

Data, Data Everywhere— Not a Report in Sight!

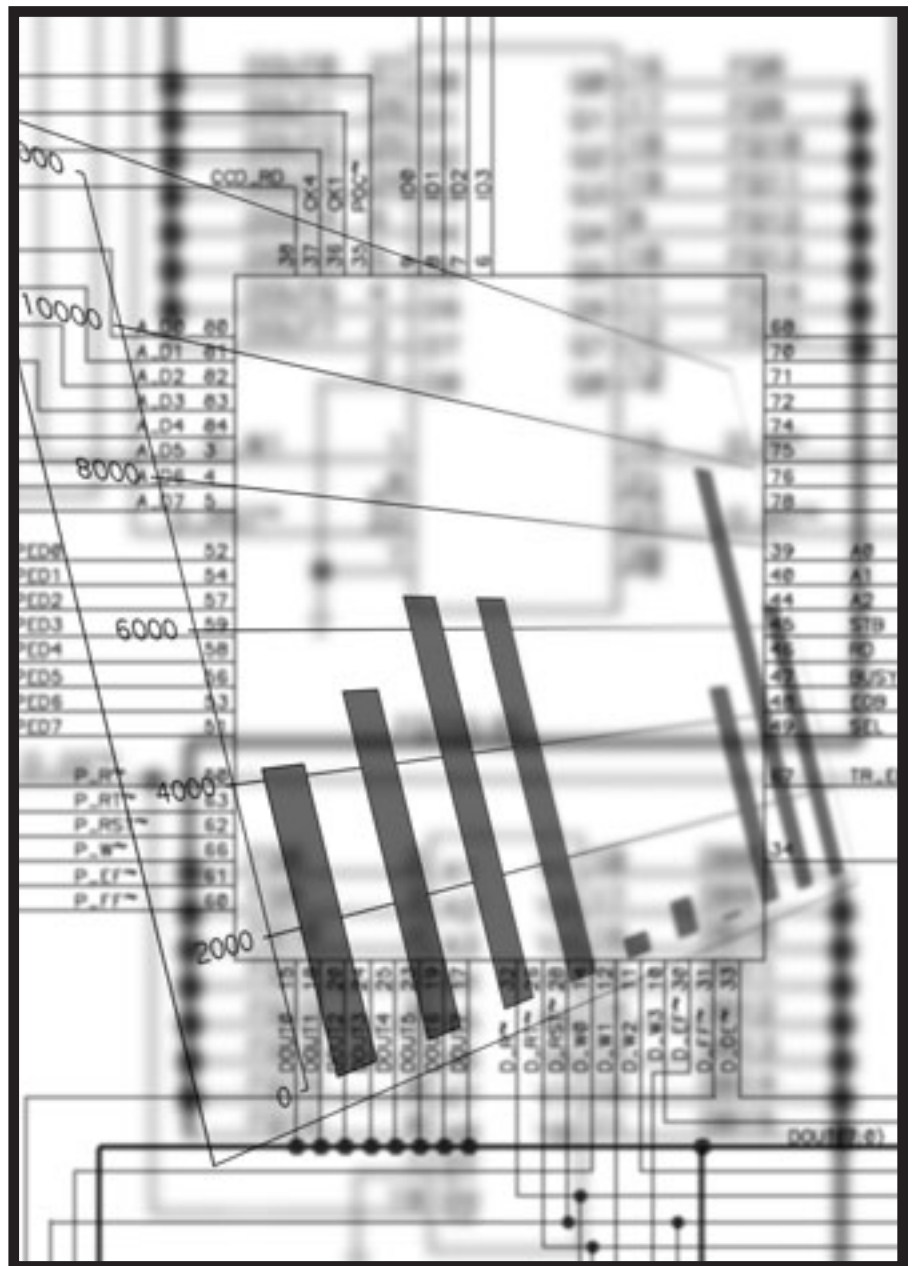
Six proven steps drive data into information

By **Wendy Norman**

Many colleges and universities find it difficult, time consuming, and costly to get the information they need from their administrative systems. These systems, designed to manage student, financial, human resource, and payroll processes, are often complex and rigid. Institutions find themselves flooded with data but often unable to access the information they require for sophisticated reporting. This article outlines an effective methodology for extracting data from complex administrative systems and turning it into useful information that supports core business requirements. Read on to learn about the six steps of data warehouse development that result in valuable, long-term reporting solutions.

