

# Úvod do problematiky dlouhodobé ochrany digitálních dokumentů - díl 4.

**Jan Hrabal, Zdeněk Hruška**

V závěrečném dílu našeho seriálu se tentokrát zaměříme na důvěryhodnost digitálních repozitářů a jaké nástroje jsou k dispozici pro její hodnocení. Protože důvěryhodný repozitář je jedním z důležitých cílů cesty k dlouhodobé ochraně dat.

Důvěryhodný digitální repozitář je takový, který vykonává přesně to, k čemu se zavázal. To znamená, že pokud repozitář tvrdí, že uložená data dlouhodobě bezpečně ochrání, zajistí jejich autenticitu, integritu a zpřístupní je, tak musí mít nástroje a postupy, kterými to zajistí. Důvěryhodnost ale není jen o technickém zabezpečení, nebo uložení dat – repozitář musí mít zajištěné financování, zkušené a schopné zaměstnance, podporu zřizovatele a další záležitosti pro bezproblémové fungování.

## **Deset principů důvěryhodnosti digitálního repozitáře podle Jana Hutáře<sup>1</sup>**

Pro důvěryhodné digitální repozitáře je nezbytné:

1. Zavázat se k neustálému opatrování a správě digitálních objektů pro určitou cílovou komunitu.
2. Zajistit svou existenci - musí mít vyřešené financování, personální otázky, apod.
3. Osvojit si a dodržovat potřebná smluvní a zákonná práva a dostát všem plynoucím závazkům.
4. Disponovat efektivní rámcovou strategií.
5. Získávat a ukládat digitální objekty na základě stanovených kritérií, které odpovídají cílům a schopnostem instituce.
6. Neustále udržovat a zajišťovat integritu, autenticitu a využitelnost digitálních objektů, které trvale uchovává.
7. Vytvářet a uchovávat potřebná metadata o událostech souvisejících s uloženými digitálními objekty v průběhu jejich uchování, jakož i metadata o samotném vytvoření digitálních objektů, podmínkách zpřístupnění a kontextu využití digitálních objektů.
8. Naplnit nezbytné požadavky na zpřístupnění objektů určité komunitě.
9. Mít strategii pro plánování ochrany a souvisejících procesů včetně záchranných prací v případě katastrofy.
10. Mít technickou infrastrukturu adekvátní pro účel neustálé údržby a zajištění digitálních objektů.

## **Nástroje na hodnocení digitálních repozitářů**

Jaká cesta vede k tomu, aby byl repozitář důvěryhodný? Zjednodušeně řečeno - děláte všechno pro to, abyste ochránili data, využíváte k tomu zkušenosti a znalosti (mezinárodní)

---

<sup>1</sup> HUTAŘ, Jan. Proč jsou české digitální repozitáře „nespolehlivé“?. In: *Archivy, muzea a knihovny v digitálním světě 2007* [online]. Praha, 2007 [cit. 2013-01-27]. Dostupné z: <skip.nkp.cz/KeStazeni/Archivy07/Hutar.ppt>.

komunity a řídíte se standardy. A máte ke všemu dokumentaci - ta je důležitá proto, aby bylo všechno hezky zaznamenáno a zaměstnanci ji mohli využít, kdykoliv je potřeba. A v okamžiku, kdy si myslíte, že to všechno máte aspoň trochu v pořádku, tak si projdete auditem podle nějakého standardu/nástroje.

Pomocí auditu se doloží správné vnitřní fungování a uspořádání repozitáře. Auditem a zhodnocením digitálního repozitáře se rovněž prokáže jeho důvěryhodnost navenek, díky čemuž se zvýší jeho reputace u uživatelů a dalších institucí, a rovněž u jeho zřizovatelů. Samotná snaha repozitáře splnit kritéria důvěryhodnosti může následně vyústit v kvalitnější služby.

Existuje řada nástrojů pro hodnocení důvěryhodnosti repozitáře, přičemž my zde blíže představíme tři z nich: DRAMBORA, Data Seal of Approval (DSA) a ISO 16363. V Německu se používá NESTOR, ale ten byl vytvořen s ohledem na tamější právní, finanční i kulturní prostředí a pro ČR není proto vhodný.

