“KOHA-llaboration:
Utilizing the resources of the open-source community to migrate the ANS library catalog.”

By: Elizabeth Hahn,
ANS Librarian

As the title suggests, this paper will examine the adoption of the open-source software Koha as the ILS of the ANS Library and how collaboration was essential at every step along the way. After a very brief introduction to who we are and what the collections contain, I will move into a discussion of why this project was necessary and how it was implemented, all while demonstrating the many levels of collaboration throughout the whole process.

To begin, the ANS is a non-profit, museum institution, with one of the largest collections of numismatic literature in the world, with well over 100,000 items (SLIDE). As our mission statement highlights (SLIDE), the library supports the equally extensive collections of coins, medals, and related objects that are housed in the ANS vault and come from all over the world and any time period. Like all libraries, we face similar issues of constant budget concerns and preservation of materials, all while trying to stay current with developing technologies and making our collections freely available to an ever wider general public. (SLIDE) Here you see one view of the library as you enter, and behind the glass is the climate controlled rare book room. And like any library, central to accessing the collections is the catalog, which has taken a variety of formats throughout the history of the ANS Library, which began more than 150 years ago.

Until the 1980s, items in the ANS library were cataloged with the standard card catalog, which still lives in the library conference room, as you see here (SLIDE). Although it has not been kept current for some time, it stands as a relic to the history of how far the library has come
and the importance of having such a tool. The initial move from the card catalog to the first manifestation of an electronic catalog took place from 1997-1999, when the bulk of the card catalog was migrated to electronic format. At that time, the library began using a system called LOAC (which stands for Library Ordering, Accessioning and Cataloging; essentially all the functions it was responsible for). (SLIDE) This custom piece of software was unique as it was written specifically for the ANS on a PRIME mini-computer. As the title suggests, the system was used for all library activities, and the cataloging component used Marc fields, as you see here in this example of an item record. This is an archaic system that exists as a DOS program running under Windows and as such responds only to keyboard commands. In 1997, Marc Magician was adopted as the primary cataloging software and LOAC remained in use for everything else except cataloging although it still housed the previously cataloged items. After records were cataloged in LOAC, they were then uploaded to the online catalog (SLIDE). This catalog was basically a relational database with a very simple web interface, as you can see, and worked best with simple, and often single, word searches rather than full name or title ones. It did not do well with diacritics, sorting results was not an option, and there was a delay of at least 3 weeks before newly cataloged items could be uploaded (and sometimes even months). As well, the catalog was only for the library, as there were separate databases for both the numismatic object collections and the archives.

One other issue that went along with this (and that spawned a parallel project that is still continuing), is the fact that the ANS library collections did not have call numbers or barcodes. Previously, the organization of the collections followed general subject divisions and then books were arranged within each section alphabetically by author, sometimes title, and occasionally by publisher or another distinction. The need for a new online catalog reminded us of the need to refine our classification system.
So, we have answered the “why” part of our migration project goals (SLIDE) – namely, LOAC was not efficient, not user-friendly, and had several delays in getting the information to the public and especially internally. There were other logistical issues such as backing up the data, which required physically moving the hard-drive to the vault at least once a month. So in the Fall of 2009, we began discussions about “what” our options were and “how” we would go about this. After receiving an initial quote that required at minimum $20,000 for the first year and migration process, discussions turned very quickly to how we could achieve this in-house as much as possible and what open-source options might be available. Both Koha and Evergreen were part of this discussion, and additional quotes were sought for outsourcing the IT part of installing the free software and more importantly migrating our data from a very out-dated format. We also initiated a call number and barcoding project, relying heavily on the use of interns and volunteers (which continues today).

So how exactly would we do this? As a non-profit with the library acting as only one arm of the whole Society, it was necessary to first gain the approval of our Board and draw up various quotes and estimates on time and budget required. We were encouraged to cut costs as much as possible. As a solo librarian with only a part-time cataloging assistant, I already knew that both time-wise and technically speaking, I could not do this alone. Luckily for me, my cataloger had a strong IT background and we also had at our disposal the Society’s IT associate. The strong community of Koha users was also one reason we decided to try that option, and it turned out that not only was it useful in sorting out problems, but we were able to contribute as well by reporting numerous bugs throughout the process (many of which were unique to our catalog and which have been fixed and included in the updated version of the software).

