From Data to Decisions

Lessons from School Districts Using Data Warehousing

From Data to Decisions

Lessons from School Districts Using Data Warehousing

Tamara Mieles and Ellen Foley

Annenberg Institute for School Reform at Brown University

January 2005

Acknowledgments

We learned a great deal during the course of this project and owe many people thanks. First, we gratefully acknowledge all the individuals who participated in the interviews on which the study is based. We cannot name them, due to our promise of confidentiality, but we appreciate their candor and willingness to share their expertise. Their emphasis on the importance of human capacity allowed this study to go beyond the technical aspects of data warehousing.

We would like to express our sincere thanks to those who, at various stages of the study, gave us direction, advice, and resources, including Gloria Frazier, Ellen Guiney, John Kline, Kristin Arnold, William Miles, Bob Palaich, Elliott Levine, Robin Wilner, Brad Duggan, and Muriel Leonard.

Much gratitude is also due to our reviewers, Scott Robinson and Dale Kalkofen, who improved the quality and utility of the content with their valuable feedback, and to our thoughtful colleagues at the Institute, Mike Grady and Peg Votta, whose ideas and energies are incorporated within this report.

We would like to extend a more personal thanks to Marla Ucelli, Carol Walker, and Scott Martin for their assistance, support, and patience in response to the authors' inquiries and requests throughout the study.

Finally, we are pleased to acknowledge the many talents of Margaret Balch-Gonzalez, Haewon Kim, Susan Fisher, Mary Arkins, and Amy Rittenhouse that enabled the editing, design, production, and dissemination of this study.

This study was made possible by the generous support of the Rockefeller Foundation.

Contents

Introduction

In the last decade, more and more researchers, practitioners, and policy-makers have been paying attention to the role that school districts can play in promoting school reform. Such attention has led to greater demands on districts, especially urban districts, to transform themselves in order to educate not just *some* children well, but *every* child well. To succeed in this new and ambitious task, districts must take on new roles and perform others far more effectively. We call this new kind of district a *smart district*[1] (see sidebar on page 2).

Most school districts recognize that one thing they need to become "smart" about is using data, if they are to improve student achievement and meet reporting requirements in the era of state accountability systems and the No Child Left Behind Act. Districts now collect plenty of data. But their ability to analyze what the data mean or to use the data for informed decision-making is limited. School districts are often saddled with antiquated systems that produce fragmented and inaccurate data. Many potential users lack access to the data, or the data arrive too late to have an impact on what happens in the classroom.

District leaders are faced with the need to upgrade technology, streamline data collection, and address issues of data access and use. But most have had little experience in these areas, and few resources are specifically designed to help district leaders implement such a data management program. The emerging field of *knowledge management* – well known in business circles – has not yet been extensively applied to education. Valuable information is widely available in *technology resources for schools,* but it is rarely geared to the specific needs of school district leaders. And there are few descriptions of what school districts need to consider when choosing and implementing technology tools.

Some districts have successfully adopted a tool known as *data warehousing* to provide access to information stored in different locations and formats and to "manage the knowledge" essential to improving schools. Data warehouses can improve outcomes by contributing to better decisions and identifying where to target resources for highest impact. They can help districts take measures to improve student achievement based on facts, rather than assumptions, and address issues of equity that persist in urban school communities.

To help districts make use of this emerging knowledge, the Annenberg Institute for School Reform, in a study concluded in 2004, looked at eight districts in various stages of implementing data warehousing. These districts – among the first to apply data-warehousing technology to education – have generously shared their experiences and

[1] More information and further resources on the topics *in italics* can be found in Resources on pages 30–32, which is organized in sections around each of the topics.

"Smart Districts" and Data

The concept of "smart districts" was developed by School Communities that Work, an Annenberg Institute task force on redesigning urban school districts. The task force defined a smart district as one that focuses on results, equity, and community engagement in order to provide high-quality, equitable educational opportunities to all children in all schools, with no significant differences in achievement based on race, ethnicity, or family income.

The task force identified three essential functions of a smart district: providing schools, students, and teachers with needed support and timely interventions; ensuring that schools have the power and resources to make good decisions; and holding people throughout the system accountable by developing and consistently monitoring appropriate indicators of school and district performance and practices.

Performing these functions well depends on having and using good data. Districts collect an abundance of data. But unless practitioners have the tools and knowledge to analyze data and use it to inform decisions, all the data in the world will do little good.

The Annenberg Institute believes that data warehousing, if designed and implemented thoughtfully, has the potential to be an important tool for decision-makers in a "smart district."

For more information about smart districts, see Resources I on page 30.

hard-won lessons. This report presents the findings from the study, describes in detail the experiences of several of the districts, and provides a listing of useful resources. It also highlights some initial lessons about what district leaders have learned about implementing a successful data warehousing plan.

By describing how the districts chose and implemented data warehousing and the successes and challenges of their work, the Annenberg Institute hopes to help other districts use this technology in their efforts to transform themselves into "smart districts." Implementing a data warehouse will not lead immediately and directly to better and more equitable student outcomes. But data-warehouse technology is a powerful tool that can provide educators and other district, school, and community leaders with greater access to information and new opportunities to create and act on knowledge. By making better decisions, districts can improve practices that influence teaching and learning and, ultimately, student achievement.

About This Study: Gathering Practical Knowledge from Districts

Many district leaders recognize the value of converting their district's large volumes of unorganized data into knowledge they can use to make better decisions – but they don't always know how. While extensive resources on technology are readily available to school districts, they can be confusing and difficult to sort through because of the quantity of technical details and language.

Recognizing the need for practical information about data warehousing, the Annenberg Institute designed a study that could help districts make use of the knowledge emerging from pioneering districts that have had success with the data-warehousing approach. The study specifically aimed to:

- examine districts that were using data-warehousing technology to store, analyze, and access data and find out about their experiences with developing, implementing, and using a data warehouse;
- identify lessons about using data warehousing in education to inform other districts that might be considering implementing such a tool;
- highlight data warehousing as a way to promote knowledge management in school districts by supporting an integrated approach for accessing, managing, and reporting data, especially data related to improving instruction and student achievement.

For the study, we interviewed educators and education technology experts on the strategies and challenges of data warehousing. We looked for districts that had implemented data-warehousing technology to improve instruction, operations, and decision making in the district, with the long-term goal of improving educational experiences for students.

The study looked at eight districts – three of which had won national recognition for using technology effectively – all in the functioning stages of data-warehouse rollout. The districts, mainly urban, ranged from 13,000 to 274,000 students, from 20 to 250 schools, and from 900 to 17,000 instructional staff. The participating school districts were located in California, Florida, Illinois, Massachusetts, Maryland, New Jersey, Virginia, and Wisconsin. We have not named most of the individuals and districts participating in the study; the interviews were conducted under a promise of confidentiality to maximize respondents' candor and willingness to share their expertise. (See Appendix A on page 33 for more details about the study and other research on data management in education.)

About This Report: Sharing Practical Knowledge with Districts

The report offers district leaders practical information about data warehousing from eight districts that have had success with this approach. It begins by reviewing some of the technology and culture issues that data warehousing has the potential to address and defining what a data warehouse is. The report then addresses key issues in operating a data warehouse, including development and implementation, use and maintenance, and the impact it is likely to have on a district and on its ability to use data to improve teaching and learning. Detailed portraits of three of the study districts accompany the text. Finally, the report offers a set of lessons gleaned from the experiences of the eight districts in the study.

Making Use of Data: A Persistent Problem, a Promising Solution

An Abundance of (Unusable) Data

Many districts, especially in the No Child Left Behind era, have a wealth of data but little actual knowledge of what the data mean. While they collect, store, and report data extensively, very rarely is any of it used to analyze student needs or quality of instruction, help form strategic plans or policy, or determine whether resources are distributed equitably.

Various technological and cultural issues are responsible for what might be called "data-free" decision making.

Scattered and Antiquated Technology Infrastructure

Most district data are held in separate, "legacy" systems and are inaccessible in a comprehensive, useful, and timely way. For example, in many districts, student demographic data and test scores are collected and stored by different departments, using different software applications. Personnel data, including teacher certification and other background information, is held by the human resources department, often only in paper files, while data about teacher participation in professional development is collected by central instructional staff. School expenditure information is maintained and accessed exclusively by the finance office.

Clear links between any of these key data areas – student, teacher, school, and district – are nearly impossible to make. It is not unusual to hear administrators' stories of hours spent weeding through paper files and reports to connect attendance records, special education records, and teacher information to understand the performance of a particular subgroup of students. Nor is it unheard of for a principal to act on incomplete or inaccurate information when assigning teachers to classes, even when necessary data, such as certification areas, student results, teaching history, and efficacy have been collected and analyzed by other parts of the school system.

Data Accessible Only through "Gatekeepers"

In most school districts, only a few people – usually data analysts – have direct access to data. Some limits on data access are necessary for reasons of confidentiality and security, but often the "gatekeepers" are simply people who know how to operate technical programs to "crunch" numbers. Generally, the analyst will respond to "static" requests by district leaders. The analyst is rarely engaged in what the original question was that drove the request, often leading to incorrect or incomplete data being provided.

