

ExLibris Rosetta

Dlouhodobá ochrana digitálního obsahu

Elektronická data se dnes stávají klíčovou součástí sbírek národních a univerzitních knihoven. Cílem knihoven a vědeckých institucí zůstává rozvíjet a podporovat růst znalostí a zajišťovat dostupnost vědeckých dat a dat kulturně historické hodnoty dalším generacím. Správa a ochrana velkých objemů digitálních dat v řadě různých formátů však staví vědecké instituce a knihovny před nové úkoly.

Rosetta je nástrojem, který má institucím pomoci s plněním těchto úkolů. Jedná se o systém, který byl vyvinut ve spolupráci s vůdčí institucí tohoto oboru – Národní knihovnou Nového Zélandu. Dnes se pro něj rozhodují jak velké vědecké instituce (v poslední době například ETH Zurich, KU Lueven, SLUB Dresden, nebo německé konsorcium GOPORTIS), tak instituce, která mají za úkol ochranu a zpřístupnění rozhodujících částí digitálního kulturního dědictví (velké národní knihovny a archivy, v poslední době například Bavorská státní knihovna). Tyto instituce oceňují mimo jiné precizní podporu uznávaných standardů, zaměření jak na ochranu bit-streamu, tak na ochranu obsahu, stejně jako flexibilitu a otevřenost systému. A také to, že Rosetta je light-archivem – tedy archivem, který umožňuje kromě archivace i šíření digitálních objektů.

Jak již bylo zmíněno, Rosetta plně podporuje všechny současné standardy pro dlouhodobou ochranu digitálních dat, ať už se jedná o referenční model OAIS, standard PREMIS nebo o formát METS pro strukturální metadata. Je navržena tak, aby maximálně automatizovala všechny procesy archivu a funkce referenčního modelu OAIS (ISO 14721). Požadavky vycházející z OAIS jsou implementovány v informačním systému Rosetty a jsou maximálně automatizovány, a to včetně implementace modulu plánování dlouhodobé ochrany, který nemá na trhu srovnání.

Systém, plně v souladu s normou OAIS, poskytuje dlouhodobou ochranu informačního obsahu a nikoli pouze trvalé uchování bitů. Jedná se však o ten nejzákladnější požadavek OAIS, který ale přesto v řadě jiných systémů není naplněn. Takováto řešení pak nejsou schopna zajistit trvalou použitelnost informačního obsahu archivovaných dat, pouze archivují informační balíčky tvořené bity a tyto bity jsou schopna v původní podobě uživateli po určité době dodat. Dlouhodobý archiv (tak jak je chápán v normě OAIS) musí podporovat celý životní cyklus archivovaných dat a informace o jejich změnách musí uchovávat spolu s daty. V dlouhodobém archivu se počítá s tím, že s uloženými objekty bude třeba neustále pracovat; každý jednotlivý soubor, který archivní balíčky tvoří, musí být dostupný pro budoucí validace a migrace a metadata archivních objektů musí být možné aktualizovat. Proto Rosetta implementuje data model založený na standardu PREMIS, který umožňuje samostatně přistupovat k intelektuálním entitám, jejich reprezentacím, souborům a bit-streamům, které je tvoří.

Základním požadavkem na systém pro dlouhodobé uchování dat je také flexibilita a otevřenost. Proto Rosetta poskytuje uživatelům standardní otevřená rozhraní pro integraci externích služeb a softwaru stejně jako pro integraci s ostatními systémy instituce. Rosetta je navržena tak, aby mohla snadno a efektivně fungovat konsorciálně nebo v komplexních organizacích, kde existuje mnoho sbírek, dodavatelů dat a dokumentů s různými nároky na zpřístupnění. Technická architektura Rosetty umožňuje snadno škálovat až k obrovským objemům ukládaných nebo zpřístupňovaných dat. Rosetta je schopna jakékoli typy dokumentů s jakýmkoli metadaty a lze ji integrovat s libovolným storage řešením.

Co se zpřístupnění týče, Rosetta není postavena jako dark-archive, naopak. Cílem institucí, které Rosettu používají, je zajištění trvalé dostupnosti svěřeného obsahu a jeho poskytnutí budoucím uživatelům v nezávisle srozumitelné podobě. Rosetta je integrována s discovery systémem Primo, který poskytuje uživatelsky přívětivé rozhraní. Kromě toho Rosetta umožňuje data publikovat do jiných portálů pomocí standardních protokolů.

Díky výše uvedeným i dalším zde nezmíněným vlastnostem Rosetty stále roste počet jejích uživatelů a spolu s ním roste i šance, že jednotlivé instituce uspějí se své snaze o uchování digitálních dat pro příští generace.

Marek Melichar, Eliska Pavlásková, Multidata Praha, s.r.o.