Choosing the Right Reporting Vehicle

Administrative systems are high-powered data engines. At first glance, they might seem the obvious choice for data extraction. Using your administrative system for reporting requirements, however, is much like driving a race car to the corner store to pick up a carton of milk. The primary function of an institution's administrative system is to support the critical day-to-day transactions, such as registering students and tracking grades. The information will change on a daily basis, and you may have difficulty replicating reports that were run in the past or doing trend analyses and forecasting. As well, running complex reports on the system not only com-



promises the technical environment but also potentially renders it useless for other users while the reports are being generated.

For longer-term performance and trend reporting you want to choose reliable transportation that can efficiently carry-out your errands, year after year, while saving wear and tear on your speedster. Designing a solid reporting solution that runs beside your administrative system will save you time, money, and frustration.

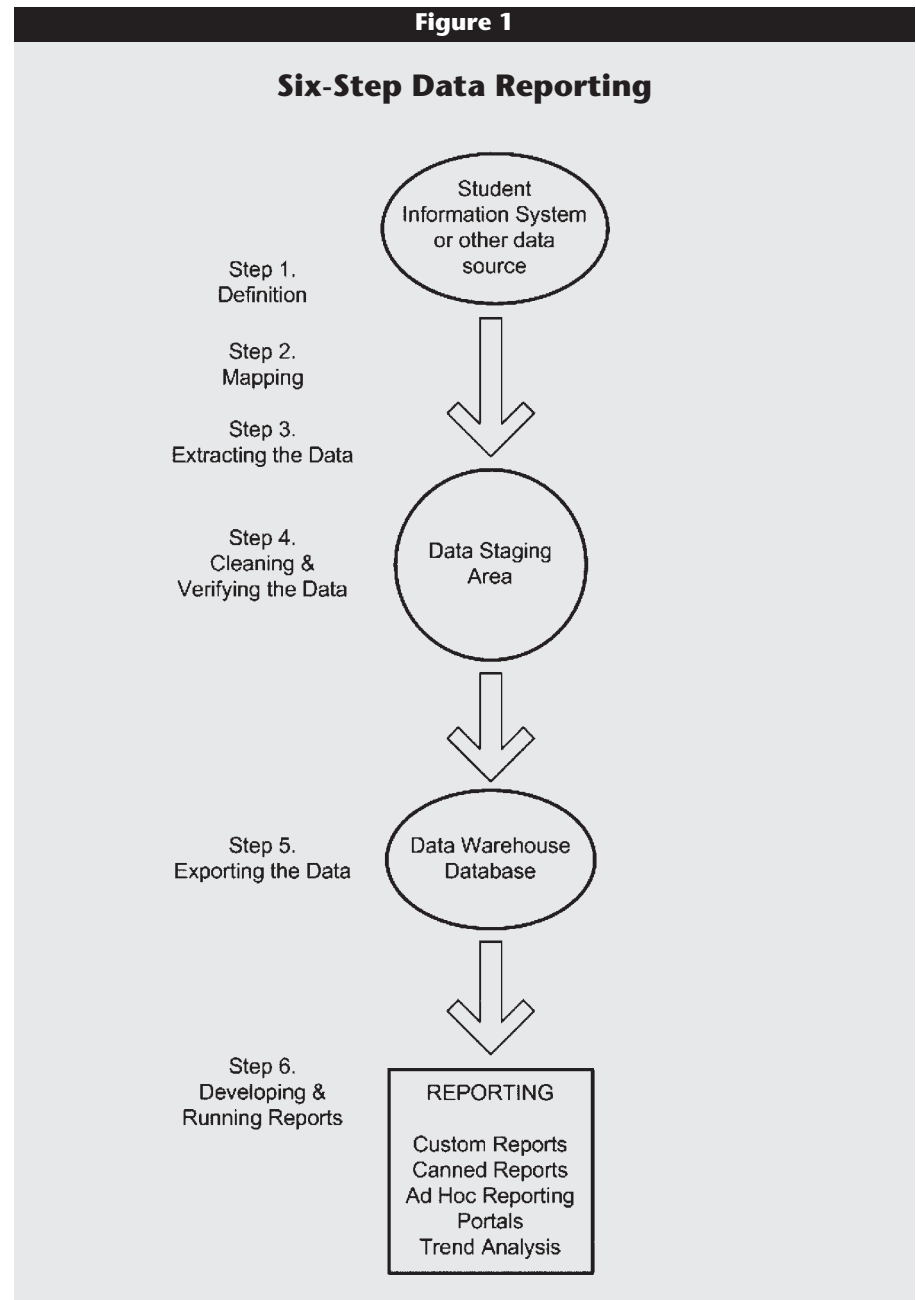
The Road to Reporting: Six Proven Steps

Reporting projects are typically thought of as technical projects, but in reality they rely most heavily on the knowledge and involvement of the business users. The Centre for Education Information Standards and Services (CEISS),¹ has helped dozens of colleges and universities successfully navigate the reporting route. Through identifying needs and mapping requirements, it is possible to build a data warehouse that can support complex reporting. And it doesn't have to cost a lot of money. The six steps to reporting described below chart a process that actively involves the business users in designing the solution and shines a light on some "potholes" to avoid. At CEISS, we follow a generic process, as illustrated in Figure 1. This article shares what we have learned along the way.