Opportunities are limited for the leaders themselves to investigate and answer questions that are raised in the process of examining data. Using data for developing policy and making decisions is difficult, since data are accessible to only a few and cannot be immediately regrouped and reanalyzed as new questions come to mind.

Inaccurate, Late, or Cumbersome Data

Teachers' most common complaints about data revolve around issues of relevance and usefulness. Recently, more and more school districts are providing data about student performance to schools and teachers, but it often arrives after students have moved on to the next grade or next school and is reported in unwieldy formats. Teachers, especially those who are unaccustomed to data analysis, find the information irrelevant to their teaching and other classroom experiences.

This issue is exacerbated by typical data collection methods. Many important data – about student performance, for example – are not "dynamic." That is, the data are often collected at one point in time during the year or, in the case of teacher information, are updated only when the teacher's status affects his or her pay. Nor is the timing of data collection aligned. Within and across departments, data are collected at different times during the school year. The data are often subject to frequent changes that are not captured in annual or semiannual measures. Again, this lack of up-to-date information hampers the use of data for knowledge management and for policy and decision making.

To make important decisions, district and school leaders, including teachers, need to be able to see both the "big picture" and detailed portraits of how the students, teachers, and schools are doing. But technology and access issues create roadblocks to meeting accountability demands and improving students' educational experiences.

Data Warehousing: A Promising Solution

The term *data warehousing* has, up to now, been primarily associated with business and industry. Known as the "engine that drives business intelligence," data warehousing has been promoted as a technology to integrate, analyze, and measure data on products, services, and organizational effectiveness.

K–12 education leaders have increasingly become proponents of this technology. By making data more integrated, accessible, and current, data warehouses have emerged as a powerful tool for districts to address the problems mentioned in the previous section and to improve policy and decision making. In school districts, there is much data to correlate, integrate, and analyze, usually housed in different departments that don't "communicate," either technologically or in the more traditional sense. With the current state and federal accountability demands, including the rigorous reporting and data analysis requirements of NCLB, information management technologies also respond to an urgent need for integrating, analyzing, and reporting data on student achievement and school performance.

The goal of data warehouses is not only to link data and improve accessibility, but also to promote the application of data in decision making. As important as the technical aspects are, pioneering school districts are also seeking to understand how a data warehouse can improve decision making to increase student achievement and better support school-based staff. Data warehouses provide school districts the potential to support, align, and transform their organization's instructional and operational practices.

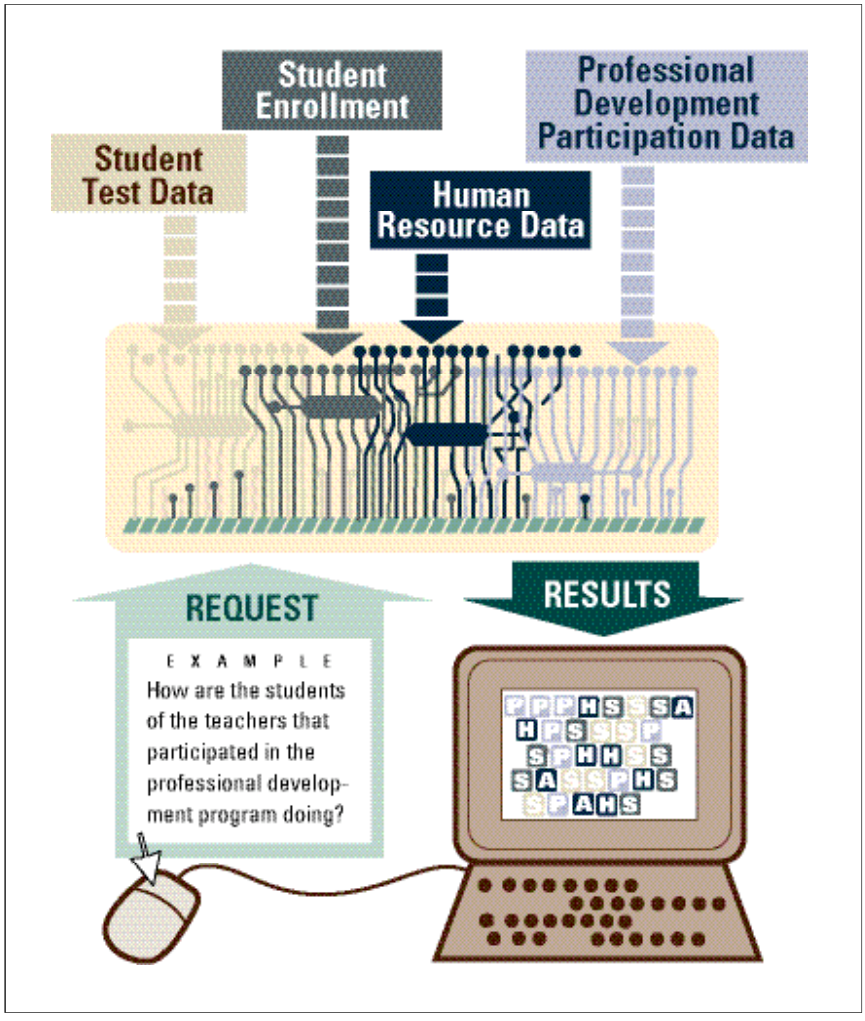
In the rest of this report, we will look at the experiences of the eight districts in our study as they worked to develop and effectively use data warehousing to meet these goals.

What Is a Data Warehouse?

A data warehouse is a comprehensive database that allows access through a server to virtually all an organization's information, which may be stored on different computers in different locations and in different formats. By helping organizations connect, integrate, and analyze data, a data warehouse can be integral in managing knowledge and supporting data-informed decision making.

The following is a non-technical description of a data warehouse, from Elliott Levine's (2002a II-A[2]) "Building a Data Warehouse."

> Sharing information via data warehousing is like sharing on the floor of the United Nations. Each database is like a different country's ambassador – each speaks a different language and cannot fully understand what the other ambassadors are saying. Technology makes it possible for the U.N. ambassadors to communicate clearly using an extensive network of translators. . . . Speech is translated into several common languages so all can participate. Data warehousing achieves the same level of communication by "translating" the information into a format that can be understood by all of your databases. (p. 1)



[2] The Roman numeral and letter following each cited work show where to find the reference in the Resources section on pages 30–32. For instance, the full reference for this citation – Levine 2002a II-A – is in section II (Data Warehousing), part A (General Resources), on pages 30–31.

Deciding to Invest in a Data Warehouse

The development and initial implementation of a data warehouse is a period of significant activity, involving district leadership as well as representatives of many key district constituencies. In this section, we describe how the eight study districts set about implementing their data warehouse programs, from the decision to invest in it to the time, people, and costs involved.

Goal Setting

In all eight districts, the decision to invest in a data warehouse resulted from a complex combination of factors, including greater demands for data from accountability initiatives; a greater recognition of the potential of technology; and the interest of key leaders such as the technology director, the superintendent, or, in one case, a key external partner.

The underlying goal was using data for decision making. As a technology director explained:

> As director, I met with consultants and talked about the idea with the superintendent. We all started to ask ourselves how we can use data at the school and at the district levels to see how kids are doing. . . . We had lots of disparate data and we needed to bring it together.

A director of information technology echoed the same need:

> We're not any different from most school districts. We have a large volume of data. Most of the data is used for operations – for example, reporting to meet operational needs of schools. In that vein, we collect attendance and suspension data. . . . But you have all this data and you want to make useful information from that. That's why we started thinking about using a data warehouse.

In many cases, individuals within the district had been interested in improving data access and use for years before the decision to actually invest in a data warehouse was made. The key difference was having a person with the cachet or authority to champion the process.

Leadership as Catalyst

As with most districtwide initiatives, the superintendent of the school district had a critical role in catalyzing the development of the data warehouse and, later, in maximizing its use. Superintendents who consistently communicated the need for the data warehouse and provided the needed support (sometimes shifting fiscal and human resources)

were often spoken of as "strongly supportive" in the early stages. One superintendent offered a broad perspective of what was needed to lead a "successful" technology effort:

> The major challenge is always the stability to stay the course. In typical districts, boards are unstable; leadership changes. You can't get them from here to there with all that. Often just for lack of attention projects are not moving ahead. Without stability you'd sink. It's very easy to give up really working to have data-driven instructional programs because it takes years to encapsulate that kind of data and spread thinking to school people.

Leaders who engaged in "selling and marketing" within their districts – helping other educators understand the benefits of the system within *their* context – made implementation easier and boosted the use of the technology. Since a technology tool does not, on its own, make itself useful, leaders focused on the people, their needs, and the improvement of the schools in the system. In all of our eight study districts, school boards and superintendents modeled the use of the technology and communicated to the schools and broader community the importance of making decisions based on accurate and timely data.

Combining Internal and External Expertise

One of the early decisions a school district must make about investing in a data warehouse is whether it should be developed in-house or with the support of external vendors. In seven of the eight districts involved in the study, the answer was some combination of internal and external expertise.

