

## Použití metadat ve veřejné správě

Efektivní správa obrovského množství údajů veřejné správy si žádá vývoj a nasazení robustních, udržitelných a interoperabilních metadatových struktur. Vládní instituce potřebují metadata k řízení, pochopení, zpřístupnění a uchování jejich životně důležitého informačního majetku v průběhu času a napříč oblastmi využití. Různé komunity odborníků nahlíží na metadata rozdílně, jelikož každá komunita je jinak používá. Odborníci z ICT nahlíží na metadata jako na data, která popisují data a datové systémy: to znamená strukturu databází, jejich vlastnosti, umístění a použití. Na druhou stranu odborníci informačního managementu pohlíží na metadata jako na strukturované informace, jež popisují a/nebo umožňují vyhledání, řízení, kontrolu, pochopení a uchování dalších informací v průběhu času. Jinými slovy, metadata dokumentují obsah, kontext a strukturu informačních zdrojů s cílem podpořit trvalé využívání těchto informačních zdrojů.

### Cíle

Vládní instituce jsou nyní silně závislé na využívání informačních a komunikačních technologií a data nebo digitální informace jsou v současné době devizou veřejné správy. V každé zemi informace veřejného sektoru představují významný prvek celkové informační a znalostní ekonomiky. Tyto informace vytvářejí základnu pro vládní rozhodnutí. Poskytují virtuální paměť o vládních rozhodnutích učiněných v minulosti a o její činnosti, což umožňuje vládě vlastní hodnocení. Informace veřejného sektoru jsou často nedostatečně využitým národním strategickým zdrojem. Přitom je životně důležité, aby informace veřejného sektoru byly efektivně řízeny způsobem, který umožní maximální přínos z této obrovské národní investice.

Využití metadat v oblasti veřejné správy zahrnuje následující:

- Správu dat (např. statistických dat)
- Vytváření a zpřístupňování zdrojů (např. standardizované popisy vládních on-line zdrojů a služeb e-governmentu)
- Uchování záznamů (dokumentů)
- Správu a zpřístupnění geoprostorových dat
- Ochranu soukromí
- IRM Řízení práv k informacím (např. pro správu duševního vlastnictví a řízení bezpečnosti)
- Digitální uchování (tj. dlouhodobého a nepřetržitého zajištění životně důležitého informačního majetku)
- Dokumentování úrovně souladu informačních zdrojů s normami přístupnosti konsorcia World Wide Web (W3C).

Neúplná a nedostatečná metadata jsou výsledkem špatné správy a nedostatečného využití informačního majetku veřejného sektoru, což vede k plýtvání veřejnými prostředky. Neúplná a nedostatečná metadata mohou mít za následek například:

- Neschopnost včas najít a/nebo sdílet důležité informace z důvodu nedostatku adekvátních popisných metadat;
- Nemožnost přečíst/využít digitální informace kvůli nedostatečným technickým metadatům o struktuře a technických vlastnostech digitálních informačních objektů;
- Neschopnost přisoudit význam nebo hodnotu informacím z důvodu nedostatku kontextuálních metadat, a
- Neschopnost ověřit pravost a důvěryhodnost informací.

### Zkušenosti

V propojeném světě, kde 'ucelená vláda' je naprosto nezbytná, potřebují být metadata obvykle sdílená/vyměnitelná a znovu použitelná různými subjekty k různým účelům, obvykle pomocí automatizovaných systémů. Vzhledem k tomu, že stroje vyžadují jasně definované algoritmy, aby úspěšně

## Odkazy, 1. část

AGLS Metadata Standard  
<http://www.naa.gov.au/agls>

Davies, John, *Semantic Web Technologies: Trends and Research in Ontology-based Systems*. Wiley, 2006. ISBN 0470025964

Day, Michael, 'Metadata', v: *DCC Digital Curation Manual* (2005). Tato část manuálu je dobrým úvodem do problematiky metadat a uchování metadat, zejména v souvislosti se správou digitálního obsahu. Poskytuje definice, diskutuje význam a použití metadat.

