

Tomorrow's Successful Research Organizations Face a Critical Challenge

R&D collaboration is becoming the difference between success and failure

BY SCOTT DEUTSCH

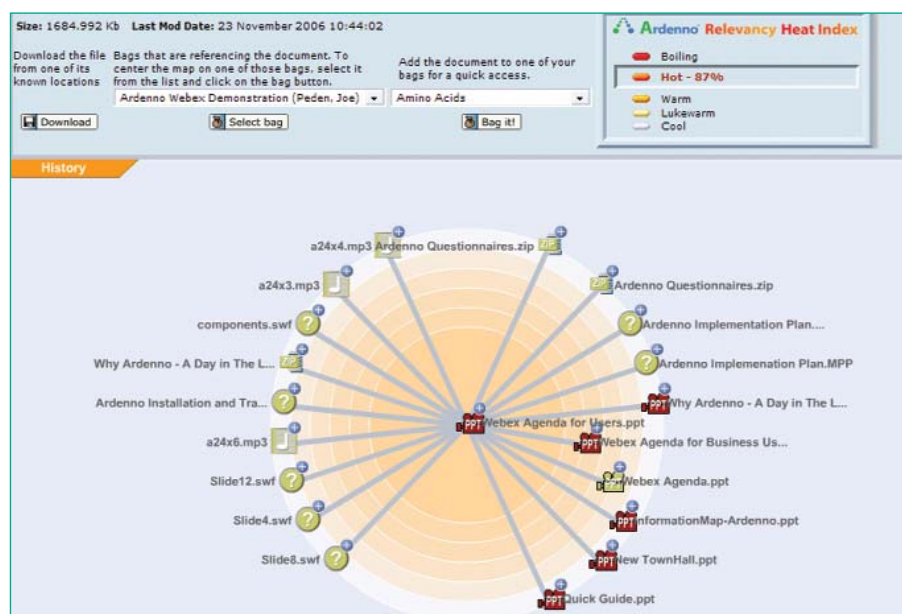
Market research studies have found that, of a researcher's time, approximately 31 percent is spent communicating with others, 24 percent is wasted waiting for decisions to be made or for something to happen, and 34 percent is spent performing administrative tasks — leaving just 11 percent of a researcher's time to devote to actual research. Further studies have found that over 80 percent of this research time (11 percent) is spent doing work that is the same or very similar to work previously done, and that actually only 2.2 percent of a researcher's time is spent performing novel research. Obviously, improving the productivity of research is a critical challenge.

Research organizations are continuing to build complex systems with the hope of a productivity return. Unfortunately, these systems are not oriented toward helping the researcher. Research is a disciplined world made up of creative individuals who collaborate with others to solve problems that can lead to sizable rewards. The problem is that document management systems, portal/intranets and custom-developed database solutions don't take into account how a researcher really works day-to-day.

One reason why these individuals are unable to collaborate better and smarter is that they simply cannot find all of the available and relevant information they need to make better use of their research time.

Inefficiencies occur for many reasons:

- ▶ inadequate data management
- ▶ poor access to information
- ▶ inadequate collaboration at all levels, from project teams to corporate
- ▶ minimum control of unstruc-



THE ABILITY to present various information sources with a multi-dimensional visual representation provides researchers with a single view into all available research content, virtually eliminating the need to search through various subfolders for available content.

- ▶ tured data
- ▶ a lack of metrics to manage laboratories and projects
- ▶ storing data on researchers' workstations
- ▶ information and knowledge loss due to staff changes
- ▶ inflexible database systems
- ▶ proprietary file formats

Given that computing power and software sophistication have increased dramatically in recent years, let's look at why these issues still exist.

Unstructured data challenges

Many database systems have been deployed that have helped to manage structured data, but they have consummately failed to manage unstructured data. Since, in the world of R&D, unstructured data makes up approximately 75 percent of all data, this presents an enormous problem. Historically,

approaches to dealing with unstructured data have included the following:

▶ Archival and data management systems have tried to convert all of the disparate research data into a single unified proprietary format, with an eye on protecting data for years to come. However, these solutions did not fit the creative and collaborative needs of the research community. Archival systems and processes are not geared toward actually helping the laboratory perform better or toward maximizing the efficiency of research. They merely ensure that the data is secure. As a result, these initiatives were doomed to fail before they ever began.

▶ Many companies have tried using corporate structured portal solutions and their complex document management systems. Unfortunately, the research community doesn't fit a solution whose focus is document manage-

ment. A portal/intranet environment that also needs to support the needs of sales, marketing and human resources doesn't offer a good solution.

► Electronic lab notebooks, although interesting to some, have generally not performed as expected due to their extreme structure. They also add an additional administrative layer that contributes to the researchers' already overburdened workload while providing minimal added value.