Step 1. Definition

The first stretch of road on the reporting map is defining your needs and knowing where the data "lives." It is critical to spend the time up-front clarifying requirements and ensuring that the identified data elements will serve your long-term plans. If you have multiple user groups, you'll want to build consensus and early buy-in to guarantee that the data defined will meet their reporting expectations.

Real-Life Example. CEISS built data warehouses at 22 institutions in order to establish common reporting standards. The group achieved consensus on collective reporting requirements, then



evaluated what data would best serve those needs. By defining their needs as a first course of action, the group was able to identify a flexible foundation that enabled a common, consistent approach to reporting across all institutions. Well-defined, future-focused data standards were the backbone of this multi-institution approach.

Tip: Think Ahead. Aim for true information that is neither an interpretation nor a subjective definition. For example, the classification of part- and

full-time students is always controversial and has changed over the years. Instead of tracking part- or full-time enrollments, capture individual enrollments and then roll them up as required to match the changing definitions. This way, you're starting with a valid "granule," and you can apply whatever aggregation suits you at the moment.

Step 2. Mapping

With the detailed requirements and desired data elements in place, you can begin to locate, or map, the source for

each element. Mapping is not a particularly glamorous step in the reporting process, but it's an important house-keeping item that will result in a final document—your roadmap.

The mapping process may involve accessing multiple sources and using manual or electronic means of gathering information. This can range from finding information on recipe cards to more complex software systems. The final deliverable should be a document that, in column format, has the target data element defined, the name of the source system that stores the element, and the name of the corresponding source element. It should also include any transformation instructions that need to occur, such as changing numerical values to character values.

Real-Life Example. Ohio Dominican University worked through the mapping process with CEISS and invested the time up front to understand their business processes and information needs. They saw this as a critical first step and the foundation on which a reporting solution would succeed—or fail. Said Wes Stahler, former ODU senior developer,

The first thing I would do is make sure that you understand what data you have, where you're putting it when it comes in, and how you're comparing it. You need to understand your business processes and practices first. Look at the data you're putting in and compare it with the information you need to make good business decisions.

The value of data warehouses is vested in the process and methodology used to construct and extract the information. It's easy to get carried away with the technology and become overwhelmed by mounting costs. Data warehouses don't have to be expensive or complex. Choose what's appropriate for your institution based on your information needs, not necessarily the latest technology.

Tip: Breadcrumbs Won't Work. Make sure you have a written document to work from—your roadmap. You can easily get lost without it and will find that



The mapping process may involve accessing multiple sources and using manual or electronic means of gathering information

you waste time and resources if you don't have a plan to follow. Similarly, bring the right people on your trip. The mapping process can't be done without the involvement of key business users who know how the information has been captured and used in the past and how and where it's currently stored.

Step 3. Extracting the Data

At this stage of your journey, you can now move the data from the administrative system into a holding or "staging" area. Data are selected for a set time period, typically a long enough interval to give you sufficient data to really work with. Once the process is more automated, data extraction can be daily, quarterly, semi-annually, or annually. Some manipulation may occur, but for the most part, the objective is to place the required data in a separate, secure location in preparation for a detailed investigation. By using a staging area, you're setting up a safe test bed where you can tinker with and refine what you have extracted before moving it to the warehouse.

Tip: Favor the Familiar. It's important to make the staging area resemble a format that is familiar to the end users. If the format is similar to the system in use, it's easier for users to go back to the source system and track anomalies and resolve issues. This also helps users readily accept the new reporting solution.

Step 4. Cleaning and Verifying the Data

The bulk of the work happens in this step. Now that the data has been successfully loaded into the staging area, it can be analyzed to assess its quality and completeness. This is usually the most important and time-consuming step of the entire project and relies heavily on the people who understand the business and related data within the administrative systems. You will need to ensure that you involve your business users in identifying and qualifying the data to promote ownership and system adoption.

