

HOUGHTON, Bernadette. Trustworthiness: Self-assessment of an Institutional Repository against ISO 16363-2012. *D-Lib Magazine* [online]. 2015, 21(3/4): - [cit. 2015-06-02]. DOI: 10.1045/march2015-houghton. ISSN 1082-9873. Dostupné z: <http://www.dlib.org/dlib/march15/houghton/03houghton.html>

Důvěryhodnost: Sebe-hodnocení institucionálního repozitáře dle ISO 16363-2012

BERNADETTE HOUGHTON
Deakin University, Geelong, Australia
bernadette.houghton@deakin.edu.au

Abstrakt

V dnešní době je téměř každý vytvořený dokument a téměř každý výstup z výzkumného projektu digitálním objektem. Ne všechno má být uchováno navždy, ale materiály akademické nebo historické hodnoty by se měly uchovat pro budoucí generace. Ochrana digitálních objektů je větší výzva než ochrana papírových položek - hardware se stává zastaralým, nový software nahrazuje starý, úložná média degradují. V posledních letech došlo k významnému pokroku v rozvoji nástrojů a standardů pro ochranu digitálních médií, zejména v souvislosti s institucionálními repozitáři. Doposud nejvíce přijímaným standardem je Trustworthy Repositories Audit and Certification: Criteria and Checklist (TRAC), který se vyvinul v normu ISO 16363-2012. Deakin Univerzity Library provedla sebe-hodnocení na základě kritérií ISO 16363. Tato zkušenost vyvrcholila v tuto zprávu, poskytující zhodnocení ISO 16363 a proces hodnocení, a poučení pro ostatní, kteří zvažují podobnou akci.

Úvod

Dlouhodobá ochrana je poměrně mladý obor, ale bylo již dosaženo významného pokroku při vytváření nástrojů a standardů pro lepší podporu snah o uchovávání. Rostoucí zájem zde byl zvláště v oblasti auditu a certifikace. Vzhledem k rostoucí závislosti na institucionálních repozitářích v uplynulém desetiletí (OpenDOAR, 2014) se pro badatele stává stále důležitější, aby mohli důvěřovat zabezpečení výstupů svých výzkumů. Tato potřeba byla oficiálně uznána Deakin University Library v roce 2013 a byl spuštěn projekt, který dodržuje nejlepší praxi digitální ochrany a který vyústil v založení výzkumného repozitáře - Deakin Research Online (DRO).

DRO byl založen v roce 2007 v rámci Australian Government's Research Quality Framework (Department of Education, Science and Training, 2007). Jeho hlavním cílem bylo usnadnit ukládání výzkumných publikací pro účely výkazu a archivace. Jako základní prvek byl z velké části vybrán software Fez/Fedora, díky jeho funkcím souvisejícím s dlouhodobou ochranou, výhodou je schopnost verzování a podpora JHOVE a PREMIS. V průběhu let byly do DRO přidány nové funkce a pracovní postupy se měnily v návaznosti na změny požadavků na ukládání a tvorbu výkazů. Vzhledem ke specifickým aspektům repozitáře DRO bylo realizováno několik ad hoc miniprojektů, ale až do roku 2013 se neuskutečnilo žádné zhodnocení nových funkcí nebo pracovních postupů na základě nejlepší praxe v dlouhodobé ochraně.

Přehled literatury o dlouhodobé ochraně naznačil, že externí certifikace důvěryhodnosti digitálních repozitářů ze strany orgánů jako je Center for Research Libraries (CRL) začala získávat určitou popularitu mezi většími mezi-institucionálními repozitáři (CRL, 2014). Pro Deakin University Library nebyla v této fázi externí certifikace vzhledem k jejím nákladům ideální volba; nicméně vlastní zhodnocení podle standardů používaných pro akreditaci bylo považováno za proveditelné.