### **DRAMBORA**

Systém DRAMBORA vyvinuly společně v roce 2007 iniciativy Digital Preservation Europe (DPE) a Digital Curation Centre (DCC). Slouží pro interní hodnocení a správcům repozitářů umožňuje posoudit schopnosti repozitáře, identifikovat silné a slabé stránky. V českém prostředí jej využila např. Národní knihovna ČR, ÚVT Univerzity Karlovy v Praze nebo Národní úložiště šedé literatury. Její verze 2.0 je dostupná online přes webové rozhraní, její dřívější verze 1.0 je stáhnutelná a využitelná offline. DRAMBORA má jen anglickou verzi.

### **Data Seal of Approval (DSA)**

Směrnice DSA byly původně sestaveny pro repozitáře výzkumných dat, ale jsou využitelné i jinými institucemi. Určují 16 cílů, které musí splňovat nejen repozitář, ale i jeho správci a uživatelé, pokud chtějí získat pečeť DSA. Je tu tedy zdůrazněna vazba repozitáře na komunitu. Při hodnocení se na každý požadavek aplikuje pětibodová stupnice (0 - vůbec nesplněn, 5 - úplně splněn). Pro udělení DSA certifikátu musí repozitář splnit třetí nebo čtvrtý stupeň u každého požadavku.

Archivovaná data musí být dle principů DSA nalezitelná na internetu, jsou spolehlivá, mají použitelný formát, lze na ně odkázat, jsou přístupná a dodržují příslušné právní normy zohledňující autorská práva a ochranu osobních údajů. Tyto podmínky reflektují původní zaměření standardu na repozitáře vědeckých dat, ale samozřejmě se dají zobecnit i na jiná úložiště.

### **ISO 16363**

Tato ISO norma se vyvinula ze standardu TRAC (Trustworthy Repositories Audit & Certification) a lze ji využít jak k internímu auditu a sebehodnocení repozitáře (více informací a praktické rady jsou dostupné zde: <http://www.iso16363.org/preparing-for-an-audit/>), tak k externí certifikaci, kterou bude provádět tým vyškolených profesionálů.

Norma hodnotí repozitář z několika úhlů pohledu - zaměřuje se na infrastrukturu, zaměstnance, financování, smlouvy a licence, plány a strategie. Dále se řeší management

digitálních objektů a tedy jejich příjem, zpracování, vytváření AIP, systémy uchovávání a ochrany a přístup uživatelů k datům. V poslední páté části je dotazována technická infrastruktura, použité technologie a zabezpečení proti systémovým i lidským chybám, proti přírodním katastrofám i cíleným útokům na repozitář.

Výhodou pro naše prostředí je fakt, že norma byla v roce 2014 přeložena do češtiny a je tedy dostupná jak pro případné audity, tak i pro plánování a budování nových repozitářů.

## **Evropský rámec pro certifikaci digitálních repozitářů**

Protože nástrojů na hodnocení důvěryhodnosti repozitářů je celá řada, komunita se dohodla na určité použitelnosti jednotlivých nástrojů pro různé repozitáře.

Memorandum o porozumění, díky kterému byly vytvořeny tři certifikační úrovně, bylo podepsáno v roce 2010. Na memorandu spolupracovaly tyto organizace:

- CCSD - tvůrci norem OAIS a ISO 16363
- Data Seal of Approval - autoři stejnojmenné pečeti důvěryhodnosti DSA
- DIN Working Group - skupina, která stojí za německou normou NESTOR

### **Úrovně certifikací:**

#### **Základní certifikace**

Jedná se o nejnižší úroveň certifikace, která je primárně určena pro menší repozitáře, ale absolvovat ji musí všechny, které aspirují na vyšší stupně. Pro získání této úrovně je nutné projít interním auditem v rámci Data Seal of Approval (DSA). Více o DSA se můžete dozvědět na webu [DSA](#). Český web o DSA a překlad podmínek auditu spatřil světlo světa v roce 2014 díky práci Andrei Fojtů, Elišky Pavláskové, Jana Hutaře a Marka Melichara a je k dispozici na <http://dsa.cuni.cz/>.

#### **Rozšířená certifikace**

Je určena pro středně velké repozitáře, které již získaly základní certifikaci DSA. V této fázi je čeká interní audit podle norem ISO 16363 nebo DIN 31644, výsledky auditu je poté nutné zveřejnit a musí projít externím veřejným hodnocením, na němž se podílí odborná veřejnost.