It was often complicated to determine how autonomous some districts were because of the many different kinds of relationships they had developed with these vendors. For example, two of the districts in our study had won sizable multiphase grants from IBM. While one district received ongoing technical assistance and support with analysis, cleaning, and storing of data and technical training, the other was receiving consultation on developing, implementing, and managing organizational change strategies for the next phase. Some districts received a combination of both technical and process supports as needed or as negotiated in contracts. A few of the districts custom-built their own data warehouses (e.g., AS400, MySQL), while the majority bought software and, at times, worked with vendors such as IBM or Oracle.

Before a final contract was made with a vendor, a complex set of activities took place. Some districts requested a demonstration and conducted interviews to determine if the technology or service was right for their particular needs. One of the districts piloted a product on site, starting in a few schools. More often, district leaders visited a neighboring school district or obtained references and relevant information about technology providers through colleagues.

The one school district that did not buy off-the-shelf products or contract for services explained that it was an expense the district could not undertake. After several years of

attempting to build a relationship with vendors, they decided that no one could address their needs as well as they could internally. Therefore, the district and technology leaders were forced to develop their internal capacity – district staff with the expertise to support the implementation of the data warehouse.

Whether the data warehouse was homegrown or not, the technology personnel we talked to emphasized the importance of a high-capacity internal team. The team's ability to work across different departments in the district and to establish on-site internal support for data-warehouse users was essential to successful implementation.

Time

Most of the districts in our study launched the data warehouse in a year or less. The one district that spent more than a year on the initial launch encountered problems with the external vendor it had contracted with to support the data warehouse. The company had experience with businesses implementing data warehousing but "didn't understand [the district's] educational needs." Additionally, there were some district failures, such as lacking the right set of questions to identify the data they might need to include in the data warehouse.

In all cases however, the development of the data warehouse did not stop after the initial launch. While the initial launch made *some* data and reports available to *some* people, the types and amount of data available and the groups the data was accessible to grew in every one of our study districts. This expansion was determined both by design – most of the districts planned to add data or users over time – and by demand; as people learned what the data warehouse could do, they made more and more demands about what should be included and who should have access.

Planning, People, and Processes

Whether they worked with a vendor or not, every district in our study spent much of its time evaluating its needs, researching, and planning before getting a data warehouse up and running. In six of the eight study districts, the planning involved a team of staff with different roles in the district.

The committees looked different in each district. Some included the superintendent and the technology director, along with directors from the various central office departments, external consultants, and, sometimes, personnel from the curriculum and instruction, assessment, and/or research departments. Sometimes there were committees led by one central office leader that included a business specialist and "customers," such as principals, teachers, parents, and, in one district, students. Some had several user-based committees; others used just one committee of "architects" that included key personnel and a few school staff. The teams aimed to include input from all those who would use and be impacted by the new technology. School-based educators, administrators, and, in some cases, outside consultants, became collectively responsible for these efforts.

In most of the districts, working teams utilized the school district's strategic plan, standards, or framework as a starting point for their discussions to determine needs and use of the technology. A few other districts first looked, with principals, at school technology plans to start assessing needs. Usually, the charge of the team was to define short- and long-term needs and prioritize multiphase rollout, by making decisions about:

- what questions the data warehouse should "answer";
- what existing data to gather, based on the questions;
- what new data to collect, based on the questions;
- what the priorities were and what the timeline should be to collect data (e.g., ongoing vs. right now);
- what outdated systems to replace or get rid of;
- what individuals, departments, or groups would ultimately get access, have ownership, and be responsible for ensuring the quality of various source data and acting on the data.

Most of the beginning work centered on developing questions aligned with the districts' goals. This effort helped clarify the data needs and relevant issues for technology personnel. These work sessions or meetings continued throughout the development process, with plenty of opportunities for feedback on draft models of the tool.

As one of the respondents told us:

> It's really important to get all of the key stakeholders at the table from the very beginning . . . as some kind of governance or advisory group, with reps from teachers and principals. Getting them talking and being part of the planning process helps set priorities. [The] Tech[nology staff] can do what you want it to do, but you have to be able to tell it what your needs are.

According to a few technology and district leaders, having school-based staff at the table early on turned out to be the most effective way of engaging other school-based staff. Principals and, in a few cases, instructional coaches and teachers involved in this phase not only helped with long-term planning and making the warehouse and support tools even more useful, but became the best advocates, communicating and demonstrating the power of the tool and pushing for use at the school and classroom level.

Development Costs

Estimating startup costs for a data warehouse system based on the districts we studied proved to be extremely difficult. Only six of the districts were able to estimate costs for the development phase of their data warehouses, and in these six districts the estimates varied widely – from $150,000 to $600,000.

Several factors made it hard for respondents to answer questions about startup costs. First, the startup costs depended on many variables: the size of district, the availability

of grant funding, whether they used "packaged" software, and the kind of external consultant or partnership. Second, it was difficult to separate out startup costs from total costs. Two of these six districts could not "parse out" development costs, explaining that the entire project would cost millions of dollars – from $2 to $10 million. One district received a $2-million grant for consulting, matched with $2 million for equipment and personnel. Finally, because of the complexity and scale of data-warehousing projects, many districts were unsure about what costs to include or about how to calculate costs for such inputs as wiring, hardware, contractor dollars, or training. None of the districts in the study calculated the cost in staff time for developing (or for maintaining) the data warehouse.

In a software review, Jeffrey Wayman, Sam Stringfield, and Mary Yakimowski (2004 IV-B) consider all these variables and provide thoughtful guidance for projecting the costs of developing, implementing, and maintaining a technology system. They conclude that the most expensive approach is building one locally.

The Payoff: Improved Output, Not Directly Reduced Costs

It is important to note that data warehouses do not directly save districts money. Rather than reduce costs, they improve output by improving decision making and targeting resources for highest impact.

Our study respondents agreed that, with the exception of automating some reports that had traditionally been done "by hand," there were "no real cost savings" to implementing a data warehouse. As a testing and evaluation coordinator told us:

> Quicker access to information makes it so [users] might work less. It's a tool to support. It helps them be more efficient, gives them better access. But my office has put in more resources, not fewer resources.

ELK GROVE SCHOOL DISTRICT

Elk Grove (CA) School District has been developing its own data warehouse since 2001. District leaders had spent years reviewing offers from vendors and seeking a product that could adapt to change quickly. Yet they continually found that off-the-shelf systems were extremely expensive and would not meet their needs effectively. Deciding to develop their own homegrown system, technology leaders took many steps to build their internal capacity, including attending a week-long Data Warehouse Institute conference, where they later sent their programmers and new staff. Working closely with the research and evaluation department, they built a strong foundation for developing a student information system called SISWEB.

The Web-based portal was released to teachers and administrators at the start of the 2003–2004 school year. The student information system was intentionally created to align with the district's goals and standards. The system includes data such as standardized-test scores, demographics, special education information, attendance, student grades, and scheduling information. At the district level, primary users are the technology and research and evaluation departments.

The other 3,500 people who have access to the Web portal include central office personnel, principals, vice principals, counselors, instructional coaches, and teachers. Plans are in the works to expand the student information system and link it to human resource, financial, and transportation data and to opportunities for professional learning and to "offer services to parents, including secure access to their children's grades, homework assignments, and attendance records" on SISWEB (EGUSD, n.d. II-B).

IMPACT

SISWEB is now being used to set targets and goals for raising achievement. Data from the Web site is used as the basis for individual teacher-student goal-setting conferences at many sites. Students meet with the teacher to examine their recent test data, to set goals for the next administration and to identify specific teacher and student actions that will support improvement. Instructional coaches use SISWEB data when they meet with grade-level teams or subject-area departments at the secondary level, for a variety of purposes. The data might be used to monitor the general effectiveness of instruction, to identify specific instruction strands that need to be reinforced, or to identify professional learning needs.

In collaboration with the research and evaluation department, technology staff have trained everyone how to use the Web site and how to use data. Their special education, research and evaluation, curriculum, and professional learning departments, in particular, use this Web site for data interpretation and for tying the analysis of that data back to instruction. Instructional coaches use this data to work with teachers on improving academic achievement. Overall, the district makes the data warehouse part of the professional learning for teachers.

> Teachers can . . . sort test data by the type of test, race, gender, and other categories. Information is available only on a need-to-know basis, so staff members can only view information about their students. . . . [The tool also helps] principals and teachers set performance targets for their schools. It allows educators to see by name which students have not passed either portion of the California High School Exit Exam and who are performing at the "far below basic," "below basic," "basic," "proficient," and "advanced" levels on the California Standards Test. This helps teachers with classroom lessons to help all students learn the academic skills they need to pass the exit exam and improve their performance until they are "proficient" or higher. (EGUSD, n.d. II-B)

SISWEB data is used by principals, as they meet with teachers, to review the progress of students. Discussion in these meetings often focuses on which students are not making progress and why. The principal then problem solves with the teacher to identify additional resources or support that might be necessary.

Teachers and administrators involved in the district's intervention programs use SISWEB data to monitor student progress. The data enables the team to identify students who are not making expected progress and revise the intervention plan. SISWEB also features cross-tab reports that allow teachers and administrators to use longitudinal data on cohorts of students for program assessment and planning.

