pro uživatelskou komunitu, by se měla dobře zamyslet na tím, které znaky těchto materiálů jsou pro danou komunitu důležité, a pokusit se zaznamenat tyto znaky jako významné vlastnosti.

Práva. Informace o **právech** nejsou pro archivaci jedinečné, ale znalost toho, co je dovoleno s objektem dělat, je pro archivační proces velmi důležitá. Měly by být zaznamenány jakékoli známé informace o **právech**, včetně stavu autorskoprávní ochrany, licenčních podmínek a speciálních povolení.

2. KLÍČOVÉ KONCEPTY PREMIS

2.1. Sémantické jednotky

Datový slovník PREMIS definuje sémantické jednotky, nikoli elementy metadat. Rozdíl je nepatrný, ale podstatný. Sémantická jednotka je část informace nebo znalosti. Element metadat je definovaný způsob reprezentace dané informace v metadatovém záznamu, schématu nebo databázi. PREMIS nespecifikuje, jak mají metadata být v systémech reprezentována, definuje pouze to, co má systém obsahovat a měl by být schopen převést

do systému jiného. Pokud chcete být PREMIS perfekcionista, musíte přemýšlet v pojmech sémantických jednotek. Pro ostatní jsou elementy metadat dostatečně přesným termínem.

Názvy sémantických jednotek PREMIS jsou tzv. „camelCase“ řetězce. To znamená, že slova nejsou oddělena mezerami, ale začínají velkými písmeny: `objectIdentifier`, `relatedEventIndetification`. V tomto dokumentu jsou vytištěna fontem Verdana.

2.2. Kontejnerové jednotky a podjednotky

Některé sémantické jednotky jsou definovány jako kontejnerové jednotky, což znamená, že samy neobsahují hodnotu, ale jejich účelem je sdružovat příbuzné sémantické jednotky. Například, kdykoli v PREMIS zaznamenáte identifikátor, musíte uvést, o jaký typ identifikátoru se jedná (např. DOI, ISBN, místní). Kontejner `objectIdentifier` je použit ke sdružení dvou podjednotek `objectIdentifierType` a `objectIdentifierValue`.

Kontejnerové jednotky poskytují Datovému slovníku hierarchickou strukturu, která je reflektována v číslování sémantických jednotek.

- 1.1 `objectIdentifier` (P, O)
 - 1.1.1 `objectIdentifierType` (P, NO)
 - 1.1.2 `objectIdentifierValue` (P, NO)

Na tomto úryvku z Datového slovníku na první pohled vidíme, že povinný (P) a opakovatelný (O) element `objectIdentifier` neobsahuje sám hodnotu, ale slouží jako kontejnerová jednotka pro dílčí elementy `objectIdentifierType` a `objectIdentifierValue`. Protože typ a hodnota jsou uvnitř kontejnerové jednotky neopakovatelné (NO), pokud bychom chtěli zaznamenat dva různé identifikátory, museli bychom zopakovat celou strukturu kontejnerové jednotky.

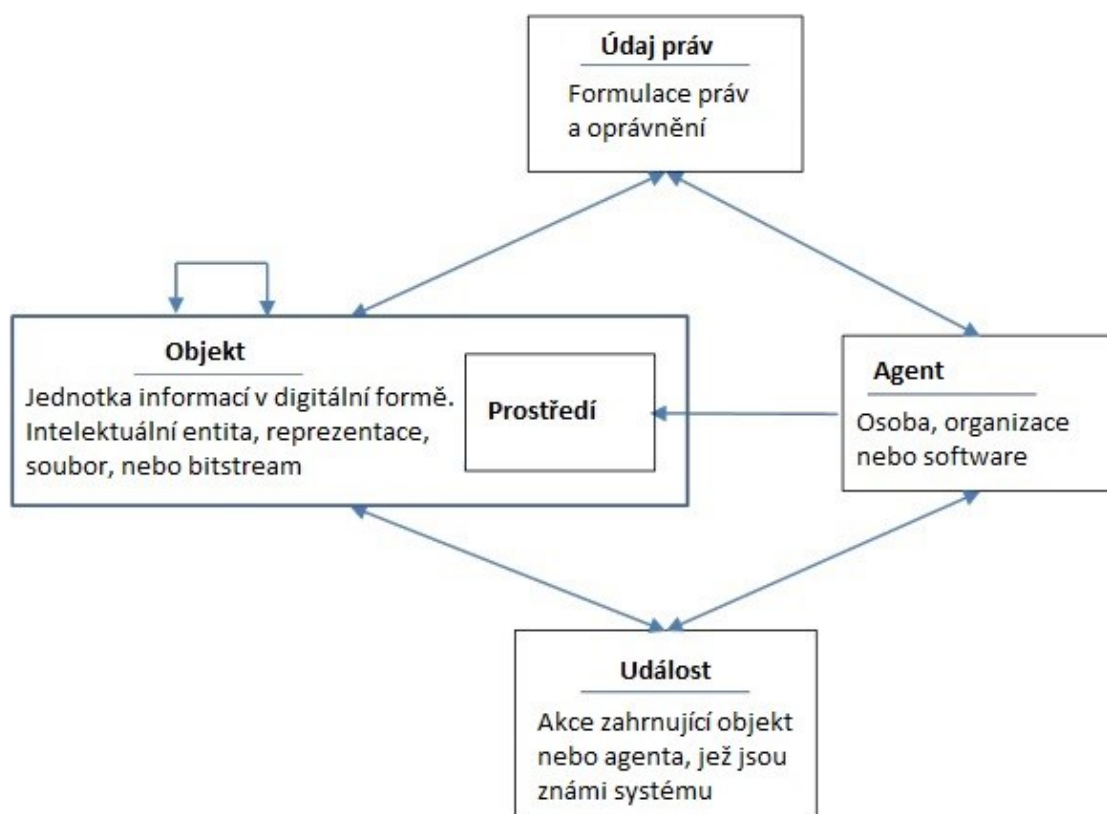
2.3. Kontejnerové jednotky rozšíření

Kontejnerová jednotka rozšíření je speciální typ kontejnerové jednotky, který pod sebou nemá definovány žádné podjednotky. Je navržen jako místo, kam jsou zaznamenána jiná než PREMIS metadata. Tímto způsobem může být PREMIS rozšířen tak, aby zahrnoval i metadata, která jsou jinak mimo oblast působnosti nebo nejsou zahrnuta do Datového slovníku z jiného důvodu. Poslední část názvu kontejnerových jednotek rozšíření je „Extension“. Například, technická metadata zaměřená na konkrétní formát nejsou zahrnuta do PREMIS, protože jsou mimo oblast působnosti, ale jsou to informace velmi důležité pro digitální archivaci. Místo pro zaznamenání technických metadat definovaných jinými schématy je v kontejnerové jednotce rozšíření `objectCharacteristicsExtension`. Může jít například o datový slovník Z39.87 pro bitmapové obrázky (jeho metadatové schéma je nazýváno „NISO Metadata pro obrázky v XML schématu—MIX“).

