The Classroom Sentinel: Supporting Data-Driven Decision-Making in the Classroom

Mark K. Singley & Richard B. Lam
IBM T.J. Watson Research Center
1101 Kitchawan Road
Yorktown Heights, NY 10598
914-945-2210
{ ksingley, rblam } @us.ibm.com

ABSTRACT
Whereas schools typically record mounds of data regarding student performance, attendance, and other behaviors over the course of a school year, rarely is that data consulted and used to inform day-to-day instructional practice in the classroom. As teachers come under increasing pressure to ensure success for all of their students, we are attempting to provide tools to help teachers make sense of what is happening in their classrooms and take appropriate proactive and/or remedial action. One such tool is a Web service we've dubbed the Classroom Sentinel. The Classroom Sentinel mines electronic gradebook and other student information system data sources to detect critical teaching and learning patterns and bring those patterns to the attention of the teacher in the form of timely alerts. In this paper, we introduce the notion of classroom patterns, present some examples, and describe a framework for alert generation and delivery.

Categories and Subject Descriptors
K.3.1 **[Computers in Education]**: Computer Uses in Education – *computer-managed instruction*, I.2.1 **[Artificial Intelligence]**: Applications and Expert Systems – *office automation*, H.4.1 **[Information Systems Applications]**: Office Automation – *workflow management*, K.4.3 **[Computers and Society]**: Organizational Impacts – *reengineering.*

General Terms: Design, Human Factors

Keywords: classroom pattern detection, data-driven decision making, data integration, alert generation, teacher cognition.

1. INTRODUCTION
In order to comply with mandates calling for more accountability regarding student achievement, most notably the No Child Left Behind Act of 2001 [6], K-12 (i.e., compulsory education) teachers are feeling increasing pressure to ensure that all of their students succeed in the classroom. Our work is predicated on the observation that, whereas schools typically collect reams of data regarding student academic performance, attendance, and other kinds of behavior over the course of the school year, little of that data is used in a proactive way by teachers to influence their day-to-day instructional decision-making in the classroom. This strikes us as a ripe opportunity to bring information technology to bear in helping teachers become more responsive and reactive to the emerging and ongoing student performance trends in their classrooms.

WWW 2005, May 10-14, 2005, Chiba, Japan.
ACM 1-59593-046-9/05/0005.

How can information technology be used to help schools attain the goal of improved achievement for all students? There are a variety of possible approaches, many of which have already been tried. One possibility is to build instructional software that teaches core academic subjects such as math and reading directly to students, but such solutions almost by definition bypass the teacher and are not well-integrated into the classroom. Adoption of this type of courseware in traditional classrooms has been spotty at best [7]. More recently, there have been numerous data visualization efforts that seek to provide school administrators with school and district-wide summaries of student performance, e.g. the School Report Card [1]. Such systems definitely have their place, but they provide macro-level views that are very blunt instruments for affecting change in day-to-day classroom activities. We have chosen instead to focus on a point of leverage that has largely been ignored to date: empowering the teacher in the classroom. We provide professional tools to the teacher to help him or her become a more effective manager of student information and a more data-driven day-to-day instructional decision-maker.

The Classroom Sentinel is a Web service we have built that mines electronic gradebook, attendance, and other student information system (SIS) data sources to detect critical teaching and learning patterns in the data. Once a pattern is detected, its existence is brought to the attention of the teacher in the form of a timely alert, at which time the teacher may choose to take some appropriate responsive action or not. The Classroom Sentinel is designed as a middleware Web service that can draw on arbitrary third-party software vendor data sources as long as the structure of the data source and some form of access is made available to the service.

The Classroom Sentinel is one component of the Teacher's Workbench [5], a Web application that provides an integrated suite of tools to help the teacher manage the classroom. In addition to the detection of student performance patterns and the delivery of alerts, Teacher's Workbench includes an electronic daily planner which supports linkages to instructional resources, and a student profiler which integrates and organizes demographic and longitudinal student information from a variety of data sources. In this paper, however, we are focusing on the Classroom Sentinel component of the system.