CHALLENGES

At the time of the interviews, the Elk Grove School District was working on the issue of real-time data and the "horsepower of the system." To prepare a longitudinal report or view for a school with 3,000 students took time. If a student takes a test, the teacher should be able to see how they did right away (data updates ranged from daily to weekly). Another challenge was getting skeptical teachers to see the tool as useful. The district has been steadily providing extensive training to administrators, teachers, and school staff on the use of the technology and, more importantly, the use of the student data. A final challenge is that the student population, mostly English language learners, has been increasing every year; in 2003 there was an increase of 3,500 students.

Key Features of a Data Warehouse

The districts we interviewed wanted data-warehousing software with specific technological features that would adapt to their context and fit their needs. They wanted help not only with integrating, storing, and accessing data, but also with meeting the internal and external reporting and management needs of a variety of users.

The specific features of each district's data warehouse were determined by who was to access it, and why, and by the district's long-term goals, resources, and technology situation. For example, one district's data warehouse that was developed primarily for central office leadership to monitor school and student performance differed widely from another district's data warehouse that focused primarily on use at the school and classroom level.

In this section, we discuss the features of the varied data warehouses we encountered in our study. We have organized the discussion by categories established by technology reviewers writing about different aspects, products, and instructional management systems designed to access, manage, and present data well (Slowinski 2002 IV-B; Wayman, Stringfield & Yakimowski 2004 IV-B).

Format

The data warehouses in our study worked from PC-based servers and interfaced with users through the World Wide Web (both intranet and extranet). No challenges with this format were cited. However, those in the early development stages had issues with cleaning and formatting data, as well as comfort level with using the Internet, "a new environment" for some. As a chief financial officer explained, "There were many learning curves that needed to be overcome."

Access

Access to the data warehouse was usually provided to staff focused on system goals and business needs, including the superintendent, chief financial officers, accountability and testing divisions, research and evaluation personnel, instructional technology, and information technology divisions. Most school districts also had secure Web-based portals for principals to access teacher, student, and school information. Secretaries often had access for data input tasks and to be able to download information and print reports on request.

Teachers: Teachers' direct access to the data warehouse was less common. In most cases, access for teachers was a key part of the implementation plan. However, at the time of the study, while four districts provided teachers with direct access to various types of student data through the data warehouse, four districts did not. Security and confiden-

tiality measures needed to be added before classroom teachers and other large user groups could access the data warehouse.

In the meantime, data from the warehouse were made available to teachers in other ways. Some teachers accessed various types of student data (e.g., report cards or achievement scores, by a combination of factors) through a shared account with a principal or through hard-copy reports, as this planning director explained:

> Teachers do not have direct access; principals are the only ones with passwords. We are hoping to open it to teachers. Principals will sit down with a team of teachers and run the reports. Teachers might get it in paper copy. They are more interested in reports on students, so they know what students need what kind of instruction.

The school-level people (coaches, facilitators, or principals) who had access to the data warehouse used it to supplement professional learning and to help teachers look more deeply at student achievement and inform their classroom practices.

Parents and students: Although every individual we spoke to supported the notion of providing access to the broader community, there were still issues with providing direct access to the data warehouse to parents, students, and the general public. Most districts were in earlier stages in the process of resolving and ensuring the security of the information and providing access to computers and the Internet. A technology director explains:

> We're one and a half years in. It's designed to serve the entire [district] community – first teachers, then students and parents, with full rollout in years four to five. Our dream is to be able to have parents look at student information online. But there are a lot of implementation obstacles to consider, including parent access, how some families have easy access to the Internet [and some do not] and how to deal with that. Also, who in a student's family would have access. We'd have to develop a fairly complex framework to deal with that because of very complex family structures. We need a different infrastructure to deal with that.

In only three of the study districts could parents use the Internet to access their children's grades, attendance, and other information. One of these districts, which had been using the data warehouse the longest – almost ten years – had successfully provided access to students on school grounds and at home. In fact, one of the district's Web portals, originally designed for guidance counselors and students, became so widely used that it was opened up (with established norms and security measures) to principals, teachers, and parents. They now have close to 90,000 users of this portal.

User-friendliness

Our eight study districts reported that making a data warehouse user-friendly is time-consuming but well worth the trouble. "The only way people use it is if you make it as easy as possible." Technology personnel, especially, commented on designing an intuitive, Web-based model that could be accessed with little background knowledge or

training. A "customer service" orientation was prevalent among many, but not all, of the districts. A director of research said:

> I had an epiphany around this time. For years I supported schools by providing them with reports and I was still not seeing data used the way I wanted it to be. It hit me then that if [analyzing and reorganizing data was] what they really wanted to do, they would work in this office! It was a primal issue. I recognized that we needed to make this very simple for teachers or they will just not do it. An ATM was our driving model. We had to make sure the interface to users was easy to access.

Many of the school districts realized that an "ATM" approach would only help to increase access and the use of data and enhance the technology in the long run. At least two of the districts continued to ask school-based user groups for feedback on the functionality of the data warehouse to improve the frequency and content of training. Although we did not ask about the amount of training available (weekly, monthly, as needed), we did hear from those we interviewed that providing adequate training was an ongoing and sometimes challenging aspect of the work, as we will discuss in the next section.

Customization and Reporting

Most technology offices used popular reporting and analysis tools that accompany IBM or Oracle data-warehouse packages. Others had added connectable systems such as Brio (now Hyperion), Cognos, and Microstrategy. All of the districts used the data warehouse to generate standard reports (e.g., attendance, student achievement), often making them available through the intranet as downloadable PDF files.

Beyond the standard reports, a user's freedom to drill data or conduct deeper analysis was determined by the options the technology personnel provided. Of course, many of these options were dependent on an individual's role in the district (e.g., counselors, school psychologists, principals) in which it was essential to build in security measures to ensure the privacy of students' or teachers' personal records. Technology directors set most other parameters based on feedback from users to modify reports or to change query options on the tool.

These "customized menus" were established only in a few districts. In one district a teacher with little training could use the data warehouse to see all her students' test scores, and then drill down to connect those scores "in endless combinations" with twelve other categories such as ethnicity and proficiency. The same system allowed users to compare and contrast data and examine long-term trends. So, although a teacher would not be able to add a new data item, the flexibility of the system allowed for a wide variety of spontaneous queries and the printing of helpful reports. Using a standards framework, another district made it possible for teachers to print out samples of their students' work and share lesson plans with other teachers.

Notably, most of the school districts were customizing features and reports (e.g., graphs, charts, outlines) as they were expanding and connecting different sorts of departmental and longitudinal data (e.g., human resources information with student achievement).

BOSTON PUBLIC SCHOOLS

Boston Public Schools (BPS) developed a data warehouse from one of its "Six Essentials": to examine student work and data to drive instruction and professional development. Initially, the FastTrack system allowed users to point and click their way through analysis of student test scores, grades, and other information. The application was downloadable to schools, but the data were static. Student mobility and other changes in enrollment made frequent maintenance and repeated downloads necessary.

BPS then developed the MyBPS portal, a Web-based system – updated daily – for examining student data by classroom (for teachers) and school (for principals, teachers, and other school-based staff). MyBPS presents basic data, such as attendance, report card grades, and schedule information, as well as Massachusetts Comprehensive Assessment System (MCAS) scores by student and by item. The system was developed over a two-year period with the help of a local academic who took a leave from his position to work in the school district.

IMPACT

Using this information, teachers can answer key questions such as "How did my students perform on the MCAS English Language Arts test?" That question can be further investigated by, for example, generating a list of students who performed at a particular level (e.g., proficient; needs improvement); comparing students' scores by a particular characteristic (e.g., race, ethnicity, gender); examining students' item responses; and viewing the actual test questions that were consistently challenging. (The text of every MCAS question, as well as students' individual responses to both multiple-choice and open-ended items, is available on the system). A principal or schoolwide data team member can also examine MCAS progress of a particular group of students not identified in standard data collection, the chess team for example, by identifying the students in the system. They can compare scores by class and grade to district or state scores on each MCAS.

All of the standard reports are listed and titled by the question they answer, making the analysis of school and classroom data very simple. However, there is a trade-off between simplicity and flexibility. Users that would like to create their own reports – for example, to compare scores on the MCAS Mathematics Test with other math assessments the district administers – are unable to do that.

MyBPS is a five-year plan that began in 2002. It is designed to serve the entire Boston community – first teachers, then students and parents, with full rollout in years four and five. The special education management system is the last mainframe system to be converted into MyBPS and is scheduled for release in spring 2005. With additional use, the demand for information to be accessible through MyBPS has grown. As one respondent said, "Because people got a taste of 'moving mountains with technology,' everybody wants a piece of it."

CHALLENGES

The primary challenge was helping teachers and principals see how the system could support their work. BPS has addressed this challenge primarily through training and collaboration among the technology, research and evaluation, and instructional staff. The Boston Plan for Excellence, the district's local education fund, has been supportive of these efforts since the development of the FastTrack system and continues to publish newsletters, provide professional development for coaches and other teacher trainers, and communicate the advantages of using data to teachers and principals.

Supporting and Promoting Use of the Data Warehouse

Building and maintaining a data warehouse involves many technical challenges, such as cleaning and maintaining data and keeping pace with technological developments and needs as usage grows over time. However, perhaps the biggest challenge in implementing a data warehouse is getting the target audience – teachers, principals, parents and/or students – to actually use the system. As one of our respondents commented, "You could have award-winning software, but if it's not being used, who cares?"

