Ako na certifikáciu repozitára podľa Data Seal of Approval

Cieľom príspevku je oboznámiť s kritériami hodnotenia digitálnych repozitárov podľa certifikácie Data Seal of Approval (DSA). Certifikát slúži ako „záruka" dôvery pre dané kľúčové skupiny, ktoré využívajú či odovzdávajú dáta do cieľového repozitára. DSA predstavuje pomerne jednoduchý spôsob auditu a certifikácie, najmä pre začínajúce repozitáre či repozitáre menšieho rozsahu. Popísané sú pôvodné a pripravované Zásady DSA a podrobne spísané sú potrebné podklady a doklady, ktoré bude repozitár potrebovať pre svoju úspešnú certifikáciu.

Kľúčovým prvkom certifikácie *Data Seal of Approval* je dôvera. Najčastejšie ju definujeme ako spoľahlivosť, vieru či predpoklad, že osoba alebo organizácia bude dodržiavať rámec spoločných hodnôt a predstáv. V kontexte digitálnych repozitárov sa za dôveryhodný digitálny repozitár považuje ten, ktorý v súlade so svojim mandátom a poslaním poskytuje cieľovej skupine spoľahlivý a dlhodobý prístup (dnes i do budúcna) k organizovaným digitálnym zdrojom. Certifikát potom slúži ako „záruka" dôvery pre dané kľúčové skupiny, ktoré odovzdávajú či využívajú dáta do cieľového repozitára. [MIRANDA, 2015]

Auditné a certifikačné nástroje

Auditné a certifikačné nástroje sú tiež skvelou pomôckou pre vyhodnotenie procesov a činností, ktoré sa dotýkajú celého životného cyklu digitálnych objektov v repozitároch či archívoch. *Európsky rámec pre audit a certifikáciu repozitárov* (2010), ktorý bol podpísaný CCSDS/ISO Repository and Audit Certification Working Group/RAC), Data Seal of Approval Board a DIN Working Group "Trustworthy Archives – Certification", stanovil tri typy auditov a certifikácie digitálnych repozitárov:

- **Základná certifikácia** (Basic certification) – pridelená repozitárom, ktoré naplnia Zásady Data Seal of Approval (DSA).
- **Rozšírená certifikácia** (Extended certification) – priznaná repozitárom, ktoré splnili Základnú certifikáciu a za externého preskúmania uskutočnili štrukturovaný a verejne dostupný „samo-audit" podľa ISO 16363 alebo DIN 31644, prípadne zažiadali o nestorSeal (vychádza z DIN 31644).
- **Formálna certifikácia** (Formal certification) – repozitár nechal posúdiť svoje procesy a činnosť externým nezávislým auditom za účelom certifikácie podľa ISO 16363 alebo DIN 31644.

Pôvodné Zásady DSA

Definovanie kritérií certifikácie DSA vychádzalo z národných a medzinárodných dokumentov pre archiváciu digitálnych dát. DSA Zásady tak predstavujú akési minimum vyextrahované z rôznych relevantných zdrojov [DILLO, 2015].

DSA predstavuje pomerne jednoduchý spôsob auditu a certifikácie, najmä pre začínajúce repozitáre či repozitáre menšieho rozsahu. V počte 16 smerníc, resp. zásad dôveryhodných repozitárov sa sústreďuje okolo troch základných prvkov:

I. Producenti dát (3 zásady):

1. Producent dát vkladá dáta do digitálneho repozitára spoločne s informáciami, ktoré ostatným subjektom umožňujú zhodnotiť ich kvalitu a to, nakoľko zodpovedajú etickým a iným normám platných pre danú disciplínu
2. Producent dát odovzdáva dáta vo formátoch doporučovaných digitálnym repozitárom.
3. Producent dát odovzdáva dáta spoločne s metadátami vyžadovanými digitálnym repozitárom.

