So you think your K-12 school district needs a data warehouse?

By Gee Kin Chou

Many K-12 school districts think they need a data warehouse. There are now many vendors to choose from. Some offer customized solutions, some are "off-the-shelf"; there are hosted services while others run on a local server. All vendors claim a "user friendly" interface so why do you hear stories of systems gathering dust through lack of use? What should you be looking for? How do you start? How can you ensure your technology spending is going to make the biggest difference where it matters?

Pinpoint what problem you are trying to solve

First you need to step back and ask yourself what is driving your effort? Is it frustration with the difficulty in meeting regulatory reporting requirements? Is it lack of reliable information to make decisions? Is it the feeling of flying blind – student performance is not improving but I can't pinpoint the reason. Implementing any significant piece of technology is time consuming and expensive. You can only do so much at any one time. So you need to be sure that what you are doing is going to solve your most urgent problems.

The superintendent and the teachers will want to see different kinds of data

I have often found that the motivation to implement a data warehouse is driven by a general drive to implement "data driven decision making". But you have to ask yourself what data are we talking about and what decisions are we trying to make? Different groups will give different answers to this question. While the desire is to meet the needs of every user group, deciding which groups get priority will largely determine how the data warehouse and the other parts of the data infrastructure should take shape.

Teachers are focused on the students they are teaching

A classroom teacher focuses on maximizing the performance of the students in her current class roster. She needs to know the current learning needs of each individual student and how she can tailor her instruction to improve the performance of each student. Her data needs are relatively narrow. She needs to know how each student is doing and what she needs to teach to help each student achieve his/her learning objectives. The key data sets are assessment results, perhaps grades, a current roster, and a well-organized and effective curriculum that is linked to standards.

Principals need to look at both their students and their teachers

Principals and department heads need similar data to teachers, only they need to see across the entire school. Their goal is to maximize the performance of all their students, and all their teachers. They may need extra cuts of the data to help them evaluate staff performance as well as student performance. And they may want to compare the performance of the students in their school with students from other schools to find opportunities to learn from others.

District leaders want to see trends over time and across the district

The academic leaders (e.g. chief academic officer, associate superintendent of curriculum & instruction, content supervisors/directors etc) need to evaluate the efficacy of each program by looking longitudinally over time and across schools and teachers. Usually, their focus is not individual students or specific classes or schools. They need to look district-wide across a diverse set of data that might influence the outcomes of each program or affect the performance of a specific student subgroup. They will be looking for indicators of strengths or weaknesses in the curriculum, or in the professional development program, which can only be gauged by looking at the performance of specific student groups over several years.

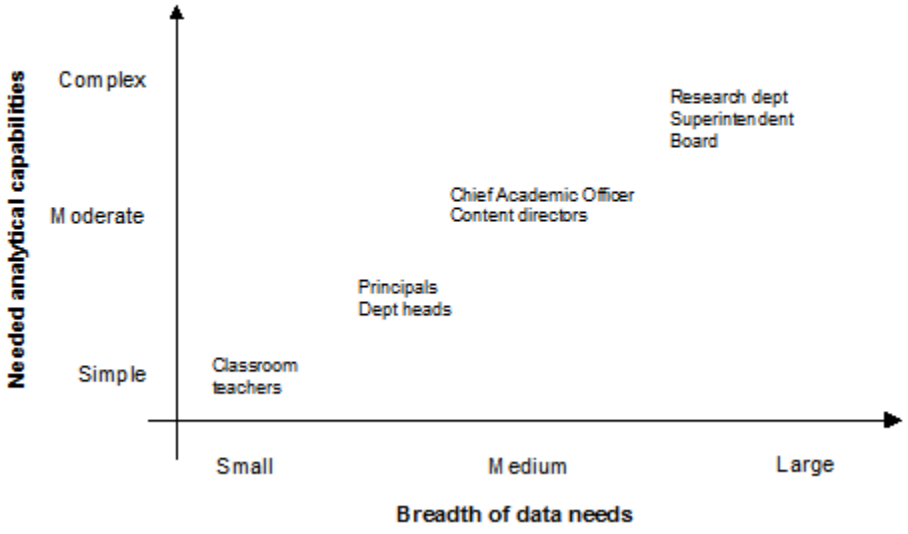

Figure 1: Different user groups have different needs

The superintendent must be able to answer almost any question about the district

The superintendent and board members have data needs that are much less predictable. They will be looking for ways to improve the performance of the

