

September 2007

ATG Interviews David McKnight

Katina Strauch

Against the Grain, kstrauch@comcast.net

Follow this and additional works at: <https://docs.lib.purdue.edu/atg>



Part of the [Library and Information Science Commons](#)

Recommended Citation

Strauch, Katina (2007) "ATG Interviews David McKnight," *Against the Grain*: Vol. 19: Iss. 4, Article 17.

DOI: <https://doi.org/10.7771/2380-176X.5107>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

Director, Schoenberg Center For Electronic Text and Image (SCETI)
of the Van Pelt Library

by Katina Strauch (Editor, *Against the Grain*) <kstrauch@comcast.net>

ATG: Hello. We have just learned that the U. of Pennsylvania plans to begin digitizing hundreds of one-of-a-kind books from its collections. Can you tell us how many books you are planning to digitize?

DM: At the present time Penn is going to focus on its legacy collection of digital facsimiles which runs to about 350 rare titles. We will also be working on one periodical *The Occident*. Thereafter we will use **Olive** as part of future digital library collection production. The total number of volumes is difficult to assess.

ATG: Will you sell copies of these books on Amazon or Barnes and Noble or will you make them freely available?

DM: Library content will be available, generally for free, we intend to make them accessible from the Library's Website. This does not preclude other third party arrangements for sale on Amazon or print-on-demand options.

ATG: When will these items be available? Is there a Website? And when will it be live?

DM: We have just acquired the software (August 2007), we expect it to take about four months to "get to know" the system and integrate the software into our larger digitization plans for the Library. We expect to go live in the first quarter of 2008.

ATG: Did you get a grant for this project? Is it an ongoing project or a pilot?

DM: We used grant money to assist us in purchasing the software. Source: **Mr. Bernard Goldstein**. We will use a variety of sources of funding for ongoing production of texts including grants, endowments and private gifts.

ATG: How will the digitization be carried out? We understand that you will be using Olive software. Tell us more.

DM: Very briefly, the great appeal of **Olive** software is that it automates the generation of XML encoded printed works. These works can range from Newspapers (in our case historic newspapers), journals, printed books, technical reports, theses, etc. The work flow is select items for scanning, scan (in house) prepare images for ingestion into the **Olive** XML Distiller and publish to the Web. We are transforming the native **Olive** XML into METS (Mime Encoding Transmission Standard) files which is the benchmark for digital objects in the academic digital library realm. METS is a widely recognized standard and endorsed and supported by the **Library of Congress**.

ATG: Will you use student help or professionals?

DM: We have a mixed model for staffing:

work study students will be trained to use the software and professional staff will manage scripts and the technical environment.

ATG: Would you comment on how this might fit in with digitization efforts by other large ARL libraries? Also how this might fit in with Google's various projects?

DM: By acquiring **Olive** we will be able to produce fully compliant XML versions of our library content. I believe that this gives an advantage over the **Google Project** in three distinct ways: it will facilitate more granular searching, XML is recognized as an excellent candidate for long term preservation, and it will enable us to output our texts in a variety of formats from pdf, print-on-demand, ability to download handheld devices etc. Although we will not be able to compete with **Google** and their ARL partners in terms of quantity, I think we'll be able to offer quality.

ATG: Are you inputting information on OCLC for the items that you are scanning?

DM: Yes cataloging records for our digital facsimiles will be uploaded to OCLC.

ATG: Are there plans for an online "warehouse" if you will of materials like these that have been digitized? It would be wonderful to have them all in one place.

DM: Again yes, we are building a comprehensive data repository that will include all

of Penn's digital objects: texts, manuscripts, images, sound files, etc.

ATG: What advice would you give libraries that are contemplating this kind of endeavor in the future?

DM: Creating digital library programs in house is an expensive proposition. To build a robust system you need imaging resources, technical staff, programmers, system and Web designers. If you have the resources in hand or possibly you are capable of sharing resources in a collaborative environment then you might consider acquiring **Olive**. If you manage a large print collection that you would like to digitize then I believe **Olive** should be considered as the software of choice, there are few other products like it. Once the mystery of hand character recognition is solved this will be a boon to special collection libraries like **Penn's** which has thousands of hand written documents. In my life time maybe.

ATG: And, finally, tell us about yourself. How did you become interested in this project? What is your training? And what do you do for fun?

DM: I am an Academic librarian and have worked in the field of rare books and digitization for ten years. I have a Master of Arts degree. For fun I am an avid book collector. There's still nothing like a book! 🐾

FOR MORE INFORMATION CONTACT:

David McKnight, Director
Schoenberg Center For Electronic Text and Image (SCETI)
Van Pelt Library, 3420 Walnut Street, Philadelphia, PA 19104-6206
Phone: 215 746-5829 • Fax: 215 573-9079
Email: <dmcknigh@pobox.upenn.edu>
Website: <http://dewey.library.upenn.edu/sceti>

Rumors
from page 42

a cover letter, a résumé or curriculum vitae, and names and contact information for three references. All nominations, applications, and inquiries should be sent, in confidence, to: **Kit J. Nichols**, Senior Consultant <sara@auerbach-assc.com> (Electronic submissions are preferred.)
www.arl.org

And speaking of the **2007 Charleston Conference**, we will have a presentation which will describe how the **US Army Gen-**

eral Library Program supports **American soldiers in Iraq** and around the world. The program will be delivered primarily by **Karen Murphy**, the librarian who selects the books provided by the Army and other services for our soldiers. My son, **Raymond**, is an avid reader (he knows where every book is in our house and we have thousands!) and he certainly visited the library many times when he was stationed in Iraq so I'll look forward to seeing this presentation. AND — related — during **SLA's Annual Conference** in Denver, a group of professionals who work tirelessly to be sure that the men and women serving to protect

continued on page 56