II. Dátový repozitár (10 zásad):

4. Digitálny repozitár má zreteľne stanovené poslanie (mission) v oblasti digitálnej archivácie a uplatňuje ho.
5. Digitálny repozitár venuje dostatočnú pozornosť dodržiavaniu právnych predpisov a zmlúv, a to vrátane tých, ktoré sa vzťahujú k ochrane osôb.
6. Digitálny repozitár aplikuje zdokumentované procesy a postupy pre správu uchovávania dát.
7. Digitálny repozitár má plán dlhodobej ochrany digitálneho obsahu v ňom uloženého.
8. Archivácia prebieha v priebehu celého životného cyklu dát a podľa jasne stanovených postupov
9. Digitálny repozitár preberá od producentov dát zodpovednosť za sprístupnenie digitálnych objektov.
10. Digitálny repozitár umožňuje používateľom nájsť a použiť dáta a trvalo na ne odkazovať.
11. Digitálny repozitár zabezpečuje integritu digitálnych objektov a metadát.
12. Digitálny repozitár zabezpečuje autenticitu digitálnych objektov a metadát.
13. Technická infraštruktúra výslovne podporuje úlohy a funkcie popísané v medzinárodne uznávaných archívnych štandardoch ako je napr. OAIS.

III. Cieľová skupina (3 zásady):

14. Používateľ dát dodržuje prístupové pravidlá stanovené digitálnym repozitárom.

15. Používateľ súhlasí s pravidlami pre zdieľanie a správne využívanie znalostí a informácií všeobecne uznávanými v danej oblasti a riadi sa nimi.
16. Používateľ rešpektuje repozitárom stanovené licenčné obmedzenia týkajúce sa využitia dát.

Certifikát je udeľovaný na 2 roky a má veľmi silné postavenie najmä medzi európskymi inštitúciami a konzorciami (CESSDA), ktoré participujú na významných európskych projektoch (CLARIN, DARIAH, EUDAT). V súčasnej dobe (október, 2016) je certifikovaných 62 prevažne európskych repozitárov a ďalších cca 50 čaká na vyhodnotenie [DILLO, 2016]. V Českej republike sú zatiaľ len dve inštitúcie a 3 repozitáre, ktoré získali certifikát DSA. Prvým repozitárom bol *LINDAT/CLARIN* ako súčasť *Ústavu formální a aplikované lingvistiky* Univerzity Karlovej (UK). Druhou v poradí bola *Mapová sbírka v rámci Digitálneho univerzitného repozitára* UK. Treticu úspešných repozitárov uzavrel *Český sociálněvědní datový archiv* (ČSDA), Akademie věd ČR. Podľa informácií na oficiálnych stránkach DSA, nie je zatiaľ na Slovensku certifikovaný repozitár podľa DSA.

Nové Zásady DSA

Hlavnými „nedostatkami", ktoré bývajú certifikácii DSA vytýkané sú práve hodnotenie na základe dôvery a obsahovo sa prekrývajúce niektoré smernice. Práve druhá výčitka by však vďaka spojeniu DSA a WDS[1] komunity mala byť čoskoro minulosťou. Pod záštitou RDA/WDS[2] záujmovej skupiny k certifikácii digitálnych repozitárov (RDA/WDS Interest Group on Certification of Digital Repositories) vznikla pracovná skupina, ktorej cieľom bolo počas 18 mesiacov definovať skutočné kľúčové charakteristiky, zefektívniť hodnotenie a zvýšiť dopad na cieľovú komunitu.

WDS, podobne ako DSA, funguje na sebahodnotení a následnej revízii vedeckou komunitou WDS. WDS ale na rozdiel od DSA, ponúka certifikáciu len pre svojich členov. Zahŕňa nielen dátové centrá (najmä repozitáre), ale aj dátové služby. Mnoho z týchto dátových služieb sa skladá z niekoľkých zložiek (dátové centrá, analytické centrá, produktové strediská atď.) a tie majú svoju vlastnú organizačnú štruktúru. Okrem toho dátové služby často majú blízke vzťahy s ICSU vedeckých zväzov a ich asociácií. Sústreďuje sa predovšetkým na priestorovú vedu a vedu o Zemi (Space and Earth Sciences). Certifikát je udeľovaný na 3 až 5 rokov a doposiaľ ho získalo viac ako 70 členov.

Pripravované pravidlá DSA pre roky 2017 – 2020, ktoré by mali byť pravdepodobne do konca roku 2016 oficiálne predstavené, vznikali v spolupráci *Data Seal of Approval Board* a ICSU *World Data System*. Cieľom bolo sprehľadniť pôvodné kritéria a identifikovať skutočné kľúčové charakteristiky dôveryhodných repozitárov. Samotný počet zásad a ich podstata, resp. hodnotené kritéria sa nemenia, sú však prehľadnejšie a logicky zoradené. Pôjde teda (s možnými drobnými úpravami) o týchto 16 okruhov + 2 doplnkové kritéria [DSA-WDS, 2016]:

