

## Code4Lib Journal

Örömmel jelentem, hogy megjelent a Code4Lib Journal új, 35. száma.

<http://journal.code4lib.org/issues/issues/issue35>

A számnak van egy magyar vonatkozása is. *Sennyey Pongrácz* a tavalyi, debreceni Networkshopon tartott A Hálózat, a Felhő és a Könyvtár című nyitóelőadásában beszámolt a texasi St. Edward's egyetem könyvtárában végzett munkájuk megközelítésmódjáról (az előadás megtekinthető a Videotóriumban:

<http://niif.videotorium.hu/hu/recordings/12854>).

Az előadás nem tért ki technikai részletekre, ám szerzőtársaival most rávilágítanak a megoldásuk technológiai hátterére:

<http://journal.code4lib.org/articles/12191>.

Két cikk foglalkozik irattári, kéziratári, levéltári jellegű dokumentumok és katalógusok kezelésével. Az egyik a Bentley Történeti Könyvtár projektjét mutatja be, melyben különböző munkafolyamatokra kifejlesztett szoftvereket (ArchiveSpace, Archivematica és DSpace) integrálnak. A másik egy számomra különösen érdekes projektet mutat be. Talán ismeretes, hogy a Harvard egyetem egész felépítése és ezen belül a gyűjteményi rendszere is erőteljesen széttagolt. Az számomra is meglepetés volt, hogy közel 40 egyetemi levéltár/irattár/kéziratár működik a kampuszon – persze sok szempontból koordináltan. A Harvard a 90-es évek közepén az EAD (Encoded Archival Description) egyik korai implementátora volt, és az összes levéltárban egységesen tértek át az EAD használatára, valamint egy egyedi EAD-kezelő szoftvert is kifejlesztettek. A szoftver sikeres volt, 20 évig szolgálta az egyetemet, de továbbfejlesztés évek óta nem történt. Tavaly úgy határoztak, hogy az idő során felgyülemlett katalógusokat (több mint 3000 segédlet, közel 1 GB-nyi összerjedelem) az időközben feltűnt, és sztenderdizálódott ArchiveSpace szoftverrel fogják áttölteni. Az áttöltés megkezdésekkor derült ki, hogy az EAD mind szemantikus mind formális tekintetben elég

megengedő, és ugyanazon jellegzetességet sokféleképpen lehet rögzíteni, amivel az egyes levéltárak éltek is, követve saját évtizedes vagy éppen százados katalógizáló hagyományait. Az ArchiveSpace azonban rigorózusabbnak bizonyult, mint maga a szabvány, és van egy kellemetlen tulajdonsága: a segédleteket csak egyben hajlandó importálni, ha a több ezer rekord közül (emlékeztetőül: az EAD egy hierarchikus leíró szabvány, vagyis a segédlet a felső szintektől (pl. egy professzor teljes hagyatéka, vagy akár egy kar teljes iratanyaga) fokozatosan halad az alsóbb szintekig), akár csak egy is „rossz”, az importálás egyetlen rekord esetében sem lesz sikeres. A szerzők a cikkben azt a folyamatot írják le, hogy hogyan sikerült mégis áttölteniük a rekordokat (belinkelve az általuk vagy mások által írt programok forráskódjait is).

Végül szeretném kiemelni *Ruth Kitchin Tillman* szerkesztőségi jegyzetét, amit a könyvtáros szakmai etikáról írt. A szakma belső szabályaihoz való ragaszkodás, különösen a nagyvállalati törekvések és az aktuálpolitika fényében számos etikai kérdést vet fel, amivel szembe kell nézni, és amelyek cselekvésre kell sarkalljanak.

Ahogy eddig is, most is szeretném megragadni az alkalmat arra, hogy felhívjam a Katalist olvasóinak figyelmét arra, hogy a Code4Lib Journal folyamatosan várja a reménybeli szerzők jelentkezését. Aktuális jelentkezési határidőnk április 14. Az érdeklődők a további részleteket itt találják:

<http://journal.code4lib.org/call-for-submissions>.

A teljes tartalomjegyzék:

Editorial: Introspection as Activism, or, Getting Our Houses in Order

<http://journal.code4lib.org/articles/12232>

Ruth Kitchin Tillman

Those of us in libraries like to trace our history to Alexandria or to the French governmental system of record-keeping, but the construction of the modern GLAM world is far more recent, almost as new as coding.

It has evolved almost as rapidly. And its future is on us, whether we choose to passively accept a status quo others build or to act and grow and develop ourselves and our workplaces.