Making sure the system is used involves massive efforts to ensure that data are accurate and accessible, communicate what the data warehouse can do, and train the target users to access it. We describe these challenges below.

Quality and Cleaning of Source Data

All our study districts sorely underestimated the time and resources required to clean source data for use in the data warehouse. Cleaning data involves collecting or extracting data from the different legacy systems, entering the data accurately, standardizing the data, and then importing the data to the data warehouse to be analyzed.

What can make this such a daunting task – what one technology director called a "hidden gotcha" – is the quality of the data in the legacy system or older database to be transferred to the data warehouse. If the source data are incorrect, the output from the data warehouse to the end user will also be incorrect. Linking the differing data systems, therefore, involves ensuring the integrity of the original data, because the more "bad data" there is, the longer it takes to clean.

Even when working with an experienced vendor, those we interviewed were troubled by this cumbersome job. One of the technology personnel told us:

> Initially everything worked quite well. [The vendor] has a lot of history and were very good consultants to us. Where we started to run into problems was getting data from where it resided to cleanse it and put it in the data warehouse. . . . [The vendor] couldn't catch bad data. For example, in one case several years in a row had the same exact data. They looked at it as "There is data here. The report must be correct." But I knew just by looking that it wasn't right.

Other common data issues involved missing fields or differing information on key student characteristics such as gender, race, or ethnicity; multiple student records; and the identification of enrolled students. The experience of cleaning data led some of our study districts to establish data collection norms and develop "data dictionaries," where source variables were clearly defined.

The tedious task of going through multiple years of data must be addressed systematically, and that means abandoning manual vetting. The districts who were further along in developing their data warehouse, for example, were more likely to have automated and nightly maintenance schedules (e.g., every twenty-four hours). As access expands, so does this imperative. As one data-warehouse analyst told us:

> Data quality is another challenge. . . . What happens is that now the public has access to good and bad data. You've got to fix bad data. . . . In a way, it is good to have people point to weakness – good at forcing all stakeholders to look at themselves.

Communication

Once the data were clean and the tool was ready, some districts learned that communicating about what the data warehouse could do was crucial. One of our study districts targeted teachers first (in contrast to the other districts' top-to-bottom approach) and rolled it out quickly to them in less than a year. In that district, the superintendent mandated the use of the data warehouse as a "communication tool." The district also encouraged use in other ways, such as mandating that elementary school report card information be submitted through the Web-based portal.

> For certain tasks, principals and teachers have to use [the data warehouse]. They may never come onto [it] otherwise. Sometimes principals have bulletins, only on [the data warehouse]. Departments post forms that people might need to fill out, only on [the data warehouse.] You hear often in the district, go to [the data warehouse].

Our respondents from this district felt that forcing school-based staff members to access the Web portal for particular tasks encouraged them to explore what else the data warehouse had to offer.

In another district, teachers were required to use the system for their yearly goal setting. "Since they have to use it to set goals, they have to learn how to use data," said a technology leader from this district. They relied on principals to communicate professional learning opportunities (trainer-of-trainer meetings; technology training on site; pre-service, after-school, and Saturday seminars) and distributed newsletters throughout the district to make sure all audiences were aware of dates and the kind of training being offered. Communication efforts alone did not bring forth more users, but they were a critical step toward doing so.

Training

Training personnel – especially school-based staff – to use the data warehouse was the most frequent challenge cited by the people we interviewed. Our study districts were challenged by the variety of users and their varying level of computer and analysis skills. Some teachers and principals, for example, don't know the difference between scrolling and clicking or "what a mouse is"; others become "power users" with only a brief introduction. And whatever their technical abilities, many teachers and instructional leaders are not accustomed to asking powerful questions about the results of their teaching.

In many ways, the desire to foster use of the data warehouse forced central office departments (including information technology, research and evaluation, and curriculum and instruction) to work together in ways they never had before. For example, in one of our study districts, the information technology division worked in conjunction with the research office to integrate training on how to use the technology with information about how to interpret the data generated. Both groups worked closely with instructional staff, especially school-based coaches and other professional development providers. The coaches would then convert their learning into professional development or in-service opportunities for teachers to discuss findings and act on data.

While these kinds of collaboration were common in the districts in our study, none of the efforts were easily done. Many of the districts had to reshape their efforts several times in an effort to get the right balance between technology training, data interpretation, and application to practice. Finding this balance is an ongoing task.

Even with these efforts, our respondents felt there was much work to do to encourage school-based personnel to use the data warehouse. We heard consistently that principals and teachers were reluctant to embrace the new technology. Some technology personnel felt this was due to discomfort with change. No one wants to "throw away scotch tape and highlighters," preferring to conduct analysis and retrieve and connect data manually, working in familiar but time-consuming ways. A central office leader also hypothesized that this discomfort stemmed less from the technology than from a reluctance to engage in the stories that the data tell, which often reveal a less than flattering portrait of the results of school and teacher practice.

Issues with teachers' training and lack of use or access of an information management tool can be related to their ability and willingness to use the data. Denis Doyle (2003 IV-A) suggests two reasons teachers resist using data: the fear that data will be used against them in punitive ways and the "loathing" of burdensome data that are not useful or grounded in their daily work (p. 3). Lisa Petrides and Thad Nodine (2003 II-B) suggest that these issues are more related to work practices within the school or district or, "information culture," than to the technology effort (p. 11).

The few school-based users we spoke with said reluctance to use the data warehouse had less to do with willingness and more to do with relevance and quality training. One principal explained his experience – "I discovered that training is only as good as the people giving the training and how they relate to principals."

This same principal appreciated all the new efforts and innovations but also was frustrated with the district being "tied to old ways."

> [They] put the administrator's bulletin in a folder in a computer on the server (in this portal). But the district still sends out a paper copy, allowing people to fall back on old ways!

Supporting the Maturing of the Technology and the School System

It is important to note that the more successful a warehouse is with the users, the more maintenance it requires. As use increases, costs go up and more resources are demanded. District help desks, if they exist, are often understaffed. As one technology director told us:

> I've directed this office since 1981. Back then, we delivered very simple reports to very few people. Now there is an explosion of data needs and the people we want to deliver the information to is teachers. Now we are dealing with thousands of users. The size of the population has changed.

The challenge of keeping up with demand is exacerbated by seemingly constant budget issues in many urban districts. An individual we interviewed was concerned the next phase in her particular district would not happen "with the money crunch." A principal we spoke to commented, "We have technology coordinators that could act as coaches but when there are budget cuts they are the first to go."

The Impact of Data Warehousing on Districts

Greater data availability, accessibility, and connectivity in our study districts – made possible by the data warehouse – impacted individual students, teachers, and schools in countless ways. Some of the examples we heard in our interviews:

- Teachers used assessment information to identify skill areas that needed attention.
- Principals created grade-level or schoolwide analyses of test results to spark discussion about patterns of performance.
- Superintendents and leaders of central office departments tracked which teachers and schools were referring more students to special education and then asked why.
- Instructional coaches used the data warehouse to work with teachers to improve academic achievement.

Such impacts varied from district to district and by the structure of the data warehouse. We have highlighted some of these impacts in the profiles that appear throughout this document. We have not summarized impacts on school-based staff across our study districts because they were difficult to ascertain, due to the composition of our study sample. Most of the respondents were central office personnel, two were principals, and only one was a teacher.

The larger impact that the highlighted examples point to is the data warehouse's contribution to creating a "data culture" in a school district. While all the districts in our study had collected and reported on voluminous amounts of information prior to implementing a data warehouse, the ease and breadth of access made data more integral to the life of the district. In many districts, the data warehouse was combined with more systematic efforts to monitor and evaluate performance through a continuous improvement approach, such as "plan, do, study, and act."

The ability to connect data that used to be stored in "ten different databases" has allowed districts to generate useful information related to virtually any question that comes out of such a process. As one superintendent told us:

> District personnel have been able to generate and view reports leading to rich discussions around factors that directly impact student achievement. We can now look at these trends and get a better picture of what is happening over time. Schools and central office personnel have become much more data driven and the conversations more focused on trending information.
>
> For example, in response to the new NCLB requirements, we have used the data-warehouse system to help identify student achievement trends [on standardized tests] for various ethnic and socio-economic groups. This information was shared with the

schools and the divisions of curriculum and instruction, special services, and staff development to guide staff development opportunities, curriculum and instructional changes, and student/teacher support services.

Several respondents noted that the availability of so much data had altered the culture of the district. They felt that more collaborative, data-informed decisions were being made. As one research-unit member told us:

> We have a value statement about managing by fact. Before the data warehouse we managed by opinion, because we had no facts to bring to a decision-making process. The data warehouse supported the realization of that value.

BROWARD COUNTY PUBLIC SCHOOLS

Broward County Public Schools (BCPS) was the first school system to develop a data warehouse. The data warehouse was made possible through an initial grant from IBM in 1995 and was piloted in three schools in 1996. It became available to all schools and district offices in 1998. As of 2004, the data warehouse includes information for over one million current and former students, including student demographics, enrollment, attendance, special programs, test scores, course transcripts, and transportation. It also includes data on over 83,000 current and past employees and information on courses, lesson plans, school performance, finance, and maintenance.