I. Organizačná infraštruktúra

1. Mandát/rozsah (Mission/Scope)
2. Licencie (Licences)
3. Kontinuita prístupu (Continuity of access)
4. Dôvernosť/etika (Confidentiality/Ethics)
5. Organizačná infraštruktúra (Organizational infrastructure)
6. Odborná pomoc (Expert guidance)

II. Manažment digitálnych objektov (Digital Object Management)

7. Integrita a autenticita dát (Data integrity and authenticity)
8. Rozhodovanie (Appraisal)
9. Zdokumentované postupy uchovávania dát (Documented storage procedures)
10. Plán dlhodobej ochrany (Preservation plan)
11. Kvalita dát (Data quality)
12. Pracovné postupy (Workflows)
13. Vyhľadanie a sprístupnenie dát (Data discovery and identification)
14. Využitie dát (Data reuse)

III. Technológie

15. Technická infraštruktúra (Technical infrastructure)
16. Bezpečnosť (Security)

Body 17. Dotatočné informácie (Additional information) a 18. Spätná väzba uchádzača (Applicant feedback) novej certifikácie DSA nie sú hodnotiace kritéria repozitárov v pravom zmysle slova. Ponúkajú možnosť dopísania doplňujúcich informácií relevantných pre hodnotenie, pričom ich nie je možné/vhodné dopísať k jednotlivým zásadám a spätnej väzby celkového priebehu certifikácie DSA z pohľadu uchádzača.

[1] World Data System (WDS) predstavuje interdisciplinárny zbor Medzinárodnej rady pre vedu (International Council for Science – ICSU, ktorá bola vytvorená v roku 2008. Podrobnejšie informácie sú dostupné na stránkach WDS – https://www.icsu-wds.org/organization.

[2] RDA/WDS Certification of Digital Repositories IG je medzinárodnou záujmovou skupinou združujúcou odborníkov na dlhodobú ochranu digitálnych dokumentov a certifikáciu digitálnych repozitárov pod záštitou RDA (Research Data Alliance) komunity. Podrobnejšie na stránkach RDA: *https://www.rd-alliance.org/groups/rdawds-certification-digital-repositories-ig.html.*

Princípy a priebeh DSA

Pre Zásady DSA je kľúčových 5 základných princípov, ktoré definujú vhodne archivované dáta:

- je možné ich nájsť na internete
- sú sprístupnené v súlade s platnou a relevantnou legislatívou, rešpektujúc ochranu osobných dát a práva duševného vlastníctva
- sú dostupné vo využiteľnom formáte pre cieľovú skupinu (designated community)
- sú spoľahlivé
- je možné na ne odkazovať (trvalé identifikátory).

Zásady sú zacielené na tri typy zúčastnených strán (stakeholders):

- producenti dát, ktorí zodpovedajú za samotnú kvalitu digitálnych dát
- dátový repozitár, ktorý zodpovedá za kvalitu uchovávania a sprístupnenia dát
- používatelia, ktorí nesú zodpovednosť za správne využívanie dát.

Priebeh auditu a certifikácie podľa DSA je možné rozdeliť na tri hlavné časti:

- sebahodnotenie prostredníctvom online webového nástroja DSA
- preskúmanie odborníkmi – hodnotenie na základe dôvery
- udelenie pečate DSA.

Nutná dokumentácia pre jednotlivé hodnotené Zásady DSA[3]

Hneď v počiatkoch sebahodnotenia pomocou online nástroja DSA je nutné definovať kontext repozitára. Tu je dôležité definovať hodnotený repozitár a špecifikovať úroveň hodnotenia, napr. či chceme certifikovať repozitár ako celok alebo len špecifickú zbierku digitálnych či digitalizovaných dokumentov. Dôležité je tiež špecifikovať mandát – je možné spomenúť aj to, či repozitár vznikol v rámci určitého projektu. Pokiaľ nie, je nutné vypracovať stručný popis, účel a charakter projektu v anglickom jazyku priamo do kontextu repozitára alebo pomocou odkazu na externú stránku. Pokiaľ áno, je dobré odkázať na webovú stránku či webové sídlo príslušného projektu. Pre lepšiu orientáciu a prehľadnosť pre hodnotiteľov je nutné sa držať terminológie OAIS definovanej v ISO norme 14721:2012. Pre dôveryhodnosť repozitára je potrebná aj informácia o tom, či repozitár využíva externých partnerov (tzv. outsourcing). Pokiaľ je outsourcing využívaný len pre niektoré časti chodu repozitára (napr. pre hardvér – cloudové úložisko) je nutné túto skutočnosť popísať v rámci jednotlivých zásad. Samozrejmosťou sú URL adresy zbierky či repozitára. Samostatný odkaz na rozhranie v anglickom jazyku nie je nutnosťou, ale výhodou.

