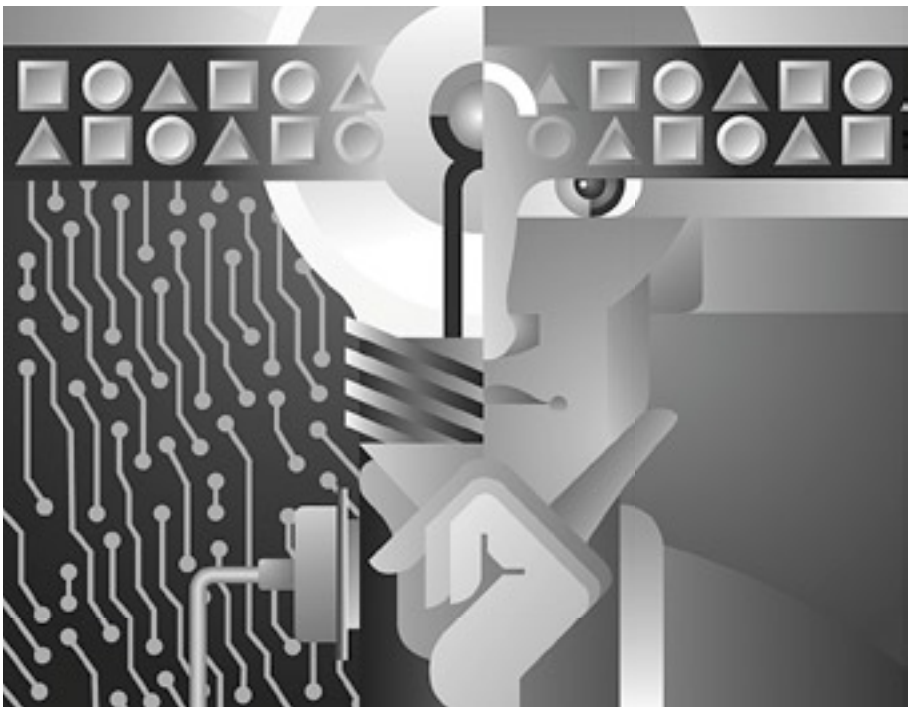
Business Intelligence: The Smart Way to Track Academic Collections

Business intelligence can expand access to collections, improve their management, and provide easy generation of database reports

By **Debra Kelly**



University collections are a vital source of knowledge for researchers and often a source of pride for universities. Without an effective way to manage and query them, however, collections often go underutilized. Parts of a collection can remain untapped for years, and the larger it grows, the more difficult management becomes. Unfortunately, improving collections management usually comes at a cost—at a time when budget cuts have forced most departments to reduce spending, any plan for improving management must include how those improvements can increase efficiency and drive down costs in the long run.

At the University of California, Berkeley, in an effort to ensure that the academic community obtained the greatest value from our collections, the Museum Informatics Project (MIP)—a department of Information Systems and Technology (IST)—launched a business intelligence initiative designed to provide collections managers with the ability to improve collection tracking, make ad hoc queries to their collection databases, and write their own reports without MIP intervention. Such a system would enable collection managers to dramatically improve efficiency while eliminating a costly burden for MIP.

Finding a Solution

During the 1990s, MIP developed the databases and management systems for several collections. At that time, accessing the data was very cumbersome. For each new report requested, MIP staff had to design, code, and test a new computer program to extract and format the data—a time-consuming process that often resulted in long delays between the request for a report and its delivery. Frequently, problems arose after delivery with report details or formatting, further delaying access to the desired information. Collections managers and staff lacked tools for accessing information on their own. As a result, researchers could not get timely answers to their questions.

In 1994, I was asked to evaluate and recommend a software solution offering some form of nontechnical collections management reporting capability. The goal was to enable users to write their own reports and do their own day-to-day ad hoc queries to get the answers they needed immediately. Such a system would have the added benefits of enabling users to ask very specific or unusual questions they would never have brought to MIP and freeing MIP staff to work on other complex technical projects for the collections.

The key requirements I identified at that time were similar to what an institution might require today:

- Ability to run under multiple operating systems on multiple platforms
- Ability to access data from several sources, including the most popular database management systems

- Ability to produce complex reports using compound queries
- Easily used with minimal staff training
- Knowledge of SQL or the structure of the database not required
- Ability to adjust and format individual report components, such as font, point size, location of columns and rows, contents of headers and footers
- Ability to store queries, report parameters and formats for easy recall
- Ability to export report results to other software, such as Microsoft Excel
- Ability to include user-input runtime parameters
- Minimal support and intervention by MIP
- Reasonable costs and licensing requirements

I examined 30 software packages, ranging from general reporting solutions such as industry leader Crystal Reports (from Crystal Decisions at the time), which offered far more capabilities than were needed, to very specific collections-management packages, which unfortunately would not have integrated with our existing databases. Other vendor packages were available for only one platform or required end users to understand the internal complexities of the database structure.