Pokud jste s PREMIS důvěrně obeznámeni, pak jistě víte, že Datový slovník PREMIS byl navržen tak, aby byl kompatibilní se značkovacím jazykem XML. Sémantické jednotky PREMIS mohou být implementovány jako XML elementy; kontejnerové jednotky jsou elementy, které obsahují pouze další elementy, kontejnerové jednotky rozšíření jsou kontejnerové jednotky pro elementy definované externím schématem. Více o PREMIS a XML najdete v sekci 5.1 níže. Pozdější snahy vyústily v ontologii PREMIS OWL, která se používá v aplikacích propojených dat (Linked Data). První verze ontologie byla kompatibilní s PREMIS verzí 2.2 a ke dni této práce je téměř hotová revize verze 3.0. Informace o ontologii PREMIS OWL jsou dostupné na adrese www.loc.gov/standards/premis/ontology.

3. DATOVÝ MODEL PREMIS

Jeden z hlavních principů PREMIS je úsilí o přesnost popisu. Ve verzích 1 a 2 PREMIS definoval pět druhů předmětů (nazývaných entity): **intelektuální entita**, **objekt**, **agent**, **událost** a **práva**. (Diagram Datového modelu pro verze 1 a 2 naleznete na webové stránce PREMIS.) Ve verzi 3, která byla vydána v červnu 2015, byl Datový model PREMIS revidován a **intelektuální entita** se stala další kategorií **objektu**.



Obrázek 2: Datový model PREMIS, verze 3

3.1. Objekty

Objekty jsou tím, co je uloženo a spravováno v archivním repozitáři. Většina PREMIS je věnována popisu digitálních objektů. Mezi informace, které mohou být zaznamenány, patří:

- unikátní identifikátor **objektu** (typ a hodnota),
- informace o neporušenosti jako kontrolní součet a algoritmus, jak ho odvodit,
- velikost **objektu**,
- formát **objektu**, který může být specifikován přímo nebo pomocí odkazu na registr formátů,
- originální název **objektu**,
- informace o jeho vzniku,
- informace o inhibitorech,
- informace o významných vlastnostech,
- informace o jeho **prostředí** (viz níže),
- kde a na jakém médiu je uložen,
- elektronický podpis,
- vztahy s jinými **objekty** a jinými typy entit.

PREMIS definuje čtyři různé typy **objektů** a požaduje po uživateli, aby je rozlišovali. Jsou to **bitstreamy**⁴, **soubory**, **reprezentace** a **intelektuální entity**.

Objekt typu **soubor** je přesně to, co si představíme, tedy počítačový soubor, např. PDF nebo JPEG.

Objekt typu bitstream je podmnožinou **objektů** typu **soubor**. **Bitstream** je definován jako data (bity) uvnitř souboru, která a) mají z hlediska archivace společné vlastnosti; a za b) nemohou stát samostatně bez přidání hlavičky souboru nebo jiné struktury.

Pokud máte například soubor ve formátu AVI (audio-video interleaved), zřejmě budete chtít odlišit audio **bitstream** od video **bitstreamu** a popsat tyto **bitstreamy** zvlášť.

Objekt typu reprezentace je sada všech **objektů** typu **soubor** potřebná k interpretaci **intelektuální entity**. Řekněme, že chcete archivovat webovou stránku, např. domovskou stránku vaší instituce v nějakém určitém čase. Je velmi pravděpodobné, že domovská stránka, kterou vidíte ve vašem prohlížeči, je vlastně složena z mnoha různých souborů – jeden nebo více HTML souborů, několik obrázků ve formátu GIF nebo JPEG, možná i audio nebo animace ve formátu Flash. Zřejmě také užívá stylování k vytvoření podoby, kterou vidíme. K zobrazení domovské stránky v prohlížeči jsou potřeba všechny tyto soubory. Z toho vyplývá, že pokud chceme v repozitáři archivovat interpretovatelnou domovskou stránku, musíme vědět o všech těchto souborech a jak je propojit. **Reprezentace** nám

⁴ Doslovně by se termín dal přeložit jako bitový tok.

umožňuje nejen identifikovat sadu souvisejících souborů, ale také popsat vlastnosti celku (např. celé webové stránky), který se může lišit od svých částí.

Intelektuální entita je definována jako sada obsahu, který je považován za samostatnou intelektuální jednotku určenou ke správě a popisu: například určitá kniha, mapa, fotografie, nebo databáze. PREMIS obvykle nedefinuje popisná metadata týkající se **intelektuálních entit**, protože existuje mnoho standardů pro popis metadat, z kterých lze vybrat. Ve verzích 1 a 2 mohly být **intelektuální entity** odkazovány pouze identifikátorem, protože byly považovány jen za konceptuální. Ve verzi 3 však může být **intelektuální entita** popsána popisnými metadaty, které nejsou součástí PREMIS, nebo pomocí archivačních metadat jako **objekt** v rámci PREMIS. Sémantické jednotky použité v **intelektuálních entitách** jsou většinou stejné i pro **reprezentace**.

Podle PREMIS by měl být **objekt** v archivačním systému propojen s **intelektuální entitou**, kterou reprezentuje, zahrnutím identifikátoru **intelektuální entity** v metadatach **objektu**. Například při archivaci výtisku Buddhism: The Ebook: an Online Introduction lze použít ISBN jako odkaz na popis **intelektuální entity** v PREMIS popisu knihy.

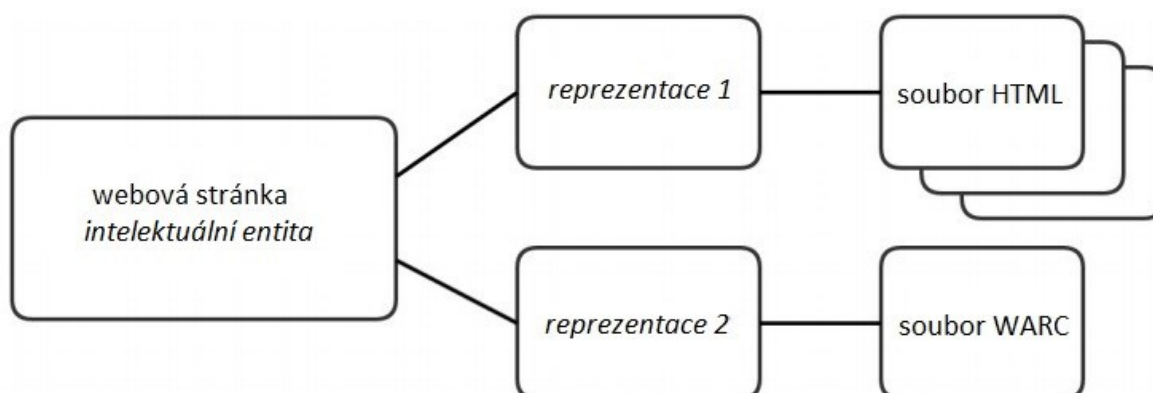
Některé sémantické jednotky definované v PREMIS datovém slovníku jsou aplikovatelné pro všechny čtyři typy **objektu**, zatímco jiné jsou aplikovatelné pouze v jednom nebo dvou typech **objektu**.