district to all stakeholders: the mission of every district is to maximize the performance of every student but there may be factors outside the classroom that influence student performance. For example, the job satisfaction and retention of teachers, the financial health of the district, the efficiency of the transportation system etc all can contribute to student performance. They need a way to look at data in all areas of the district's operations and determine how to do things better. They may need to look for correlations between seemingly unrelated sets of data. And they need to respond to questions from outside agencies or the media, and provide background to support governance or financing initiatives.

So the needs of the district can be segmented into the needs of each discrete user group, as shown in Figure 1.

More computing power usually means higher training costs

In addition to defining the breadth of data each user group needs to access, and the kinds of analyses they need to perform, there are several other considerations that will help clarify what kind of system the district needs.

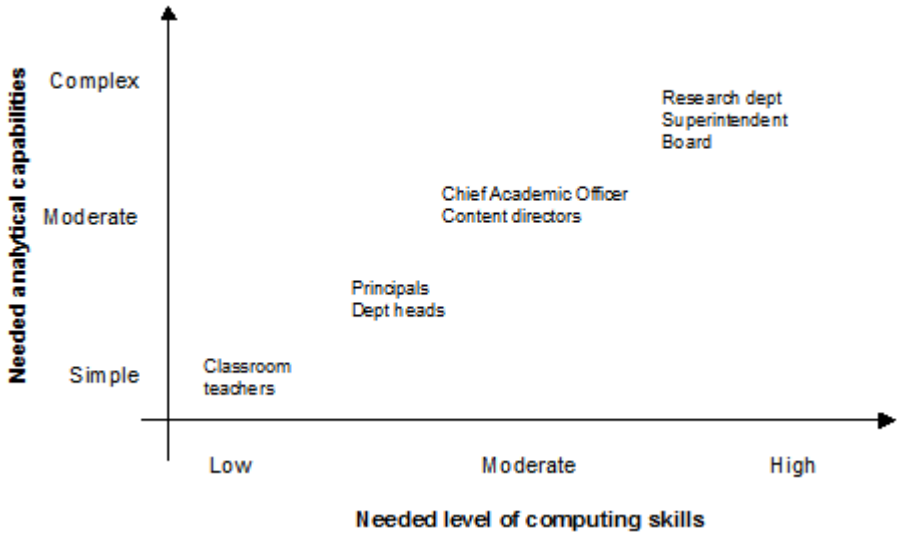

Figure 2 More complex analyses demand higher level computing skills

Nearly all vendors describe their applications as "user friendly". However, users with varying levels of computing skills will have differing expectations of "friendliness". What is user friendly to someone in the research department may be bewildering to a classroom teacher. The most powerful tools enable almost total flexibility, allowing users to see any relationships among the accessible

data. However, most classroom teachers will not have this expectation. They will appreciate tools that allow them to see most of what they need very easily, and not expect much more. In general, more powerful analytical capabilities require a higher level of computing skills. This relationship is shown in Figure 2.

The benefits of computing power have to be weighed against the costs of training. This is a particularly important consideration for classroom teachers who are the largest user group; spending the time and resources upfront to pinpoint what the teachers need (and don't need) will enable the selection of a tool for teachers that does not require an inordinate amount of professional development time.

Principals and many district leaders, including the superintendent, will need to answer relatively complex questions that will require the capabilities of the more powerful tools; however, they may not have the skills, or have the desire to learn the skills, to use these tools. This situation needs to be accommodated. One solution might be to rely on the research and IT department to create reports that they specify. Another solution might be for an skilled user within the district to create report templates that can be "served" up on the district's intranet – every individual user that can access the template will see data that is specific to his/her user sign on. They won't have the freedom to perform their own analyses, but this may be an acceptable tradeoff.