The data warehouse is designed to accommodate the needs of users with different technical skills: from power users, who can create their own queries and reports or download data to their own computer, to those who just want to read a regularly produced, "static" report. At least one person in every school is trained to use the query and report development software Hyperion (formerly known as Brio). With its vast data stores and flexibility, BCPS's data warehouse is widely accessible to district employees, students, and parents. Through one of its most unique aspects, the "Virtual Counselor," students, parents, and other qualified users can access, via the Internet, individualized information such as graduation requirements (including credits needed and community service hours accumulated), bus routes, grades, absences, and health information.

IMPACT

Combined with the "Sterling process," a four-step continuous improvement process (plan, do, study, act), the data warehouse has infused data into decision making throughout the district. For example, the district's senior management team modeled for principals and other school leaders how to examine school-performance reports. The change was slow, but has taken hold. Superintendent Frank Till (2004 II-B) wrote, "Eventually, principals began to understand how the use of data could improve instruction and student performance. As results began to surface, more and more principals came around to believing in the process and adopting the use of data into their individual school plans" (p. 14). Every school now includes in its school plan a "Sterling project" – a major area for reform – determined by examining data from the data warehouse.

District personnel use the data warehouse to examine the results of particular programs and initiatives. Block scheduling – the course scheduling strategy that results in longer classes – was shown to be effective in improving student performance by comparing the test scores, grades, and attendance data of students in block and traditional schedules.

Teachers can also relate data about results to their practice. The data warehouse provides detailed information on student performance on the Florida Comprehensive Assessment Test. If data show poor outcomes on a particular skill – reading graphs, for example – teachers can modify their instruction accordingly.

When teachers have access to reports through the district's intranet, they save preparation time for parent-teacher conferences and have more meaningful conferences. These reports can be used in addition to the report card as another measure of student performance and progress. Teachers can show parents historical test scores and discuss strategies with them to strengthen areas of concern.

Through Virtual Counselor, parents are able to check progress or get regular updates whenever they want to know how well their child is doing in school. Attendance records, test scores, class rank, and more can all be found online. Virtual Counselor can be used at school, office, home, or anywhere there is a computer with Internet access. All parents or students need to do is log on to the Web site using their passwords and, a few clicks later, they have the information they are looking for. If they find something of concern, they can schedule an appointment with a school guidance counselor.

"Virtual Counselor was designed to provide students and their parents with a way to check their own records and answer basic questions a guidance counselor would typically be asked," explains Phyllis Chasser, senior data-warehouse analyst in BCPS's Education Technology Services, "questions like 'what's my grade-point average or class rank?' 'Did I pass the FCAT?' 'What classes do I still need to take

to graduate?' and 'Did the service learning hours I turned in last week get posted yet?'"

The data warehouse is used to catch behavior patterns while there is still time to use preventive measures to correct problems and keep them from becoming more serious. The biggest example is the reduction in dropouts by monitoring absenteeism and test scores and using a variety of intervention services to keep students in school. Another example is the reduction in the number of students retained in grade. Due to proactive monitoring of performance over years, students receive additional assistance to help them succeed at grade level and avoid the stigma of being retained. In Florida students lose their drivers' licenses if they have a specified number of unexcused absences. Students and their parents receive a warning letter as they approach that level. If they exceed the allowed absences, the state is notified. This effort used to require a high degree of manual effort. The end result was an excessive amount of time and money spent on relatively inaccurate data and reports.

CHALLENGES

BCPS has used data warehousing longer than any other district, but it still faces challenges. The quality of the data is a key challenge. As the BCPS data-warehouse technician told us, "You quickly realize the weaknesses in your data when you earnestly begin to use them."

Data access is another continuing challenge. With wider accessibility comes greater use and its associated demands. More and more departments want their data, or more of their data, included in the data warehouse; users seek out support and advice about interpretation of data. According to Chasser, "Our motto is: Ease of Access = Utilization of Data. How effectively we are able to provide painless, seamless, simple, and quick access to meaningful data will dictate the success of our data warehouse."

Lessons for School Districts

The pioneering districts in this study have blazed a trail for other districts interested in developing data systems that are more connected and accessible and easier to use. The lessons that follow were drawn from the collective experiences of the eight districts in the study.

➔ Inventory the content and quality of existing source data.

One of the key lessons about building a data warehouse was: Don't start your process by building the data warehouse – there is much pre-work to be done. Prior to purchasing software or contracting with a vendor or consultant, districts interested in data warehousing should assess the data they have and determine what other data they want. Many technology experts suggest that a school district first take an inventory of not only what data and technology it already has, but also of what it doesn't have and needs. This process helps decide whether and how to engage outside help and, more importantly, what the scope of the data-warehouse project will be, including costs, time, and energy (Deck 2000 II-A; Kongshem 1999 II-A; Levine 2002a II-A).

During this assessment stage, it is most important to look closely at the quality of the data in each of the originating systems. Are the data up-to-date and are terms standardized (for example, does the district have – or could it compile from different sources – a data dictionary)? Are there already processes in place to ensure the data are accurate and not duplicated? Most of our study districts spent an unexpected amount of time cleaning data; it is critical that source data are inventoried during the development and launching of a data warehouse.

➔ Determine what you want to know and how to capture it in the data warehouse.

Our informants urged other districts to work from their own strategic plans or other vision-oriented documents to be clear about what type of data to store in and, ultimately, extract from the data warehouse. This challenging process necessarily involves stakeholders from throughout the district.

What we heard many times from those we interviewed was how hard, but how essential, it was to anticipate "the right questions." Prior to developing the data warehouse, districts must also figure out how to answer their key questions. This involves identifying data from both existing and new sources and may also involve broadening the definition of what constitutes data. Getting beyond standardized-test scores is important for supporting data-informed decision making. Datasets might include: allocation of instructional materials; participation in professional development; provision of curriculum and teacher supports; demographics of students, staff, and schools; lesson plans; and survey data.

Involve end users and provide ways for them to give input and feedback.

Experienced data-warehouse builders always have in mind the interests of their end users – those who will actually make use of the system once it is running. The consistent involvement of end users is pivotal during the planning phase, but is also important as the warehouse is developed and maintained. Teachers will not have the same concerns and requests as principals or parents or superintendents, making this process time-consuming and "messy" because it generates more data. But knowing the different perceptions and needs of users in the beginning and getting their feedback along the way pays off. The data generated are more relevant and useful, so re-formatting and re-coding are minimized.

In addition to the internal committees described earlier in this report, the study districts also used e-mail, listservs, and discussion boards and held activities such as focus groups, interviews, and meetings with different stakeholder groups. Each of these strategies was employed to garner user feedback. Understanding and articulating users' sometimes-conflicting ideas helped the technology team improve their implementation plans.

Take time to research and select a vendor and appropriate software.

Working within a school district's budget, technology directors and programmers in the study districts were often the ones researching software products and services before presenting ideas on which the superintendent could base and justify purchasing decisions. Making time to do research,[3] sample the product, view a demonstration, and interview existing users of the tool or service is an essential part of determining the right products for your district. Not surprisingly, our study participants suggested that districts seek guidance from nearby or similar districts to learn about key steps they took and challenges they faced in developing and implementing a data warehouse.

Any software purchased or developed should be compatible with the *School Interoperability Framework (SIF)* or, at the least, ODBC[4] compliant and XML[5] enabled. SIF is an effort of the major educational software producers. Its goal is to assure that all new education software is able to connect seamlessly with other technology systems, new or old. (For a discussion of SIF and how it benefits data warehousing, see Appendix C on page 37.)

[3] The Wayman, Stringfield, and Yakimoski (2004 IV-B) article is an invaluable resource for this step.

[4] ODBC stands for Open DataBase Connectivity, a Microsoft standard method for accessing different database systems such as Oracle or SQL from Windows. "The goal of ODBC is to make it possible to access any data from any application, regardless of which database management system is handling the data" <www.webopedia.com/TERM/O/ODBC.html>.

[5] XML stands for Extensible Markup Language, widely used on the Internet to define other markup languages and help diverse users exchange data. See Appendix C on page 37 for more information.

➔ Get the data warehouse up and running first, then phase in access for different user groups.

Most of our study districts built and began rolling out their data warehouse in a year or less. According to Jeffrey Wayman, Sam Stringfield, and Mary Yakimowski (2004 IV-B), a "rapid, successful implementation" (p. 8) was a key strategy for maintaining the momentum of the project and maximizing the organization's effectiveness in implementing the technology tool.

Given that one of the main goals of district leaders was to achieve a "district view of data," study participants recommended a multiphase rollout that starts with the top-level decision-makers and gradually adds access for principals, teachers, and ideally, parents and students. Only one of our study districts has succeeded in rolling out to teachers first, rather than starting with top administration. Only now is this school district in the process of making the tool user-friendly for district and instructional administrators so student and school data can be compared more holistically across the system.

➔ Look beyond the technology and model a data culture.

Data warehousing is a powerful technology. But using it to its fullest advantage challenges the traditional cultural norms of school districts. The collaboration, proactive planning, and two-way communication required to develop a data warehouse run counter to the traditional district culture, which tends toward isolation, reactivity, and entrenched hierarchy.

