Referát o článku The Rijksmuseum Collection as Linked Data

Filip Slivka

Autoři článku popisují současný stav sbírky amsterdamského muzea Rijksmuseum publikované jako linked data. Je představena historie tohoto projektu, proces digitalizace, jednotlivé charakteristiky dat, statistiky datového souboru a podrobnější popis vybraných modelů dat.

Rijksmuseum jakožto nejnavštěvovanější muzeum v Holandsku je jedno z mála muzeí na světě, která zveřejnila svou sbírku umění jako propojená data. Motivací k tomuto kroku bylo hned několik. Sbírka muzea čítá přes milion objektů, z nichž jen zlomek může být vystaven ve stejnou dobu. Pro zveřejnění všech objektů začalo muzeum sbírku digitalizovat a publikovat ji online, což by dle očekávání mělo zvýšit počet návštěvníků muzea. Dalším přínosem je postavení uměleckých děl do širšího kontextu díky propojeným datům a linkování na externí datasety.

Muzeum zaměstnává katalogizátory, kteří registrují objekty do systému pro správu muzejní sbírky a fotografa, jenž pořizuje snímky objektů ve vysokém rozlišení. Na začátku roku 2016 čítal systém přes půl milionu objektů, pro jejichž popis je k dispozici 597 atributů. Ne všechny z těchto atributů jsou však použity na výstupu systému, některé z nich totiž obsahují citlivé data, jako jsou údaje o pojištění. Transformace dat z interního systému do RDF je uskutečňován pomocí XSLT souboru. Všechny relevantní metadata jsou mapována na vlastnost z Europeana Data Model. Pro některá pole v systému jsou dostupné pouze textové hodnoty, ostatní jsou specifikovány pomocí konceptů. Ty jsou manuálně přidávány pracovníky muzea, kteří je přiřazují buď z interního tezauru muzea, externích datasetů nebo kombinací těchto dvou možností. Všechny použité koncepty mají unikátní identifikátory. Například termín purl:PEOPLE.5706 označuje Rembrandta nebo aat:300015050 označuje olejomalbu. Na výstupu aplikačního rozhraní muzejního systému jsou data načítána do grafové databáze. Aktualizace této databáze provádí zaměstnanci každý měsíc.

Jak je již zmíněno výše, verze sbírky muzea v propojených datech je modelována podle Europeana Data Model (EDM). V EDM jsou využívány elementy z již existujících modelů, jako je například Dublin Core. Struktura modelu je vyjádřena v RDF Schema, využitím konstruktů subclass a relací subproperty. Avšak ne všechny konstrukty zahrnuté v EDM vyhovují datům o muzejní sbírce. Pro tyto účely podporuje EDM úpravu a rozšiřování modelu. URI jednotlivých objektů jsou založeny na schématu purl a skládají se z pěti elementů: prefix purl, typ datasetu, kód země, ID organizace a ID objektu. Pro metadata o sbírce je používán model Vocabulary of a Friend (VOAF). Objekty mohou být propojeny mezi sebou pomocí predikátů dcterms:hasPart a dcterms:isPartOf. Tyto relace jsou například využívány pro propojení jednotlivých fotografií s jejich albem.

Muzeum poskytuje na svých webových stránkách API pro vývojáře, díky čemuž byly data použity v mnoha výzkumných projektech a bylo nad nimi vytvořeno nespočet různých aplikací. Jiná muzea (například Smithsonian American Art Museum) se odkazují ve svých databázích do tezauru Rijksmuseum. Pro online návštěvníky je k dispozici sofistikovaná

webová aplikace, kde je možno prohlížet sbírku muzea a vyhledávat pomocí mnoha různých parametrů. Propojená data mají pro muzeum značný přínos a je na tento dataset nahlíženo jako na součást kulturního dědictví.

Reference

DIJKSHOORN, Chris, Lizzy JONGMA, Lora AROYO, Jacco van OSSENBRUGGEN, Guus SCHREIBER, Wesley ter WEELE a Jan WIELEMAKER. The Rijksmuseum Collection as Linked Data. Semantic Web journal [online]. IOS Press [cit. 2016-11-26]. ISSN 1570-0844. Dostupné z: http://www.semantic-web-journal.net/system/files/swj1210.pdf