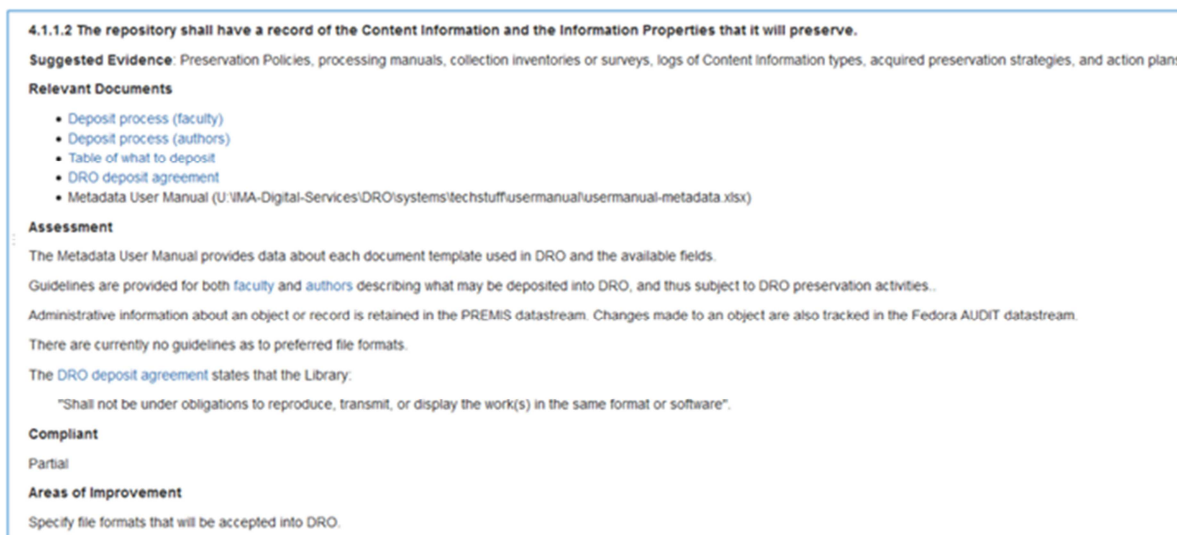
Literární přehled naznačil, že kontrolní seznam Trustworthy Repositories Audit and Certification (TRAC) (CRL, OCLC, 2007) byl nejrozšířenějším kritériem pro posouzení důvěryhodnosti digitálních repozitářů. TRAC, původně vytvořený Research Library Group (RLG) a National Archives and Records Administration (NARA) se vyvinul v *Audit a certifikace důvěryhodnosti digitálních repozitářů: Doporučené praktiky CCSDS 652.0-M-1* (Audit and Certification of Trustworthy Digital Repositories: Recommended practice CCSDS 652.0-M-1) (Consultative Committee for Space Data Systems (CCSDS), 2011), který byl v roce 2012 formalizován jako ISO 16363 (International Organization for Standardization, 2012). ISO 16363 byl pro účely Deakin University Library vybrán jako vhodný základ sebe-hodnocení repozitáře DRO.

Předběžný výzkum projektu začal v květnu 2013, s jedním zaměstnancem (autorem) přiděleným k projektu na částečný úvazek. Předpokládaný termín dokončení byl zpočátku stanoven na srpen 2013. Avšak vzhledem k jiným pracovním prioritám nezačalo skutečné hodnocení až do července 2013. Závěrečná zpráva byla vypracována v prosinci 2013.

Položení základů práce

Jakmile projekt započal, byl proveden průzkum za účelem zmapování zkušeností ostatních. Zvláštní význam měly audity repozitářů HathiTrust, Portico and Scholar's Portal provedené CRL (CRL, 2014). Bylo zjištěno, že každý z přezkoumaných repozitářů provedl kromě externí certifikace provedené CRL také vlastní ohodnocení. Webová stránka CRL obsahuje kontrolní seznamy a šablony, jež mohou být použity jako základ pro vlastní audit (CRL, 2014). Nicméně tyto byly považovány za zastaralé, jsou na základě kritérií TRAC z roku 2007 spíše než nedávné ISO 16363.

Jako mechanismus pro dokumentaci sebe-hodnocení byla zvolena wiki platforma. Webové stránky CRL sice obsahují šablonu určenou pro dokumentaci, nicméně volba wiki se ukázalo být jako docela moudré rozhodnutí. Některé z kritérií ISO 16363 vyžadovaly dlouhé odpovědi, které nezapadaly do tabulek příliš dobře. Wiki byla nastavena pro jednotlivá kritéria: “Navržené příklady”, “Relevantní dokumenty”, “Zhodnocení”, “Míra shody”, “Vyhovění požadavkům” a “Možnosti pro zlepšení”. Tato etapa byla dobrou příležitostí pro seznámení se s kritérii a dokumentací potřebnou pro další krok hodnocení.