Některé sémantické jednotky jsou definovány tak, aby zaznamenaly **prostředí objektu**, tedy to, jaký hardware a software je potřeba k zobrazení **objektu**, a k jakým jiným **objektům** má vazby. Ve verzích 1 a 2 je tato informace součástí popisu **objektu**. Ve verzi 3 může být na **prostředí** jako zvláštní typ **intelektuální entity** odkázáno ze **souborů**, **reprezentací** a **bitstreamů**, které **intelektuální entitu** používají. Třeba PDF soubor může být zobrazen v několika verzích Adobe Acrobat a Adobe Reader, ale stejně tak i v jiných open source a komerčních programech. Každý z nich podporuje různé operační systémy a požaduje specifický hardware (rychlost procesoru, paměť a disk). Jelikož Adobe Reader není aplikací, která funguje samostatně, ale jde o rozšíření prohlížeče, je závislý na různých verzích jednotlivých prohlížečů; například Reader 9 pro operační systém macOS vyžaduje verzi 2.0.4 nebo novější prohlížeče Safari. Informace o **prostředí** je zásadní pro určité archivační strategie, ale je složité a časově náročné ji určit, a je zřejmě nejlepší, aby byla zaznamenána i v centrálním registru, jako je PRONOM spravovaný Národním archivem Velké Británie.⁵ PREMIS repozitářům umožňuje odkazovat na informace uložené v externích registrech, pokud je to výhodnější než lokální uložení.

Abychom ilustrovali vztahy mezi **objekty**, znovu si představme výše zmíněnou webovou stránku. Webová stránka je **intelektuální entita**, která může být vyjádřena několika různými **reprezentacemi**. Jedna **reprezentace** se, jak již bylo řečeno, skládá z mnoha jednotlivých

⁵ The National Archives of the UK

souborů (HTML stránka, obrázky, stylpis atd.). Nicméně webová stránka může být také archivována jako jediný soubor webového archivu (WARC) (viz www.digitalpreservation.gov/formats/fdd/fdd000236.shtml). Je to odlišná **repräsentace** stejné **intelektuální entity**, která však zahrnuje jen jeden soubor, jak vidíme na obrázku:



Obrázek 3: Příklad vztahů objektu v PREMIS

3.2. Události

Entita událost sdružuje informace o akcích, které ovlivňují **objekty** v repozitáři. Přesný a důvěryhodný záznam **události** je zásadní pro zachování digitálního původu objektu, který je zase důležitý pro demonstraci autenticity **objektu**.

Mezi informace, které mohou být o **události** zaznamenány, patří:

- unikátní identifikátor **události** (typ a hodnota),
- typ **události** (vznik, příjem, migrace atd.),
- datum a čas, kdy **událost** proběhla;
- detailní popis **události**,
- kódovaný výstup **události**,
- detailnější popis výstupu,
- **agenti**, kteří mají vztah k **události**, a jejich role,
- **objekty**, které mají vztah k **události**, a jejich role.

V každém systému repozitáře musí být rozhodnuto o tom, které **události** zaznamenat jako trvalou část historie **objektu**. PREMIS doporučuje vždy zaznamenat akce, které změní **objekt**, a poskytuje kontrolovaný slovník důležitých typů **událostí**, aby podpořil konzistentní zaznamenávání těchto **událostí**.

Tento slovník je na <http://id.loc.gov/vocabulary/preservation/eventType>.

3.3. Agenti

Agenti jsou činitelé, kteří mají role v **objektech událost**, prohlášeních **práv** (viz 3.4 Práva) a **prostředí**. **Agenti** mohou být lidé, organizace, softwarové aplikace nebo hardware. PREMIS definuje pouze minimum sémantických jednotek, které jsou nezbytné k identifikaci **agentů**, protože existuje několik externích standardů, které mohou být použity pro jejich detailnější popis. V repozitáři může být používán zvláštní standard pro zaznamenání přidáných informací o **agentech**, nebo identifikátor **agenta**, který odkáže na informace zaznamenané externě.

Datový slovník zahrnuje:

- unikátní identifikátor **agenta** (typ a hodnota),
- název **agenta**,
- určení typu **agenta** (osoba, organizace, software, hardware),
- verze **agenta** (pro software a hardware),
- obecná poznámka o **agentovi**,
- **události** asociované s **agentem**,
- **práva** asociovaná s **agentem**,
- **prostředí** asociovaná s **agentem**.

Kdykoli je **agent** zmíněn v souvislosti s **událostí** nebo prohlášením **práv**, měla by být zaznamenána také role **agenta**. Jakýkoli **agent** může mít mnoho rolí. Např. autor tohoto textu by mohl být autorem a držitelem práv jednoho díla, autorem (ale ne držitelem práv) druhého díla a deponentem třetího díla. Podle modelu PREMIS by mu v repozitáři byl přidělen unikátní identifikátor a tento identifikátor by byl odkazován ve všech záznamech **událostí** nebo prohlášení **práv**, ve kterých by byl **agentem**, spolu s jeho rolí v daném kontextu.

3.4. Práva

Většina archivačních strategií zahrnuje vytváření identických kopií a odvozených verzí digitálních objektů, což jsou akce, které by měly být podle autorského práva prováděny pouze držiteli práv. Entita **práva** sdružuje informace o právech a oprávněních, které náleží k objektům v repozitáři, aby bylo možné dělat vše pro zachování **objektů**. Každý údaj PREMIS **práv** dává na vědomí dvě věci: akce, které mohou být v repozitáři provedeny, a zdůvodnění nároku na tato práva.

Například, v repozitáři se může nacházet naskenovaná verze knihy publikované roku 1848, která je tedy v „public domain“⁶. S touto digitální verzí se díky jejímu právnímu statusu může

⁶ Jedná se o uvolnění autorských práv po uplynutí doby jejich ochrany. Ta v České republice trvá 70 let od smrti autora.

dělat cokoli. V jiném repozitáři se nachází objekt zkopírovaný z publikovaného CD, kde licence⁷ umožňuje vytvoření záložních kopií, ale omezuje přístup a užití.

Informace, které mohou být o **právech** zaznamenány, zahrnují:

- unikátní identifikátor údaje **práv** (typ a hodnota),
- zdůvodnění nároku na tato práva, zda se jedná o autorská práva, licenci, statut nebo např. institucionální pravidla,
- aplikovatelné detailnější informace o autorských právech, podmínky licence nebo statutu,
- akci nebo akce, které údaj **práv** umožňuje,
- omezení akce nebo akcí,
- termín grantu nebo omezení, nebo časový úsek, po který údaj platí,
- **objekt** nebo **objekty**, ke kterým se údaj vztahuje,
- **agenti** zapojení do údaje **práv** a jejich role.

