GIS in the Pittsburgh Public Schools

Chris Temple
Pittsburgh Public Schools

Glenn Ponas
Ken Sochats
Visual Information Systems Center
University of Pittsburgh

ABSTRACT

Pittsburgh Public Schools (PPS) has formed a collaborative arrangement with the Visual Information Systems Center (VISC) of the University of Pittsburgh to research and develop creative solutions that address the issues raised by the No Child Left Behind Act. The joint PPS/VISC project, called VIPER (Visualizing Information for PPS Evaluation and Research) has developed research and communication tools that seamlessly integrate GIS and other visualization technologies. VIPER provides relevant, timely and understandable information to teachers, administrators, parents, and students for decision making related to district, school, class, teacher, and student performance. Through VIPER, GIS tools are available not only for research and administration, but also in the classroom. PPS has acquired a site license for ESRI's ArcGIS so that every district stakeholder has access to these tools. This paper provides the rationale for the VIPER project, reviews the project's history and developments, and outlines the plans for future development.

INTRODUCTION

The Pittsburgh Public Schools (PPS) serves approximately 31,500 students in 86 schools within the City of Pittsburgh's 94 neighborhoods. With a 2005 operating budget of $530 million, the district employs over 5,100 employees, including over 2,500 teachers and 210 administrators. The district faces the same educational challenges as other major urban school districts, including a 65% poverty rate, a large population of students with special education needs, and differing learning styles and needs of diverse racial and ethnic groups. No Child Left Behind-driven accountability measures indicate substantial gaps in achievement for each of these subgroups, and have increased the pressure on District leaders to determine the factors that contribute to these gaps and implement prescriptive measures to close them.

In the past decade, there has been an average decrease of 2% per year in the number of students enrolled in PPS. The reasons for this decline are numerous, and include an overall decline in city population, and an increase in the number of charter schools and enrollments in private schools. The decline in membership has increased the number of empty seats in the district's schools. In response to declining enrollment, the District instituted a resource realignment process through which eleven schools have been closed in the past 6 years. Even with the closings, the district schools in 2004 were operating at approximately 74% of capacity, and some schools were operating at only 40% of capacity. As a result, there has been an increased interest in developing new methodologies to implement long-term strategic resource realignment (i.e., school closings) based on data and predictive modeling.

In support of student learning and strategic decision-making, PPS has built and implemented an educational data enterprise environment and comprehensive, solutions for collecting, aggregating, storing, and disseminating information and resources to all district stakeholders. Data from student information, instructional management, learning resource management, PeopleSoft, transportation, and other systems flow into common data repositories where they can be analyzed using a variety of applications. These solutions run on a state-of-the-art technology infrastructure, including fiber-based gigabit WAN connectivity, gigabit connectivity switched to every desktop, wireless access, world-class lights-off data center, multiple server clusters, and over 50 terabytes of NAS and SAN storage.

THE VIPER PROJECT

Concurrent to the development of a state-of-the-art technology environment, PPS added "home-grown" and "off-the-shelf" informational and data analysis applications to its solutions portfolio. These applications were chosen based on their ability to serve specific functions. Two examples are:

- Real Time Information (RTI)
  A home grown application, RTI manages Enroll-Reenter-Withdraw (ERW) transactions and permits teachers and administrators to view pertinent demographic, assessment, attendance, and special education data and standard reports in tabular format.

- SchoolNet Account (Account)
  A purchased product, Account provides standard and customizable reports on a subset of student achievement and demographic data in both tabular and chart formats.

The solutions portfolio of applications proved to be effective provided that they were used for their designed purpose. As the application user base became more mature, however, there emerged a greater demand for solutions that stretched the boundaries of these applications to the point that their functionality was not sufficient. Two examples of this phenomenon are increasing student achievement relative to the No Child Left Behind legislation and "resource realignment" (the PPS term for school closings). The solutions portfolio was found to be deficient in four areas:

1. The applications were not robust
   Analysis relative to NCLB and resource realignment required new analysis techniques beyond those native to the solutions portfolio. These analysis techniques would have to be hand-coded into existing software by either in-house programming staff or by vendors at substantial expense.

2. The applications were not flexible
   NCLB and resource realignment necessitated the integration of information from

disparate data stores. Due to the nature of software development for the K-12 vertical, integration of new data elements cannot be done “on-the-fly”, and would be possible only after customization of Extract-Transform-Load (ETL) processes.

3. They lacked spatial analysis capability
   Both NCLB and resource realignment required the capability to represent spatially the locations of students and to overlay layers of data onto those locations. In most K-12 applications, spatial data, such as addresses, can be sorted, searched, and filtered, but not represented spatially

4. The individual applications were unable to communicate with each other
   In order to facilitate solutions to NCLB and resource realignment solutions, the solutions portfolio needed a layer of interoperability, in which data subsets could be transported to other analysis tools for subsequent analysis without losing their identity and integrity. Applications in the K-12 vertical are built mostly on proprietary databases, and lack the interoperability found in most other verticals.

The PPS Office of Information and Technology (OIT), led by then Chief Technology Officer Elbie Yaworsky, recognized the emerging issues and in late 2003 invested over $450,000 in direct and indirect funding into research and development. This amount represented approximately 7% of the total operating budget of OIT in 2004. OIT created a Strategic Data Initiatives (SDI) group within the Office of Information and Technology. The group consisted of a program manager (Glenn Ponas), a GIS solutions architect (Christopher Temple), and the funding of the Visual Information Systems Center of the University of Pittsburgh School of Information Science. This cooperative team developed the Visualizing Information for PPS Evaluation and Research (VIPER) project. The team was tasked with developing, testing, and validating a new generation of software for visualizing student data. The software would need to have the following characteristics:

1. Robustness
   The toolset should allow additional analysis techniques to be added easily without hard-coding.
2. Flexibility
   As new data is added to centralized repositories, transparent management tools should permit the seamless integration of the new data into the analysis tool
3. Spatial Analytical Methods
   Schools and school districts should have access to the same types of spatially-accurate visualization tools as city planners and other social scientists.
4. Seamless Interoperability
   Data elements should not only retain their integrity and identity, but should be transportable via transparent interoperability mechanisms between multiple analysis tools.