Once we had decided to try Koha, the next several months involved many hours working on migrating the data and trying to explain to our IT associate why it was important to do things
a certain way in order to adhere to library standards. More hours were spent tweaking the data, submitting bug reports, testing out the different features of the software, and coming up with a new name for the catalog (DONUM). I won’t go into much detail about all the problems we encountered, but much of it had to do with the way our data was stored, and also with the fact that we wanted to customize Koha to accommodate certain specific features, especially the indexing of articles. The summary of the reasoning behind many of the early bug reports is that the ANS library has a great many analytic records which store host publication information in the 773 field. It was important to us that the 773s showed up and were searchable, since there are too many records to add notes to each one. The ANS also had never had item records, and the precise holdings of the library were too unclear for an automated "item addition" process to be acceptable. As a result, we wanted to show locations out of the 852 field for those bib records which do not have items attached. Finally, the ANS has an extensive collection of auction catalogs, and it was important to us that we were able to search by auction date.

These major bugs were resolved and our barcoding and call number project has been slowly allowing us to add item records to the records.

In addition to difficulties with the customization, there were other problems on the most basic level, namely designing the OPAC. As some of you might know, the installation comes with a generic Koha design, which we wanted to re-design to fit more along with the ANS website (and even just adding our logo and title). This required learning basic html, which I learned to do with the help of Google.

After everything was up and running, we ultimately decided to outsource hosting and servicing our catalog to ByWater. This was actually more economical and made more sense for our own limited library staff than keeping it all in-house. I think it is worth noting as well that all the other librarians that I spoke with who use Koha have some outside support, such as LibLime
or ByWater. Having a tech staff that is familiar with library standards and expectations greatly improves the quality and workflow of using the new software. I admit too that it is nice to know that ByWater is available with customer service 24/7 should anything come up after-hours and having tech support from a company that is familiar with Koha, as ByWater intimately is, means less time that we have to spend learning and researching solutions ourselves.

Also instrumental in this migration process was the library community, both in and out of Koha. I attended the ARLIS conference for the first time last year (in Boston) and was drawn to a session about open-source, and had some very useful conversations with new colleagues who continued to assist via email and phone with questions and comments from their own experiences. (Special thanks to Francine Snyder of the Guggenheim Museum).

Although we are still very much making changes to the data, submitting bug reports, and working on the call number and barcode project, the general culmination of our efforts were rewarded on April 19, 2010, when we went live with the new version of the catalog (SLIDE). It continues to be well received, and above all, makes life for our library staff much easier (some things I didn’t mention include the various modules of Koha, such as accessioning, which are easy to use and much more efficient and effective than our previous system. More importantly, we are able to conduct better searches of the variety of languages stored in our records. Our old catalog would display a work in Arabic (transcribed) like this (SLIDE), while in the new one (SLIDE), not only was the transcription cleaner, but there was the option to include even the original Arabic script (although not yet searchable). This was done both by either downloading records via Z39.50, an option available in Koha, or adding the text manually into the 880 field (by changing the keyboard language). And another example here demonstrating the vast improvement of display from our old catalog records here (SLIDE) – and this is a work in Russian – and the new record here (SLIDE).
In sum, a daunting task from the start was made possible with collaboration at every step along the way. From the initial brainstorming and discussions among our minimum library staff, to involving other staff of the Society, as well as the ever-important Board of Directors, to finally solving problems and reporting bugs through the Koha community, we were able to transform an outdated database into a fully functioning and attractive ILS.

As librarians, our primary goal is to make our library collections, and the information contained in them, accessible to researchers in the most efficient and complete way possible. There is much to be gained by establishing a network of support and communication, and I believe that collaboration and communication of projects and information are important steps in opening this communication and ultimately making all of our lives easier.

Thank you.