The overall goal of the Classroom Sentinel is to improve day-to-day instructional decision-making by providing teachers with a finer-grained, more timely understanding of the ever-changing patterns of student proficiency in their classrooms. Teachers currently use student performance data rather passively to support the assignment of grades rather than actively to inform future instructional practice. Given the large number of students they teach and the many demands made on them, teachers typically have little time to pore over their gradebook and planbook data searching for interesting

patterns to inform future practice. Furthermore, they have little access to longitudinal and/or demographic data about their students to lend perspective on what's happening and to calibrate instructional goals. To remedy this situation, the Classroom Sentinel actively mines the teacher's data and delivers notifications, thereby amplifying the teacher's ability to analyze and understand student performance in a timely way. The teacher is put in a position to respond quickly when it counts, rather than at the end of an entire unit of instruction, end of semester, or end of year when it is too late to affect outcomes.

A variety of technical challenges confront the continued development and deployment of the Classroom Sentinel. As mentioned previously, the Classroom Sentinel is part of a larger Web application known as Teacher's Workbench, which has been deployed in prototype form in six schools across two large urban school districts. A successful deployment involves reading and integrating disparate data sources, many of which may not be fully "ready" for integration, performing efficient event-driven data analysis for large numbers of teachers and students, and perhaps most importantly, defining critical yet useful and usable teaching and learning patterns that teachers can understand and act upon.

Teacher's Workbench bears a family resemblance to a number of other commercial and academic systems that offer classroom management functionality, such as the Microsoft Class Server [3], and Chancery Student Management Solutions [2]. These systems typically provide strong support for administrative data reporting, and may provide tools for teachers to peruse the data upon which administrative reports are based. However, to our knowledge these systems do not provide the kind of active data mining, automatic notification, and resource retrieval provided by the Classroom Sentinel.

2. ANATOMY OF AN ALERT

Figure 1 shows an example of a pattern detected by the Classroom Sentinel and the corresponding alert delivered to the teacher through Teacher's Workbench. All alerts are delivered in the "Heads Up" section of the Teacher's Workbench application, where the alerts are organized by class and student. Alerts appear in a list that is refreshed on a periodic basis in response to the availability of new data and the subsequent detection of new instances of patterns. In this particular alert, a pattern has been detected in the performance of a particular student, Carl Farella (this alert is based on actual data collected at one of our partner's sites but all of the names have been changed). The system has detected that this student has done reasonably well on homework and other assignments (average of 75) but has done significantly worse on tests (average of 46). In addition to reporting the basic pattern, the alert offers a couple of possible explanations, namely, that Carl has test anxiety or possibly poor test-taking skills, and offers links to a number of resources for remediating these problems. Finally, the alert offers the teacher the opportunity to email this alert to the student and/or his parents, and include links to either all or some subset of the resources.