The motto for this phase might be "expect the unexpected," as no matter how much attention has been given to data entry, unexpected issues will likely arise during this step. You might be tempted to believe that your data is in good shape, but this is rarely the case, so you should be cautious about its true quality. Determine what threshold of data is acceptable and decide what process you will follow to get the information you need. For example, if one of your systems hasn't collected the gender information you need, will you call students to validate data, or will you make inferences from names? You must also recognize when enough is enough and be careful about spending 90 percent of your time trying to validate the last 10 percent of your elements.

Real-Life Example. The Bellevue School District in Washington state created a substantial data warehouse to meet increased accountability reporting requirements. CEISS built a data warehouse for them that extracts information from student information databases into one centralized reporting system. Bellevue maintains data at the district level, but found they had to go back to the individual schools to address anomalies.

Critical success factors in this project were the strategies and processes Bellevue established to verify their data and address these anomalies with multiple sources in a timely manner.

Tip: Garbage In, Garbage Out. Be sure to build ownership with your users by involving them in the quality process and addressing issues of importance. Their involvement is crucial to ensuring clean and valid data. If you don't input good-quality data, users will not believe in the system and its ability to produce valuable reports.

Step 5. Moving the Data into a Relational Database

After the data has been cleaned and its quality assured, the data can be moved into the warehouse. If database constraints apply, now is the time consider them. You will want to conduct further testing to check for quality and anomaly issues. For example, you will need to test that students are under or over a certain age. Not many two-year-olds have completed a degree in philosophy! As well, you'll need to check on course completion dates and ensure that students receive credentials after the course has finished, not before it has begun. In this step, you need to build in as many business-rule quality checks as you can think of to ensure that users will be committed to using a reliable, dependable source of data.

The data is then exported into the database that it will permanently call "home." If you have done a thorough job in Step 4, this step is usually quite simple. You're almost there!

Tip: Form Versus Function. You don't want to apply database constraints or other structural requirements until your data is cleaned and you have reached Step 5. It is important to treat structure and content requirements as separate processes. Leave the structural problems until now, and focus on data content and quality during Step 4.

Step 6. Developing and Running Reports

True analysis can now begin with clean and organized data. You're ready

To develop ownership and perceived value, you'll need to build in the appropriate training and implementation support

to report! There are many mechanisms for reporting, and reports may have been designed during the definitions phase of Step 1. Typically, reporting suites include canned or predeveloped reports that can be run easily, data summaries that can answer an array of "typical" questions, and ad hoc queries to answer specific questions. Reporting can also encompass full-blown data portals that may provide ultimate flexibility to users who manage their own unique information needs.

For those involved in the project thus far, there will be much cause for celebration. Before you schedule the party, don't overlook the importance of developing and implementing a full rollout and training plan for your reporting warehouse. Users might now face a suite of new tools and be reluctant to make the move or give up their old tools. To develop ownership and perceived value, you'll need to build in the appropriate training and implementation support. You will want to ensure that day-to-day reporting needs are met through the reporting warehouse and not by the administrative systems or by standalone systems that duplicate a version of the same information.

Real-Life Example. The Ministry of Advanced Education (MAVED) in British Columbia has made the most of reporting by mapping a process to extract existing data for enrollment reporting. MAVED developed a common set of information standards and definitions and a single data warehouse where all postsecondary institutions can deposit and then access their data with flexible reporting and analysis tools. The ministry decided to develop standard reports that ran against that data warehouse

to attain enrollment information. They now have de facto enrollment reporting for the entire province (state) and are able to make decisions based on consistent, comparable, and accurate data.

Tip: Bringing the Project to Life. Your project now needs a face and personality. You might consider developing a name for the reporting tool, a project logo, or a Web site of standard reporting or executive dashboards that track trends and growth. To gain participation and user adoption, you will need to illustrate how the new system can provide benefits, bringing in other information and performing integrated reporting not possible before.

Driving Data into Information

Understanding your data and turning it into information can help inform decision making and provide valuable input into the strategic direction of your institution. A data warehouse is a viable reporting solution that can support your business users without being overly expensive or complicated.

Remember that a data warehouse is not merely a technology project. It involves people and processes and depends on the quality of data that the business users identify and input. Without the involvement of your business users, there is no quality data. And, without quality data, there will be no perceived value or ownership in the data warehouse. Solid planning, forethought, and collaboration will serve you well on your road to reporting. Happy reporting trails to you! *e*

Endnote

1. Since 1996, the Centre for Education Information Standards and Services (CEISS) has provided research, information and technology solutions, and consulting expertise to improve the performance of schools, colleges, and universities. CEISS is an independent, publicly funded organization.

Wendy Norman (wnorman@ceiss.org) is Director, Partnerships and Shared Services, for CEISS Research and IT Solutions in Victoria, British Columbia, Canada.