Not everyone needs to see data updates immediately

Time is another factor that needs to be considered. Classroom teachers are focused on their current roster. They need to see changes within 24 hours; in particular, they need very quick turnaround of classroom assessments so that they can adjust instruction for each student. As the focus broadens to encompass the entire school or the entire district, users tend to be less concerned about immediate events but rather on longer term trends. The superintendent and the cabinet may not need data that are frequently refreshed, but they may need to go back further in history.

You should envision an entire "data architecture", rather than a single product

In general, no single data tool / data warehouse is optimal for all user groups. Products that have the breadth and depth of data for the research department and the superintendent, and the flexibility for ad hoc analyses, tend to be too difficult to use for the typical classroom teacher. At the other end of the spectrum, tools designed specifically for teachers tend to lack the capabilities that district leaders require. The objectives of each user group are different and the "one size fits all" strategy ultimately tends to satisfy nobody. Rather than talking about implementing a data warehouse, districts should be thinking about an entire "data architecture".

Districts should establish a vision that encompasses the needs of all user groups, establish a target architecture that realizes this vision, and then implement the architecture piece by piece according to priorities. Figure 3 shows what the data architecture might look like. There are products that can make each user group very happy, but districts that rely on a single data-reporting tool for teachers, principals, and all central office users usually are disappointed and frustrated. Note that while many districts will speak in general about implementing a "data warehouse", strictly speaking, it is not unusual for the classroom teachers only to have access to a smaller data mart, rather than the data warehouse itself. This "split" architecture often can better meet the diverse needs of teachers, principals and the central office.

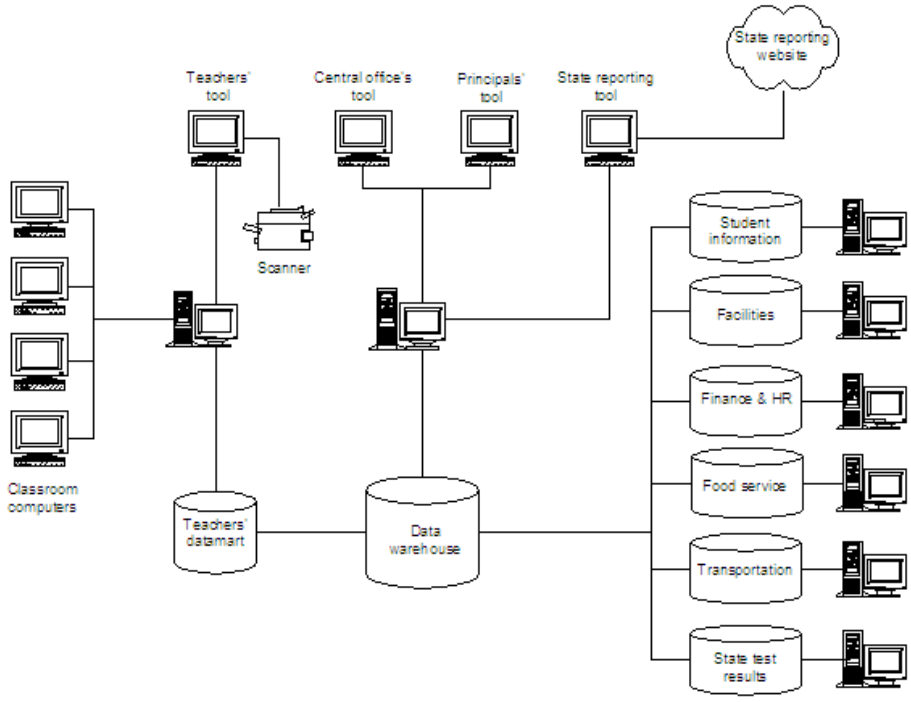

Figure 3: Sample data architecture

Don't try to do everything at once

Districts should be prepared to implement the entire data architecture in phases over a number of years. They need to consider which pieces of the total picture will have the largest and quickest payback; there may be a rationale for giving priority to a relatively simple system that teachers use every day rather than a more powerful system that can provide the answers to all the superintendent's questions. In the end, it might be more cost effective for the research department to manually find the answers to the more complex, but infrequently asked questions, rather than spending the money on technology for the superintendent and other senior executives to find the answers themselves. Keep in mind that in addition to the initial "hard" costs (e.g. software licensing

fees, cost of new hardware etc), districts must budget for the initial and recurring "soft" costs associated with every new system; these include the costs of training, content development, project management and annual maintenance or subscription fees.

