Protecting your image

Choosing the right informatics system will ensure that all of your data is in the picture.

BY MICHAEL COSGROVE

Few in the life sciences industry dispute the immense value to drug discovery and development of integrating image data with other experimental data. Interpreting image data can provide insights into biological modes of action and the behavior of functional groups for leads or drug candidates. According to one estimate, as much as 70% of all drug discovery research today generates image data.

A study might start with analysis of a tissue sample and use various techniques including microscopy. The same study might entail taking the research to a cellular level, applying antibody-based staining to look at individual cells. In conjunction with this analysis, a researcher might run 2-D gels to identify activated proteins within samples. At both the tissue and cellular levels, a researcher will develop pathology images.

All of the images will be either generated digitally or scanned so that they can later be accessed digitally. But after numerous images of different types and formats are created, the issue then becomes how to collect all this data into a single shared system where it can be easily annotated, searched, and extracted.

Images, informatics, and infrastructure

Image data presents particular challenges to database management because image files tend to be much larger than other types of files (sets of images can be gigabytes vs mega- or kilobytes) and require a software system specifically designed to quantify aspects of the image and manage files effectively. Researchers at many institutions routinely store image data on individual CDs, hard drives, or a server, only to find that files organized in folders by date are hard to retrieve later and difficult to compare with image and non-image data from other experiments. The cost of inefficient image management can be measured in lost productivity and missed opportunity.

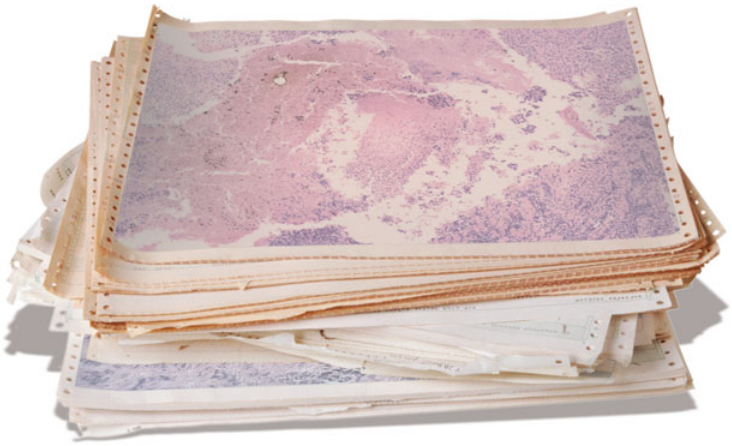


Image-handling solutions have emerged to manage the database requirements involved in the storage, management, retrieval, and analysis of image data. But these new systems aren't all alike, and companies planning to purchase a system for image informatics should evaluate the software solutions carefully with their workflow practices in mind.

Boosting productivity

To find a specific image for a researcher, imaging lab managers routinely search CDs by date, sorting through as many as 100–150 images and opening files one by one—a time-consuming practice with several drawbacks. Beyond locating the original image, what researchers really want to know is under what conditions they have seen the image before. Similarities of images across different tumor or tissue samples, compound candidates, chemical structures, and biological behaviors can lead to invaluable insights.

For instance, if a particular antibody seems useful against tumor tissue, it would be good to know what other tissue samples have been stained by this and other antibodies that might have shown similar activity. Using an image informatics system, a researcher can search through previous data on antibody stains and similar experiments in a matter of minutes.

Some software solutions collect image data but don't render it searchable. With the right image informatics system, however, researchers can set their own search parameters and rapidly browse hundreds of images.

One pathologist at a large pharmaceutical company recently explained that to compare different antibody responses, he lays out all his pictures on a table and then moves them around, sorting for differences and similarities. He worried that with a software application he wouldn't be able to physically rearrange images. But researchers need not fear losing flexibility and hands-on control of their images with a software solution. By using an innovative image navigation feature, they can easily organize images according to user-selected parameters and regroup them quickly.

For this reason, when evaluating whether an image informatics solution will help boost productivity in a lab or organization, it is important to ask several questions about the current lab setup:

- How long does it typically take to find a single image or a set of similar images?
- How often is an important experimental analysis found only after tedious searching in notebooks—or not found because manual searching is too laborious?
- Can you interrogate image data by asking where you have seen the same image before?

