Digital Asset Collection Management in Support of Agriculture-Oriented Educational Activities

Brian T. Watson - Management Information Specialist, The University of Georgia
Charles T. Bargeron - Computer Services Specialist IV, The University of Georgia
Robert D. Hamilton, III – Coordinator: Office of Information Technology, The University of Georgia
G. Keith Douce – Professor: Extension Entomology, The University of Georgia
Edward A. Brown – Extension Coordinator: Plant Pathology, The University of Georgia
David J. Moorhead – Professor: Extension Forest Resources, The University of Georgia


ABSTRACT:

Almost every educator associated with agriculture, forestry, environmental science, or other similar field relies, to some extent, on visual resources to assist in the transfer of ideas. Photographic slides, for example, can be used to illustrate specific points in a classroom, during a professional presentation, in a publication, or on the Internet. Many educators have accepted and are actively making use of the new technologies that allow them to develop distributed/distance education programs, publish ideas to the World Wide Web, give a presentation using a dynamic electronic slideshow, and collaborate in real time or at near real time with their peers or students. While the technology is available and the motivation is present, one key element that is missing in the equation is ready access to resources that are already in digital format, of sufficient quality, and ready for use. This session will be devoted to the issues associated with the planning, development, implementation, and evaluation of a large-scale digital media repository specifically to support educational/outreach activities. The larger aspects of cost, potential impact, foreseen benefit, process development, access methodology, and multi-disciplinary coordination will also be considered.

INTRODUCTION:

KnowPics is a collaborative endeavor between the Internet Imaging System and the Bugwood Network, both of which are working groups within The University of Georgia's College of Agricultural and Environmental Sciences. These two groups have exemplified leadership in the evaluation and adaptation of technologies as they relate to academic, extension, and outreach activities. The Bugwood Network has produced an extensive collection of online resources for the forestry and entomology communities. To date, Bugwood's efforts have resulted in an archive of more than 6500 high-resolution digital images and a vast library of fact sheets, newsletters, and publications. The Bugwood Network is internationally recognized for both the quality of its content as well as for its technical sophistication. The Internet Imaging System is most widely recognized for its development of the Distance Diagnostics through Digital Imaging (DDDI) system. DDDI allows county extension faculty to submit electronic samples, complete with digital images, of such things as plant diseases or insects for timely identification by geographically diverse specialists. A recent survey indicates suggest that the DDDI project has saved Georgia farmers and homeowners in excess of $17 million in its first two years of implementation.

Not only do the Internet Imaging System and the Bugwood Network have common missions and goals, the respective skills of project personnel and the resources at their disposal seem to compliment each other as well. Personnel working with the Bugwood Network have strong World Wide Web development and database design skills. They also have extensive digital image acquisition and storage expertise. DDDI personnel are well versed in Web-to-database connectivity; web programming; and digital image acquisition, manipulation, and transmission. Both groups have shown interest in the development of a Web-based, large-scale digital asset archive. It was agreed that the most effective way to make such an archive a reality would be to pool the talents, experience, expertise, and resources of these two groups.
While keeping in mind that the ultimate goal of this endeavor is the construction of an archival system capable of handling a variety of digital media types and formats, it was decided that the prototype system should be geared toward digital images alone. In this way, the system's functionality and navigational framework can be designed without having to consider the special circumstances that may arise when handling other media types. Once the system has been beta-tested with its core functions, user interface, and advanced features in place, modifications can be made to accommodate the variety of file types and metadata that are anticipated.

SITUATION:

Agricultural educators, like their counterparts in other fields, are beginning to embrace technology in conveying and reinforcing their respective messages. In the academic environment, instructors are using electronic slideshows while presenting lectures, support materials are being made available on CD-ROM and on the World Wide Web, and entire courses are being migrated to the Web. Similarly, extension and research scientists are using electronic presentations to share information with their peers and their respective clientele, and Web-based collaboration tools are facilitating communication between specialists without regard to location. Extension faculty members continue to discover new benefits of the Web's rapid information distribution capability.

The acknowledgement of the effectiveness of electronic media and presentation tools by educators may have an interesting impact on the evolution of education, knowledge-sharing, and information dissemination. In a time when educators often face smaller budgets, decreased resources, and more students and clientele, the adoption of appropriate technologies presents a promising alternative for the efficient delivery and maintenance of one's programs. This movement, however, does present logistical issues for support organizations within the academic environment. In this particular case, the opportunities for utilizing electronic media rely heavily upon the ability to digitize physical media and/or have digital asset collections readily accessible.