1. Producent dát vkladá dáta do digitálneho repozitára spoločne s informáciami, ktoré ostatným subjektom umožňujú zhodnotiť ich kvalitu a to, nakoľko zodpovedajú etickým a iným normám platných pre danú disciplínu

Nové kritérium: 6. Odborná pomoc

Toto kritérium zisťuje, do akej miery repozitár podporuje producentov alebo poskytovateľov obsahu pred a počas procesu vkladania dát do repozitára.

Je možné popísať workflow a postavenie producenta či producentov dát v rámci workflow. Pokiaľ napr. ide o prípad sprístupňovania "in-house" digitalizácie, producentom dát je v taktom prípade samotný repozitár. Je vhodné popísať prípadne odkázať na metodológiu vkladania dát, definovať kompletný balík informácií, ktoré majú byť v repozitári uložené (ku príkladu informovať o zdokumentovanom výbere preferovaných a akceptovateľných formátov súborov, frekvencii importu dát, spôsobe importu (dávkový, ručný apod.) preferovaných popisných, technických a ďalších typov metadát). Je dobré doložiť aj dôkazy o tom, že zber alebo tvorba dát vznikla v súlade s etickými a právnymi normami (napr. zmluva s firmou zabezpečujúcou externú digitalizáciu, zmluva s producentom dát, odkaz na legislatívu v prípade povinného zberu dát a pod.). Je potrebné odkázať na producentov dát a ich pridružené organizácie a doložiť, že dáta sú výsledkom činnosti daného producenta obsahu (napr. inštitúcie alebo konkrétnej osoby). Pokiaľ nie sú odkazované dokumenty dostupné v anglickom jazyku, je nutné vytvoriť ich stručný popis (obvykle postačí 10 základných viet) a odkaz na dokumentáciu v originálnom znení.

Podľa nových kritérií DSA je tiež relevantné, či repozitár využíva odbornú pomoc, či sa zapája do národných a medzinárodných projektov a či sa zaujíma o svoju cieľovú skupinu (získavanie spätnej väzby od používateľov, napr. prostredníctvom dotazníkových šetrení).

2. Producent dát odovzdáva dáta vo formátoch doporučovaných digitálnym repozitárom

Nové kritérium: 8. Rozhodovanie *(náväznosť na nové kritérium 11.)*

Nové kritérium: 12. Pracovné postupy *(čiastočne; náväznosť na nové kritérium 4., 8., 9., 16.)*

Toto kritérium rozvíja predchádzajúci bod. Dôveryhodný repozitár je si vedomý základného obmedzenia dlhodobej ochrany – t. j. skutočnosti, že nie je možné dlhodobo ochrániť všetko. Keďže existujúce technológie a dátové (súborové) formáty sú vystavené riziku zastarávania, je pre repozitár dôležité mať prijímané formáty pod maximálne možnou kontrolou. Preto použí-

[3] V čase písania tohto príspevku neboli ešte nové kritéria oficiálne schválené. Pre lepšiu orientáciu v starých a nových zásadách DSA sú uvedené najprv pôvodné Zásady a pod nimi odkaz na nové Zásady DSA.

vateľom a producentom dát poskytuje záruky dlhodobej ochrany u vybraných formátov. Je vhodné stanoviť preferované formáty (ideálne neproprietárne), u ktorých repozitár dokáže zabezpečiť čo najdlhšiu zobraziteľnosť a použiteľnosť. Takýto zoznam vybraných formátov by v rámci transparentnosti mal byť zverejnený na stránkach inštitúcie, ktorá repozitár spravuje. Je potrebné popísať aj spôsoby, akými repozitár zabezpečuje kontroly kvality, aby producenti dát poskytovali dáta len v preferovaných formátoch (napr. u tzv. master copies, prípadne user copies). Zároveň repozitár informuje o tom, ako je narábané s digitálnymi objektmi, ktoré sú v iných než preferovaných formátoch (napr. zamietnutie importu, absencia garancie dlhodobej ochrany apod.). Je vhodné tiež popísať, či repozitár požaduje po producentoch dát podrobnejšie informácie o formátoch súborov a prípadne nástrojoch či metódach, ktoré boli pri tvorbe súborov použité.