► Search engines can help to locate information, but can do little more. They are not research collaboration tools, and this limited functionality means that users spend a lot of time finding and organizing. Although cheap and easy to deploy, most leading solutions offer little or no security features. Some even publish details, such as search configurations, pages viewed and Web sites accessed by you and your users, to third-party databases. This clearly is not desirable in the world of intellectual property. According to Google, poor searching costs a company at multiple levels through time wasted tweaking and weighting documents to satisfy requirements. A recent study by the Nielsen Group measuring users' portal and search performance found, on average, almost \$1,500 of time wasted per user.

Increasing productivity

Research organizations are realizing that, in order for their teams to work collectively and not simply as groups of individuals, they must create an automated knowledge management environment that fosters sharing and collaboration without adding any additional administrative overhead to the individual researcher. Given the vast array and magnitude of available data, collaboration is becoming the difference between success and failure in beating the competition to market. Solving this key problem starts by looking at fundamental business processes that hinder effective research. The root causes of the problems must be addressed.

Let's take a look at some key individ-

ual components that can increase research productivity.

► Minimize duplicate research effort

The amount of redundant effort that actually occurs surprises virtually every company that chooses to address the issue. It is a latent problem that could be greatly reduced with little effort.

Improving awareness of similar research efforts across the entire organization (including multi-locations) enables companies to rapidly attack the problem of redundant research. Metadata tagging content in an automated fashion is a key enabling technology.

► Eliminate "my data" information silos

Conventional strategies for storing R&D data have forced investigators to operate in a "my data" environment. Files are usually saved on the researcher's desktop or stuck in a sub-directory on a server, where no one can find them except the author. This business challenge can be solved by securely storing newly created files, making native collaborative file copies available in a separate location for shared access.

The importance of physical proximity is forcing this issue to greatly diminish. After all, if you can recruit or retain a key employee without physical boundaries, your resource talent pool grows exponentially. Who really cares if they live in Beijing, San Paulo, Hyderabad, San Jose, Manchester or Tokyo? This business approach is not feasible without a strong, collaborative knowledge management framework for the researcher. As the workforce undergoes widespread change, the ability to provide a constructive collaborative working environment is becoming a workplace requirement. As broadband and wireless networks continue to grow and proliferate, the workplace will reshape itself rapidly. Increasingly, these changes will enable people to customize their workplace, information sources, tools, learning options and community networks. It is vital that organizations solve the "my data" problem now.

► Improve team collaboration

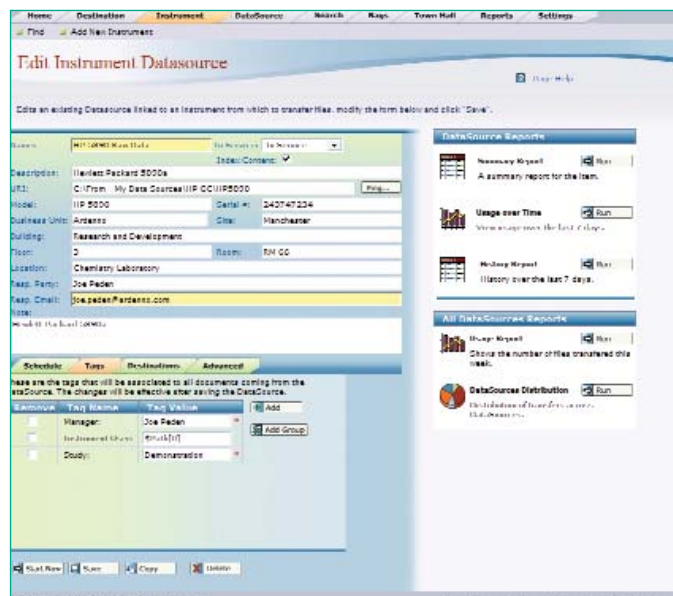
One of the biggest issues facing

research teams is the loss of key members and the addition of new members. Teams lose momentum when senior staff members transition off a project and are required to support new team members during their ramp-up period. One company stated that, "When a key researcher left in the middle of the project, the new researcher had almost a whole year getting up-to-speed." Companies can solve this business challenge by providing an active and accessible data repository of structured and unstructured data files that enables new members to rapidly ramp up.