Většina informací je navržena tak, aby byly prakticky použitelné (tzn. zaznamenány v kontrolované formě, se kterou může počítačový program pracovat). Prohlášení o **právech** v PREMIS je uplatněním práv, nikoli záznamem informací, ze kterých mohou být práva vyvozena. Tzn. PREMIS nedefinuje detailní informace o autorech, datu a místu publikace a notifikaci autorských práv, jaké jsou definovány např. ve specifikaci copyrightMD vytvořené v California Digital Library (<https://cdlib.org/groups/rights-management-group-copyrightmd/>). Účelem copyrightMD je pomoci lidem o právech rozhodovat, zatímco účelem entity **práva** je poskytnout systémům ochranných repozitářů prakticky použitelné informace.

4. DATOVÝ SLOVNÍK

4.1. Ukázkové heslo Datového slovníku pro jednoduchou sémantickou jednotku

Tabulka 1 níže obsahuje ukázkou hesla Datového slovníku pro sémantickou jednotku *size*⁸, která je součástí nebo je podjednotka kontejneru nazvaného *objectCharacteristics*⁹. *Size* sama nemá žádné podjednotky. Heslo Datového slovníku zahrnuje definici elementu a důvod (zdůvodnění), proč je zařazen mezi základní metadata PREMIS, stejně tak jako příklady a poznámky toho, jak hodnota může být získávána a užívána. To vše by mělo pomoci uživatelům používat tento element správně.

⁷ Jedná se o tzv. „shrink-wrap“ licenci, kterou kupující přijímá rozbalením obalu.

⁸ Do češtiny lze přeložit jako „velikost“.

⁹ Do češtiny lze přeložit jako „vlastnosti objektu“.

Dva řádky tabulky, „kategorie objektu“ a „použitelnost“, jsou použity k tomu, aby ukázaly, jestli je sémantická jednotka vhodná k popsání **intelektuální entity, reprezentace, souboru a/nebo bitstreamu**. Jak zde vidíme, size se vztahuje pouze k **souborům a bitstreamům**. Dále se zde nachází sady pravidel používání: datový typ, opakovatelnost a závaznost.

Datové omezení určuje hodnoty, kterých mohou sémantické jednotky nabývat. V tomto případě hodnotou velikosti musí být celé číslo. Dalším typickým případem datového omezení může být, že hodnota musí být převzata z řízeného slovníku. Pokud existuje zavedený řízený slovník, je v datovém omezení odkázán. Některé termíny ze slovníku jsou uvedeny jako příklady v Datovém slovníku:

<http://id.loc.gov/preservationdescriptions/>

Mohou být použity jiné řízené slovníky, v takovém případě musí být název slovníku zaznamenán. V Datovém slovníku nejsou pro názvy slovníků definovány žádné sémantické jednotky, ale PREMIS XML schéma pro ně poskytuje prostor, a pokud jsou identifikovány URI identifikátorem, ten bude obecně reflektovat název řízeného slovníku.

Opakovatelnost naznačuje, jestli mohou být sémantické jednotky opakovány.

Závaznost naznačuje, jestli je hodnota sémantické jednotky povinná nebo volitelná. Závaznost může být matoucí, protože PREMIS jasně uvádí, že repozitář nemá povinnost ukládat žádnou konkrétní informaci. Povinná sémantická jednotka neznámá, že musí být zaznamenána a uložena v repozitáři. Repozitář ale musí být schopen vygenerovat hodnotu sémantické jednotky, pokud je potřeba, jako pro výměnu s jiným repozitářem. Např. v nepravděpodobné situaci, kdy v repozitáři není uloženo nic kromě obrázků TIFF 6.0, není nutné zaznamenat informaci o formátu pro každý objekt. Nicméně repozitář by obsahoval informaci, že Soubory jsou obrázky TIFF 6.0, a mohl by ji poskytnout, pokud by bylo třeba. (Viz sekce 5.2. Soulad s PREMIS.) Některé sémantické jednotky jsou povinné uvnitř kontejneru; pokud kontejner není povinný, sémantická jednotka nebude použita, pokud není použit ani kontejner.

Tabulka 1: Výňatek sémantické jednotky size z Datového slovníku

Sémantická jednotka	1.5.3 size		
Sémantické podjednotky	Žádné		
Definice	Velikost souboru nebo bitstreamu uložených v repozitářích, udáváno v bytech		
Odůvodnění	Size se používá pro zajištění, aby bylo z úložiště získáno správné číslo bytů a aby aplikace měla dostatek prostoru přemísťovat nebo zpracovávat soubory. Může být také použita při vyhodnocování nákladů na provoz úložiště.		
Datové omezení	Celé číslo		
Kategorie Objektu	Intelektuální jednotka / Reprezentace	Soubor	Bitstream

Použitelnost	Není použitelný	Použitelný	Použitelný
Příklady		2038937	2038937
Opakovatelnost		Neopakovatelný	Neopakovatelný
Závaznost		Volitelný	Volitelný
Poznámky o vzniku / údržbě	Automaticky získáno v repozitáři.		
Poznámky k použití	Pro definování této sémantické jednotky v bytech není nutné zaznamenat jednotku rozměru. Nicméně, pro účel výměny dat by měla být jednotka rozměru ustanovena nebo vyrozuměna oběma stranami.		

4.2. Ukázka hesla z Datového slovníku pro kontejnerovou jednotku

Tabulka 2 zobrazuje začátek hesla Datového slovníku pro jednotku `objectCharacteristics`, která je kontejnerovou jednotkou zahrnující jednotku `size`. Je vidět, že se jedná o kontejnerovou jednotku, protože obsahuje sémantické podjednotky a datový typ je „kontejnerová jednotka“. Všimněme si, že zahrnuté sémantické podjednotky nemusí mít pod sebou další jednotky nebo mohou být samy kontejnerové jednotky, jako např. `format`¹⁰.

Tabulka 2: Výňatek sémantické jednotky `objectCharacteristics` z Datového slovníku

Sémantická jednotka	1.5 <code>objectCharacteristics</code>		
Sémantické podjednotky	1.5.1 <code>compositionLevel</code>		
	1.5.2 <code>fixity</code>		
	1.5.3 <code>size</code>		
	1.5.4 <code>format</code>		
	1.5.5 <code>creatingApplication</code>		
	1.5.6 <code>inhibitors</code>		
	1.5.7 <code>objectCharacteristicsExtension</code>		
Definice	Technické vlastnosti souboru nebo bitstreamu , které jsou použitelné na všechny nebo většinu formátů.		
Odůvodnění	Existuje několik důležitých technických vlastností, které se používají u objektů jakéhokoli formátu. Detailní definice vlastností zaměřených na formát je v tomto Datovém slovníku vyloučena, přesto ale tyto vlastnosti mohou být zahrnuty v <code>objectCharacteristicsExtension</code> .		
Datové omezení	Kontejner		
Kategorie Objektu	Intelektuální jednotka / Reprezentace	Soubor	Bitstream
Použitelnost	Není použitelný	Použitelný	Použitelný
Opakovatelnost		Opakovatelný	Opakovatelný
Závaznost		Povinný	Povinný

¹⁰ Do češtiny lze přeložit jako “formát”.