The applications that resulted from the this collaboration would need to allow end users to create dynamic visualizations of data including GIS-based representations, multi-dimensional spatial and thematic analyses, and visual statistics The visualizations needed to be viewable at varying degrees of granularity and needed to paint robust pictures of student achievement, enabling practitioners and researchers to map the relationships between the myriad variables affecting student learning and to determine the impact of potential school closings.

THE STRUCTURE OF VIPER

Strategic and tactical educational decisions occur at five levels: the district, school, classroom, home, and community. Educators, leaders, parents, guardians, community members all require access to pertinent information via dynamic interfaces that are easy to use and that provide them with visualizations of data that facilitate their respective decision-making processes. Often, these decisions are made in the context of educational research findings in collaboration with university research teams. In order for the application suite to be successful, it would need to meet the analysis needs of all stakeholder and user groups at a level of interface complexity that matched their technical skills. As a result, the collaborative team sought two distinct but related application interfaces.

Research Analysis: The “Power/Expert/Research User” (PERU) Interface

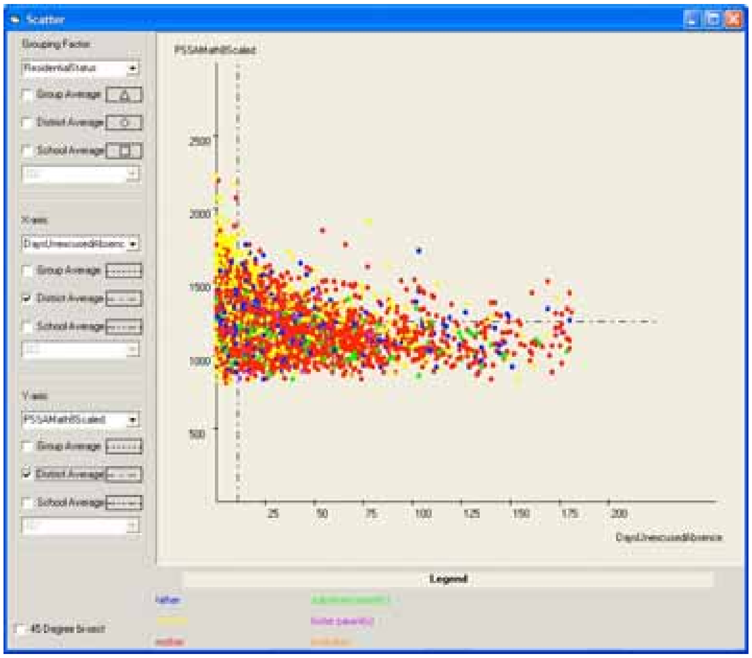
The design of the PERU interface began with the assumption that this group of users would A) require maximum functionality and access to all available data, and B) have the necessary content domain and research methodology knowledge base to use the interface effectively and efficiently, and to make scientifically sound conclusions. Ponas, in his doctoral dissertation (2005) describes the interface as a tool for the emerging “Dynamic Iterative Action Research” paradigm. The process can best be described as an information “pull”, where the content domain expert uses his or her knowledge of the domain to guide the iterative processes of drill-down, aggregation, visualization, and element-level analysis. This methodology has many benefits, including but not limited to:

- Through this iterative process, the user “discovers” empirical evidence of relationships that heretofore may have been either hidden to statistical methodology or only assumed or implied by anecdotal evidence.
- Both trends and outliers are clearly visible through visualization, and the fact that data retains its integrity and identity allows users to explore *why* outliers exist and what factors contribute to a given data distribution.
- Multiple factors can be visualized together and separately, providing insight into the nature of the complex interrelationship between educational variables.
- Spatial and thematic analysis becomes a major component in the educational research process.

Example: Adequate Yearly Progress (AYP) in Attendance

The combination of a content domain expert, a sound research paradigm, and a robust, flexible, and interoperable interface provided fascinating and important insights into the educational process. Consider the current challenges of NCLB as it relates to attendance. Schools must meet an Adequate Yearly Progress goal of 95% attendance rate per year (AYP Attendance). This policy is based on the premise that getting students to attend school regularly is a major step toward improving performance. Does that premise hold true in a given school or district? Does the student who meets the 95% attendance goal (approximately less than ten absences in a year) perform better by virtue of that attendance? What other variables might influence achievement other than attendance?

A simple visualization provides some insight into this question. In the scatter plot below, mathematics achievement on the 8th grade Pennsylvania State System of Assessment (PSSA) standardized test (y-axis) is graphed against the number of absences a student has in the testing year (x-axis). Each dot represents a student. In this visualization, the points are grouped (colored) by residential status, where yellow = two parent household, red = one parent-mother only household, blue = one parent-father only household, green = substitute parent, purple = foster parent, and orange = institution. The vertical dotted line indicates the district average of 9 days absent per year. The horizontal line indicates the district average score for the assessment (~1,200).



In this analysis, the vertical line at 9-days absent also indicates the 95% attendance threshold; students to the left of the line (low absenteeism) meet AYP attendance goals, and students to the right (high absenteeism) do not. On one hand, given the preponderance of students with high absenteeism who perform poorly on the 8th grade math assessment, it is appropriate to say that high absenteeism has an effect on performance. On the other hand, the data indicates that students with low absenteeism are spread across the performance spectrum. Simply coming to school, then, does not guarantee improved performance.

Another interesting result of this analysis is the breaking of the “urban myth” surrounding performance in single-parent household. Traditional belief structures within society are children are better off living with their mothers than their fathers. The distribution of red (single parent mother) compared to blue (single parent father) in this data set sets off a “red flag” in the mind of the content domain expert, spurring further analysis The expert user applies standard statistical measures to the subsets identified through visualization, and finds that students from single-parent father homes perform on average some 50 points better than those in single-parent mother homes.