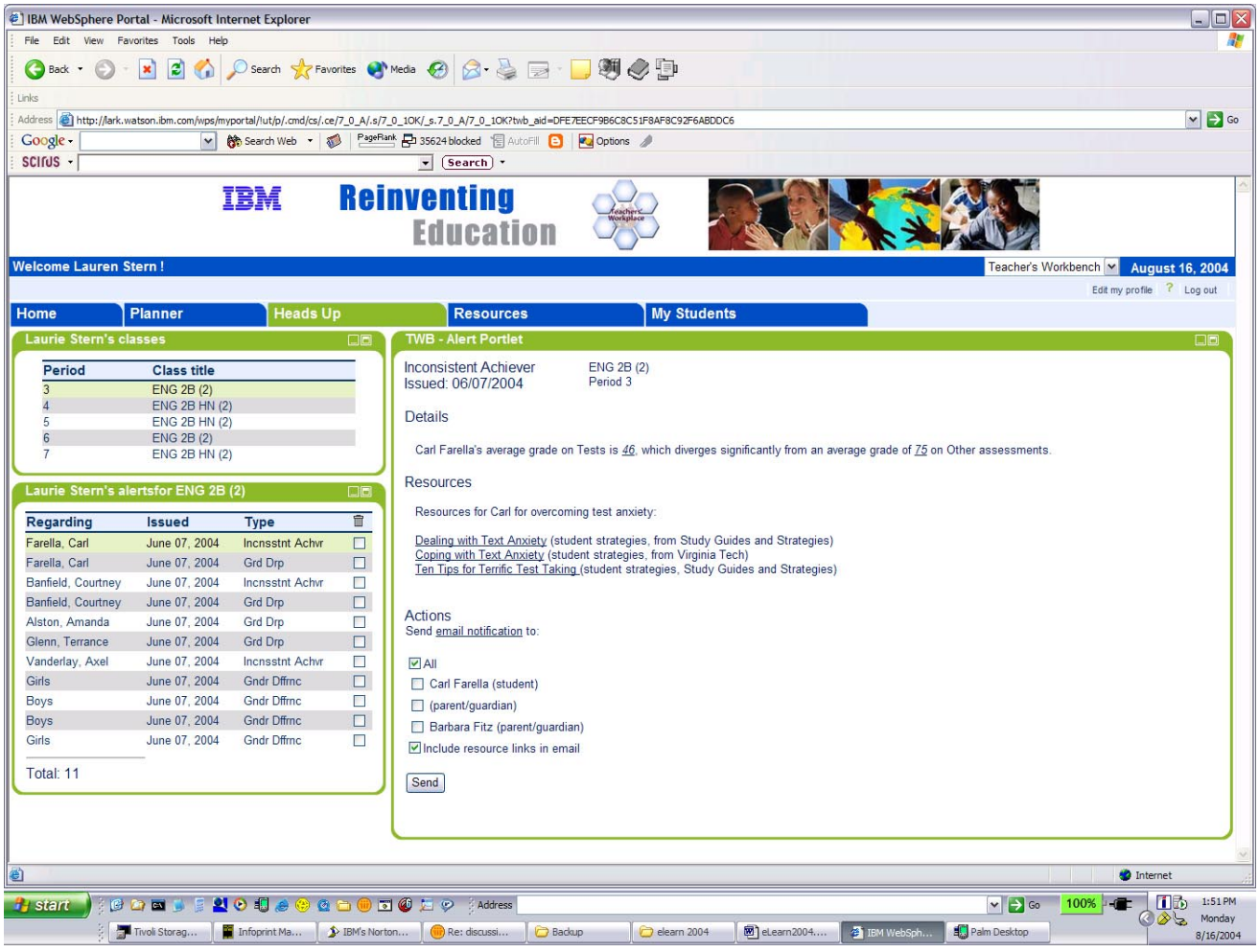

Figure 1. Teacher's Workbench alert delivery.

From this example, it can be seen that an alert is composed of three main parts: 1) an observed pattern of performance, 2) a set of possible explanations, and 3) a set of possible responses. Given the relative paucity of data available for our analyses, and the philosophical impossibility of making definitive statements about causality in non-experimental classroom situations, the best we can do is offer possible explanations for the patterns observed. It is up to the teacher, who has a much better grasp of the situational and contextual factors, to choose the best possible explanation offered, or to reject them all and posit another.

In addition to these three basic parts, defining a new alert type also involves specifying the following:

Event triggers. In order to maximize the efficiency of the system and minimize the amount of data analysis required, alert types are associated with various event triggers. One such trigger would be the appearance of new assessment data. In our example, the "Inconsistent Achiever" pattern would be re-matched whenever the grades for a new assessment become available to the system.

Precedence relationships. Some patterns are preferred over others as the best characterization of what is happening in the classroom. For example, one of our patterns, "Student Performance Drop," detects when a student's performance has fallen precipitously. Another, "Difficult Assessment," detects when the average grade on an assessment is significantly lower than the averages of other assessments. These are potentially competing characterizations of the same phenomenon, and when both are detected, "Difficult Assessment" is more parsimonious and should be preferred.

State relationships. Once an alert type has been issued for a particular student, the system allows other alert types to be turned on or off in recognition of the student's new "state." As a somewhat degenerate case of such a state relationship, the delivery of an alert type for a certain student may disable the delivery of any more alerts of that type for that student (this would be the case for our example alert). As a somewhat more interesting case, after a "Student Performance Drop" alert has been issued for a particular student, a whole new set of alert types may become activated that serve to scrutinize the student's behavior carefully. In this way, an "at-risk" student can be put on watch and monitored closely.