Poznámky k použití

Se sémantickými jednotkami zahrnutými v kontejneru `objectCharacteristics` by mělo být zacházeno jako se skupinou informací příslušejících k jedinému objektu v jednom `compositionLevel`. Vlastnosti objektu mohou být opakovány, pokud byl objekt vytvořen aplikací dvou nebo více kódování, např. jako komprese a šifrování. V tomto případě by každé opakování `objectCharacteristics` mělo vyšší `compositionLevel`.

Pokud je aplikováno šifrování, blok `objectCharacteristics` musí obsahovat sémantickou jednotku inhibitor.

Bitstream vnořený do souboru může mít odlišné vlastnosti než soubor. Pokud jsou tyto vlastnosti relevantní pro archivaci, měly by být zaznamenány.

5. PREMIS V PRAXI

5.1. PREMIS v XML

Předpokládá se (nikoli vyžaduje), že pokud je PREMIS použit pro výměnu dat, je zobrazen v XML. Podpora PREMIS poskytuje schéma v XML, které odpovídá Datovému slovníku, a poskytuje přímý popis **objektů**, **událostí**, **agentů** a **práv**. Obrázek 4 zobrazuje ukázkou metadata PREMIS v XML schématu.

V praxi většina archivačních systémů už používá XML formáty pro import a export dat. Většinou je využíván standard METS (Metadata Encoding and Transmission Standard) spravovaný Knihovnou Kongresu, a to jako XML kontejnerová jednotka pro spojení různých typů metadat dohromady. Je možné použít PREMIS ve standardu METS, ale toto užití není zcela jednoduché, a to ze dvou důvodů. Za prvé, METS rozděluje informace do různých sekcí podle toho, jestli se jedná o technická metadata, metadata práv, nebo metadata provenience. Podle Datového slovníku má schéma PREMIS sekce pro **objekty**, **práva**, **události** a **agenty**. Je tu tedy určitá shoda mezi těmito dvěma strukturami, není ale přesná, a to především u informací o **agentovi**. Za druhé, PREMIS a METS se někdy překrývají; například, oba standardy definují označení pro uložení kontrolního součtu. Pokud použijeme oba standardy dohromady, musíme se rozhodnout, jestli zaznamenáme tyto průniky elementů v sekcích PREMIS, METS, nebo v obou.

Pokud by každá implementace rozhodla jinak, vedlo by to k různým způsobům reprezentace dat a ztížení interoperability. Proto uživatelé vyvinuli doporučení (best practice) pro užívání PREMIS a METS dohromady. Průvodce pro využití PREMIS s METS pro

výměnu dat¹¹ je dostupný na webových stránkách Podpory PREMIS: <http://www.loc.gov/standards/premis/guidelines2017-premismets.pdf>.

```
<event>
  <eventIdentifier>
    <eventIdentifierType>DAITSS</eventIdentifierType>
    <eventIdentifierValue>10012</eventIdentifierValue>
  </eventIdentifier>
  <eventType>Format Validation</eventType>
  <eventDateTime>2008-05-06T10:40:22-04:00</eventDateTime>
  <eventOutcomeInformation>
    <eventOutcome>Invalid</eventOutcome>
    <eventOutcomeDetail>
      <eventOutcomeDetailNote>ill-formed DateTime value</eventOutcomeDetailNote>
    </eventOutcomeDetail>
  </eventOutcomeInformation>
</event>
```

Obrázek 4: Ukázka PREMIS v XML

5.2. Soulad s PREMIS

Specifikace PREMIS obsahuje sekci, která říká, co pro repozitář znamená odpovídat PREMIS. Formulace souladu, kterou PREMIS Editační komise poprvé vydala v říjnu 2009 a revidovala v roce 2015, specifikuje principy použití, úrovně souladu a poradenství k implementaci. Je dostupná na adrese: <http://www.loc.gov/standards/premis/premis-conformance-20150429.pdf>.

Principy použití

- 1) Pokud repozitář implementuje (uchovává nebo exportuje) datové elementy, které mají význam sémantických jednotek PREMIS, datové elementy by měly mít stejnou definici, datové omezení a použitelnost jako sémantická jednotka definována v PREMIS. Pokud má metadatový element stejnou definici jako sémantická jednotka, ale nemá stejný název, v repozitáři musí být vybudováno mapování mezi metadatovým elementem a příslušnou sémantickou jednotkou PREMIS.
- 2) Pokud je do repozitáře implementována sémantická jednotka PREMIS, její opakovatelnost a závaznost může být přísnější, ale nikoli volnější než PREMIS požaduje. To znamená, že opakovatelná sémantická jednotka může být implementována jako neopakovatelná, ale ne naopak, a povinná jednotka nemůže být změněna na volitelnou.

¹¹ Anglický název: Guidelines for Using PREMIS with METS for Exchange

- 3) Implementace musí obsahovat povinné sémantické jednotky pro všechny entity Datového modelu (**objekt**, **událost**, **agent** a **práva**) podporované repozitářem. Všimněme si, že povinné sémantické podjednotky jsou zahrnuty, pouze pokud je realizován rodičovský kontejner.
- 4) Pokud jsou z repozitáře exportovány informace pro využití jiným repozitářem, musí dodat hodnoty pro všechny sémantické jednotky, které jsou v Datovém slovníku povinné. Nicméně je to poměrně flexibilní, protože repozitáře nemají povinnost podporovat sémantické jednotky povinné pro typy entit, které nepodporují. Jinými slovy, v repozitáři můžeme volně rozhodovat, zda budeme nebo nebudeme podporovat **agenta** PREMIS, ale pokud ho podporovat budeme, pak je agentIdentifier¹² povinný. Obdobně nemusí konkrétní repozitář podporovat **objekty bitstream**, a v takovém případě nemusí poskytovat identifikátor **bitstreamu**, který by byl jinak povinný.

Úrovně souladu

Úrovně specifikují tři možnosti uplatňování souladu s PREMIS v repozitářovém systému:

- **Úroveň 1.** Repozitář je schopen namapovat archivační metadata na PREMIS.
- **Úroveň 2.** Repozitář je schopen exportovat archivační metadata jako PREMIS.
- **Úroveň 3.** Repozitář používá PREMIS jako interní schéma takovým způsobem, který nevyžaduje žádné další mapování nebo konverzi.