Obrázek 1: Příklad wiki dokumentace u sebe-hodnocení
Zobrazit [větší verzi](#) Obrázku 1.

Jakmile byla skladba wiki hotova, byl proveden předběžný sběr dokumentů za účelem identifikace dokumentace, která by poskytla důkazy o rozsahu, v jakém DRO splňuje každé z kritérií. Příslušné zásady a postupy jak na univerzitní, tak na knihovní úrovni byly vyhledány a navázány na wiki. To vyžadovalo dobré povědomí o organizační odpovědnosti z prostředí univerzity i knihovny, jako i obeznámenost se správou webových stránek univerzity. ISO 16363 poskytuje “navržené příklady” pro každé z těchto kritérií. Brzy však bylo zřejmé, že mnohé z navrhovaných příkladů nebyly pro DRO relevantní. Od té doby byly dokumenty shromažďovány na základě autorovy znalosti místního prostředí. Předběžný sběr dokumentace nezabránil v potřebě hledat další dokumenty v průběhu hodnocení, ale ušetřil nějaký čas.

Poté, co byly sesbírány předběžné dokumenty, bylo každé kritérium přezkoumáno vůči příslušné dokumentaci. Byla zajištěna nezbytná další dokumentace a příslušný personál byl obeznámen s pracovními postupy a zásadami. Poté se zhodnotilo, jak DRO dodržoval kritéria, a identifikovaly se oblasti, které byly nutné vylepšit. Všechna zjištění byla postupně dokumentována ve wiki, přičemž každému kritériu byl přidělen údaj “úplná shoda”, “částečná shoda”, nebo “neshoduje se”.

Zhodnocení ISO 16363

ISO 16363 obsahuje 105 kritérií, které pokrývají 3 oblasti. Kritéria jsou celkově velmi obsáhlá, ale některá mají neadekvátní nebo chybějící vysvětlení a mohou tak být volně interpretována. U jednotlivých kritérií dochází k jakémusi překryvu a ne všechna kritéria tak budou použitelná pro všechny repozitáře. Oněmi oblastmi jsou:

- **Organizační infrastruktura**, jež zahrnuje správu, organizační strukturu a životaschopnost repozitáře, personální obsazení, odpovědnost, zásady, finanční udržitelnost a právní záležitosti
- **Management digitálních objektů**, jež pokrývá akvizici a zpracování obsahu, plánování a postupy dlouhodobé ochrany, informační management a přístup

- **Infrastruktura a řízení bezpečnostních rizik**, zahrnující infrastrukturu a otázky bezpečnosti.

Vzorová dokumentace ISO 16363 pro každé kritérium by měla být chápána pouze jako obecné vodítko. Pokrytí každého repozitáře se zde bude lišit; dokumentace, která je relevantní pro jeden repozitář, nebude relevantní pro jiný. Shromáždění dokumentace byla zřejmě časově nejnáročnější část sebe-hodnocení Deakin University Library, protože dokumenty byly roztroušeny v různých oblastech univerzity.

Nejednoznačná terminologie, kterou uplatňuje ISO 16363, způsobila zmatek. Například ISO 16363 zmiňuje linkovací službu, jež odkazuje na perzistentní odkazy nebo materiály v repozitáři. V případě naší praxe se však tato terminologie běžně používá k odkazu na linkovací služby, jakými jsou WebBridge a SFX.

ISO 16363 je založen na referenčním modelu OAIS (Open Archival Information System) (CCSDS, 2012). Je třeba mít na paměti, že OAIS je konceptuální model, který se nemusí vždy být v souladu se skutečnými nebo požadovanými pracovními postupy. Například OAIS zachází se SIP (vstupní informační balíček), DIP (výstupní informační balíček) a AIP (archivní informační balíček) jako s oddělenými entitami. To však neznamená, že dodržení OAIS modelu vyžaduje oddělené SIP, DIP a AIP balíčky. Teoretický charakter OAIS se odráží v ISO 16363.

Podstoupení vlastního hodnocení dle ISO 16363 není jednoduchý počín a pravděpodobně přesahuje možnosti menších repozitářů. Každá organizace, která podstoupí sebe-hodnocení dle ISO 16363, by si jej měla co nejlépe přizpůsobit vzhledem k jejím vlastním okolnostem a finančním možnostem. Pokud se provedení ISO 16363 stále jeví jako příliš mamutí počín, pak lze využít alternativní nástroje, jako třeba *Levels of Digital Preservation* vyvinutý konsorciem NDSA (Owens, 2012).