Example: Safe Walking Routes to School

Other questions of interest to expert users involve where to place safe walking routes to school.   In any given neighborhood, what are the effects of natural and man-made boundaries on the patterns of students walking to school? How does crime in a given area affect the safety of students as they walk to school? Can a district make more effective uses of its crossing guards to enhance student safety? Using the ESRI ArcGIS component of the PERU interface, the expert user can display multiple layers of information in a spatial visualization that guides the decision-making process. The figure below demonstrates this analysis relative to three schools.

ML King Elementary School, Allegheny Traditional Middle School, and Allegheny Traditional Elementary school all sit in the park setting of Allegheny Center on the North Side of Pittsburgh. Surrounding the park on the north, east, and west are residential communities. To the south are commercial development, sports stadiums, parking, and the Allegheny River. Crime hotspots in this area can be shown by overlaying city crime data onto the map. High traffic streets can be displayed.   The expert user understands the multiple and sometimes conflicting constraints associated with creating safe walking routes from neighborhoods to schools, and can use the tool to model walking distance, times, and potential numbers and locations of crossing guards. By applying traffic flow by time-of-day data, it also allows for better decisions on the placement of school bus embarkation and debarkation locations that permit easier parent pickup and smooth rush-hour traffic flow around the busy Allegheny Center Drive.

This visualization, when layered with census data, can also provide insight into population rate of growth or loss in the communities served by the schools in the picture, and how that rate of growth or loss will influence enrollment in the schools. ML King Elementary is a "feeder pattern" elementary school, meaning that its students come from the community that surrounds it. Allegheny Traditional Elementary is a magnet school, meaning that students from across Pittsburgh can apply for admission. The student capacity of King Elementary is approximately 800 students, and currently operates at less than 50% capacity. The student capacity of Allegheny Elementary is approximately 300 students, and currently operates at 100% capacity. By using population trend data from the census data layers, enrollment trend data from PPS, and housing trend data from the city, county, and state, the expert user can develop a GIS-based model that predicts future changes in capacity and enrollment. This data would prove essential in strategic decisions about whether to consolidate both elementary programs in the same building, thereby creating substantial building and resource savings and creating the potential to double enrollment in the Allegheny Traditional Middle program, which is also currently at full capacity and has a waiting list that is double its current capacity.

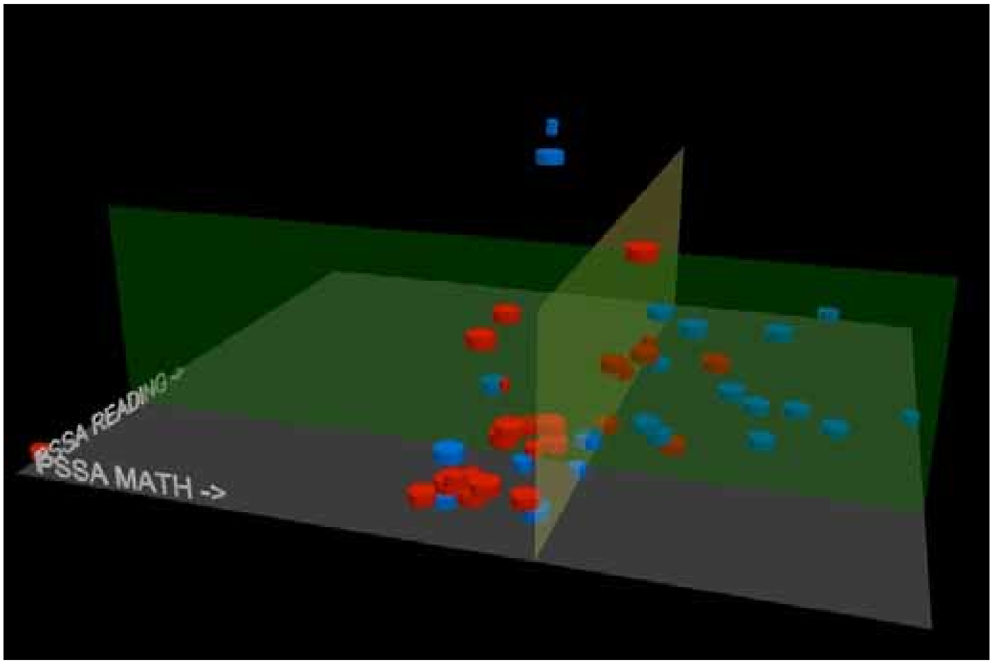
Example: Interrelationship between Performance in Math, Performance in Reading, Attendance, age of students, and Demographic Attributes

In the current high-stakes environment of tying federal funds to student performance, it is critical for districts to clearly understand the ever-changing nature of the dynamic performance, demographic, and attendance attributes of students and to leverage that understanding in support of increased student learning. The ArcScene application provides the expert user with a platform to observe the interrelationships between multiple variables simultaneously.

ArcScene, an application developed by ESRI as part of its 3-d Analyst extension, is primarily used for landscape and building modeling. In this case, however, Christopher Temple re-applied this tool for the display of non-geographic data in an interactive three-dimensional environment. This task was accomplished by creating a map projection landscape in which traditional latitude, longitude, and elevation coordinates were replaced by student data attributes to create an event layer buffered by a distance determined by the value of a specific student attribute and extruded by still another attribute. The result is six dimensions of information represented in three dimensions of space.

The replacement of traditional geographic information with "geography of performance" data is demonstrated in the three-dimensional scatter plot below. It contains data on PSSA math performance (x-axis), PSSA reading performance (y-axis), and number of absences (z-axis). The cylinders represent the age of students at the time of the test, with larger cylinders of larger diameters representing older students. The colors represent race, where red = African-American and blue =white. The sample size for this visualization is fifty ninth grade students chosen at random from one high school. All students took the test in the preceding year as 8th graders. The number of absences represents the total absences during the 8th grade school year. The semi-transparent green x-z and y-z planes indicate the district average PSSA math and PSSA reading scores respectively. The grey x-y plane is a datum plane that provides a visual reference point for the expert user to perceive the depth of the z-axis. The origin for this scatter plot is in the bottom-left of the datum plane.