#### **Formální certifikace**

Repozitář projde externím nezávislým auditem a získá certifikát podle normy ISO 16363 nebo DIN 31644. Tato úroveň je určena pro velké repozitáře (v Česku by tento audit byl určen pro instituce typu Národní knihovna ČR, nebo Národní archiv ČR), protože získání certifikátu je náročné na lidské i finanční zdroje.

### **Situace v ČR**

A jaké jsou zkušenosti s audity v České republice? Národní úložiště šedé literatury (NUŠL) provedlo v letech 2009 - 2011 tři hodnocení pomocí nástroje DRAMBORA a identifikovalo hlavní hrozící rizika pro svůj repozitář. Více o zkušenostech s DRAMBORA a výsledcích uskutečněných auditů si můžete přečíst [zde](#) a [zde](#).

Jediným repozitářem, jež je u nás v současnosti certifikovaný podle jedné z výše uvedených norem, je LINDAT/CLARIN, který získal v roce 2014 pečeť DSA. Jedná se o digitální repozitář vědeckých dat z oboru lingvistiky a oborů příbuzných, provozuje ho Karlova univerzita a je určen pro vědce, kteří chtějí ukládat, uchovávat a sdílet svoje data i výsledky své práce. Web repozitáře je zde: <https://lindat.mff.cuni.cz/>.

LTP repozitář Národní digitální knihovny se připravuje na budoucí audit pomocí ISO 16363. V současné době probíhá sběr a vytváření nezbytné dokumentace, bez níž je jakýkoliv audit nemyslitelný. O tom, co je rámcově hotovo a co je ještě potřebné dokončit, se můžete dočíst v prezentaci Jana Mottla ze společnosti AIP Safe, která zazněla na letošním Inforu: <http://www.inforum.cz/pdf/2015/mottl-jan.pdf>. Otázkou je, na kolik jsou tyto informace relevantní a na kolik odráží skutečnost, vzhledem k faktu, že pocházejí od soukromé firmy a nikoli přímo od Národní knihovny.

## Použité zdroje a další materiály pro zájemce

CUBR, Ladislav a Jaroslava HAVLOVÁ. Důvěryhodný digitální repozitář. In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha: Národní knihovna ČR, 2003- [cit. 2013-02-24]. Dostupné z:

<[http://aleph.nkp.cz/F/?func=direct&doc\\_number=000014621&local\\_base=KTD](http://aleph.nkp.cz/F/?func=direct&doc_number=000014621&local_base=KTD)>.

HRUŠKA, Zdeněk. Audit digitálních repozitářů. *Duha* [online]. 2013, roč. 27, č. 4 [cit. 2015-06-02]. Dostupný z WWW: <<http://duha.mzk.cz/clanky/audit-digitalnich-repozitaru>>. ISSN 1804-4255.

HUTAŘ, Jan. Proč jsou české digitální repozitáře „nespolehlivé“?. In: *Archivy, muzea a knihovny v digitálním světě 2007* [online]. Praha, 2007 [cit. 2013-01-27]. Dostupné z:

<[skip.nkp.cz/KeStazeni/Archivy07/Hutar.ppt](http://skip.nkp.cz/KeStazeni/Archivy07/Hutar.ppt)>.

NÁRODNÍ ULOŽIŠTĚ ŠEDÉ LITERATURY. *Audit Národního úložiště šedé literatury (NUŠL) v NTK s použitím nástroje DRAMBORA: Třetí audit 2011*. Praha, 2011. Dostupné z:

[http://repozitar.techlib.cz/record/389/files/idr-389\\_1.pdf](http://repozitar.techlib.cz/record/389/files/idr-389_1.pdf)

ROSENTHAL, Colin, Asger BLEKINGE-RASMUSSEN a Jan HUTAŘ. Průvodce plánem důvěryhodného digitálního repozitáře (PLATTER). 1. vyd. Praha: Národní knihovna České republiky, 2009, 51 s. ISBN 978-807-0505-694. Dostupné z: <<http://www.ndk.cz/platter-cz>>.

<http://dsa.cuni.cz/>

<http://www.repositoryaudit.eu>

<http://www.digitalpreservation.cz/search/label/certifikace>

<http://www.slideshare.net/AndreaFojtu/dsa-prezentace-seminardigitalizace-2-kopie>