Tyto úrovně jsou dále rozděleny do kategorií: implementace pouze entity **objekt** nebo implementace entity **objekt** a jedné nebo více dalších entit. Jinými slovy, v repozitáři musí být implementovány alespoň informace o **objektu**, ale podpora všech typů entit definovaných v datovém modelu PREMIS není povinná. Také není povinné ukládat metadata interně pod názvy sémantických jednotek PREMIS ani používat hodnoty, které odpovídají datovým omezením PREMIS. Jinými slovy, nezáleží na tom, jakým způsobem je v repozitáři hodnota PREMIS uložena – může být uložena pod stejným nebo odlišným názvem, namapována z jiné hodnoty, uložena pomocí odkazu do registru, pomocí odvození, může být nastavena jako výchozí nebo být uložena jakýmikoli jinými prostředky. Jestliže repozitář může poskytnout odpovídající hodnotu PREMIS, pokud je potřeba, je v souladu s PREMIS.

Na druhou stranu, čím více sémantických jednotek repozitář podporuje, tím větší přínos z využívání PREMIS získá. Datový slovník PREMIS byl vytvořen tak, aby identifikoval základní informace, které většina repozitářů bude potřebovat k dlouhodobé archivaci digitálního obsahu. Zodpovědní správci archivního repozitáře by se měli pozorně podívat na PREMIS a pokud by se rozhodli neimplementovat jakoukoli část Datového slovníku, měli by pro to mít dobrý důvod.

¹² Do češtiny lze přeložit jako „identifikátor agenta“.

6. DOPLŇKOVÉ INFORMACE O PREMIS

Webová stránka Podpory PREMIS (<http://www.loc.gov/standards/premis/>) má detailní informace pro každého, včetně odkazů na PREMIS Implementers Group (PIG), PREMIS implementation fairs¹³ a tutoriály, schémata, nástroje a novinky. Podpora PREMIS má také sekci „PREMIS Resources“, která odkazuje na literaturu o PREMIS a příbuzných tématech (<http://www.loc.gov/standards/premis/bibliography.html>). Některé z užitečnějších zdrojů pro běžného čtenáře jsou následující:

O archivačních metadatech obecně:

LAVOIE, Brian a Richard GARTNER. *Preservation Metadata (2nd edition)* [online]. Oxford: Digital Preservation Coalition, 2013 [cit. 2018-04-14]. ISSN 2048-7916. Dostupné z (DOI): <http://dx.doi.org/10.7207/twr13-03>.

O implementaci PREMIS:

DAPPER, Angela, Rebecca Squire GUENTHER a Sébastien PEYRARD, ed. *Digital preservation metadata for practitioners: implementing PREMIS* [online]. Cham: Springer, 2016 [cit. 2018-04-14]. ISBN 978-3-319-43761-3. Online ISBN 978-3-319-43763-7. Dostupné z (DOI): <https://doi.org/10.1007/978-3-319-43763-7>.

O změnách na verzi 3.0 (webinář):

DAPPER, Angela. Digital Preservation Metadata and Improvements to PREMIS in Version 3.0. In: *PREMIS: Preservation Metadata Maintenance Activity* [online]. May 27, 2015 [cit. 2018-04-14]. 70 min. Dostupné z: www.loc.gov/standards/premis/v3/tutorial.html.

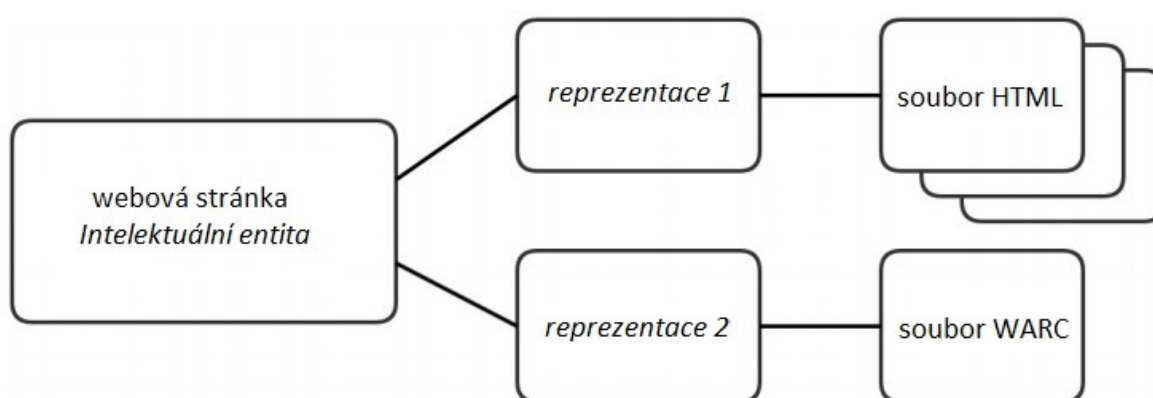
¹³ Jedná se o název mezinárodního workshopu, který se standardem PREMIS zabývá.

Příloha A: Příklady

A.1. Příklad objektu

Tento příklad popisuje, jak může být webová stránka modelována v hypotetickém archivačním systému. Příklad obsahuje obecný diagram, dále detaily, jak mohou být sémantické jednotky příkladu naplněny, a nakonec několik ukázek XML, abychom viděli, jak mohou být sémantické jednotky do repozitáře implementovány. Příklady u každé úrovni **objektu** neukazují všechny relevantní sémantické jednotky.

V tomto příkladu se webová stránka nazývá „PREMIS Website“. Je to **intelektuální entita**, která má být uchována. Existují dvě **reprezentace** této webové stránky. V tomto příkladu je jedna **reprezentace** archivní kopie a druhá uživatelská kopie.



Intelektuální entita

Úroveň **intelektuální entity** v tomto příkladu stanoví jednu významnou vlastnost této webové stránky. Ta řídí chování obou **reprezentací**. Je samozřejmě mnohem pravděpodobnější, že jakákoli významná vlastnost bude mnohem méně spojitá, ale příklad slouží jako ukázka toho, jak mohou být významné vlastnosti využity na úrovni **intelektuální entity**.

Sémantická jednotka PREMIS	Intelektuální entita 1
1.1 objectIdentifier	
1.1.1 objectIdentifierType	IEPID
1.1.2 objectIdentifierValue	17415491
1.4 significantProperties	
1.4.1 significantPropertiesType	chování
1.4.2 significantPropertiesValue	Externí odkazy se otevírají v novém okně prohlížeče

Reprezentace

Úroveň archivace popisuje institucionální pravidla pro tento **objekt**. Úroveň archivace může být nastavena pro **intelektuální entitu**, **reprezentaci** nebo **soubor**. V tomto případě byla nastavena pro **reprezentaci**, protože instituce si přeje, aby se o každou **reprezentaci** pečovalo jinak. Instituce se v tomto příkladu rozhodla, že **reprezentace**, která obsahuje soubor WARC (**reprezentace 2**), je považována za archivní kopii. Bude dlouhodobě zachována a budou z ní vytvářeny uživatelské kopie. **Reprezentace 1** je uživatelská kopie, jejím cílem je být tou **reprezentací**, kterou využívá veřejnost k prohlížení. Její hodnota není dlouhodobá, bude uchována, dokud bude využíván současný zpřístupňovací mechanismus.