Business intelligence solutions, a new breed of software at the time, offered a different approach, simplifying access to multiple existing databases through natural language queries (see the sidebar on business intelligence tools). Of the business intelligence software packages I researched, I selected six for in-depth evaluation and testing because only they were available for platforms other than Windows. Some of the issues I weighed in my evaluation included the ease of installation, the robustness of the software, the quality of technical support and documentation, the size and number of installations licensing the software, the length of time the vendor had been in business, ease of use for both the manager and end-user modules, cost, and, most important, the opinions of the collection managers, who got a demonstration of each of the six products.

After careful consideration, I recommended we license BusinessObjects from Business Objects (then a relatively new company, it acquired Crystal Decisions in 2003), for a variety of reasons:

- It installed easily.
- The product functions were accessible through point-and-click pull-down menus or icons.
- I learned to create reports easily with a minimal amount of help from technical support.

At the time, the full client version (report processing occurred on the user's computer, with database access via a database server) was available for Windows-based PCs, Apple MacIntoshes, and Sun Solaris UNIX workstations. Only one of the other products I evaluated worked on all three platforms, and it could not generate complexly formatted reports. With BusinessObjects I could easily replicate one collection's report that had previously required an SQL script of several statements, export the results of the first script to another software product, and use a second script to format the results. I was unable to produce similar results in any of the other software packages I evaluated.

The manager's modules hid the database structure from the users. Several

Business Intelligence Tools

Also known as Decision Support Systems (DSS) or Executive Information Systems (EIS), business intelligence makes large amounts of raw data more accessible, understandable, and useful by providing computer-based tools for users to directly process, organize, manipulate, integrate, and analyze that data. Business intelligence tools help users turn basic data into information and knowledge.

Dashboards: Software that provides real-time digital visual indicators of how well predetermined aspects of an organization are working. Think of the monitors on the dashboard of a car.

Data Mart: A subset of a data warehouse that focuses on a particular aspect of an organization's activities.

Data Mining: Software that analyzes data looking for patterns and relationships that can be used to predict future trends.

Data Warehouse: A comprehensive database containing information that has been extracted, cleaned up, filtered, reorganized, and integrated from several electronic sources of data. At Berkeley, the Natural History Museums have contributed extracted data, formatted according to the Darwin Core data standard, to an online data warehouse that lets users query the holdings of all the museums through a portal.

ETL: Extract, transform, and load—shorthand for the software and processes needed to find, cleanse, and process data into a data warehouse, data mart, or other integrated database or system.

OLAP: On Line Analysis Processing—tools that allow for multidimensional analysts of data.

Portal: A Web site that provides access to a structured set of online resources, such as a search engine, a news service, a company home page, or other online service that a user wants to have access to on a daily basis.

Scorecards: Software that provides visual digital measurements of the factors identified by an organization as critical to its success.

of the other products required users to understand the database structure and to know SQL and how to join database tables to retrieve the requested data. BusinessObjects included access to the most popular database management systems, and its price was midrange for comparable products.

MIP licensed the two manager's modules, Supervisor and Designer, as well as several end-user Reporter modules. The Supervisor module handles access and security down to the value level, with the ability to group users by department and role and to restrict access to some of the functions of the reporting product. A user might be able to refresh the data in an existing report, for example, but not to create new reports. The supervisor also manages a database repository behind the scenes that holds security information about the users, the reporting databases, and the documents and reports created by the users. If needed, a user can be assigned a limited supervisor role in order to handle security for a single collection.

The Designer module creates universes used by the reporting software to generate SQL statements needed to retrieve the data from the collections' databases. A universe—the equivalent of a database—is used to model the tables, columns, and join structures necessary to retrieve data. The universe contains definitions for classes, the equivalent of tables, and for objects, the equivalent of columns, which must be configured before users can query the database. An object can consist of concatenated columns, and a class can include more than one table. These classes and objects can be named to reflect the terminology understood by the collection staff, rather than technical names used in the database.