All participants articulated the potential of data-warehouse technology to address the broader goal of creating a data culture. Data warehousing was seen not simply as a technological concern or a matter of installing a gadget for only a select few to use. One technology director captured it this way, "The technical part – that's fairly standard. The toughest piece is changing the culture of the way people do things, to make them believe that this is really worth it and let's put the resources into it."

Experienced districts stressed looking beyond technological capabilities to get the best out of the data warehouse. This requires "data becoming knowledge" (Petrides & Nodine 2003 II-B) and, therefore, must involve action – not only on the part of information technology departments that concentrate on the technological aspects and content of the tool, but also the people whose responsibilities usually involve training or assisting users on how to interpret data (e.g., research, evaluation, and accountability departments) and, in turn, those who help teachers make the data relevant to their classroom practice, choice of curricular materials, and teaching strategies (e.g., curriculum and instruction departments). For districts committed to this particular kind of technology initiative, the emphasis on building time for this work and building the internal capacity to do it cannot be stressed enough. All of our informants agreed on the importance of investing the "time, money, and training of staff" to plan, develop, and support use of the data warehouse.

Creating and Acting on Knowledge

Data warehouses are not a panacea. Implementing a data warehouse will not lead immediately and directly to improved results at the classroom level. But data-warehouse technology can contribute to school and district improvement in important ways. The ultimate goal is to use this technology to build a collaborative culture that manages by fact, rather than by assumption, and seeks to address issues of equity that persist in urban school communities.

None of the districts we spoke with can claim they have reached this "holy grail," as one district leader dubbed it. Knowledge management and continuous improvement must be built into the data warehouse. And because technology on its own is just a tool, it is people who must use it, pushing for deep changes based on data and working as catalysts of district improvement efforts (Petrides & Nodine 2003 II-B; Slowinski 2002 IV-B).

All of the districts interviewed in this study continue to pursue systems that provide educators and other district, school, and community leaders with greater access to information and new opportunities to create and act on knowledge. Their experience so far highlights both the challenges and the exciting potential of this technology. Their hard-won lessons will help them and other districts improve the practices that influence teaching and learning and, ultimately, student achievement.

Resources

I. Smart Districts

The concept of "smart districts" was developed by School Communities that Work, an Annenberg Institute task force on redesigning urban school districts. The task force defined a smart district as a high-performing community of schools that focuses on results, equity, and community engagement in order to provide high-quality, equitable educational opportunities to all the children in all its schools, with no significant differences in achievement based on race, ethnicity, or family income.

Annenberg Institute for School Reform. 2002. "Results and Equity." In *Portfolio for District Redesign.* Providence, RI: AISR, Brown University. Available online at <www.schoolcommunities.org/portfolio/index.html#results>.

Ucelli, M. R., and E. L. Foley. "Results, Equity, and Community: The Smart District," *Voices in Urban Education* no. 5:5–10. Available online at <www.annenberginstitute.org/VUE/fall04/Ucelli.html>

II. Data Warehousing

A data warehouse is a database that allows access through a server to virtually all an organization's information, which may be stored on different computers in different locations and in different formats. By helping organizations connect, integrate, and analyze data, a data warehouse can be integral in managing knowledge and supporting data-informed decision making.

A. General resources

Deck, S. 2000. "Data Warehousing: This Goes on Your Permanent Record," *CIO Magazine* (November 1). Available online at <www.cio.com/archive/110100/permanent.html>.

Education Commission of the States. 2002a. *Next Generation Models of Education Accountability: The Quality Improvement Model.* Denver, CO: ECS.

Education Commission of the States. 2002b. *No Child Left Behind Issue Brief: Data Driven Decision Making.* Denver, CO: ECS.

Kongshem, L. 1999. "Smart Data: Mining the School District Data Warehouse," *Electronic School* (September). Available online at <www.electronic-school.com/199909/0999f1.html>.

Levine, E. 2002a. "Building a Data Warehouse," *American School Board Journal* (November). Available online at <www.asbj.com/2002/11/1102technologyfocus2.html>.

Palaich, R. M, D. G. Good, C. Stout, and E. Vickery. 2000. *Smart Desktops for Teachers.* ECS Issue Paper. Denver, CO: Education Commission of the States. Available online at <www.ecs.org/clearinghouse/18/47/1847.htm>.

Schwartz, Wendy. 2002. "Data Driven Equity in Urban Schools," *ERIC Digest* no. 177.

Why Total Cost of Ownership (TCO) Matters: A Report and Estimating Tool for K–12 Districts. 2003. Washington, DC: Consortium for School Networking.

B. Case studies

Consortium for School Networking. 2003. *Vision to Know and Do: The Power of Data as a Tool in Educational Decision Making.* Washington, DC: Consortium for School Networking. Available online at <www.3d2know.org/publications/index.html>.

Elk Grove Unified School District (n.d.). "On-Line Student Performance Data for Teachers – SISWEB Fact Sheet," <www.egusd.k12.ca.us/parents/siswebfact.htm>.

Lafee, S. 2002. "Data-Driven Districts," *The School Administrator* (December). Web edition: <www.aasa.org/publications/sa/2002_12/Lafee.htm>.

Petrides, L. A., and T. R. Nodine. 2003. *Knowledge Management in Education: Defining the Landscape.* Half Moon Bay, CA: Institute for the Study of Knowledge Management in Education, 1702-1717.

Till, Frank. 2004. "Working toward a Data-Driven, People-Centered District," *Voices in Urban Education* (Fall), 11–15.

III. Knowledge Management

Knowledge management is an emerging field that describes how organizations can develop, share, and act on knowledge (including quantitative data) more efficiently and effectively. This information, while commonplace in business circles, is not yet easily available to educators – and is only recently being applied to education.

Knowledge management can be defined as the sum of practices, behaviors, and structures that enable organizations to know and to share what they know – to get the right information to the right people at the right time in the right context for the right purpose. This definition was developed by Kris Kurtenbach and Gloria Frazier of Collaborative Communications Group.

Petrides, L. A., and S. Z. Guiney. 2002. "Knowledge Management for School Leaders: An Ecological Framework for Thinking Schools," *Teachers College Record* 104:8.

Petrides, L. A., and T. R. Nodine. 2003. See Resources II-B.

Senge, P., N. Cambron-McCabe, T. Lucas, B. Smith, J. Dutton, and A. Kleiner. 2000. *Schools That Learn: A Fifth Discipline Fieldbook for Educators, Parents, and Everyone Who Cares About Education.* New York: Doubleday/Currency.

IV. Technology Resources for Schools

A wealth of valuable information is available to school districts today about technology for education, although the information can be difficult to sort through because of the quantity of technical details and language.

A. "Dos and don'ts" advice of technology experts

Deck, S. 2000. See Resources II-A.

Doyle, D. 2003. "Data-Driven Decision Making," *T.H.E. Journal Online* (May). Available online at <www.thejournal.com/magazine/vault/articleprintversion.cfm?aid=4419>.

Kongshem, L. 1999. See Resources II-A.

Levine, E. 2002a. See Resources II-A.

Levine, E. 2002b. "Technology Nightmares," *Electronic School* (January). Available online at <www.electronic-school.com/2002/01/0102f6.html >.

B. Reviews of vendors, products, and services

Gordon, E., and B. L. Bridglall. 2003. "Toward a Relational Management System for Education Management System for Education," *Pedagogical Inquiry and Praxis* 4 (February), Teachers College, Columbia University, and the College Board.

Slowinski, Joseph. 2002. "Data-Driven Equity: Eliminating the Achievement Gap and Improving Learning for all Students." Unpublished Manuscript.

Wayman, J. C., S. Stringfield, and M. Yakimowski. 2004. *Software Enabling School Improvement through Analysis of Student Data.* Baltimore: Center for Research on the Education of Students Placed at Risk, John Hopkins University.

V. Schools Interoperability Framework Initiative (SIF)

Brittain, D. 2003. "My Database Will Call Your Database, but Can They Talk?" NCLB supplement. *T.H.E. Journal Online* (May). Available online at <www.thejournal.com/magazine/vault/A4418.cfm>.

Levine, E. 2002a. See Resources II-A.

Mercurius, N., G. Burton, B. Hopkins, and H. Larsen. 2004. "New IDEAS: Putting the School Interoperability Framework to the Test," *T.H.E. Journal Online* (January). Available online at <www.thejournal.com/magazine/vault/articleprintversion.cfm?aid=4621>.

Trotter, A. 2005. "Software Framework Opens Up Data-Sharing," *Education Week* 24, no. 21 (February 2).

"The Work behind SIF's Framework." 2002. *T.H.E. Journal* (March). Available online at <www.thejournal.com/magazine/vault/A3948.cfm>.

Appendix A

About the Study

To identify districts using data warehousing, we contacted educators and education technology experts in organizations that work closely with school districts on technology issues (such as the American Productivity and Quality Center, Just for the Kids/National Center for Educational Accountability, and the Public Education Network). We asked these informants to review the study design and protocols and to identify "model" districts already implementing a data warehouse. Specifically, we asked our initial informants to identify districts that intentionally sought to implement data-warehousing technology to improve the instructional and operational data practices of the district and, ultimately, to improve the educational experiences of students, teachers, and staff. We sought out districts that selected data warehousing as a platform and as a decision support tool that would address the planning, decision making, and customer service of the district.