Reprezentaci 1 byla přisouzena hodnota „bitové archivace“. Podle institucionální politiky jsou vytvářeny pouze záložní kopie a není nutná žádná analýza rizik ani následná emulace nebo migrace.

Reprezentace 2 byla na druhé straně zařazena jako plná „logická archivace“. Ta je institucí definována jako **reprezentace**, která podstoupí relevantní analýzu rizik a na kterou je v případě potřeby aplikována emulace nebo migrace, aby byla zajištěna archivace pro budoucnost.

PREMIS Sémantická jednotka	Reprezentace 1	Reprezentace 2
1.1 objectIdentifier		
1.1.1 objectIdentifierType	REPPID	REPPID
1.1.2 objectIdentifierValue	17415492	17415495
1.3. preservationLevel		
1.3.1 preservationLevelType	bitová archivace	logická archivace
1.3.2 preservationLevelValue	nízká	plná
1.3.3 preservationLevelRole	schopnost	záměr
1.3.4 preservationLevelRationale	institucionální politika	institucionální politika
1.3.5 preservationLevelDateAssigned	23/02/2015	23/02/2015
1.13 relationship		
1.13.1 relationshipType	derivace	derivace
1.13.2 relationshipSubType	má zdroj	je zdrojem
1.13.3 relatedObjectIdentifier		
1.13.3.1 relatedObjectIdentifierType	REPPID	REPPID
1.13.3.2 relatedObjectIdentifierValue	17415495	17415492

Ukázka XML pro archivní úroveň by mohla vypadat například takto:

```
<premis:preservationLevel>
  <premis:preservationLevelType>logical preservation</premis:preservationLevelType>
  <premis:preservationLevelValue>full</premis:preservationLevelValue>
  <premis:preservationLevelRole authority="preservationLevelRole"
    authorityURI="http://id.loc.gov/vocabulary/preservation/preservationLevelRole"
    valueURI="http://id.loc.gov/vocabulary/preservation/preservationLevelRole/int">intention
  </premis:preservationLevelRole>

  <premis:preservationLevelDateAssigned>2015-02-23
</premis:preservationLevel>
```

Soubory

Reprezentace 1 zahrnuje počet jednotlivých souborů, z nichž každý je popsán jako **objekt soubor** s vlastní sadou metadatových elementů.

PREMIS Sémantická jednotka	Soubor 1	Soubor 2
1.1 objectIdentifier		
1.1.1 objectIdentifierType	FILEPID	FILEPID
1.1.2 objectIdentifierValue	17415493	17415494
1.5 objectCharacteristics		
1.5.1 compositionLevel	0	0
1.5.2 fixity		
1.5.2.1 messageDigestAlgorithm	SHA256	SHA256
1.5.2.2 messageDigest	d2bed92b73c7090bb3 0a0b30016882e7069c 437488e1513e9deaac be29d38d92	074862dbfa0806ef 5a26c3ff748d394e 79728e9c957ff8c19 8ae13214c9cfec0
1.5.2.3 messageDigestOriginator	NRI	NRI
1.5.3 size	4859385	123451
1.5.4 format		
1.5.4.1 formatDesignation		
1.5.4.1.1 formatName	Extensible Hypertext Markup Language (XHTML)	JPEG File Interchange Format
1.5.4.1.2 formatVersion	1	1.02
1.5.4.2 formatRegistry		
1.5.4.2.1 formatRegistryName	PRONOM	PRONOM
1.5.4.2.2 formatRegistryKey	fmt/102	fmt/44
1.5.4.2.3 formatRegistryRole	specifikace	specifikace
1.13 relationship		
1.13.1 relationshipType	strukturální	strukturální
1.13.2 relationshipSubType	je zahrnut v	je zahrnut v
1.13.3 relatedObjectIdentifier		
1.13.3.1 relatedObjectIdentifierType	REPPID	REPID

1.13.3.2 relatedObjectIdentifierValue	17415492	17415492
---------------------------------------	----------	----------

Ukázka XML pro objectCharacteristics **objektu 1** by vypadala takto:

```
<premis:objectCharacteristics>
  <premis:compositionLevel>0</premis:compositionLevel>
  <premis:fixity>
    <premis:messageDigestAlgorithm>SHA256</premis:messageDigestAlgorithm>
    <premis:messageDigest>d2bed92b73c7090bb30a0b30016882e7069c437488e1513e9deaacbe29d38d92
      </premis:messageDigest>
    <premis:messageDigestOriginator>NRI</premis:messageDigestOriginator>
  </premis:fixity>
  <premis:size>4859385</premis:size>
  <premis:format>
    <premis:formatDesignation>
      <premis:formatName>Extensible Hypertext Markup Language</premis:formatName>
      <premis:formatVersion>1</premis:formatVersion>
    </premis:formatDesignation>
    <premis:formatRegistry>
      <premis:formatRegistryName>PRONOM</premis:formatRegistryName>
      <premis:formatRegistryKey>fmt/102</premis:formatRegistryKey>
      <premis:formatRegistryRole>specification</premis:formatRegistryRole>
    </premis:formatRegistry>
  </premis:format>
</premis:objectCharacteristics>
```

A.2. Příklad události

Tento příklad popisuje **událost** provedenou repozitářem na objektu. Obvyklá funkce repozitáře je identifikovat formát souboru objektu a ověřit, zda odpovídá standardům pro tento formát, zároveň uvést míru shody nebo nalezené výjimky. Tento příklad demonstruje validaci formátu **souboru 1** z předchozího příkladu **objektu**.