The interface uses drag-and-drop functions to define the universe components and the necessary joins in the database. If there is more than one path to a particular table, the Designer has several functions that can be used to resolve the ambiguity for the end user. All these definitions can be stored in the repository set up by the Supervisor. Finally, the software will automatically download the latest version of a universe to the end user's computer when the user first accesses the universe after a change.

The Reporter module is used to create and submit queries to the database and then format the results. Report formatting is extremely flexible, with the ability to modify fonts, point sizes, column and row positions, report layouts, line breaking, sorting, and headers and footers. Users can conditionally hide data, generate charts, fold report details so that only the summary data shows, insert pictures and bitmaps into reports, wrap text, and create crosstab reports. The query mechanism is point-and-click and drag-and-drop, with the ability to use subqueries, correlated subqueries, union, minus, or intersect queries; to drill down through data hierarchies; to set up runtime prompts; and to combine data from more than one data source in a single report.

The drawbacks of BusinessObjects were that some of the other products I evaluated had more functionality and better documentation. Some instructions in the user manuals were not detailed enough, while other sections seemed to be translated badly into English. Ten years later, however, the missing functionality is there, and the documentation is more detailed and clearly written.

Over the years, Business Objects dropped full client Reporter support for the Mac and UNIX platforms but developed an Internet server–based product called WebIntelligence that can be run using a Web browser on any of the three platforms. Currently, the Supervisor and Designer modules only work on Windows-based platforms. Other products include Broadcast Agent, a batch scheduler; Publisher, a report bursting application; Auditor, an auditing product; Data Integrator, an ETL (extract, transform, and load—see the sidebar) product; access to online analysis processing (OLAP); BusinessQuery, a product that facilitates queries directly from and into Excel; and Application Foundation, a dashboard and scorecard generator.

The cost of running BusinessObjects includes the initial purchase of licenses for the manager's modules and a license for each named user of the Reporter module. Annual maintenance and support costs are 15 to 20 percent of license fees, depending on the level of support needed. Several levels of support are available, from Web-based only to 24 × 7 phone access and on-site assistance. In-house support is handled by one supervisor and two designers working with five universes. In 1994, we negotiated an educational discount to lower the cost of our initial installation. MIP also paid for training for a single supervisor and designer. Maintenance and support fees for 2005 are $8,100 for intermediate level support for 1 Supervisor, 2 Designers, WebIntelligence, Broadcast Agent, 16 Reporter licenses, and 1 BusinessQuery license. Initial licensing fees are usually negotiable depending on the size and business of the purchasing organization.

MIP implemented BusinessObjects, creating the initial reports to replace the reports users were accustomed to receiving, and also provided training and support. We were surprised, however, by the relatively low training requirement. Even the least computer-literate staff members found it easy to access standard reports, and those who needed to write reports or make ad hoc queries learned to use the system quickly. In one case, the History of Art Department collections curator learned to write her own reports with just an hour's experience with self-paced training software.

The Applications

Over the years, MIP has implemented business intelligence for three collections that are independent of the library system: the History of Art Department's Visual Resources Collection, the Botanical Garden's live plant collection, and the Berkeley Language Center's audio recordings collection. Each implementation is different and reveals the different ways collection managers can use business intelligence to obtain practical, long-term benefits, from saving time to justifying staffing.

The History of Art Department Visual Resources Collection

The History of Art Department collection currently includes approximately 380,000 slides and 280,000 photographs of works of art from all over the world, dating from prehistoric times to the present. Prior to implementing the current Sybase database and collection management system, and BusinessObjects-based query and reporting system five years ago, data were stored in a simple FoxPro database developed by the collection curator, with limited reporting capabilities. Today, items are catalogued directly into the collections management system, called seurat. Digital images are then created and linked to the catalog data by storing the URL and file system path for the image in the database. As the staff re-catalogues older slides, a migration utility transfers the data from the old FoxPro system format to the new seurat system structure. Eventually, most of the slides will be digitized and linked to the catalog data. Today the seurat database contains more than 35,500 accession records documenting 18,000 works of art.

The system also stores the orders placed by faculty to add new slides and photographs to the permanent collection. Seurat automatically generates slide and photograph labels and includes the ability to adjust the amount and types of data, the format of each label line, and the location of the symbol that indicates the orientation of the slide when projected. A barcode is added to track circulation of collection items. The database contains profiles of personnel authorized to access the collection, links to art history terminology authorities, links between courses and lectures and the specific images used by the faculty in teaching, information about the sources for collection images, and data about the methods and equipment used to generate digital images of the collection slides.