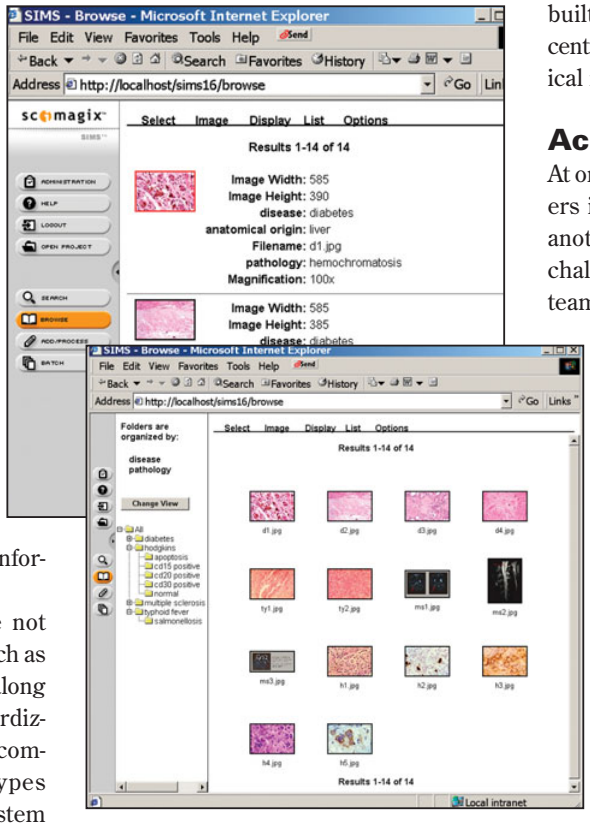
PHOTO: PHOTODISC

- How often do you replicate experiments because you can't quickly locate a record of original work?
- How many hours do you spend annotating data by hand when software tools could help automate this process?

Data set diversity

In most cases, image data is just one piece of the puzzle; researchers don't often make decisions until they consider all the relevant data sets. Increasingly, the number of different software platforms and applications—each generating different data types—is causing significant integration challenges. Here again, an image informatics solution can help.

Scientists can scan data that are not conventionally considered images, such as chromatography traces, and enter it along with image data, effectively standardizing different data types and making comparisons across multiple data types possible. An image informatics system built on an open architecture platform centralizes data for review and solves a critical integration problem.



Across space and time

At one pharmaceutical company, researchers in the toxicology lab commented on another kind of integration hurdle: the challenge of collaborating with research teams across multiple geographic locations. Part of a "global toxicology group" spanning four states and two countries, their challenge is to turn the idea of a global group into a reality. These days, collaboration is increasingly global and is becoming more prevalent across different divisions within the company. The toxicology department is talking to the pathology department more than it ever did before. With a Web-based image informatics solution, a researcher in Camden, NJ, can talk to a fellow researcher in Santiago, Chile, while each of them looks at the same data.

It still happens much too frequently that researchers within the same company aren't aware of related research already performed by others in the company. A central image informatics system loaded with image data from the entire enterprise makes it easy to see that someone has already evaluated a particular compound under various conditions or the toxicity of a certain drug candidate.

Marshaling resources

Eliminating compounds early from the drug discovery process and moving on to more promising targets can save enormous amounts of time and resources. Effective image data management can help this process.

Researchers at Pfizer Global R&D (Ann Arbor, MI) noticed a similarity between the protein expression profile of a new drug candidate and that of a compound previously tested and found to be toxic. This correlation raised concerns about the safety of the candidate drug. When tested, the new compound also proved to be toxic. As a result of using visual data, the research project was immediately refocused, saving money, time, and resources.

Because image data often needs to be modified from its original state, it presents special challenges for compliance with GXP and 21 *CFR* Part 11 requirements. A system that allows for image modification while retaining the original image requires unique software capabilities in a regulated environment. Life science companies depend on the software system they have selected to ensure that electronic records and signatures submitted to the FDA are compliant with regulations.

Although some of the GXP and many of the 21 *CFR* Part 11 regulations can be met through software use, no software package alone can completely satisfy all regulatory requirements. An image informatics software provider should offer guidance about regulatory compliance for image data to pharmaceutical and biotech companies and should actively partner with companies to help define work practices and software use that together enable a compliant system.

The bottom line

Drug discovery and development efforts can receive a significant boost from implementing an image informatics solution that centralizes image data on an enterprise-wide open-architecture platform. Such a system allows researchers to integrate image data with non-image data and can significantly boost drug discovery and development efforts. With the ability to search images according to the parameters they choose, researchers can save enormous amounts of time and realize real productivity gains. The ability to interrogate image data makes it possible to evaluate the effects of the same or similar compounds or biological processes across different experiments, tissue types, and/or therapies. An image informatics solution that allows researchers to "see" their data leads to new insights that can advance research at every stage.

Michael Cosgrove is director of enterprise systems at Scimagix, Inc. (San Mateo, CA). Send your comments or questions about this article to mdd@acs.org or to the Editorial Office address on page 3. ■

Image informatics checklist

Consider the following features when evaluating a system:

- ☐ The system was designed by life science and image management experts to improve existing workflow practices, especially design support.
- ☐ The database is searchable by both content and context.
- ☐ It is a Web-based, enterprise-wide system for "anytime, anywhere" access.
- ☐ It uses an open architecture that is fully interoperable with legacy and existing applications from multiple vendors.
- ☐ It accepts images of all types.
- ☐ It is secure.
- ☐ It is a flexible program with parameters that researchers can configure.