► Automate metadata and ontologies

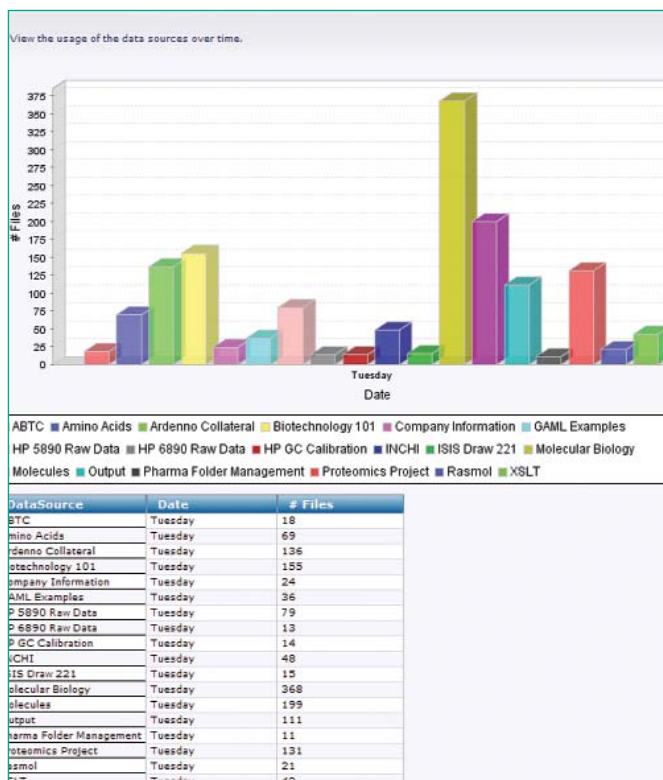
Metadata is information about information. More precisely, it's structured information about a data file. The addition of metadata generally uses a controlled vocabulary, an ontology that provides context for a document or a file, so it provides more scope to locate information with the best possible recall. The promise of metadata is to enable semantic searching. Until recently, there has not been an easy way to add the necessary metadata tags to data and files. The problem has been that previous approaches required the end-user to apply tags manually. This has been the main reason why simple tagging systems have failed to take root within the corporate environment. In addition to automated tagging, it is important to allow users to add their own unique meta tags and notes on-the-fly to increase overall semantic searching capabilities. Allowing people to find information the way they want it, rather than the way structured applications manage it, is the key to empowering semantic searching and knowledge filtering.

Another key approach to applying metadata tags is through the use of taxonomies or ontologies. Ontologies classify information into logical multiple tiers and categories. They supercharge semantic search and retrieval capabilities. Unlike search technologies alone, ontologies reveal the overall structure of a knowledgebase in a hierarchy that adds tremendous relevancy and visibili-



ABOVE: Adding metadata to all datasources and instruments provides a superior framework for semantic searching and data retrieval and eliminates the need for individual researchers to add their own metadata to each file.

RIGHT: Having the ability to quantitatively measure instrument utilization and data output metrics helps improve overall research management capabilities.



ty to the user community. The user navigates through sub-categories to narrow the search, a process that helps to avoid false hits outside the area of interest.

▶ **Semantic search**

It is vital to have extensive enterprise searching capabilities, and it's best if you can use a scalable search engine with tools and configuration flexibility to meet the needs of research and development. This search engine should be able to exploit metadata, ontology tags and indexing capabilities to enable advanced semantic searching, which is needed to support the specific requirements of the research community. Researchers should be able to perform the most sophisticated searches in the context of their own work and projects. Users also can use the metadata tags to enable more creative semantic searching. As a researcher begins typing a search phrase into a tag field, the search engine displays the possible matching tag label and values using its auto-complete features. It's important to note that you should not need to know or learn a query language such as SQL!

▶ **Automate data source management**

The term "data sources" is used to describe all sources of laboratory data and files. Normally, these are folders stored on the company network, as well as in local folders including numerous "my data" repositories. The polling of these datasources should be automated for any new or changed data and files. When new files are located, the new and changed files should be indexed and automatically moved to designated destination(s).

▶ **Improve equipment and resource utilization metrics reporting**

Research projects are often difficult to resource cost-effectively because there is minimal utilization metric reporting available. Research project bottlenecks are mostly overcome by over-buying resource-constraining instruments. This business challenge can be addressed by providing integrated summaries and detailed resource usage reporting metrics by project, study, experiment, instrument, datasource, and so forth. Managers are enabled to better justify equipment purchases and to minimize

project bottlenecks due to constrained staffing or financial resources.

Summary

Gartner said it best when talking about tomorrow's successful companies: "Tomorrow's knowledge worker will be the intellectual driver for successful companies, empowered by individualized tools, knowledge, informational sources, social networks and employment styles."

R&D management has the responsibility to empower their teams to support their employment styles. Researchers are looking for and need research-oriented solutions that add value to their day and minimize administrative workload. This is not a database problem, or a problem that databases or document management systems can solve. Success starts with improving the accessibility and data relevancy of structured and unstructured data and opening the door to data sharing and research collaboration. **SC**

Scott Deutsch is VP of Marketing and Business Development at Ardenno. He may be reached at editor@ScientificComputing.com.