Je třeba zdůraznit, že sebe-hodnocení není audit. V nezávislém externím auditu si auditor udržuje od repozitáře profesionální odstup, jsou zapotřebí důkazy o dodržování zásad a postupů a je zapotřebí provádět testování, které by potvrdilo, že funkce repozitáře související s dlouhodobou ochranou jsou v činnosti tak, jak bylo plánováno. Sebe-hodnocení v Deakin University Library bylo podstoupeno autorem, který s DRO úzce pracuje, a v podstatě bylo omezeno na přezkoumání příslušné dokumentace a postupů. Jako první sebe-hodnocení to bylo považováno za vhodnou strategii. Budoucí činnosti sebe-hodnocení by však měly podstoupit přísnější testování základní funkčnosti DRO.

Zjištění

Zhodnocení ukázalo, že DRO splňuje většinu kritérií, aby mohl být považován za důvěryhodný digitální repozitář, ale jak se očekávalo, je zde prostor pro zlepšení. Přesněji řečeno, DRO splňuje 67 ze 105 kritérií, částečně splňuje 32 a nesplňuje 15 kritérií; jedno kritérium bylo nepoužitelné. Tabulka 1 shrnuje tyto výsledky napříč třemi oblastmi ISO 16363.

Tabulka 1: Přehled shody ISO 16363

Oblast ISO 16363	Úplná shoda	Částečná shoda	Neshoda
Organizační infrastruktura	13 (54 %)	8 (33 %)	3 (13 %)
Management digitálních objektů	36 (62 %)	20 (34 %)	2 (3 %)
Infrastruktura a řízení bezpečnostních rizik	18 (82 %)	4 (18 %)	—

Je třeba upozornit na skutečnost, že ne všechna kritéria by měla být hodnocena jako by měla stejnou váhu. Zatímco ISO 16363 nepřidělí stejnou váhu každému kritérii, některá kritéria jsou z hlediska důvěryhodnosti repozitáře důležitější než jiná. Prohlášení, že "repozitář splňuje 60% ISO 16363" může být velmi zavádějící, jestliže jedno zásadní kritérium pro důvěryhodnost repozitáře není splněno - například absence měření integrity dat (ISO 16363, 3.3.5).

Sebe-hodnocení naznačilo, že dlouhodobá ochrana směřuje u DRO k ad hoc přístupu, což již bylo známo. To je dáno do značné míry kvůli stále se vyvíjející povaze oblasti dlouhodobé ochrany, a většina repozitářů bude v současnosti v podobné situaci. Tento přístup je doposud považován za vhodný; do budoucna se však očekává, že v souladu s dostupností nových nástrojů a pokročilejších znalostí o dlouhodobé ochraně bude přijat více pro-aktivní přístup. Většina zjištěných skutečností se nyní začleňují do strategického plánu dlouhodobé ochrany Deakin University Library.

Dalším naším krokem je realizace praktických strategií navržených ve strategickém plánu dlouhodobé ochrany za účelem co nejlepšího využití postupů v oblasti dlouhodobé ochrany.

Shrnutí a závěry

Celkem vzato, sebe-hodnocení bylo časově náročné a bylo to pro nás také cvičení na přidělování zdrojů, ale výsledkem bylo odhalení několika slabých oblastí v nastavení a pracovních postupech DRO. Od té doby byl strategický plán upraven k vyřešení těchto nedostatků, a tím se zvýšila důvěryhodnost DRO jako repozitáře. Předpokládá se, že sebe-hodnocení bude prováděno pravidelně v intervalu 3 let; neočekává se však, že budou tak zátěžové na zdroje jako počáteční zhodnocení, jestliže byla velká část energie věnována k identifikaci a nalezení příslušné dokumentace.

Důrazně se doporučuje, aby podobné sebe-hodnocení repozitářů provedly v určité fázi i ostatní knihovny, a použily nástroje a hloubku hodnocení v závislosti na velikosti knihovny a úrovni dostupných zdrojů.

Vědci očekávají, že jejich publikace budou v institucionálních repozitářích zabezpečené a správci repozitářů musí zajistit, že repozitáře budou tato očekávání splňovat. Sebe-hodnocení je jeden z mechanismů, díky němuž mohou zjistit, zda je jejich repozitář skutečně důvěryhodný.