3. Producent dát odovzdáva dáta spoločne s metadátami vyžadovanými digitálnym repozitárom

Nové kritérium: 14. Využitie dát *(náväznosť na nové kritérium 2.)*

Nové kritérium: 11. Kvalita dát *(náväznosť na nové kritérium 7., 8., 12.)*

V tomto bode je nutné poukázať na to, akým spôsobom digitálny repozitár podporuje producentov dát pri odovzdávaní metadát, a to nielen popisného charakteru. Ideálne je, pokiaľ repozitár má aspoň rámcovo stanovené požiadavky na:

a) popisné (informácie nutné pre jednoznačnú identifikáciu a vyhľadanie, prípadne aj ozrejmenie významu daného dokumentu)
b) štrukturálne (popis vzťahov medzi jednotlivými prvkami skupiny vzájomne súvisiacich dát)
c) technické (určenie špecifických vlastností súborov – farebný profil, formát súboru a pod.)
d) administratívne metadáta (popis práv duševného vlastníctva, podmienky využitia a prístupu a pod.).

Skúma sa, či sú požadované popisné metadáta relevantné pre používateľov? Dôležité je aj popísať, akým spôsobom repozitár zabezpečuje kontrolu kvality dodaných dát. Sú k dispozícii štandardné či vlastné nástroje kontroly dát ako na strane producentov dát, tak aj na strane repozitára? V prípade, že kvalita metadát nedosahuje požadovanú úroveň, je potrebné popísať ako repozitár v danom prípade postupuje.

4. Digitálny repozitár má zreteľne stanovené poslanie (mission) v oblasti digitálnej archivácie a uplatňuje ho

Nové kritérium: 1. Mandát/rozsah

Aj keď sa môže zdať, že toto kritérium nie je nijako významné, opak je pravdou. Vo svojej podstate ide o východiskový bod činnosti repozitára. Repozitár musí vedieť aké má poslanie a ktoré ciele má napĺňovať (mission statement). Je nutné tiež upozorniť na platnosť a váhu daného dokumentu tým, že repozitár odkáže na autoritu daného poslania (napr. zriaďovateľ, financujúca organizácia apod.). V prípade, že repozitár niekedy v budúcnosti nebude schopný daný mandát naplniť, je nutné určiť, čo sa bude diať s dátami? Je preto potrebné mať vypracovaný tzv. nástupnický plán (succession plan).

5. Digitálny repozitár venuje dostatočnú pozornosť dodržiavaniu právnych predpisov a zmlúv, a to vrátane tých, ktoré sa vzťahujú k ochrane osôb

Nové kritérium: 4. Dôvernosť/etika *(náväznosť na kritérium 2., 12.)*

Transparentnosť repozitára je daná nielen dodržovaním interne definovanej dokumentácie, ale aj akýchkoľvek právnych prepisov (na národnej i medzinárodnej úrovni), štandardov (ISO, de facto štandardov atď.), zmlúv, ktoré ovplyvňujú chod repozitára. Je vhodné začať popisom samotného právneho/organizačného statusu, v rámci ktorého hodnotený repozitár existuje. Pokiaľ je to možné, je vítané zverejniť zmluvy s producentmi dát alebo aspoň šablóny uzatváraných zmlúv. Akým spôsobom repozitár postupuje v prípade citlivých a osobných údajov (napr. zverejňovanie zmlúv, uchovávanie logov o prístupoch do repozitára, sprístupnenie dát dôverného charakteru)? Cestou anonymizácie či zabezpečenia prístupu k nim? Sú zamestnanci na riešenie takýchto situácií dostatočne preškolení?

6. Digitálny repozitár aplikuje zdokumentované procesy a postupy pre správu uchovávania dát

Nové kritérium: 9. Zdokumentované postupy uchovávania dát *(náväznosť na nové kritérium 15., 16.)*

Tento okruh certifikácie odkazuje na schopnosť repozitára dlhodobo ochraňovať digitálne dáta. Je dôležité mať zdokumentovanú stratégiu dlhodobej ochrany, spôsob bitovej ochrany dát (zálohovanie - offline a/alebo online, viacnásobné kópie), kontrola celistvosti dát (napr. kontrolné súčty), spôsob obnovy dát (zodpovednosť, postup). Repozitár by si mal byť vedomý rizík, ktoré mu hrozia (v okolí či v rámci repozitára). Ideálne je využiť niektorú z techník manažmentu rizík (napr. pomocou DRAMBORA nástroja). Dôležitý je aj určitý druh monitoringu najnovších a zastaralých technológií a médií.