Aggregation relationships. If similar patterns have been detected for more than one student, it may be useful for the teacher to receive an aggregate alert that lists all students that exhibited the pattern rather than individual alerts for each student. In this way, the teacher can take remedial action on an entire subgroup of students at once. The system supports the specification of rules for the aggregation of alerts.

User-defined parameters. The system allows users (teachers) to turn alert types on and off, and to specify various parameters and thresholds. Figure 2 shows the Teacher's Workbench Alert Settings panel that teachers use to control the alert generation process.

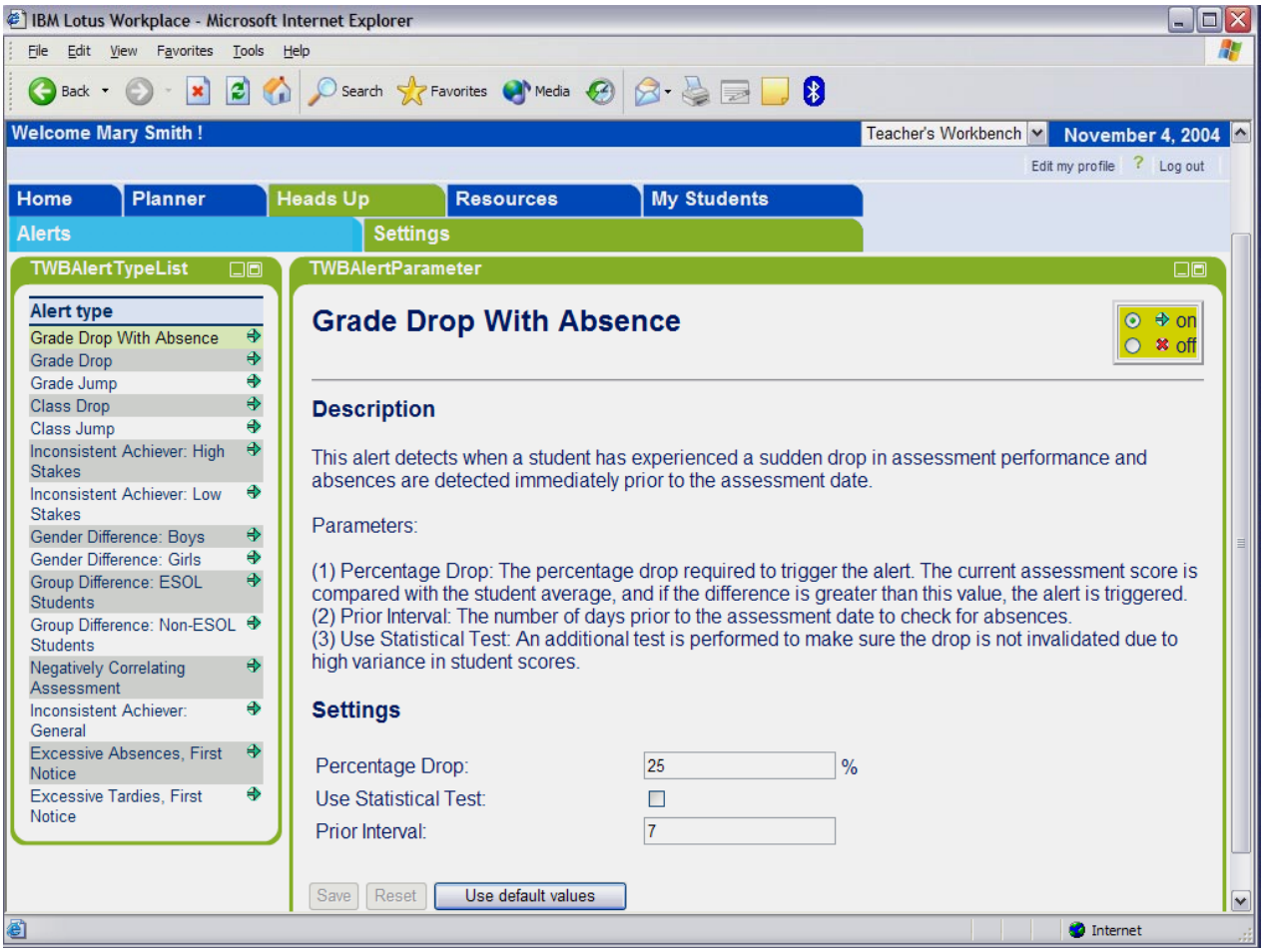

Figure 2. Alert generation control panel.

