Getting it all together: building an institutional repository collection of ICAHIS (International Conference of Animal Health Information Specialists) papers

Vicki F. Croft, Head, Animal Health Library, Washington State University, Pullman, WA USA

Kay E. Vyhnanek, Scholarly Communication Librarian, Washington State University, Pullman, WA USA

Abstract: Since its first meeting in 1992, the International Conference of Animal Health Information Specialists (ICAHIS) has generated a large collection of informative papers, largely inaccessible even in this age of Internet access. Development of the Washington State University (WSU) Libraries' institutional repository, the WSU Research Exchange, is a venue to provide accessibility and permanence. Working together, Vicki Croft, Head Librarian of Washington State University’s Animal Health Library, and Kay Vyhnanek, Scholarly Communication Librarian at WSU, began a project to create a Web-based collection of papers from all five ICAHIS meetings. This presentation will give descriptions of the design and planning, the process and procedures for obtaining copyright permissions from the authors, the methods employed in capturing and/or creating digital images of the papers, the creation of the metadata representing the papers, and posting the papers to the Research Exchange. The use of the collection as the basis for an ICAHIS website, the value of using the repository for this purpose, and a summary of data on access to the papers will also be presented.

Introduction:
The first International Conference of Animal Health Information Specialists, which was held in Reading, England in 1992, brought together the first time veterinary and animal health libraries worldwide. This first conference was organized by a small group of librarians, working closely with the Medical Library Association’s Veterinary Medical Libraries Section, with a grant from the U.S. National Agricultural Library, and supplemental funding from other institutions, companies, and agencies. Organizers and attendees alike were pleasantly surprised that the conference attracted eighty participants from 18 countries and 3 continents, with six invited speakers and 14 contributed papers (1).

Today’s meeting is the sixth in the series of the International Conferences of Animal Health Information Specialists, and the second to be held in conjunction with an ICML (International Congress on Medical Librarianship) conference. Following the very successful first ICAHIS, the 2nd ICAHIS was held five years later in Copenhagen, with two invited speakers and 18 contributed papers. (2) The third conference, held as part of the 8th ICML, included 11 papers. The last two conferences, held in Budapest (2003) and Onderstepoort, South Africa (2005), included 11 and 22 papers presented, respectively. (3,4) The Proceedings of the 5th conferences were published with limited distribution. (1-4) OCLC holdings indicate 61 total paper copies of Proceedings 1-5, with none for the 3rd. Unfortunately the papers of the 3rd conference were published online only and have since disappeared from the Internet. (5) The Proceedings of the 4th are also located on the Internet server for the Faculty of Veterinary Science Library at the Szent István University, Budapest, Hungary. (6)

A digitization project offering a stable, open access permanent archive would be important in preserving the content, as well as making the papers more widely available. The model of archiving grey literature, including conference proceedings, in institutional repositories is not a new one. At the 6th International Conference on Grey Literature, New York, 2004, LeFleur and
Rupp spoke about a project from Cornell to make grey literature, including conference proceedings, part of their institutional repository using DSpace. In 2004 Walters discussed Georgia Tech’s SMARTech for hosting conferences such as their School of Engineering’s Space Systems Engineering Conference. Currently more than 500 conference and workshop proceedings are part of their IR. AgEcon Search is a free-to-user subject repository containing the full text of grey literature of agricultural and applied economics that has developed over the past 15 years. It contains the full text of conference papers from more than 20 associations and other grey literature from more than 150 workshops, organizations, conferences, and the like. Plans for converting to DSpace are in progress.

The Research Exchange, WSU’s institutional repository, was chosen as the site for our papers. The Research Exchange uses the DSpace software developed by MIT and Hewlett Packard. The structure of the software allows for the development of collections within communities and provides a logical way to present the papers from the ICAHIS conferences.

Methodology:

Locating content:

The first step was to locate the papers from all of the conferences. The Proceedings of the 1st, 2nd, 4th, and 5th conferences are available at WSU in print format. In addition, Mitzi Williams of the University of Illinois was able to provide a CD of the papers from the 5th ICAHIS, published in 2007. The difficulties in obtaining the papers from the 3rd ICAHIS will be described in the discussion section.