SOLUTION:

KnowPics, as a project, can be viewed as having two separate but related components. The first component is a digitization service for converting physical slides, photographs, diagrams, etc. into electronic form. Faculty within the College of Agricultural and Environmental Sciences may have anything within their private collections scanned into electronic form as long as they a) have legal release to do so and b) are willing to provide the textual support materials of the electronic asset. The project will provide the scanning equipment, workstation, personnel, and storage space for the electronic files and the individual faculty member will retain the original in physical form.

The second component of KnowPics is a Web-based application that will provide access to the archive of electronic resources. Faculty, staff, and students will have the ability to either browse or perform a specific search for files, compare file contents, and then retrieve and download the files that they deem appropriate for their needs. Users will also be given the opportunity to create, store, annotate, and share collections of images online.

The combination of these two components results in a comprehensive system for the digitization, archival, and retrieval of digital images that can be operational at the college-level and beyond. Of course, mechanisms must be developed to provide resources to operate and expand the system.

APPROACH:

Given that this prototype (KnowPics) will be developed, at least at first, as a system for archiving and retrieving digital images, the first order of business was the determination of how and in what form these images will be delivered. As was mentioned above, the obvious delivery mechanism is the World Wide Web. By building a Web-based interface to the archive, users will be able to access the stored resources at any time from anywhere. Geographic restrictions are not imposed upon clientele using a Web-based interface. Also, access to the system is not based on specialized client software, which would have to be downloaded and installed onto every user's computer. By using common web browsers as the client, and by creating a rich html-based environment, unnecessary restrictions will not be placed upon access to KnowPics.

The issue of which image file formats to support was, in part, resolved by the choice of interface. Given that access to KnowPics will be Web-based, the most prolific image format on the web was chosen:
JPEG. It is understood that this particular format utilizes a routine that can cause the loss of detail in an image that is overly compressed. However, this format performs exceptionally well in situations where the image is viewed on-screen such as in electronic slideshows and on the Web.

In instances where the images will be used for "slick" publications and for other high-end, printed applications, a file encoded into Kodak's proven Photo CD (PCD) format will be made available. The PCD format is the choice of many photographers and graphics professionals. PCD files can be accessed in a variety of dimensions and color depths, making it very versatile, and does an excellent job of encoding and storing authentic color information. These files are typically 4 to 6 MB in size and, therefore, may take some time to transfer via the web.

By encoding the archived images into a variety of JPEG files as well as a high-quality PCD and then making all of these available through the Web interface, it is anticipated that KnowPics can provide an image appropriate for just about any need.

COST:

Three primary cost centers have been identified in implementing KnowPics: equipment, personnel, and consumables. During this project's developmental, beta-testing, and initial launch stages, these costs will be covered by existing resources contributed by the two collaborating entities and by other cooperating units and departments. As the product's life cycle matures, costs will be recovered through system membership fees and through planned publishing and duplication services. No additional funding is currently allocated to this college-level project, thus we are seeking funding to implement and operate the system at the college-level and beyond.

The initial investment in equipment for KnowPics of between $70,000 and $80,000 has been, and was expected to be, the most significant expense incurred. Two high-end scanning stations, a dedicated server, and an expandable network storage device form the core of the physical infrastructure of the system.

THE PROCESS:

The process for submitting images into the KnowPics system has, of course, evolved over time and will continue to evolve as the application is developed into a final product. The following steps outline the submission process in its current form:

1. A faculty member contributes to the system by delivering the physical materials (such as slides, negative, and prints) to a system operator.
2. The operator scans the materials and saves them in digital form. Slides can be digitized in batches of fifty, negatives by the strip or roll, and all other originals individually.
3. The operator crops the images, performs any necessary color correction, and resizes the images to a master file size of 3072 x 2048 pixels. To encode a file into the Kodak Photo CD format, these dimensions must be exact. Each image is assigned an "image number" and a visual tag containing that number is then affixed to the image. The image is then saved as an uncompressed TIFF file that is 18 MB in size using the image number as the filename.
4. Once a complete set of originals have been scanned, edited, and saved as described above, a preprogrammed script is run on the files to generate the following seven files from each TIFF image:

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>3072x2048</td>
<td>RAW</td>
</tr>
<tr>
<td>3072x2048</td>
<td>JPEG</td>
</tr>
<tr>
<td>1536x1024</td>
<td>JPEG</td>
</tr>
</tbody>
</table>
5. The seven files are saved into folders based on their respective file type and dimensions. For example, all 3072 x 2048 pixel JPEG images are stored in the same directory.