7. Digitálny repozitár má plán dlhodobej ochrany digitálneho obsahu v ňom uloženého

Nové kritérium: 10. Plán dlhodobej ochrany

Tento bod súvisí s predchádzajúcim kritériom 6. a rozširuje ho o podmienku dlhodobého prístupu k dátam, ako aj ich dlhodobej použiteľnosti. Pokiaľ má repozitár k dispozícii určitý spôsob sledovania zastaralých formátov, je nutné vysvetliť, akou cestou sa repozitár bude uberať v prípade ich zastarania – napr. migrácia, emulácia. Dlhodobú použiteľnosť ovplyvňuje nielen celistvosť formátu a objektu ako takého, ale aj jeho zrozumiteľnosť a využiteľnosť cieľovou skupinou používateľov (náväznosť na zásadu 1.). Získava repozitár spätnú väzbu od používateľov? Vie, či pre používateľov sú dané dáta v podobe, v akej sú dáta prezentované aj použiteľné? Pokiaľ nie, ako to rieši?

8. Archivácia prebieha počas celého životného cyklu dát a podľa jasne stanovených postupov

Nové kritérium: 5. Organizačná infraštruktúra *(čiastočne)*

Toto kritérium dôveryhodnosti podľa DSA odkazuje na potrebu dôkladnej dokumentácie, pracovných postupov (workflow), rozhodovacích procesov, výberu dát, prístupu k dátam. Keďže ľudský faktor je neoddeliteľnou súčasťou repozitára, musí naviac repozitár (podľa nových pravidiel DSA) preukázať, že má dostatočný rozpočet, potrebné technológie a jeho zamestnanci majú zabezpečené priebežné vzdelávanie v odbore i profesionálny rast.

9. Digitálny repozitár preberá od producentov dát zodpovednosť za sprístupnenie digitálnych objektov

Nové kritérium: 3. Kontinuita prístupu *(náväznosť na nové kritérium 1., 10. a 14.)*

Repozitár by mal popísať licenčné či zmluvné dohody s producentmi dát. Do určitej miery sa tento bod obsahovo prekrýva so zásadami 4 a 7. Pokiaľ je to možné, zmluvu je vhodné sprístupniť online a doplniť o krátky výťah v anglickom jazyku. Do tohto kritéria ale spadá aj existencia krízového plánu, ktorý prehľadnou formou dokumentuje identifikované riziká repozitára a popisuje riešenia napr. pri haváriách, živelných pohromách, katastrofách, úmyselných útokoch (vo vnútri organizácie/repozitára i zvonku), nedostatku finančných prostriedkov a pod.

10. Digitálny repozitár umožňuje používateľom nájsť a použiť dáta a trvalo na ne odkazovať

Nové kritérium: 13. Vyhľadanie a sprístupnenie dát

Repozitár by mal popísať, akým spôsobom trvale sprístupňuje a odkazuje na objekty v repozitári. Ponúka repozitár vyhľadávacie rozhranie, prvky pokročilého vyhľadávania, filtre a pod.? Sprístupňuje dáta vo viacerých formách – napr. aj prostredníctvom OAI-PMH protokolu? Má k dispozícii integrované citačné nástroje? Aké trvalé identifikátory priraďuje objektom – len interné alebo aj medzinárodne uznávané persistentné identifikátory (Handle, DOI apod.)?

11. Digitálny repozitár zabezpečuje integritu digitálnych objektov a metadát

Nové kritérium: 7. Dátová integrita a autenticita *(náväznosť na nové kritérium 8., 9., 10., 12. – 14.)*

Využíva repozitár kontrolné súčty pre verifikáciu dát? Aké typy? Prebieha pravidelné monitorovanie integrity dát a metadát? Pracuje repozitár s viacerými verziami dát? Pokiaľ áno, aká je stratégia verzovania objektov?