Next challenge: obtaining permission

The co-authors agreed that permission to include a paper in the repository would be required not only from the individual authors of the papers but also the publishers of the proceedings.

A template of a permissions form was drafted for the purpose of obtaining copyright permission from the individual authors. Each form was personalized to include the author’s name, paper title, and proceedings, and sent to authors via email. Because signatures were required, forms had to be returned by PDF via email, fax, or postal mail. Email addresses from the printed proceedings and VETLIB-L were used for initial contacts. Colleagues were contacted, as needed, for bounced requests. Google and other internet searches provided leads for others, and institutional web sites were useful, too. In a separate step, publishers of the previous conferences were contacted for permissions.

Depositing the papers:

Once permissions began to be received, the work within the Research Exchange began. A collection for each specific conference was created under the umbrella community for the ICAHIS proceedings. Using capabilities of the software, some of the Dublin Core metadata elements used within DSpace were determined and entered into the submission form for each conference. This obviated the need to complete each element for each paper and helped streamline the process of deposit. Each paper was then deposited to the appropriate collection by completing the submission form with the required data elements. Subject metadata included tags assigned to individual papers based upon a thesaurus that Trenton Boyd and one of the authors created for another project.
To track each paper through the deposit process, the details of each paper were recorded in an Excel spreadsheet. Title, author or authors, contact information, date and number of conference were all tracked along with the status within the permission process to make the actual deposit easier to accomplish. Color highlighting was used as a visual cue to indicate the step in the process each paper had reached.

Repository as collection site:

Using the Research Exchange repository as a place to bring all the papers of the ICAHIS conferences together provides more than permanence to the collection. Because the DSpace software is based on the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) (15), it is using an accepted standard for preservation, storage, and access of materials. The Protocol is a standard used by repositories worldwide that promotes gathering of data from a repository by search engines such as Google (http://google.com) and Yahoo (http://yahoo.com). In addition there are specialized search engines that are set up specifically to harvest the information from digital repositories. OAIster (16) and the Registry of Open Repositories (17) are two services based on OAI-PMH that help to expose the contents of repositories like the Research Exchange.

The repository also provides a permanent URL (handle) for each of the papers deposited there. Anyone who has been presented the dreaded 404 error after entering a specific URL knows how impermanent URL’s can be. While the specific server where the papers exist may change, users of the ICAHIS proceedings collection can count on the validity of the handle that has been assigned to each paper. An additional benefit is that of migrating the format to an updated version should that become necessary, the preservation step provided by repositories.

ICAHIS Website:

The advantages and standards behind the collection make it an ideal storage place for future papers as well as the papers from the past conferences. From the complete collection, a future possibility is the creation of a specific ICAHIS Website. By linking to the papers from all the collections the specialized Website can also include other data such as background information about the conferences and the participants, and information about forthcoming activities.

Summary Data:

Another thing that repositories provided easily is the number of times papers are accessed. In the short time that the collections have been available through the Research Exchange, each paper has garnered some impressive numbers. Individual views of items in each collection have ranged from less than 20 (only one paper out of all 100+ papers!) to 158 views. Total views for the collections range from 390+ for the smallest collection to 1240+ for the largest. Total views for the community were over 4200. On an overall basis, the papers in all the collections were viewed an average of 54 times each.

These numbers help to illustrate the advantage and value of having papers posted in a repository. While many of these papers may be available on individual servers, there was no one location where they could be accessed. Additionally, some of the papers had all but disappeared. By placing them in the repository these papers will not suffer the fate of most grey literature. Instead a significant body of information is exposed and being used by others.
Discussion:

Obtaining papers from the other conferences was easy, but not those from the 3rd ICAHIS. The 8th ICML Proceedings web site (5) no longer contains either abstracts or papers for the 8th ICML or the 3rd ICAHIS. Even the program/speaker list is incomplete. C. Trenton Boyd of the University of Missouri provided a copy of the original program so names of the presenters and papers could be identified. A Google search retrieved one paper that is part of the University of Pretoria’s institutional repository. (18) In March 2008 one of the authors contacted the Chair of the 8th ICML Organizing Committee (McSean T 2008, email, 23rd March) and requested his assistance in locating the missing 8th ICML archived papers. His email response reads, “I’m afraid I don’t have an answer......, However, I’ve copied this to three people who might know more than I do so all is not yet lost. “Fortunately one of those whom Tony contacted was Robert Kiley, who replied, “To my surprise, I seem to have a set of papers for the conference on the network. From memory I had the distinctly dubious task of trying to put together the web site. This involved converting Word files to html using FrontPage express, and then FTP’ing these to the ICML server.”[Kiley RA 2008, email, 24th March]. He went on to write “PS. I note the Internet Archive has a copy of the site. Perhaps that will prove a better source: http://web.archive.org/web/20010531153110/www.icml.org/confproc/.” (19) Gleaning content from Robert’s personal archive, the Internet Archive’s Way Back machine, and the lone paper from the UP repository, the full text of 8 of the 9 presented papers was located. The lost papers of the 3rd ICAHIS illustrated what could happen if steps were not taken to preserve the ICAHIS Proceedings content.

There was no problem in contacting the editors/publishers of the 1st, 4th, and 5th Proceedings. Each of the editors provided written permission to proceed. Obtaining permission for the 2nd ICAHIS was more difficult, as the person(s) who could provide permission was not easy to locate. Johnny Carlsson, currently senior librarian of the Karolinska Institutet (Sweden), one of the Proceedings co-editors, felt that someone from the Danish Veterinary and Agricultural Library (BVFB) should provide the permission. His co-editor, for the proceedings, Anna Gjerluf Eslau, had changed jobs one month earlier and could not be located. With Johnny’s assistance, Frede Mørch current head of the former DVLG, now the Faculty of Life Sciences Library, University of Copenhagen, was contacted, and he graciously provided the requested permission. The 3rd Proceedings, which had been only “published online” as part of the 8th ICML and then lost, presented yet another problem - the ICML is not really an organization per se. Upon locating the lost e-papers, Tony McSean and Robert Kiley were approached for permission to add them to our repository. Tony responded, “Actually, it would be hard in law to define ICML as any sort of entity whatever, let alone a rights-holder, since it only exists in the same way as children’s imaginary friends.”(McSean T 2008, email, April 1st) Upon his and Robert’s advice, permission was assumed and the next step followed: obtaining copyright from the individual contributors.

Due to the span of nearly 16 years and various life-changes for authors between the first conference and this project, obtaining permissions for all papers was not possible. Obtaining permissions from the individual authors presented some problems. There were a total of 103 papers, with 176 authors, all of whom needed to be contacted for permissions. During the time since the first conference in 1992, some of the authors/presenters had passed away, retired, changed jobs, left the profession, moved to other countries, changed names, etc. For example, a letter with just a presenter’s name and city was mailed to the small Iowa hometown of a retired librarian. The retiree (who was on a float trip in Arkansas) phoned to express her regrets that she couldn’t meet our deadline, but that permission would be granted and she would send her
signed form when she returned home. In summary, most of the authors were contacted and we were very fortunate to receive permission from the vast majority of authors, although often we had to make 2nd and 3rd requests. In fact, not one negative response was received. When personal contact was not possible or completed permission forms were not returned, the papers in the ICAHIS collections were included for the sake of a complete record. The metadata for these papers includes information on how to get in touch with us in case an author has an objection to posting the paper that did not come to our attention prior to the creation of the collection. A lesson to be learned is that creating an archive of papers more than a few years old can become difficult, as contacts and personal situations change with time. The AgEcon Search model of partnering with organization/association planners in getting the manuscripts on their web site pre-conference takes care of that problem.

Conclusion:

Locating conference papers is sometimes one of the hardest things a researcher or librarian has to do. Using repositories, either institutional or subject based, is an effective means of rescuing what otherwise might be lost literature. With the completion of this project, a significant body of information, which otherwise might have gone lost, has been brought together and made available on the Internet. Usage statistics indicate that these papers are being accessed frequently and in high numbers, providing widespread exposure not possible from the limited number of print copies. In so doing, the information is not only shared among colleagues in animal health information science through presentations at ICAHIS conferences but is freely available to anyone, anywhere on the global information highway. (20)