PREMIS Sémantická jednotka	Soubor 1
2.1 eventIdentifier	
2.1.1 eventIdentifierType	DPS
2.1.2 eventIdentifierValue	25
2.2 eventType	validace
2.3 eventDateTime	2013-07-01T17:23:25Z
2.4 eventDetailInformation	
2.4.1 eventDetail	Na souboru byla provedena identifikace formátu
2.5 eventOutcomeInformation	
2.5.1 eventOutcome	úspěch
2.5.2 eventOutcomeDetail	
2.5.2.1 eventOutcomeDetailNote	FORMAT_ID=fmt/102;IDENTIFICATION_METHOD=SIGNATURE;FILE_EXTENSION=html;DEPOSIT_ACTIVITY_ID=659202;PID=FL26274026;SIP_ID=444381;PRODUCER_ID=29408010705;TASK_ID=48;PROCESS_ID=;MF_ID=8
2.6 linkingAgentIdentifier	

2.6.1 linkingAgentIdentifierType	Lokální
2.6.2 linkingAgentIdentifierValue	DROID6_63_1
2.6.3 linkingAgentIdentifierRole	Uživatel
2.7 linkingObjectIdentifier	
2.7.1 linkingObjectIdentifierType	FILEPID
2.7.2 linkingObjectIdentifierValue	17415493

Ukázka XML pro **událost** identifikace formátu by vypadalo takto:

```
<premis:event>
  <premis:eventIdentifier>
    <premis:eventIdentifierType>DPS</premis:eventIdentifierType>
    <premis:eventIdentifierValue>25</premis:eventIdentifierValue>
  </premis:eventIdentifier>
  <premis:eventType authority="event type"
    authorityURI="http://id.loc.gov/vocabulary/preservation/eventType"
    valueURI="http://id.loc.gov/vocabulary/preservation/eventType/val"
    validation</premis:eventType>
  <premis:eventDateTime>2013-07-01T17:23:25Z</premis:eventDateTime>
  <premis:eventDetailInformation>
    <premis:eventDetail>Format identification performed on file</premis:eventDetail>
  </premis:eventDetailInformation>
  <premis:eventOutcomeInformation>
    <premis:eventOutcome>success</premis:eventOutcome>
    <premis:eventOutcomeDetail>
      <premis:eventOutcomeDetailNote>
        FORMAT_ID=fmt/102; IDENTIFICATION_METHOD=SIGNATURE; FILE_EXTENSION=html;
        DEPOSIT_ACTIVITY_ID=659202; PID=FL26274026; SIP_ID=444381;
        PRODUCER_ID=29408010705; TASK_ID=48; PROCESS_ID=; MF_ID=8</premis:eventOutcomeDetailNote>
      </premis:eventOutcomeDetail>
    </premis:eventOutcomeInformation>
  <premis:linkingAgentIdentifier>
    <premis:linkingAgentIdentifierType>LOCAL</premis:linkingAgentIdentifierType>
    <premis:linkingAgentIdentifierValue>DROID6_63_1</premis:linkingAgentIdentifierValue>
    <premis:linkingAgentRole authority="eventRelatedAgentRole"
      authorityURI="http://id.loc.gov/vocabulary/preservation/eventRelatedAgentRole"
      valueURI="http://id.loc.gov/vocabulary/preservation/eventRelatedAgentRole/imp"
      implementer</premis:linkingAgentRole>
    </premis:linkingAgentIdentifier>
  <premis:linkingObjectIdentifier>
    <premis:linkingObjectIdentifierType>FILEPID</premis:linkingObjectIdentifierType>
    <premis:linkingObjectIdentifierValue>17415493</premis:linkingObjectIdentifierValue>
  </premis:linkingObjectIdentifier>
</premis:event>
```

A.3 Příklad agenta

Tento příklad popisuje **agenta** asociovaného s **událostí** popsanou výše. V tomto případě je **agentem** software, který akci na **objektu** provedl.

PREMIS Sémantická jednotka	
3.1 agentIdentifier	
3.1.1 agentIdentifierType	Lokální
3.1.2 agentIdentifierValue	DROID6_63_1

3.2 agentName	REG_SA_DROID
3.3 agentType	Software
3.4 agentVersion	Verze 6.01
3.5 agentNote	Signature verze Binary SF v.63/ Kontejner SF v.1

Ukázka XML pro informace o **agentovi** by vypadala takto:

```
<premis:agent>
  <premis:agentIdentifier>
    <premis:agentIdentifierType>local</premis:agentIdentifierType>
    <premis:agentIdentifierValue>DROID6_63_1</premis:agentIdentifierValue>
  </premis:agentIdentifier>
  <premis:agentName>REG_SA_DROID</premis:agentName>
  <premis:agentType>software</premis:agentType>
  <premis:agentVersion>Version 6.01</premis:agentVersion>
  <premis:agentNote>Signature version Binary SF v.63/ Container SF    v.1</premis:agentNote>
</premis:agent>
```

Příloha B: Slovníček termínů

Tento slovníček obsahuje definice, které se dříve objevily v textu tohoto průvodce. Definice mohou být méně formální než ty, které se objevují v Datovém slovníku pro archivační metadata PREMIS.

agent: Osoba, organizace, nebo počítačový program, které hrají roli ve vztahu k Události nebo prohlášení o Právech.

archivační metadata: Metadata podporující aktivity určené k zajištění dlouhodobé použitelnosti digitálního zdroje.

archivační strategie: Techniky použité k zajištění dlouhodobé použitelnosti digitálního zdroje; dvě obecné strategie jsou migrace a emulace.

digitální původ: Dokumentace provozního řetězce a historie změn digitálního zdroje.

emulace: Archivační strategie zahrnující kopírování interpretačního prostředí na nový hardware a/nebo software.

entita událost: Entita PREMIS, která získává informace o aktivitách, které ovlivňují objekty v repozitáři.

inhibitory: Vlastnosti digitálního objektu určené k zákazu přístupu, použití nebo migraci dat.

intelektuální entita: Sada obsahů, která je zpracována jako jednotka pro zpracování a popis; podobá se "bibliografické entitě" v knihovnictví.

kontejnerová jednotka: Sémantická jednotka, která nemá žádnou vlastní hodnotu, ale sjednocují související podjednotky.

kontejnerová jednotka pro rozšíření: Speciální typ metadatového objektu PREMIS, který nemá pod sebou definované podjednotky, ale slouží jako prostor pro jiná metadata než PREMIS.

metadata práv: Entita PREMIS, která získává informace o právech a povoleních týkajících se Objektů v archivních repozitářích.

migrace: Archivační strategie, která zahrnuje vytváření verzí digitálních souborů v novém formátu souboru.

objekt bitstream: Typ objektu PREMIS; data uvnitř souboru, která mají společné vlastnosti z hlediska archivace a která nemohou stát samostatně.

objekt prostředí: Objekt, který je součástí technického balíku softwaru, hardwaru, a jiných závislostí potřebných ke správné interpretaci reprezentací, souborů a bitstreamů.

objekt reprezentace: Typ Objektu PREMIS; sada všech objektů soubor potřebných k interpretaci intelektuální entity.

objekt soubor: Typ objektu PREMIS; počítačový soubor jako například PDF nebo JPEG.

objekty: Digitální předměty, které jsou uloženy a zpracovány v archivních repozitářích. PREMIS definuje čtyři typy objektů: soubory, bitstreamy, reprezentace a intelektuální entity.

praktická použitelnost: Objekt je zaznamenaný v řízené formě, a proto na něm může počítačový program vykonávat akce.

prostředí: Hardware, software, a jiné objekty nutné k interpretaci objektu.

sémantická jednotka: Jednotky informací nebo znalostí.

významné vlastnosti: Vlastnosti objektu, které by měly být zpracovány během archivačních opatření.