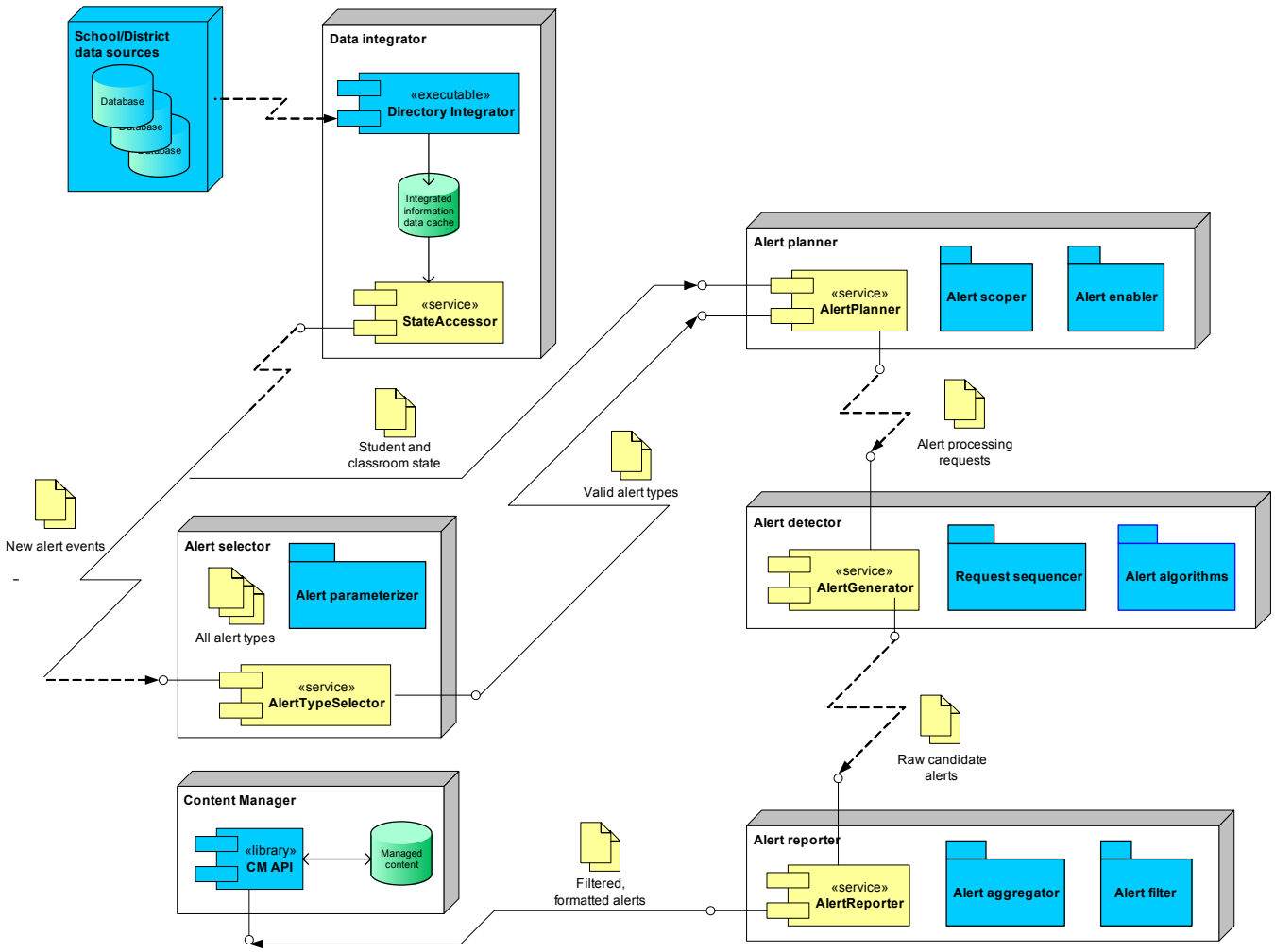

Figure 3. Alert generation framework.

