AAHSL Data Scholarship Report
Report on Research Data Access and Preservation Summit 2018

Wladimir Labeikovsky
Bioinformationist
Health Sciences Library
UC Denver | Anschutz Medical Campus
orcid.org/0000-0001-6074-3269

* Evaluation of the program

I attended the RDAP (Research Data Access and Preservation) conference in Chicago in March 2018. This annual conference is organized by the Association for Information Science and Technology (ASIS&T). The conference consists of a keynote, panels, lighting talks, group brainstorming sessions, a poster session, and an optional workshop on the final day.

This is a conference by and for practitioners: librarians, curators, and data scientists. As a non-librarian library worker, I chose this conference as an entry in to the data librarian world, its challenges, and its vision from a researcher's perspective. The conference did not disappoint in this regard.

A few themes emerged from their panels and their associated discussions:

- The range of choices and different strategies that librarians and curators can apply to offer data management support and services can be paralyzing. The community can fall into a mode of "throwing everything at the wall" or trend-following (or, as Andrew Johnson from the UC Boulder libraries described it: "Fear of missing out / Keeping up with the Joneses") without properly assessing impact and performance. A central challenge is how to focus services (e.g. focus on data reuse versus preservation) while working to build largely lacking skills among practitioners (e.g. statistics, computational, and domain expertise).

- Partnerships among and within institutions was frequently put forward as a way to compensate for shortfalls in resources and expertise. The most developed example was the Data Curation Network project, recently granted funding from the Simons foundation. On this subject, one fails to see how pooling resources can make up for the deficit if we observe that none of our institutions is really operating with excess capacity. The building up of capacity/resources ultimately must be addressed and remains a central challenge for multiple aspects of librarianship.

- A spotlight on data publication practices revealed the trend of commercial publishers advertising research data management and curation services. More than an opportunity, I think this ultimately presents a threat to the mission of librarians. At the conference we saw examples of libraries effectively outsourcing their institutional repositories to commercial services with no consideration for their specific needs. A bit of probing to representatives from the commercial entities revealed that the curation services they offer are necessarily minimal. The proper assignment of resources and responsibilities towards data curation and publishing at scale remains a dangerous blind spot in the field.

As if to preface the conference with an extreme example of data management challenges, the Chief Data Officer for the city of Chicago gave the RDAP Keynote. The thoroughness to which his office is trying to bring civic value from data management and transparency was inspiring. A common observation is how practitioners are faced with adapting and changing the practices of different constituencies (in our case: researchers and librarians).

For the last day's workshop, I chose to attend the "Building with the Carpentries session". This did not offer much new information, as I was already in the process of joining the different communication channels of the Software/Data Carpentry projects and started the instructor training process.

* Impact on job responsibilities

Aside from valuable networking and the sharing of the above impressions, lessons, and themes with my department, a few initiatives, plans, and projects have resulted directly from contacts and content acquired at this conference, including:

- A group of librarians attending the conference from the Mountain West area (roughly, Colorado, Wyoming, Kansas, Nevada) started planning a Data Librarianship unconference modeled after the MidWest Data Librarian Symposium. We successfully had the first instance in August at UC Boulder (see https://mountainwestdls.github.io) and plan to continue it as an annual or biennial meeting.

- I have joined my library's committee charged with managing our institutional repository, to draft guidelines so we can start ingesting datasets into the repository. Currently I plan to lead the curation of the submitted (or solicited) datasets from our faculty while we draft a proposal, in collaboration with other campus units, to obtain resources to scale up this effort.

- Lastly, I've started a series of workshops modeled after Library Carpentry as part of the library's internal Research Interest Group meetings in the hopes of helping and motivating our librarians to build new skills related to data management.