Today, approximately 15,000 to 20,000 collection images are circulated each semester, some of which are used for teaching. Until recently we expected to use the barcodes on the slide labels to track circulation and BusinessObjects to report on the circulation data. As the History of Art faculty convert to teaching courses using only digital images, reporting on the location of the physical slides will become less important, and BusinessObjects will probably be used to track the use of digital images by course and faculty member.

BusinessObjects provides self-service reporting capabilities, and several staff members have learned to create their own reports, perform day-to-day ad hoc queries, and refresh standard reports, such as the list of images to be photographed each day. The business intelligence system also allows staff to analyze information in an unlimited number of ways. They can run a list of faculty and graduate students who have borrower privileges; check slide or photograph orders by faculty name, status, or due date; check that all the required data about a work of art has been entered into the database for a set of images; and even check the accuracy of the data entered by the staff. When faculty members order duplicates of collection slides for their personal collections or purchase other supplies, the curator can generate invoices so that the collection can be reimbursed for staff time and resources.

The staff also generate a variety of reports, including summaries of charges by semester and faculty member and lists of call numbers by order number, title, and artist name. They can also print the drawer-front labels and interfile shelf-list cards used to identify the contents of slide cabinet drawers.

On the administrative side, BusinessObjects is used to produce an annual productivity report that tracks, for example, the number of orders per year per staff member and how many orders were completed versus outstanding. The curator also uses BusinessObjects reports to justify funding for her unit and to inform the productivity portion of staff performance reviews.

Botanical Garden Plant Collection

The Botanical Garden's plant collection dates back to the 19th century. The SAGE (System for Accessing Garden Entries) database and collection management system, which were created to replace a legacy system that was not Y2K compliant, now includes 37,000 records of plants that have been in the garden over the years. More than 19,000 are alive today.

The primary goals of this business intelligence implementation were to improve overall management of the collection and answer research questions faster. Curators use SAGE to enter data about the garden's holdings. BusinessObjects is used to report on what plants are used for particular academic courses and research projects; what plants are being propagated from existing stock and whether they are being propagated for sale, use in a course, or research; and which plants in the garden have died, enabling staff to maintain up-to-date lists of the contents of each garden bed. In addition, staff can quickly match collectors to the plants they have collected and the locations the plants came from, as well as both the taxonomic and common names of each specimen. Finally, the curator can determine which plants from a garden bed have had plant material sent to the campus herbaria (repositories for dried plant material).

On the administrative side, BusinessObjects helps the manager of the horticultural staff monitor staff performance. With BusinessObjects, he can track staff activities in a timely manner. Reports listing the living plants in the Botanical Garden Plant Collection are distributed to the University and Jepson Herbaria and the bio-collector's laboratory. Additionally, reports are generated that list which plants have been moved or have had different taxonomic names assigned to them and need new signs engraved to identify the plant in its current location in the garden.

Berkeley Language Center Collection

The UC Berkeley Language Center (BLC) maintains a collection of linguistics books, historical recordings for researchers, and practice recordings for students learning a new language. BLC contains recordings of Native American languages with few or no living speak-

ers, as well as other rare recordings. The center's entire collection is currently maintained in a Sybase database accessed by BusinessObjects. The BLC's current holdings include approximately 4,000 items, a little over half of which are recorded series totaling 14,000 individual segments.

The primary goal of this implementation was to improve how the staff tracked holdings and circulation. Until recently, BusinessObjects was used to monitor acquisitions; cataloging and circulation of collection items, including tracking checked-out, overdue, and lost items; and any fines levied. Reports tracked usage by series, language, project, individual, or department. The system also automatically tracked which books or recordings could not be checked out, preventing rare or fragile items from leaving the center.

BLC is currently converting all physical recordings to digital recordings and making them available online. This will eliminate the need for circulation reports. BLC is using BusinessObjects to track the conversion and plans to use the system to track circulation of print items and usage of nondigital media.

On the administrative side, the center can produce reports on the use of BLC rooms by course, language, and department. The system also tracks how grant funds are utilized by monitoring the hours spent on a project by language and by individual.

The Benefits of Business Intelligence

Our three business intelligence implementations share many benefits. The most obvious is that business intelligence saves time. Reports that used to take hours or days to produce now take minutes or seconds. In the History of Art Department, for example, an annual productivity report that took an entire week to compile manually now takes about two hours—and the report regularly demonstrates how business intelligence has increased overall staff productivity and accuracy.

With business intelligence, both staff and researchers can ask questions that were impractical before. For example, when a concern about Sudden Oak Death arose, the Botanical Garden was able to immediately determine which species in the garden were susceptible. Without business intelligence, obtaining this information would have required a lengthy manual process or making a request of MIP.