6. The operator then runs a program developed by project personnel that will convert the 3072 x 2048 RAW files into Photo CD (PCD) files and outputs them to their own folder. The TIFF and RAW files are then deleted from the system and the remaining seven JPEG files and the PCD file are migrated into a permanent storage location.

7. At this point, the operator uses a Web-based information system to log (or initialize) new database records for each of these images. Original materials are returned to the submitter.

8. E-mail is automatically generated to the person who supplied the images notifying them that the materials have been scanned. These individuals are also prompted to log into the KnowPics website to supply supporting textual information (such as scientific classification) and categorization for the images.

9. Once the supporting information is completed and the submitter acknowledges that they have the legal right to include these images into the system, the database record is marked as complete and the image and its information become available to the entire KnowPics community. It should be noted that an artificial limit of fifty scans is imposed on each submitter. Once each batch of fifty images is scanned and information is provided for those images, the submitter will be allowed to submit up to fifty more items. This limit will help to deter the use of KnowPics as just a scanning service by users who may be unwilling to provide information for the entire community.

It is also worth noting that, though there is no current process in place, project personnel also plan to develop a "plug-in" for Distance Diagnostics through Digital Imaging. Any diagnostician using a DDDI site developed by The University of Georgia will have the ability to forward images submitted through that system into KnowPics. These images will, of course, carry the same contextual information and photographer data as those that were scanned directly into the system.

**BENEFITS:**

Project personnel believe that KnowPics will offer a number of significant benefits to faculty, staff, and students. First of all, KnowPics will provide a digitizing service (at no cost) to the local academic community. It is not uncommon for faculty to use funds from their own budgets to purchase scanners and hire students and/or technicians to digitize slides and other materials. While this practice does illustrate that the interest does exist, it is really not practical from the broader perspective. In the KnowPics paradigm, the equipment is bought once, the service can be made available to anyone, and the resulting scans are stored in a centralized archive accessible by everyone.

Therefore, the second benefit resulting from the KnowPics project lies in the archive itself: a repository of [already] digitized images from various academic and scientific disciplines that are available on demand by anyone needing them. Educators, over their careers, often accumulate quite large collections of slides (not to mention other visual resources) in support of their activities. These collections, however, are not readily available to others. KnowPics will allow individuals to have materials scanned for their own purposes; but, as a matter of process, those same resources will now be available to anyone else.

Finally, as information technologies become more and more prevalent in the educational arena the business concept of sustainable organizational memory becomes a greater issue. For example, as faculty members retire, their "individual" slide collections are often dumped into boxes and stored with no documentation, given to someone who may or may not have a use for them, or are simply taken home by the retiree. In any of these cases, the usefulness of individual slides is lost, even if the slide
itself is not. KnowPics provides an avenue by which slides are collected into a master archive (in a state that will not degrade) along with that very important textual description of what the slide depicts. In addition, it is possible to collect and maintain data regarding the original photographer and/or the slide or slide collection's caretaker.

**EVALUATION:**

While the personnel associated with the KnowPics project have considerable experience with Web development and programming and are familiar with and capable of applying proper interface design and usability techniques, this in no way guarantees that the KnowPics interface and access methodology will be embraced by its users. Because of this, project personnel plan to formally evaluate the KnowPics design and then make modifications to that design as necessary.

As is commonly understood, an "information system" takes raw INPUT and PROCESSES that input into an OUTPUT.

![Input Processing Output Diagram](image)

A "dynamic information system" goes a step further in an attempt to become more efficient. The addition of a FEEDBACK LOOP to a system provides a channel through which output is evaluated and the inputs and processes are modified to make the output more appropriate for its purpose.

![Input Processing Output Feedback Diagram](image)

This system approach, when applied to KnowPics, will require that an evaluation (by user surveys, hit counts, etc.) of the product be conducted. The results of such evaluation will expose flaws in the interface design, the navigational framework, or in the deliverables (images and text) themselves. This feedback will allow programmers to modify the system to better suit the needs, desires, and even the expectations of the users.

**CONCLUSION:**

There is still much to be done to make KnowPics a reality and a success. Every indication, to date, is that this project will succeed as both an efficient image delivery system and as a prototype for a much larger digital asset delivery system. By initially concentrating on the greater issues of user interaction and product delivery; a stable, yet flexible, infrastructure can be erected capable of handling a diverse range of digital media objects. By being cognizant of client needs and preferences in addition to the technical details, KnowPics personnel can develop a final product that will not only meet their goals, but also some of those of the larger academic community.

**REFERENCES:**