12. Digitálny repozitár zabezpečuje autenticitu digitálnych objektov a metadát

Nové kritérium: 7. Dátová integrita a autenticita *(náväznosť na nové kritérium 8., 9., 10., 12. – 14.)*

Toto kritérium dôveryhodného repozitára sa zameriava na zachytenie „cesty" digitálneho dokumentu od vzniku až po jeho sprístupnenie používateľom, a to z hľadiska spoľahlivosti originálu (autenticity) a informácií o pôvode (proveniencie) dát. Preto je dôležité popísať, akým spôsobom dáta vznikajú (digitalizáciou, povinným elektronickým výtlačkom a pod.), ako je zabezpečená autenticita, integrita a kvalita, ako repozitár reaguje na zmenu dát (stratégia zmeny dát), napr. v prípade inej verzie objektu, zmeny formátu súboru po migrácii, poškodení súboru v dôsledku chyby na strane hardvéru alebo softvéru apod. Akým spôsobom repozitár odkazuje na metadáta, má identifikované signifikantné vlastnosti súborov? Akým spôsobom repozitár kontroluje identitu vkladateľa/producenta dát?

13. Technická infraštruktúra výslovne podporuje úlohy a funkcie popísané v medzinárodne uznávaných archívnych štandardoch ako je napr. OAIS

Nové kritérium: 15. Technická infraštruktúra

Nové kritérium: 16. Bezpečnosť *(náväznosť na nové kritérium 9., 12.)*

Je možné využiť a ďalej rozvinúť myšlienky fungovania repozitára z úvodného kritéria Kontext (viď prvý odstavec časti „Nutná dokumentácia pre jednotlivé hodnotené Zásady DSA"). Je nutné popísať do akej miery repozitár a predovšetkým jeho technická infraštruktúra rešpektuje štandardy (ISO, zaužívané štandardy a pod.). Samozrejmosťou by mal byť súlad s referenčným modelom OAIS (ISO 14721:2012). Ďalšie štandardy môžu byť špecifické pre technické riešenie (napr. ISO ISO 27001:2013 Systém manažérstva technickej bezpečnosti (SMBI)), a to najmä pre väčšie až veľké repozitárové/archívne riešenia. Výhodou je, ak má repozitár prípadne vypracovaný aj plán rozvoja infraštruktúry. Má repozitár pre svoje softvérové riešenie (či už komerčného alebo open-source charakteru) dostatočne silnú komunitu, na ktorú sa môže obrátiť pre zdieľanie vyvíjaných funkcionalít či v prípade problémov?

14. Používateľ dát dodržuje prístupové pravidlá stanovené digitálnym repozitárom

Nové kritérium: 2. Licencie *(náväznosť na nové kritérium 4.)*

Repozitár by mal jasne vedieť, aké majú digitálne objekty, prípadne kolekcie podmienky prístupu. V prípade, že niektoré dáta vyžadujú špecifický režim, je nutné ho dostatočne popísať. Je možné odkázať aj na špeciálne licencie typu Creative Commons alebo pripojiť a preložiť konkrétne licenčné podmienky (pokiaľ ich repozitár uplatňuje), ktoré pre používateľov platia. Zároveň je nutné poukázať na to, čo sa deje v prípadoch, keď dané podmienky nie sú plnené a aké opatrenia sú v platnosti, aby sa predišlo zneužívaniu podmienok prístupu.

15. Používateľ súhlasí s pravidlami pre zdieľanie a správne využívanie znalostí a informácií všeobecne uznávanými v danej oblasti a riadi sa nimi

Nové kritérium: 2. Licencie *(náväznosť na nové kritérium 4.)*

Táto zásada vo väčšine prípadov naväzuje na predchádzajúce kritérium. Obecné pravidla správania sa často vzťahujú práve na licenčné ujednania.

16. Používateľ rešpektuje repozitárom stanovené licenčné obmedzenia týkajúce sa využitia dát

Nové kritérium: 2. Licencie *(náväznosť na nové kritérium 4.)*

Toto kritérium sa opäť vzťahuje na predchádzajúce kritéria o uplatňovaní licencií, prípadne na platnú národnú legislatívu.