Most important, users like using the business intelligence system. They rely on it for innumerable daily activities and are constantly looking for new ways to use it to save time and manage their collections.

Beyond Collections Management

Beyond the benefits for collections administrators and researchers, business intelligence has the potential to benefit both the academic community and the public.

The History of Art Department, for example, is planning to use BusinessObjects to download information about works of art mentioned in course lectures into instructor and course Web sites. Once instructors request information, collection staff will run a BusinessObjects report, export the formatted data into a Microsoft Excel spreadsheet, and then upload it to a digital asset management application, Portfolio. Students can then use the course Web sites to study for exams instead of making a trip to the university library to study posted photographs. The need for digital posting of works of art for courses has become critical as the space available for posting photographs for study has decreased.

MIP is also in the process of deploying Business Objects WebIntelligence, part of the Business Objects business intelligence suite. WebIntelligence enables users to access reports and make ad hoc queries using a standard Web browser. Along with providing direct access for appropriate staff and researchers using any browser-equipped computer on or off campus, WebIntelligence will enable other campus collections that use Apple and UNIX-based computers to explore initiatives for making information available to campus researchers and students and to the public. We will be implementing WebIntelligence before the end of 2005. Reports can then be stored as HTML or PDF and posted on a campus Web site for public access or e-mailed to researchers and users who do not have access to Business Objects software.

The Keys to Implementing Business Intelligence

UC Berkeley has spent more than 10 years developing and refining the business intelligence implementations for our collections. Today, however, business intelligence is a standard enterprise application, and many academic institutions, including Berkeley, are using business intelligence for everything from admissions (Florida State University, http://fsu.edu/) to finance and human resources (Notre Dame, http://nd.edu/) to student performance (University of Texas at San Antonio, http://www.utsa.edu/). MIP has learned that business intelligence may not be the right solution for every collection, but it can be an excellent solution for many. Institutions considering implementing business intelligence for their collections can achieve benefits similar to those we now enjoy; however, several factors will determine how easily and quickly those benefits can be realized.

- *Understand the need.* Most collections can benefit from business intelligence if an adequate data management system is already in place. If there is a very large collection or multiple collections to manage, or if collections are underutilized because accessing information is difficult or staff members struggle to keep up with information requests, then business intelligence can improve access and efficiency. It is important to recognize that business intelligence is a software system that requires an existing database. If the information is not already stored in a database, then a database or data warehouse project must precede or coincide with the business intelligence project.

- *Make sure you have a willing information technology department.* In many cases, an information technology department

will support or even sponsor a self-service business intelligence project because of the desire to reduce the report writing burden on IT, allowing it to stay focused on its core mission. However, IT must be willing to develop and administer the system and provide any required training and support. Without this, the collection departments could face very expensive consulting fees.

- *Line up support from above.* Support from university officials should include more than just money. Implementation takes time, and the success of the project often depends on the willingness of the IT department, collection staff, and the departments to change processes and behavior. With active support from above, it is much easier to get universal buy-in.

- *Obtain adequate funding.* Funding is almost always an issue for collections management, and this has been true at UC Berkeley, where it has been a barrier to rolling out BusinessObjects to additional university collections. As with any IT project, the key to obtaining adequate funding for implementation and maintenance is comprehensive and ongoing analyses of total cost of ownership (TCO) and return on investment (ROI). When these take into account both the direct and indirect benefits of a business intelligence implementation, and when they are supported by clear demonstrations of the benefits, either through pilot projects or case studies, funding is often easier to obtain.

- *Encourage adventurous users.* Though not always considered before a business intelligence project, the character of the user base is important. When users are computer-savvy and willing to experiment, training time and costs are reduced, adoption is faster, and benefits of self-service access—to staff, researchers, and IT departments—are much greater. Adventurous users also tend to experiment with the system, leading to completely unanticipated benefits.

- *Find the right solution.* Today, several products are available commercially. Take great care to fully assess both current and future needs, the vendor's stability, track record, product roadmap, product support options, and licensing policies.

With its flexibility, self-service capabilities, and relative ease of implementation, business intelligence has an important role to play in the management of large academic collections. In addition to the primary drivers of staff efficiency and improved collection reporting and analysis, Web-based access to business intelligence capabilities offers institutions the potential for opening up their collections and demonstrating their value to a wider audience. *e*

Debra Kelly (dkelly@berkeley.edu) is a Museum Informatics Specialist at the University of California, Berkeley.