Na základě zkušeností Deakin University Library jsou předloženy následující návrhy pro správce repozitářů, kteří plánují podstoupit zhodnocení dle ISO 16363:

1. Proveďte sebe-hodnocení před tím, než budete zvažovat externí certifikaci. Certifikace - a re-certifikace - je drahá.
2. Získejte na svoji stranu vrcholové vedení. Jeho podpora je zásadní. Dlouhodobá ochrana je dlouhodobá záležitost.
3. Jednotlivci provádějící sebe-hodnocení by měl být dostatečně obeznámen se zásadami a postupy organizace a repozitáře.
4. Pokud nemáte čas a zdroje pro podstoupení ISO 16363 zhodnocení, zvažte zhodnocení dle Levels of Digital Preservation od konsorcia NDSA (Owens, 2012).
5. Pro dokumentaci sebe-hodnocení zřídte wiki. Udělejte to na začátku a nálezy dokumentujte za pochodu.
6. Přizpůsobte sebe-hodnocení dostupnému času a zdrojům.
7. Předem určete, do jaké hloubky bude hodnocení provedeno. Například, bude posuzovatel pouze sbírat a přezkoumávat dokumentaci, nebo bude rovněž kontrolovat, že dokumentované postupy byly dodrženy a vše v repozitáři funguje správně?
8. Při shromažďování dokumentace použijte vlastní znalosti. "Navržené příklady" od ISO 16363 jsou pouze jedna z možností.
9. Před započítím hodnocení se seznamte s kritérii. Některá dokumentace budou příslušet více kritériím, takže to ušetří čas, pokud kritéria zavčasu určíte.
10. Pamatujte, že ne všechna ISO 16363 kritéria budou použitelná ve vaší konkrétní situaci.
11. Nezůstaňte stát na místě. Dokončení sebe-hodnocení neznamena, že tvrdá práce je u konce. V průběhu se ukážou možnosti pro zlepšení. Zaměřte se na vybudování digitální odolnosti repozitáře v průběhu času.
12. Naplánujte si pravidelná, opakující se, sebe-hodnocení.
13. Zvažujete-li provedení sebe-hodnocení dle ISO 16363 někdy v budoucnu, začněte proces už teď. Zřídte wiki stránku, kde budete mezitím zaznamenávat příslušnou dokumentaci, na kterou narazíte. Sledujte problematiku dlouhodobé ochrany a dle potřeby wiki stránku aktualizujte, abyste později ušetřili čas.
14. Pokud je váš software repozitáře kompatibilní s OAIS, nepředpokládejte, že samotný repozitář taky. Pracovní postupy a nastavení repozitáře mohou, nebo nemusí dodržovat OAIS.
15. Ne všechna ISO 16363 kritéria mají stejný význam nebo úroveň rizika. Posuďte každé kritérium odpovídajícím způsobem.
16. ISO 16363 je založen na konceptuálním modelu OAIS. Neočekávejte, že kritéria budou nezbytně v souladu s nastavením a pracovními postupy vašeho repozitáře.

Reference

- [1] Center for Research Libraries, OCLC, 2007, [Trustworthy repositories audit & certification \(TRAC\) criteria and checklist](#), Version 1, Centre for Research Libraries.
- [2] Center for Research Libraries, 2014, [TRAC and TDR checklists](#), Center for Research Libraries.
- [3] Consultative Committee for Space Data Systems, 2011, [Recommendation for space data system practices: audit and certification of trustworthy digital repositories](#), Recommended practice CCSDS 652.0-M-1, Magenta book, Consultative Committee for Space Data Systems.
- [4] Consultative Committee for Space Data Systems, 2012, [Reference model for an Open Archival Information System \(OAIS\): recommended practice CCSDS 652.0-M-2](#), CCSDS, Washington, DC.
- [5] Department of Education, Science and Training, 2007, [Research Quality Framework: assessing the quality and impact of research in Australia. RQF technical specifications](#), Commonwealth of Australia.

- [6] International Organization for Standardization 2012. [*Space data and information transfer systems: audit and certification of trustworthy digital repositories.*](#)
- [7] OpenDOAR, 2014, [*Growth of the OpenDOAR database — worldwide*](#), OpenDOAR.
- [8] Owens, Trevor, 2012, [*NDSA levels of digital preservation: release candidate one*](#), Library of Congress.