In the example shown, the user can specify what percentage drop is necessary to trigger the Grade Drop With Absence alert, and whether an inferential statistical test should be applied to validate the difference. Also, the user can specify what length of prior interval the system should check for absences associated with the drop in performance.

Alerts are predefined according to the elements described above and programmed into the system in an alert library. There are currently about twenty such alerts defined operationally in our system, and we are working to implement more. We work with the teachers at our partner sites as well as IBM education consultants to generate new ideas for useful patterns, and to identify possible explanations and appropriate "best practice" resources to respond to the patterns.

3. ALERT GENERATION FRAMEWORK

Figure 3 provides a schematic view of the alert generation process and its constituent software modules. In the description below, we proceed from start to finish of the process in a roughly clockwise traversal of the figure.

First, the Data Integrator module reads gradebook, demographic, and other kinds of student data from various school district data sources. This data is reformatted and organized in a standardized data cache for easy access by the remaining alert generation modules. The Data Integrator module can provide information about the current state of student and classroom data as well as identify new classroom "events" that necessitate new alert generation requests.

Next, the Alert Selector module selects those alert types that are triggered by the current set of new events for further processing. From this set, the Alert Selector retains those that have not been turned off by the user. The remaining set is parameterized by a mixture of default and user-specified parameters as directed by the user.

The Alert Planner module takes the set of triggered, parameterized alert types and generates a queue of alert processing requests, which can be considered the "plan" for alert generation. An alert processing request specifies not only the alert type but also the "analysands" of the alert, i.e. the set of students, subgroups of students, etc. for whom the alert is applicable. This determination is made through consideration of both current student/classroom state and the "scope" of the alert type. So, for example, if a particular alert type is enabled only for a student if that student has already received an alert of another type, the Data Integrator module is queried at this point to determine whether a particular student is in

fact eligible for the alert. In this way, lists of enabled and disabled analysands are assembled for each alert type in the alert processing request. The type of analysand assembled is determined by the "scope" of the alert. All of these alert properties (enablement relationships, disablement relationships, scope, etc.) are defined declaratively when a new alert type is authored.

The Alert Planner passes the queue of alert processing requests to the Alert Detector module. This module cycles through the set of processing requests, deferring those for whom a potentially superceding request exists downstream in the queue, and skipping those for whom a superceding alert has already been generated. If neither case holds, the appropriate generation algorithm is executed for the current alert type for each of the analysands. If the pattern characteristic of the alert type is detected, a new "raw" alert has been generated.

Finally, the set of new raw alerts is passed to the Alert Reporter, a post-processing module that performs alert aggregation and alert filtering as specified by the user. The polished alerts that emerge from this module are stored in Teacher's Workbench content manager along with aligned instructional resources for eventual retrieval and perusal by the user.

4. A FLEDGLING ALERT TAXONOMY

Where do alert ideas come from? We generate new ideas for alerts in two ways:

1. Working backwards. Here we accept as a given the kinds of data that might typically be available in a current classroom. Then, with the help of the teachers and administrators at our partner sites and education consultants, we try to imagine what kinds of alerts might be interesting and/or possible given the available evidence.

2. Working forwards. Here we start with a theoretical framework that attempts to enumerate and organize all possible factors that have been shown to affect learning outcomes in children [5]. This framework includes such diverse factors as parental involvement, lesson quality, and student mastery of prerequisites. Then we ask what evidence might bear on the measurement of each factor. Finally, we consider whether this evidence is available, and if it is not, what might be necessary to collect it.

For practical reasons, we have chosen to focus primarily on the first strategy (i.e., "working backwards") in our initial deployments of the system. The alerts we have implemented are based on actual data currently collected by the districts and residing in electronic form. But in parallel, we are working with the school districts to define new kinds of alerts that necessitate new data collection and storage practices in the districts. This is the kind of work that will truly revolutionize classroom practice and lead to the most dramatic realization of the Classroom Sentinel vision.