#### Bridging Technologies to Efficiently Arrange and Describe Digital

Archives: the Bentley Historical Library's ArchivesSpace-Archivematica-DSpace Workflow Integration Project

<http://journal.code4lib.org/articles/12105>

Max Eckard, Dallas Pillen, Mike Shallcross In recent years, ArchivesSpace and Archivematica have emerged as two of the most exciting open source platforms for working with digital archives. The former manages accessions and collections and provides a framework for entering descriptive, administrative, rights, and other metadata. The latter ingests digital content and prepares information packages for long-term preservation and access. In October 2016, the Bentley Historical Library wrapped up a two-year, \$355,000 grant from the Andrew W. Mellon Foundation to partner with the University of Michigan Library on the integration of these two systems in an end-to-end workflow that will include the automated deposit of content into a DSpace repository. This article provides context of the project and offers an in-depth exploration of the project's key development tasks, all of which were provided by Artefactual Systems, the developers of Archivematica (code available at <https://github.com/artefactual-labs/appraisal-tab>).

#### The Devil's Shoehorn: A case study of EAD to ArchivesSpace migration at a large university

<http://journal.code4lib.org/articles/12239>

Dave Mayo, Kate Bowers

A band of archivists and IT professionals at Harvard took on a project to convert nearly two million descriptions of archival collection components from marked-up text into the ArchivesSpace archival metadata management system. Starting in the mid-1990s, Harvard was an alpha implementer of EAD, an SGML (later XML) text markup language for electronic inventories, indexes, and finding aids that archivists use to wend their way through the sometimes quirky filing systems that bureaucracies establish for their records or the utter chaos in which some individuals keep their personal archives. These pathfinder documents, designed to cope with messy reality, can themselves be difficult to classify. Portions of them are rigorously structured, while other parts are narrative. Early documents predate the establishment of the standard; many feature idiosyncratic encoding that had been

through several machine conversions, while others were freshly encoded and fairly consistent. In this paper, we will cover the practical and technical challenges involved in preparing a large(900MiB) corpus of XML for ingest into an open-source archival information system (ArchivesSpace). This case study will give an overview of the project, discuss problem discovery and problem solving, and address the technical challenges, analysis, solutions, and decisions and provide information on the tools produced and lessons learned. The authors of this piece are Kate Bowers, Collections Services Archivist for Metadata, Systems, and Standards at the Harvard University Archive, and Dave Mayo, a Digital Library Software Engineer for Harvard's Library and Technology Services. Kate was heavily involved in both metadata analysis and later problem solving, while Dave was the sole full-time developer assigned to the migration project.

#### Participatory Design Methods for Collaboration and Communication

<http://journal.code4lib.org/articles/12127>

Tara Wood, Cate Kompare

Website redesigns can be contentious and fraught in any type of organization, and libraries are no exception. Coming to consensus on priorities and design decisions is nearly impossible, as different groups compete to ensure their subject or specialty area is represented. To keep projects on track and on time, libraries may give a few staff members the authority to make all of the decisions, while keeping user research limited to a small number of usability tests. While these tactics are sometimes necessary, at best they can leave many feeling left out of the process, and at worst, can result in major oversights in the final design.

Participatory design methods can bring users and stakeholders into the design process and ultimately lead to a better design and less friction in the project. The authors share their experience and lessons learned using participatory design techniques in a website redesign project at a large, multi-location academic library, and how these techniques facilitated communication, shaped design decisions, and kept a complex, difficult project on track.

#### Python, Google Sheets, and the Thesaurus for Graphic Materials for Efficient Metadata Project Workflows

<http://journal.code4lib.org/articles/12182>

Jeremy Bartczak, Ivey Glendon

In 2017, the University of Virginia (U.Va.) will launch a two year initiative to celebrate the bicen-

ennial anniversary of the University's founding in 1819. The U.Va. Library is participating in this event by digitizing some 20,000 photographs and negatives that document student life on the U.Va. grounds in the 1960s and 1970s.

Metadata librarians and archivists are well-versed in the challenges associated with generating digital content and accompanying description within the context of limited resources. This paper describes how technology and new approaches to metadata design have enabled the University of Virginia's Metadata Analysis and Design Department to rapidly and successfully generate accurate description for these digital objects. Python's pandas module improves efficiency by cleaning and repurposing data recorded at digitization, while the lxml module builds MODS XML programmatically from CSV tables. A simplified technique for subject heading selection and assignment in Google Sheets provides a collaborative environment for streamlined metadata creation and data quality control.