Je zde i část věnovaná modelu OAIS, uchovávacím metadatům, standardům a formátům pro balení metadat jako je METS.

<http://www.dcc.ac.uk/resource/curation-manual/chapters/metadata/>

Dublin Core Metadata Initiative  
<http://www.dublincore.org/>

Dublin Core Metadata Initiative Glossary  
<http://www.dublincore.org/documents/usageguide/glossary.shtml>

Evans, Joanne, Barbara Reed and Sue McKemmish, 'Interoperable Data: Sustainable Frameworks for Creating and Managing Recordkeeping Metadata', *Records Management Bulletin*, vol. 18, no. 2, 2008, pp. 115-129.

Úvod do ontologií a sémantický Web.  
<http://obitko.com/tutorials/ontologies-semanticweb/>

zpracovaly metadata (tj. aby metadata byla interoperabilní), je důležité pro režimy metadat dosažení souladu s uznávanými průmyslovými standardy. Normy pro metadata standardizují jeden nebo více ze tří hlavních aspektů metadat:

1. Strukturu (jak jsou metadata strukturovaná - často na úroveň základních prvků informací nebo 'vlastností' v souladu s určitým (konkrétním) datovým modelem nebo ontologií);
2. Sémantiku (co základní prvky nebo vlastnosti metadat znamenají), a
3. Syntaxi (jak jsou metadata zapsána/vyjádřena/zakódována pomocí obvyklých značkovacích jazyků jako HTML a XML a aby hodnoty dat byly v souladu s určenými řízeními slovníky a kódováním).

Některé důležité normy pro metadata používané ve veřejné správě zahrnují:

- ISO / IEC 11179 - Standard pro metadata registrů. Metadata registrů poskytují spolehlivý zdroj informací o významu daného prvku metadat nebo části metadat. Metadata registrů v souladu s tímto standardem se skládají z hierarchie 'konceptů' se souvisejícími vlastnostmi pro každou koncepci. Každý koncept a vlastnost musí mít přesně formulovanou definici datového prvku.
- ISO 15836 - Sada základních prvků metadat Dublin Core (vládními běžně užívaný standard pro on-line poskytování zdrojů. Například australská vláda přijala standard metadat AGLS, což je rozšíření a použití tvaru Dublin Core a který je používán k popisu vládních informací a služeb).
- ISO 19115 - Geografické informace - Metadata.
- ISO 23081 - Metadata pro záznamy - Část 1: Principy a Část 2: Konceptuální a prováděcí záležitosti.
- ISO 8601 - norma pro kódování časových informací (datum a čas).
- PREMIS (Preservation Metadata Implementation Strategies) - Strategie implementace uchovávacích metadat (2005).
- METS (Metadata Encoding and Transmission Standard) - Standard kódování a přenosu metadat.
- SCORM (Sharable Content Object Reference Model) - Referenční objektový model sdíleného obsahu (pro e-learning).
- RDF (Resource Description Framework) - Rámec popisu zdroje (rodina specifikací World Wide Web Konsorcia (W3C), původně navržen jako datový model metadat, který je používán jako obecný způsob modelování informací v různých formátech syntaxe).
- MODS (Metadata Object Description Schema) - Schéma metadat popisu objektů (US Library of Congress).
- SOAP (Simple Object Access Protocol).
- OAI (Open Archives Initiative) - Iniciativa otevřených archivů (protokol pro „sklizeň“ metadat - metadata harvesting).

Je důležité si uvědomit, že ačkoli různé oblasti použití metadat se vzájemně nevyklučují, řada z těchto standardů se často v některých oblastech překrývá. Rovněž mnoho z těchto norem má specifické zaměření nebo jsou určeny pro konkrétní oblasti (např. zdravotnictví, školství), které přesahují rozsah nebo úroveň granularity příslušné mezinárodní normy.

Při návrhu a realizaci metadatových struktur se setkáváme s nejrůznějšími otázkami a úvahami, které by měly být řešeny. Mezi ně patří:

- Lidé obecně neradi vytváří metadata ručně, protože to často považují za tíživou zátěž při jejich už tak velkém zaneprázdnění. Je proto

## Odkazy, 2. část

National Archives of Australia, Australian Government Recordkeeping Metadata Standard verze 2.0, 2008.