What makes for a good alert? Of course, the most important criterion is that the alert should detect an important performance pattern in the classroom that the teacher wants to recognize and act upon. But there are other, more pragmatic, concerns that we face in the successful deployment of an alert. Any new alert proposal should satisfy the following constraints:

- data upon which it is based is available (see the "working backwards" strategy above)
- has sufficiently high frequency to justify the overhead of creation and deployment
- is not obvious (but even an obvious alert may be very useful if we can help the teacher respond to it, thus amplifying the teacher's ability to manage classroom situations)
- is not arcane or obscure (teachers must resonate to it and must understand the data behind it)
- is non-controversial (but No Child Left Behind necessitates the monitoring of certain subgroups)
- can be delivered in a timely manner so that it is actionable

Given these constraints, and with the advice and consent of our teacher partners, we have generated a preliminary set of alert ideas and organized them into a taxonomy. This preliminary taxonomy, along with a few examples in each category, is shown in Table 1.

In order to guide and prioritize our efforts, we conducted a survey of our teacher partners, asking them to rate the usefulness of 39 candidate alert ideas. Twenty-five teachers and administrators from both partner sites completed the survey, which asked respondents to rate the potential usefulness of an alert on a standard five-point Likert scale, where a rating of 1 was "not at all useful" and 5 was "very useful." The 39 candidate alerts were distributed fairly evenly across all the categories and included the examples shown in the table.

The middle column of Table 1 shows the results for each category. Interestingly, the highest-rated category was "Classroom Management," followed closely by "Student Performance." Lagging at the bottom were the categories of "Student Opportunity" and "Subgroup Performance." The highest rated alert in the entire set was "Student has an in-school suspension," a Classroom Management alert, with an average of 4.8. It appears that, aside from the increasing demands regarding student achievement, teachers are still preoccupied with the more prosaic yet essential tasks of classroom management, such as simply keeping track of the many students under their charge. Interestingly, teachers were not that interested in receiving reports about differences in subgroup performance, even though NCLB requires attention to these patterns of performance. One possible explanation is that teachers do not have differentiated instructional tools and strategies readily available yet for dealing with these kinds of differences. Another explanation is that NCLB reporting requirements have not had time to be filtered down from the administrative level to the practicing teacher level.

Table 1. Preliminary alert taxonomy with examples.

Category	Rating	Examples
Student Performance	4.1	Student grade has dropped significantly from past assessment. Student has performed below average in more than one class this week. Student did poorly on the assessment and completed less than average number of homeworks. Struggling student has performed above average on assessment.
Student Trait	3.9	Student is performing below average on exams and quizzes but average to above average on other assignments. Student is performing below average on out-of-class assignments but above average on in-class assignments. Student should perform (average, below average, above average) based upon their past standardized test scores. Student was excessively absent last year.
Student Opportunity	3.2	Student is within range to increase marking period grade by a letter on next assessment. Student performed above average on quiz and may be able to tutor other students for the upcoming test. Student has achieved a 90 or better for all quarters and is a candidate for an honors or AP program.
Subgroup Performance	3.2	English as a Second Language (ESL) students performed below average on a particular math assessment although they have above average to average scores on other math assessments (high verbal content?). Emotionally Disabled (ED) students have a higher than average absence rate. Girls did not perform as well as boys on a particular assessment (speeded?).
Teacher Performance	3.5	Last assessment has low correlation with other assessments administered this quarter. Students scored significantly lower/higher than last assessment. Students that performed average or above average on topic quiz scored below average on topic test. It has been two weeks since the last assessment. Assessments involving Standard 1.2 had the lowest student average.
Classroom Management	4.2	Student has an in-school suspension. Student has been absent/tardy n times. Student is missing a grade for an assessment and the makeup deadline is tomorrow. Student has a birthday tomorrow.

5. USABILITY AND PRACTICAL ISSUES

As mentioned above, Teacher's Workbench and the Classroom Sentinel have been deployed in six schools in two large urban school districts for about a year. Although we have struggled with a variety of technical issues such as the integration and subsequent analysis of disparate and sometimes messy school district legacy data sources, perhaps our greatest ongoing challenges are in the realm of usability engineering. We cite just a few recurring usability issues here:

First, our experience has shown that the number and frequency of alerts generated for the teachers must be finely tuned. As mentioned earlier, teachers are already stretched to their cognitive and attentional limits and have little capacity available to consider analyses of student performance. The alerts must be timely, insightful, and of a reasonable number so as not to overload the teacher. We have found that alerts that are accompanied by useful instructional resources to remediate the problem are more warmly received than those that provide no such support. As shown above, teachers have control over individual alert parameters and thresholds, which can be tuned to make the alerts more or less sensitive and thereby more or less numerous.