Výhody DSA

Medzi nepochybné výhody auditu a certifikácie DSA, ktoré však majú všeobecnú platnosť aj pre ostatné druhy certifikácií dôveryhodných digitálnych repozitárov, patrí [DILLO, 2015]:

- **Dôvera zainteresovaných strán**: pečať DSA dáva investorom a zriaďovateľom garanciu účelne investovaných prostriedkov, producentom dát istotu dobre ochránených dát a používateľom záruku vhodne spravovaných dát.
- **Zlepšenie komunikácie**: príprava sebahodnotenia (interného auditu) dáva príležitosť dôkladnej dokumentácie procesov a postupov na všetkých úrovniach repozitára ako organizácie, čím do značnej miery zlepšuje komunikáciu.
- **Zefektívnenie procesov**: sebahodnotenie povzbudzuje repozitáre ku zlepšovaniu a zefektívňovaniu procesov a postupov za účelom vyššej kvality, odbornosti a profesionality.
- **Vyššia transparentnosť**: certifikácia DSA si zakladá na verejne prístupných dokumentáciach a dôkazoch (tam, kde je to možné).
- **Zvyšovanie povedomia o dlhodobej ochrane**: súlad so Zásadami DSA preukazuje záväzok repozitárov uchovávať dáta pre budúce generácie.

Cieľom nových DSA Zásad však bolo najmä sprehľadniť pôvodné kritéria a identifikovať skutočné kľúčové charakteristiky dôveryhodných repozitárov. Samotný počet zásad a ich podstata, resp. hodnotené kritéria sa nemenia, sú však prehľadnejšie a logicky zoradené.

Nespornou výsadou DSA oproti iným certifikáciam je predovšetkým nižšia prácnosť, časová a finančná náročnosť (16 zásad DSA oproti 34 kritériám DIN 31644 alebo 109 metrikám ISO 16363).

Záver

Dlhodobá udržateľnosť repozitárov prináša so sebou celý rad náročných problémov z rôznych oblastí, či už ide o organizačné, technické, finančné, právne záležitosti. Audit a certifikácia digitálneho repozitára tak dáva istotu, že repozitár v súlade so svojim mandátom a poslaním poskytuje spoľahlivý a dlhodobý prístup k organizovaným digitálnym zdrojom cieľovej skupine – dnes i do budúcna. Plnenie týchto funkcií, resp. kritérií musí byť objektívne preukázateľné, čo v praxi znamená, že dosiahnutie dôveryhodnosti je do značnej miery závislé od auditu a certifikácie.

V súčasnosti je dostupných viacero certifikačných štandardov na rôznych stupňoch komplexnosti. Základná certifikácia, do ktorej spadá aj Data Seal of Approval, pokrýva kľúčové otázky a minimálne kritéria (identifikácia silných a slabých stránok), ktoré by dobre fungujúci, dôveryhodný repozitár mal spĺňať. Nespornou výsadou DSA oproti iným certifikáciám je však predovšetkým jeho nižšia prácnosť, ako aj časová a finančná náročnosť. Práve tá by mohla byť motiváciou ku zefektívneniu procesov a zárukou pre používateľov u repozitárov menšieho rozsahu či repozitárov, ktoré sa s problematikou dlhodobej ochrany a dôveryhodných repozitárov len oboznamujú.

Zoznam bibliografických odkazov

Data Seal of Approval : Guidelines version 2 [online]. 2013. Data Seal of Approval Board, 2014 [cit. 2016-10-23]. Dostupné na: *https://assessment.datasealofapproval.org/guidelines_52/pdf/*

DILLO, Ingrid a Lisa DE LEEUW. 2014. Ten Years Back, Five Years Forward: The Data Seal of Approval. *International Journal of Digital Curation* [online]. 2014, 10(1), 230-239 [cit. 2016-10-23]. DOI: 10.2218/ijdc.v10i1.363. Dostupné z: *http://ijdc.net/index.php/ijdc/article/viewFile/10.1.230/394*

DSA-WDS Partnership Working Group Catalogue of Common Requirements [online]. 2016. World Data Systems, February 16, 2016 [cit. 2016-10-23]. Dostupné na: https://www.rd-alliance.org/system/files/DSA–WDS%20Catalogue%20of%20Common%20Requirements%20V2.2.pdf

MIRANDA, Andrea. 2015. Důvěryhodná digitální úložiště, jejich audit a certifikace. *Knihovna (knihovnická revue)* [online]. 2015, 26(2), 49-57 [cit. 2016-10-23]. ISSN 1802-8772. Dostupné z: *http://knihovnarevue.nkp.cz/archiv/dokumenty/2015-2/miranda.pdf*

Tento príspevok vznikol v rámci grantového projektu DG16P02R044, **NAKI II** ARCLib: komplexní řešení pro dlouhodobou archivaci digitálních (knihovních) sbírek (2016 – 2020).

Mgr. Andrea Miranda, Ph.D.

andrea.fojtu@ruk.cuni.cz andrea.miranda@ruk.cuni.cz ■

(Univerzita Karlova v Prahe)