The figure below represents just one angle of view for the 3-d scatter plot, which if used by itself for analysis has the potential to skew the user's perception of spatial location. This skewing is inevitable when the human mind attempts to analyze a single flat (2-d) representation of a 3-d perspective drawing, but can be mitigated by an expert user with good spatial visualization and perception skills who can internalize the 3-d representation by viewing it from different angles. A benefit of using the ArcScene for this specific visualization is the ability of the expert user to manipulate his or her view of the visualization. The expert user can rotate the plot along any axis, zoom in and out of the plot, and "fly through" the plot, thereby mitigating potential perception and interpretation problems.



This visualization provides some fascinating insights into the interaction of the component variables. For example, one can clearly see the racial achievement gap by observing the preponderance of African-American students (red cylinders) in the area representing below average performance in both math and reading PSSA (left-front quadrant). The ages of the African American students appears to be higher (bigger cylinders) than white students. African-American students tend to cluster across all dimensions, while white students tend to be more evenly spaced across all dimensions. Number of absences appears to be a greater factor in low achievement among African-American students than among white students (red cylinders in the front quadrants are higher above the datum plane than in the back quadrants). The opposite is true for white students, where high absences appear to correlate to higher achievement scores (blue cylinders in the back quadrants higher above the datum plane than those in the front quadrants).

The three-dimensional scatter plot also provides the opportunity for the power user to turn the visualization on its axes in such a way that any one of the three axes can "disappear". The result is a two-dimensional scatter plot. For example, when viewing the visualization from a point of view that is perpendicular to the PSSA reading plane, the two-dimensional scatter plot of PSSA math to number of absences appears. Toggling between two-dimensional and three-dimensional scatter plots allows the expert user to cross reference results obtained in the 3-d view and the 2-d view to identify more quickly and clearly any effect of the third variable.

Interoperability of Multiple Analysis Tools

The previous examples demonstrate the power of individual data visualizations and interfaces in solving specific problems and answering specific research questions. Yet individual data visualizations represent only a fraction of the capabilities that a fully interoperable suite of visualization tools can provide relative to data analysis. Consider again the three-dimensional scatter plot above, particularly the two blue and three red outlier cylinders at the top that rise high above the datum plane. These cylinders represent fairly young white students and three average age African-American students with a large number of absences. By viewing the scatter plot from multiple angles, it can be determined that these students score at the district average in math, but well below average in reading. Does this suggest that for some students high absentee rate affects reading performance more directly than math performance? Are there common traits that these students have that might indicate potential causality for this disparity in performance? This visualization by itself raises but cannot answer these questions. It would be necessary to "gather up" all existing data on the students that the cylinders represent and seamlessly move those students and their data into other tools for further analysis.

This kind of interoperability is exactly what the joint SDI/VISC team created. By utilizing ODBC compliance and underlying Microsoft interoperability, the team created a data environment where all data about all students is available at all times for use by all components. Built-in data selection tools were harnessed where possible and modified or built where necessary in order to allow the user to harvest specific student sequence numbers. These numbers, as the primary identification key for the student, are transported from application to application in a custom-built "lens". The expert user can leverage all available student data and the power of the analysis tools in the new application suite.

Using the PERU interface, the expert user has a set of tools at his or her disposal to explore phenomena in a unique, powerful, and integrated manner. The diagram below represents one possible sequence in which an expert user might engage while exploring a specific problem or issue. Each figure represents a different view which is accessible at any time and at any level of granularity while working within the suite.

Figure 1 represents the ArcGIS View. Student sequence numbers are ported into the GIS interface by linking a geo-database to a Microsoft SQL database. Existing layers of spatial data in the ArcGIS tool as well as all student data can be applied simultaneously to student data layers. From the views created by these layers, questions can be both answered and additional questions generated. As is true with all views, individual or groups of data points can be identified and selected, and their sequence numbers copied to the lens. Those sequence numbers are then portable to all other applications in the suite.

Figure 2 is the Multi-dimensional Analysis View. Student points in this view are indicated by squares on a blank canvas known as a “world”. Data attributes are shown as circles. The interface places the student points in the world relative to the location of the data attributes using a mean-referencing protocol. As more data attributes are added, student points are pulled toward or pushed away from each new attribute based on the relative attraction or repulsion of the attribute relative to the data values. As is true with all views, individual or groups of data points can be identified and selected, and their sequence numbers copied to the lens. Those sequence numbers are then portable to all other applications in the suite.

Figure 3 is the Data Grid View. All available data about individual students within a given group or population is available in the data grid. This view permits sorting and filtering of student records by any attribute It also provides the expert user with a common tabular format for displaying data, which is often underestimated as a visualization tool.

Figure 4 is the Scatter Plot View. The scatter plot represented in this figure represents only one of a set of visual statistics tools available to the expert user, including bar charts, line plots, radial drawings, and other common visual statistical representations. The inclusion of group, school, and district averages as well as colorized grouping factors allows the expert user to increase dramatically the functionality of these visualizations by expanding the number of attributes represented.

Four Visualization Views of the Same 321 Student Data Records

Figure 1.

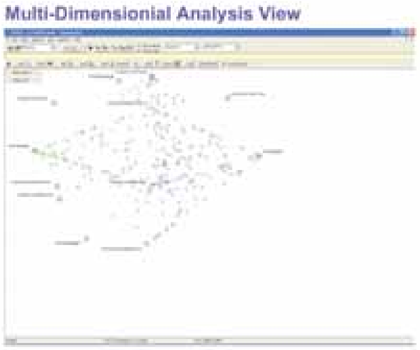


Figure 2.

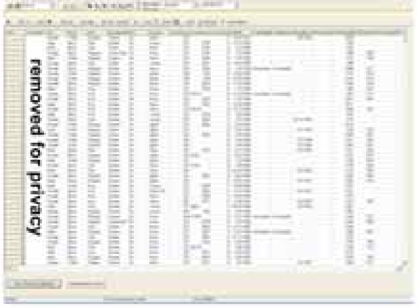


Figure 3.