Supporting Oral Histories in Islandora

<http://journal.code4lib.org/articles/12176>

Marcus Emmanuel Barnes, Natkeeran Ledchumykanthan, Kim Pham, and Kirsta Stapelfeldt Since 2014, the University of Toronto Scarborough Library's Digital Scholarship Unit (DSU) has been working on an Islandora-based solution for creating and stewarding oral histories (the Oral Histories solution pack). Although regular updates regarding the status of this work have been presented at Open Repositories conferences, this is the first article to describe the goals and features associated with this codebase, as well as the roadmap for development. An Islandora-based approach is appropriate for addressing the challenges of Oral History, an interdisciplinary methodology with complex notions of authorship and audience that both brings a corresponding complexity of use cases and roots Oral Histories projects in the ever-emergent technical and preservation challenges associated with multimedia and born digital assets. By leveraging Islandora, those embarking on Oral Histories projects benefit from existing community-supported code. By writing and maintaining the Oral Histories solution pack, the library seeks to build on common ground for those supporting Oral Histories projects and encourage a sustainable solution and feature set.

Building a Scalable and Flexible Library Data Dashboard

<http://journal.code4lib.org/articles/12152>

Nathan Mealey

Data dashboards provide libraries with the means to demonstrate their ongoing activities and usage in an engaging and communicative fashion.

Yet, due to the number of service platforms used by libraries, and the wide-ranging technical specifications they entail, bringing all of this content together in a sustainable way is a significant challenge.

This article describes Portland State University's project to design and build a data dashboard based on a scalable and flexible infrastructure that would enable them to present data in a visually compelling and dynamic interface.

What's New? Deploying a Library New Titles Page with Minimal Programming

<http://journal.code4lib.org/articles/12199>

John Meyerhofer

With a new titles web page, a library has a place to show faculty, students, and staff the items they are purchasing for their community.

However, many times heavy programming knowledge and/or a LAMP stack (Linux, Apache, MySQL, PHP) or APIs separate a library's data from making a new titles web page a reality. Without IT staff, a new titles page can become nearly impossible or not worth the effort. Here we will demonstrate how a small liberal arts college took its acquisition data and combined it with a Google Sheet, HTML, and a little JavaScript to create a new titles web page that was dynamic and engaging to its users.

OPRM: Challenges to Including Open Peer Review in Open Access Repositories

<http://journal.code4lib.org/articles/12171>

Pandelis Perakakis, Agnes Ponsati, Isabel Bernal, Carles Sierra, Nardine Osman, Concha Mosquera-de-Arancibia, Emilio Lorenzo The peer review system is the norm for many publications. It involves an editor and several experts in the field providing comments for a submitted article. The reviewer remains anonymous to the author, with only the editor knowing the reviewer's identity. This model is now being challenged and open peer review (OPR) models are viewed as the new frontier of the review process. OPR is a term that encompasses diverse variations in the traditional review process. Examples of this are modifications in the way in which authors and reviewers are aware of each other's identity (open identities), the visibility of the reviews carried out (open reviews) or the opening up of the review to the academic community (open participation). We present the project for the implementation of an Open Peer Review Mod-

ule in two major Spanish repositories, DIGITAL.CSIC and e-IEO, together with some promising initial results and challenges in the take-up process. The OPR module, designed for integration with DSpace repositories, enables any scholar to provide a qualitative and quantitative evaluation of any research object hosted in these repositories.

#### Adopting a Distributed Model for Data Services

<http://journal.code4lib.org/articles/12191>

Casey Gibbs, Marcos Hernandez, Pongracz Sennyey This article describes how the Saint Edward's University Library implemented a distributed model for the Institutional Repository.

Based on Cloud Based platforms and APIs, the Library has created an Institutional Repository that is scalable and modular, considerably lowering its implementation and maintenance costs, while lowering its technical complexity.

#### Developing an online platform for gamified library instruction

<http://journal.code4lib.org/articles/12122>

Jared Cowing

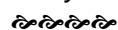
Gamification is a concept that has been catching fire for a while now in education, particularly in libraries. This article describes a pilot effort to create an online gamified platform for use in the Woodbury University Library's information literacy course. The objectives of this project were both to increase student engagement and learning, and to

serve as an opportunity for myself to further develop my web development skills. The platform was developed using the CodeIgniter web framework and consisted of several homework exercises ranging from a top-down two-dimensional library exploration game to a tutorial on cleaning up machine-generated APA citations.

This article details the project's planning and development process, the gamification concepts that helped guide the conceptualization of each exercise, reflections on the platform's implementation in four course sections, and aspirations for the future of the project. It is hoped that this article will serve as an example of the opportunities—and challenges—that await both librarians and instructors who wish to add coding to their existing skill set.

*Kellemes olvasást!*

*Király Péter*



Péter Király

software developer

GWDG, Göttingen – Europeana – eXtensible Catalog –  
The Code4Lib Journal

<http://linkedin.com/in/peterkiraly>

Válogatta: *Berke Barnabásné*