We then asked superintendents, directors, and managers from the technology, research, and testing departments in the recommended districts to participate in a forty-five-minute phone interview about their experiences implementing data warehousing. Using a semi-structured protocol (See Appendix B on page 36) to guide the conversation, interviewees were asked about their start-up and development processes, the technological features of the tool, and the impact this undertaking had on their school community. A qualitative data software package, QSR N5, was used to code, categorize, and analyze interview responses.

Of the eighteen school districts that were contacted, twelve responded and demonstrated interest in sharing their approaches and reflections on past and current work. However, during the interview process we discovered that four of the districts either did not meet our criteria or were still too early in the development stage to fully answer our questions. Therefore, we included only respondents from eight districts that were in the functioning stages of data-warehouse rollout. Three of these districts had won national recognition for using technology effectively.

The size of the participating (mainly urban) districts ranged in student population from 13,000 to 274,000, in number of schools from 20 to 250, and in instructional staff from 900 to 17,000. The participating school districts were located in California, Florida, Illinois, Massachusetts, Maryland, New Jersey, Virginia, and Wisconsin. Data from twenty respondents (from one to three people in each of the eight districts) were included in the study. Table 1 provides a brief description of each district's data warehouse.

Table 1. School District Data Warehouses Studied (Data collected fall 2003 through spring 2004)

District size and location	Year data warehouse piloted or launched	Initial rollout user group	Groups who have access	Data available	Technical information
District A: over 62,000 students, 131 schools, nearly 5,000 teachers	2001: teacher portal "MyBPS" 2004: full-fledged data warehouse with live data	Teachers	Teachers, principals, school data teams, instructional technology staff, coaches, research and evaluation staff District was working on expanding the system to include district personnel for this year and then roll out to parents next year	Student information: MCAS data, test results, report card information, attendance, schedule, assessment data	Contracted with IBM, MySQL 2000 server
District B: 274,000 students, about 250 schools, staff of over 38,000 (17,000 instructional staff)	1996: piloted unsuccessfully to teachers in 3 schools 1997 piloted successfully in 21 schools 1998: went live to entire district	First rolled out to teachers, unsuccessfully Re-introduced to top-level administrators, then launched in several pilot schools successfully	All: district administrators, central office staff, all school-based staff, parents, students	Student info: demographics, enrollment, attendance, special programs, test scores, course transcripts, transportation, more Other information about: human resources, courses, lesson plans, school performance, finance, maintenance	Contracted with IBM, AS400 server; uses Brio (Hyperion) for reporting
District C: 25,000 students, 32 schools, 2,832 staff (1,928 teachers)	2003 (September): development began 2004 (early February): first rollout went live	District administrators	District personnel, principals Preparations were underway to roll out to teachers for September 2004	Student information, finance information, data to address NCLB reporting requirements	IBM Insight at school/ server xSeries 345
District D: about 20,000 students, 26 schools, over 1,800 teachers	N/A	Division of testing, data analysts	Various central office staff divisions, including curriculum and instruction, special services and staff development, and, recently, principals	Financial and personnel info: employment/termination, attendance, demographics (ethnicity, gender, degree attained), Whole School reform model information Student info: enrollment, attendance, Whole School reform model, demographics (ethnic, gender), classifications (LEP, SPED), student performance (state- and district-mandated test)	IBM, MS SQL 2000 server
District E: About 55,000 students, 53 schools, 7,000 staff (3,000 teachers)	2001	Technology, research and evaluation staff	Central office staff, principals, vice principals, counselors, instructional coaches, teachers	Standardized-test scores, demographics, special education information, attendance, student grades, scheduling information	MS SQL 2000 server

District size and location	Year data warehouse piloted or launched	Initial rollout user group	Groups who have access	Data available	Technical information
District F: 103,000 students, about 200 schools, 12,000 staff (5,000–6,000 teachers)	2000	District administrators, principals	Principals, school secretaries, learning teams, literacy coaches Teachers can get access but data security often prevents it	Student information such as test scores, proficiency level, enrollment data, list of courses, incidents, and demographic information For principals: transportation, assessments, calendar, administrators' bulletin	Oracle, MS SQL, Hyperion (Brio)
District G: 13,000 students, 20 schools, 2,250 staff, (900 teachers)	2000/2001	District administrators, principals	District administrators, principals	Longitudinal student-achievement data, test scores, benchmark data, attendance reports, enrollment information, and demographic information Qualitative data such as annual surveys on satisfaction given to principals, students, and parents.	Contracted with IBM, IBM Warehouse, AS400 server; uses Microstrategy for reporting
District H: 66,000 students, 80 schools, 8,459 staff (4,492 teachers)	2001/2002	District administrators, principals	District staff, (e.g., testing), principals	Student test scores, demographic information required by No Child Left Behind (race, ethnicity, gender, language proficiency, citizenship status, disabilities) Preparing to expand system to link to human resources and finance information	Contracted with IBM, Oracle Warehouse; uses OracleAS Discoverer

Appendix B

Interview Questions for District Users of Data-Warehousing Products (October 2003)

1. What is the name of the vendor and the data-warehousing product(s) you are using?
2. Describe the process for choosing your current vendor. Why did you/your district decide to pursue data warehousing in the first place? What was the selection process like? How did you learn of your vendor's product? *[Prompt for colleagues, checked references, Web, other.]*
3. Describe the development and start-up process. How long did it take to develop the data warehouse? What challenges were there? *[Prompt for cost, cleaning data, data formatting, technical support.]*
4. Describe your experience using the data warehouse.
   a. Accessibility: Is it widely available or just accessible by a few users? Who uses it? *[Prompt for central office staff, school administrators, teachers, a small set of trained staff, etc.]*
   b. User-friendliness: Does it take a lot of background knowledge/training to use the data warehouse?
   c. Customization: Does it generate only standard reports or can it generate custom reports? Are teachers able to generate reports? How easy is it to modify what variables are included? Was the district able to establish a menu of customized reports to meet the needs of its different groups of users?
   d. Format: Where is the data warehoused *[Web, PC, mainframe]*? What challenges does this format present, if any?
   e. Maintenance: How often are updated data put into the warehouse? How easy is it to put district data in the appropriate format?
5. Describe the kind of data you can extract from the data warehouse. *[Prompt for compliance reports, disaggregated data, specific categories.]* What can you do with your data-warehousing system that you couldn't do before?
6. What actions have you taken as a result of information generated by the data-warehouse system? Please provide an example. *[Prompt for improved teaching and learning, professional development, and/or communication].*
7. Describe the kinds of support you get when inevitable glitches occur. *[Prompt for training, technical assistance and support, ongoing consultation.]*
8. What advice do you have for other districts considering the purchase of a data-warehouse system? *[Prompt for planning advice, questions to ask, key features to include.]*

Appendix C

The Schools Interoperability Framework Initiative

Launched in 1999 by Microsoft, the Schools Interoperability Framework (SIF) initiative has set as its mission to develop a set of exchange standards to ensure that all K–12 instructional and administrative software applications are able to share data quickly and dynamically. Compatibility between and among technology reassures school district leaders of their long-term investments in newer technology or in upgrading older operational systems.

In recent years, implementing SIF industry standards has been challenging, for three main reasons. First, the development, maintenance, and resource costs of compliance were simply too much for some smaller software companies to handle. Second, for school districts, purchasing all new SIF-compliant software was unrealistic and not cost effective. Third, although the significance of having different programs, platforms, and database designs working as a "unified system" has always been understood by many of the networked software companies (namely, those involved in the parent organization of the broad network of vendors called the Software and Information Industry Association, or SIIA), it has been difficult to engage them in the competitive K–12 industry, where preserving their software products are key to their success. (*The Work behind SIF's Framework* 2002 V; Levine 2002a II-A; Mercurius et al. 2004 V; Brittain 2003 V).

Some progress has been made; last year, the Open Group was established as a nonprofit entity independent from the SIIA. The SIF compliance certification program was launched by the Open Group in 2003 to promote a systemic and formal way to improve the interoperability of educational software systems. This progress, coupled with the focus on NCLB requirements, has some – although unfortunately not all – leading software companies in the K–12 technology industry working quickly towards SIF-enabled software. (See <www.sifinfo.org/index.asp> for more information.)

With these improvements, data-warehousing software allows diverse software technologies at the school district (student-information system, transportation, etc.) to interact and eliminates the common issue of duplicate data entry records (e.g., when a student moves from one school to another). SIF-enabled software simplifies the exchange of a wide variety of data by making all systems speak the same language, instead of "translating." The SIF specification is based on the World Wide Web Consortium (W3C)–endorsed Extensible Markup Language (XML), which is not linked to a specific operating system or platform. XML defines common data formats and rules of interaction between architectures. A critical component of SIF is a specification for a Zone Integration Server (ZIS) – the data integration broker between applications that support SIF. For data warehousing, this allows far less time spent on cleaning and standardizing information and the ability to move data into the warehouse more quickly.

Annenberg
Institute for
School Reform

Brown University
Box 1985
Providence, RI 02912

T 401.863.7990
F 401.863.1290