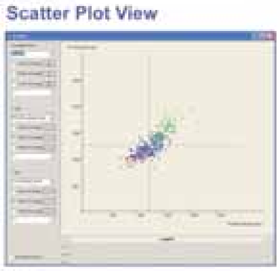


Figure 4.

Interoperability Example: Differences in Rates of Giftedness by Communities

In order to demonstrate the concept of integrated visualization analyses, consider the following example:

> *Two adjacent communities in Pittsburgh have populations differing greatly in demographic attributes of race and poverty. Representatives of the poor, predominantly African-American community have brought to the attention of district administrators that they perceive an equity problem relative to representation in the gifted education program when compared to the affluent, predominantly white community to their immediate south. The issue has stirred passions on both sides, and has created a need to determine the degree of such a disparity if one exists, the potential causes of such a disparity, and possible solution paths to correct the disparity.*

The expert user begins by spatially locating all students within the communities by using the ESRI ArcGIS tool (Figure 1 above). In this figure, triangles locate one or more students at a given address. Utilizing the data layering capability of the tool, the user overlays the variable of special education status using colors within the triangles (green = gifted, blue = not special education, red = severe special education). The disparity between the distribution of gifted students (green) is clearly evident, with a majority of students in the southern community identified as gifted, and with very few students in the northern community identified as gifted.

By turning on the aerial photography layer (not shown), the expert user recognizes that the disparity appears to occur along the man-made border of a busway that divides the communities. Crossing points over the busway occur every seven to eight blocks, with the fifteen foot high barbed-wire fence on both sides of the busway along sections between crossings After some archival research, the expert user determines that the road was once a set of railroad tracks that was the divider between a predominantly white and affluent community to the south, and the predominantly black and poor community to the north that served the affluent community. Here is a literal example of "the other side of the tracks". By overlaying school attendance patterns, the expert user determines that the busway was used as the enrollment border between two schools. By overlaying a zip code layer, the expert user determines that the enrollment borders throughout the city mirror the fifty year-old zip code borders. By overlaying demographic data such as race, poverty, and residential status, it becomes evident that the communities differ radically in makeup across all demographics.

Using the select and lens functions, the expert user moves the list of students to the Multi-dimensional Analysis Tool (Figure 2) and displays their locations relative to the attributes of number of days absent and tardy to school and performance on the math PSSA test. Clusters of students occur in this view relative to all variables. The visualization indicates that, generally speaking, these two communities have fundamentally different performance, absence rates, and tardy rates, with students in the white community consistently outperforming students in the African-American community. There are a number of outliers, however, related to black students who perform well enough on PSSA's to outperform the highest performing white gifted students. Yet none of these students are categorized as gifted. There are also a number of outliers related to white students who are categorized as gifted and perform consistently poorly on achievement tests. Are these outliers merely coincidence, or is there a set of not-yet-discovered attributes that link these students?

The group is then collected via the lens and viewed in the Data Grid View (Figure 3). The expert user iteratively sorts students by different attributes to look for patterns. Many patterns are discovered. For example, the subset of black students with high PSSA scores but not categorized as gifted also show very high grades in all subjects, exceptional behavior, and involvement in multiple after-school activities. These same students came from an elementary school that turned out few to no gifted students.

The group is again collected via the lens and viewed in the Scatter Plot View (Figure 4). The same students appear as blue (regular education) outliers in a sea of green (gifted education) in the upper-right quadrant of the scatter plot (above average math and reading PSSA scores). The expert user highlights those students and collects their student identification numbers. Armed with the information gathered during the process, the expert user can now work with the community groups and educational leaders to begin the data-informed decision-making process.

Power to the Masses: The Stakeholder Dashboard/Portal Interface

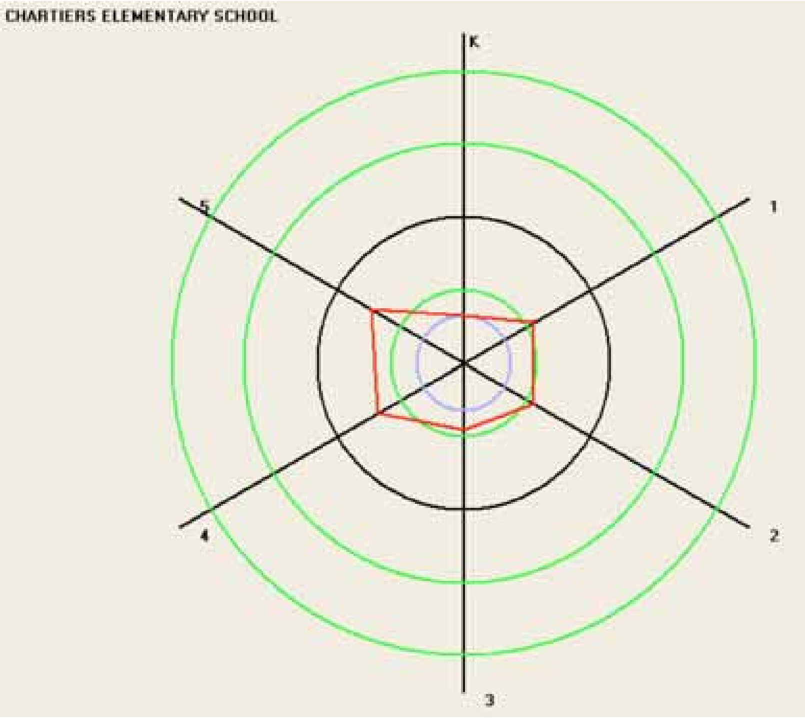
As noted earlier, power, expert, and research users have extensive content domain knowledge, practice-enhanced research experience, and a keen understanding of the potential misinterpretations that can occur in visualization-based analysis. They also are given the freedom and time to engage in lengthy and detailed explorations on specific problems. The reality for stakeholders such as teachers, administrators, parents, students, and community members is far different. Different stakeholders within and among stakeholder groups have different levels of content domain knowledge, research experience, and visualization theory expertise. Perhaps more important, most stakeholders are end users with limited time and require immediate access to specific sets of information.