Understand what you have before you buy anything new

Many districts will not be starting from a clean slate. They will need to assess how their current systems should fit into the target architecture. Usually, this process will begin with an audit of the existing systems to provide a picture of all the pieces, how they work (or do not work) together, and to highlight any issues such as age, capacity, performance, usability and data quality. Some existing pieces may fit nicely into the future architecture, while other pieces might better be phased out and replaced.

Follow a disciplined process and involve the users from start to finish

A cross-functional team that includes Curriculum & Instruction, Assessments and Technology should drive the effort. If the team is looking at tools for teachers and principals, then teachers and principals also should be part of the team.

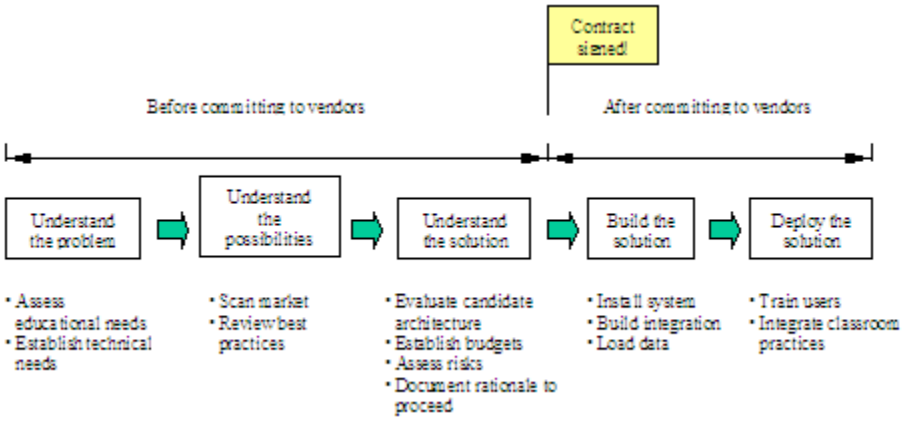

Figure 4: Process flow for a typical technology project

Before committing to any product vendor, districts must understand their educational needs and their technical environment, explore the possibilities, identify all the pieces that need to come together to make the technology effective, and plan their resources accordingly. Most projects will require significant commitments of staff time from departments in addition to Technology, and for some teachers to take on additional responsibilities, and good project planning will alert everyone to the need at the onset. The success of most technology projects depends more on good execution than on the technology itself, and districts that follow a process like the one shown in Figure 4 are much

more likely to succeed. Without this discipline and commitment of resources, technology purchases often are a waste of money.

Epilogue: It usually takes more than technology to solve the problem

Technology only helps make sense of what data are available; the pertinent data have to be gathered. For example, in order to individualize instruction, teachers need timely formative performance data reflecting the immediate learning needs of each student; the director of Curriculum & Instruction may need to identify and track the performance of specific student subgroups to evaluate its new literacy program; the superintendent may need per pupil spending for each school to analyze the economical factors affecting student performance. A data warehouse and the data analysis and reporting tools that have been the focus of this paper will only yield useful information if districts also are prepared to implement the processes and systems to gather these kinds of data.

While there are technologies that can facilitate data gathering, obtaining useful data still will depend largely on manual input from district staff: e.g. teachers will have to write test items and design learning interventions, school secretaries will have to keep student demographic data updated, the accounting department will have to establish specific accounts for revenues and expenses etc.

In the end, the right technologies will only take districts halfway; they must commit their personnel to provide reliable and effective inputs, and they must establish a culture where teachers and administrators are inquisitive and constantly using data to drive their work.