The deployment of a tool such as the Classroom Sentinel raises a number of thorny privacy issues, for both teachers and students. Historically, the classroom has been a teacher's personal and private dominion, which has been both a boon and a bane to the practicing teacher. Although the classroom can be a teacher's sanctuary, it can also be a place of profound loneliness and professional isolation. A tool like the Classroom Sentinel reaches into this private dominion and detects patterns of performance that can be served up either for the teacher's personal and private use, or shared among fellow teachers and administrators. Although sharing such information should provide rich opportunities for mentoring and professional development, the sometimes adversarial relationship between teachers and administrators may make such opportunities unwanted. Our policy has been to let the individual teacher control their own level of sharing.

Aside from teacher privacy, a tool such as the Classroom Sentinel impacts student privacy as well. If tapped into the appropriate data sources, the Classroom Sentinel can provide longitudinal information about student performance in past years, as well as a more complete picture of student performance across classes in the current year. Such information provides teachers with a much richer, contextualized view of their students, which should lead to more informed and principled interventions. But some may see value in the current situation where teachers have no preconceived notions and students are afforded a "fresh start" in each class at the beginning of every school year.

Finally, the Classroom Sentinel can provide analyses of student performance based on race, ethnicity and gender, the kinds of analyses required by the No Child Left Behind Act [6]. But our observation is that the teachers have not been particularly interested in these analyses to date. Our explanation is that there is a lack of instructional strategies and resources readily available for remediating such differences. A tool such as the Classroom Sentinel highlights the need for more basic research in this area.

6. FUTURE WORK AND CONCLUSION

The Classroom Sentinel is part of Teacher's Workbench, an IBM Reinventing Education 3 grant project. We are currently beginning the final year of a three-year research and development effort. Currently, a pilot is ongoing at two grant sites involving approximately 40 users. In terms of further development, we are enhancing the efficiency and extensibility of the alert generation framework, extending the library of alerts, and improving our instructional resource alignment strategies in order to suggest more useful and effective responses to the performance patterns detected.

We face a number of ongoing challenges in the successful development and deployment of the Classroom Sentinel. First, our Web-service middleware approach involves forging working partnerships with the vendors of the end-user applications whose data we wish to tap, most notably electronic gradebooks. Second, the IT infrastructures of the schools we work with are often not prepared to support the rich data integration we require. Third, and perhaps most importantly, we need to work with our teacher partners to define a new kind of teacher practice that is much more centered on the ongoing analysis, monitoring, and response to classroom performance patterns.

7. ACKNOWLEDGEMENTS

We thank Nancy Aitken, Rick Boehme, Lei Kuang, Kevin Ho, Cliff Kramer, Mark Laff, and Jung-Mu Tang for their extensive help in implementing and deploying Teacher's Workbench.

8. REFERENCES

[1] Baker, E.L. (1998). Understanding educational quality: Where validity meets technology. November, 1998, Educational Testing Service Policy Center Report.

[2] Chancery Student Management Solutions website: http://www.chancery.com/prod_sms_features.html

[3] Microsoft Class Server website: http://www.microsoft.com/Education/ClassServer.aspx

[4] Singley, M.K., Boehme, R., Kuang, L, & Lam, R.B. (2003). Teacher's Workbench: Pattern Detection and Resource Retrieval in K-12 Education. In Proceedings of E-Learn 2003 (Phoenix, AZ, November 2003), Association for the Advancement of Computing in Education, 115-119.

[5] Singley, M.K., Fairweather, P. & Schnitz, J. Toward a theoretical framework for classroom pattern detection and instructional alignment. In preparation.

[6] U.S. Department of Education (2003). No Child Left Behind: A Parent's Guide. DOE Publication #EAT0060P.

[7] Wenglinsky, H. (2000). How Teaching Matters: Bringing the Classroom Back Into Discussions of Teacher Quality. October 2000, Educational Testing Service Policy Center Report.