Given these differences, the joint SDI/VISC team decided upon a portal-based approach to visualization for stakeholders. Delivered through a web interface, the portal is a “push” environment, where specific visualization tools and data are automatically delivered to stakeholders based on their FERPA/HIPPA-based need to know. Through the preceding work of the power/expert user (see above), specific visualizations and relationships were identified as having potential value to a specific stakeholder group. That information is delivered to the stakeholder via the District’s web portal and dashboard environments. The visualizations are dynamic, allowing stakeholders to change the parameters or scalars in the visualization, and to alter, add, or remove data from the visualization based on their specific data needs.

Icon-like images that incorporate the results of GIS analysis were designed to present this information to administrators, principals, teachers, parents and students through their existing system web interfaces. We called these icons desktop elements because they were developed to be a component on each stakeholder’s desktop interface. To eliminate clutter, a particular element would automatically appear on a user’s desktop if it were relevant to that user and if it showed some interesting result. Any of the desktop elements that don't merit desktop space can be accessed by a navigational aid.

An example of this process was the development of enrollment patterns by grade over time. Through the power/expert user process, it was determined that schools of varying rounds (1 round = 1 class of 20 students for each grade level) had varying levels of success in maintaining steady, evenly distributed enrollment across grades. By pushing a portal-based visualization tool displaying a specific school's distribution by grade level, the principals can see this information upon login, providing them with information in a timely manner for preparing budgets and schedules for the following year based on predictive enrollment trends.

The picture below shows a desktop element that might appear on a principal's or administrator's desktop showing the result of an analysis of the feeder pattern for an elementary school. In the icon, each black, labeled spoke represents an elementary grade at the school (K-5). Each concentric green circle represents one "round". A round is the maximum enrollment (20) in a classroom for each grade. The black circle represents the school's capacity. The red polygon represents the actual enrollment in each of the grades. The blue circle represents the projected enrollment for this school using the GIS model for service area and population density.



The desktop element incorporates descriptive, predictive and normative information. The users can quickly and easily see:

1. The school has one class per grade.
2. The school's capacity is two classes per grade.
3. The size of each grade varies widely.
4. The enrollment is steadily declining.

5. Long-term enrollment is predicted to be half of current enrollment.

From a principal's perspective the desktop element provides information and highlights areas for concern. In particular, the principal must focus on:

1. Staff planning so that teachers can effectively cover the wide area of grades and subjects necessary in an out-of-balance school.
2. Resource allocation of rooms, desks, equipment, computers, etc.

From the administrator's perspective, the desktop element raises questions such as:

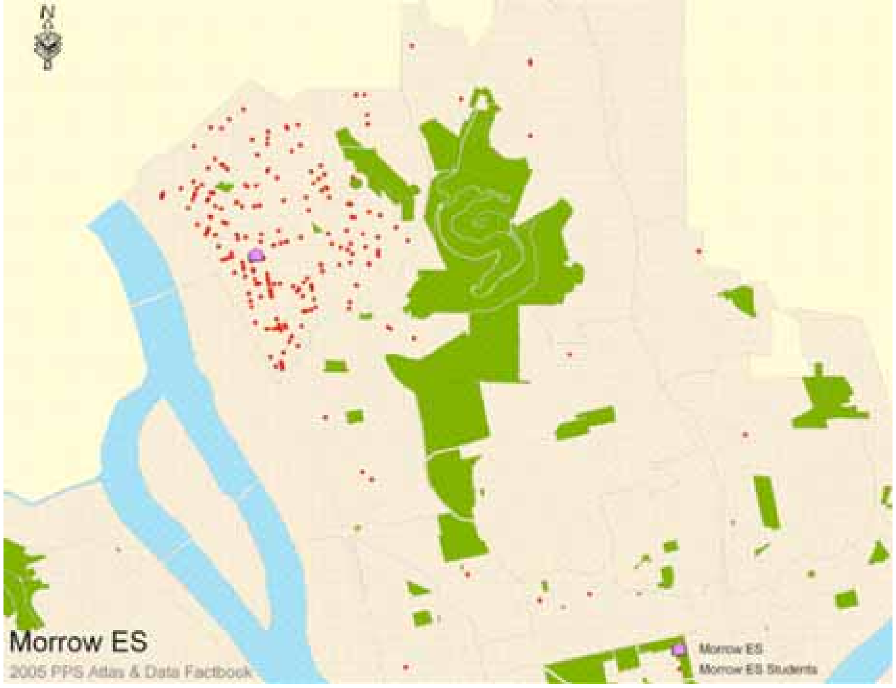
1. Should the feeder pattern for this school be adjusted to meet capacity?
2. Should the school be closed?

Should a user want more information, each desktop element can be "clicked" to lead to the underlying GIS displays and database analyses.