<http://www.naa.gov.au/recordsmanagement/create-capturedescribe/describe/RKMS/index.aspx>

PREMIS (PREservation Metadata: Implementation Strategies) Working Group, Data Dictionary for Preservation Metadata : Final Report of the PREMIS Working Group, May 2005 (United States of America)

<http://www.oclc.org/research/projects/pmwg/premis-final.pdf>

Tambouris, Efthimios and Konstantinos Tarabanis, 'An Overview of DC-Based e-Government Metadata Standards and Initiatives', Lecture Notes in Computer Science, Springer, Berlin, 2004, pp. 40-47.  
Autor: Adrian Cunningham (National Archives of Australia)

důležité, aby tvorba a správa metadat byla co nejvíce automatizovaná.

- Značnou pozornost je třeba věnovat zajištění kvality metadat. Obecně není obtížné vytvořit dobrá metadata, ale je také velmi snadné vytvořit špatná metadata. Špatná metadata jsou horší než žádná metadata.
- Při implementacích metadat je nutné věnovat pečlivou pozornost tomu, pro co mají být metadata vytvořena a udržována. Vytvoření a údržba metadat vždy něco stojí a všechny tyto náklady musí být odůvodněny podnikatelskými potřebami. Metadata by neměla být vytvořena a udržována jen kvůli nim samotným. Organizace by měly vytvořit a udržovat metadata, která jim přinesou prokazatelnou návratnost investic.
- Metadatové struktury mohou být jednoduché nebo velmi složité. Obecně platí, že čím jednodušší metadata, tím lépe, protože u složitých metadat je nákladná jejich správa a je obtížné zajistit jejich kvalitu. Nicméně, jednoduchá metadata nemusí vždy přinést požadovaný výsledek – honba za větší složitostí a přesností metadat může být odůvodněna jasně definovanými obchodními argumenty.
- Vzhledem k tomu, že informace veřejného sektoru jsou dynamické a stále se vyvíjí, je důležité metadata průběžně aktualizovat, aby odrážela měnící se povahu informačních zdrojů, ke kterým se metadata vztahují. Metadata by neměla být statická, ale spíše by měla být aktivně udržována, řízena a aktualizována.
- Platí obecné pravidlo, že technická infrastruktura podporující implementaci metadat by se měla spíše opírat o flexibilní než pevnou podnikovou architekturu. Například Service Oriented Architecture (SOA) slibuje ideální přístup k zavedení flexibilního, dynamického, interoperabilního a opakovatelného použití metadat.

## Metadata a 'Sémantický web'

Sémantický web je rozvíjející se rozšíření World Wide Webu, ve kterém je sémantika informací a služeb na webu definována tak, aby webu umožnila pochopit a uspokojit žádosti lidí a strojů o webový obsah. Tim Berners-Lee, ředitel World Wide Web Konsorcia, popisuje ve své vizi web jako univerzální médium pro data, informace a výměnu znalostí. Jádro sémantického webu zahrnuje řadu konstrukčních principů, spolupracujících pracovních skupin a celou řadu prováděcích technologií. Některé prvky sémantického webu jsou uváděny jako v budoucnu potenciální možnosti, které mají být ještě zavedeny nebo realizovány. Ostatní prvky sémantického webu jsou formálně specifikovány. Některé z nich jsou zahrnuty v Resource Description Framework (RDF – Zdroje rámcových popisů), v různých formátech výměny dat (např. RDF / XML, N3, Turtle, N-Triples) a v notacích jako RDF Schema (RBP) a Web Ontology Language (OWL), z nichž všechny jsou určeny k zajištění formálního popisu koncepcí, pojmů a vztahů v rámci určité znalostní domény.

Metadata veřejného sektor slibují, že se stanou klíčovou aktivitou pro vznik a rozvoj sémantického webu. Vládní implementace metadat by proto měla zvážit potenciální přínosy v souladu s protokoly sémantického webu jako RDF a OWL.