THE PITTSBURGH PUBLIC SCHOOLS ATLAS AND DATA FACT BOOK

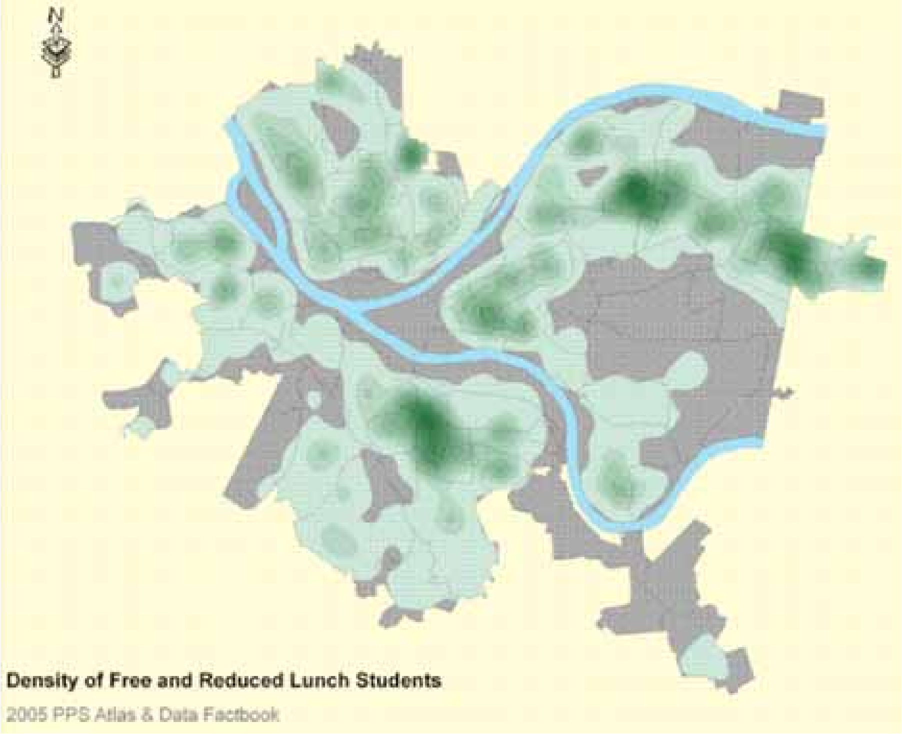
The Stakeholder Dashboard/Portal interface is an effective way of disseminating data to individual stakeholders in ways that allowed controlled access to individual student information as constrained by HIPPA and FERPA rules. There is also an on-going need to disseminate to the general public aggregated information stripped of individual student identifiers.

Each year, the project assembles and publishes The Pittsburgh Public Schools Atlas and Data Fact Book. The Atlas is produced in printed, CD and web versions. The current Atlas can be accessed at http://visc.sis.pitt.edu. The Atlas is comprised of structured sets of maps and data that chronicle the state of Pittsburgh Public School District in a standardized manner. These standards permit the district to perform longitudinal analyses on the data. The map information is actively stored in an ArcGIS geo-database attached to a standard Microsoft SQL Server database. Each Atlas contains maps and data that present information for the district, at each educational level (elementary, middle, secondary and center) and for each individual school. The map below shows the location of Morrow Elementary School and the relative location of students with respect to the school. All maps are provided at a variety of resolutions to meet the needs of various users.



2005 PPS Atlas & Data Factbook

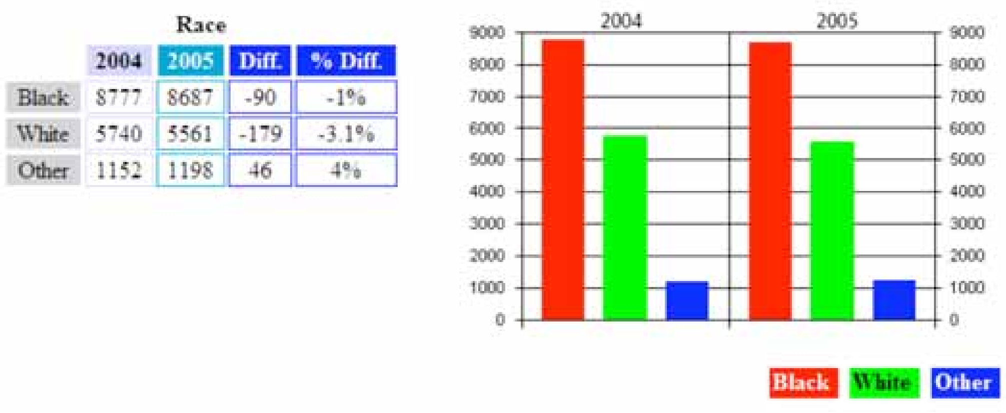
The Atlas also presents analyses of numerous student factors (e.g. race, gender, poverty, living arrangements, etc.) and their distribution in the district. Special attention is paid to those factors outlined in the No Child Left Behind Act. The map below shows the geographic density of students (darkening shades of green) in the district who participate in the Federal Free or Reduced Lunch Program. Maps are also available that show the combined effects of two or more of these factors (e.g. race and poverty).



Density of Free and Reduced Lunch Students

2005 PPS Atlas & Data Factbook

Supplementing the maps are numerous charts and tables of data relating to the district and its schools. Data tables and graphs are available for enrollment, race gender, socioeconomic status, and other relevant factors affecting student performance.



Race

	2004	2005	Diff.	% Diff.
Black	8777	8687	-90	-1%
White	5740	5561	-179	-3.1%
Other	1152	1198	46	4%

The Atlas provides as a service a number of maps that are helpful adjuncts to school and student information. Maps of streets, community centers, school district boundaries, neighborhoods, libraries, parks and other education-related facilities are included. Aerial and topographical maps are also included. The figure below shows a map of all of the public libraries within the boundary of Pittsburgh Public School District.

The current year's version of any Atlas version contains the maps and data for that year as well as the maps, data and analyses of the changes in the district over the preceding years.

GIS IN THE CLASSROOM

The next phase of using geospatial technology in PPS is the integration of GIS software tools into classroom teaching. Since the ArcGIS site license allows for the installation of the software on all PC systems in the district, the opportunity to expose students to the benefits of using GIS for such areas as:

- Basic geography instruction
- Mathematics and geometry instruction
- Basic computer and technology instruction
- Life sciences instruction (mapping specimens locations and environmental factors)
- Community and urban planning instruction

GIS software has the capacity to be an important method for maximizing the return on the district's investment in PC labs as well as our data storage resources. Information only has use if it is able to be analyzed and communicated. Students can learn with computers only if there is worthwhile material loaded on them. GIS software can help this district in both capacities.

The total costs of implementing curriculum level GIS solutions are more than just the software, however. Other direct and indirect costs include:

- One full time or two part time software support specialist(s)
- Approximately 10 Gigabytes of central server space in the preexisting server cluster to host and deliver the spatial information needed to utilize the GIS software
- The full time use of one high end PC workstation for administration of the GIS package, and for advanced administrative analysis
- Approximately 100 Megabytes of storage space on each PC that the software is loaded on
- Approximately eight hours of training time and associated workshop pay ($28/hr) for each classroom instructor who wishes to use GIS for classroom purposes

It is important to note that a comprehensive suite of tutorials, professional development materials, and online courses are included at no extra cost with the software bundle.

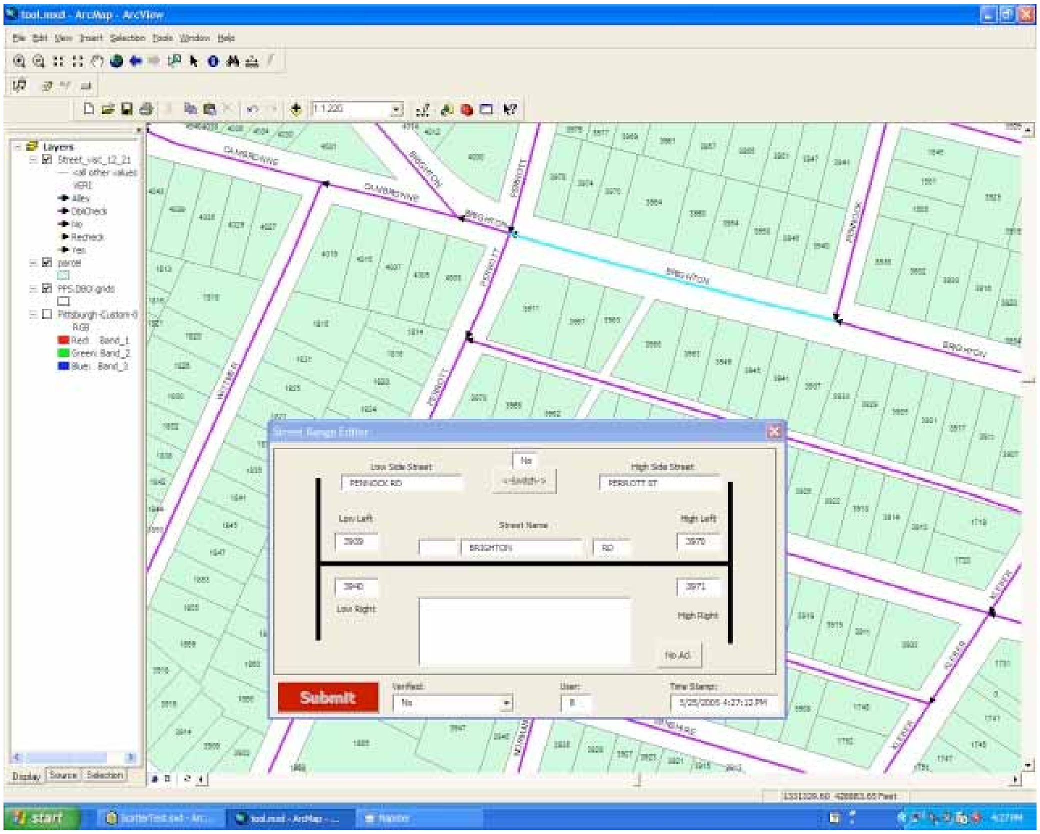
Funding for associated costs can come from general funds for technology. However, there are several attractive grant options that could fully or partially defray the cost of this program. These include:

- Pennsylvania Department of Education EETT wave 2
  http://www.pde.state.pa.us/ed_tech/cwp/view.asp?A=169&Q=100376
- The Toyota/National Science Teachers Association Tapestry Grant
  http://www.nsta.org/programs/tapestry/
- National Geographic Society Education Foundation Grosvenor Grant
  http://www.nationalgeographic.com/education/teacher_community/grosvenorgrants/
- Davis Demographics and Planning K-12 Public School District GIS Grant Program
  http://www.davisdemographics.com/grant/FAQ.htm

- National Science Foundation Geo-science Education Grant http://www.nsf.gov/pubs/2004/nsf04598/nsf04598.htm
-

In process of working through the student geo-coding process leading up to the 2004-2005 school closings, an important discovery was made concerning the available street range data. It was simply not good enough for the district's purposes. Data available from the City of Pittsburgh's Planning Division was spatially correct, but inaccurate in terms of numeric attributes. Tiger based street data from GDT was much better numerically, but too spatially imprecise for use at the neighborhood and block level. The solution to this problem came from a collaboration between the Visual Information Systems Center's Dr. Robert Regan and Pittsburgh Public School's Christopher Temple. Using the Planning Division's street shapefile as a starting point, a total revision of the street range data for the city of Pittsburgh was performed in eight months. The process began with Dr. Regan riding the streets on bicycle and reading the street ranges into a Dictaphone. Next, he transcribed his audio notations onto large format printouts, marking the high and low house numbers for each street segment as well as any changes to street names and topology. Mr. Temple then took the paper maps, and utilizing a custom ArcObjects interface built on top of ArcMap 9.0 and ArcSDE 9.0, supervised the entry of the data by PPS interns. These interns, who were Pittsburgh Public Schools seniors, did an excellent job of correcting the street data as well as gaining valuable experience in GIS technologies.

A view of the ArcObjects tool for street data entry:



Low Side Street: PENNOCK RD
No
High Side Street: PERROTT ST
Low Left: 3939
Street Name: BRIGHTON RD
High Left: 3970
Low Right: 3940
High Right: 3971
Verified: No
User: 8
Time Stamp: 5/25/2006 4:27:13 PM

SUMMARY

The joint PPS/VISC project has demonstrated the power of GIS to supply educational stakeholders at all levels analytic tools that facilitate the strategic and tactical decision making. AYP attendance, safe routes to school, and the correlation between performance and attendance are just a few examples of the flexibility of GIS to inform stakeholders and help shape educational policy.

The project has also demonstrated that GIS can be seamlessly integrated into an educational application environment. By building interoperability into the underlying data structure of the visualizations, information can be freely transferred to and from multiple visualization applications. In this way, the power of GIS is augmented by standard and novel data analysis techniques.

The bottom line is that GIS can make a substantial contribution to school districts' endeavors to ensure that no child is left behind regardless of where they are located.