



Compromising Data Warehouse Integrity:
Spatial Abstraction of Data

WHITE PAPER

Fear and Loathing in ...

Borrowing from Hunter S. Thompson, author of “Fear and Loathing in Las Vegas,” there is a growing sense of fear and loathing associated with desktop spreadsheets, especially among business executives who rely on information generated in these spreadsheets for critically important decisions. Executives fear the often-repeated phrase “We had an error in the spreadsheet,” and loathe even the annoyingly small inconsistencies in the information they receive.

Desktop spreadsheets were first introduced over 25 years ago, long before organizations created and managed databases for user-initiated reporting and analysis. Every time users copy data from an enterprise database to their desktop spreadsheet application, the integrity of the database is compromised. It is difficult to adequately control, even know, what happens on the desktops of individual users. Paraphrasing another reference to Las Vegas, “Whatever happens on the desktop, stays on desktop.”

To be sure, users love Excel®, which is an extremely flexible data analysis tool. Of course, the security of data resident on a user’s desktop is all but non-existent. Far more troubling is the abstraction of data that occurs in spreadsheet applications. Data abstraction creates inaccuracies, inconsistencies, and a loss of transparency that would never be tolerated by those responsible for database administration. Yet spreadsheet applications are used to make some of the most important decisions related to managing business performance.

The high risk and high cost of spreadsheet computing is drawing increasing attention, especially with the demands for regulatory compliance. A growing number of articles on the problems of spreadsheets conclude there is a pressing need to manage, control, and validate spreadsheet applications. Certainly, managing desktop spreadsheets is a necessary first step in addressing the problems with spreadsheets. This is a useful first step, especially with respect to the conventional belief that “Users will never give up their spreadsheets.”

It is time to challenge that conventional belief. Organizations should consider replacing desktop spreadsheets when: (1) the underlying data used in the analysis is managed in a large database; (2) it is essential to validate business rules used in analysis and reporting; and (3) decision making requires information sharing, consolidation, and collaboration among users.

eiVia has developed an online spreadsheet — more accurately, a comprehensive application development platform — that maintains data warehouse integrity, security, accuracy, consistency and transparency, while supporting user-directed analysis. eiVia’s application development platform addresses a fundamental issue with desktop spreadsheets, the spatial abstraction of data that desktop spreadsheets impose.

Spatial Abstraction of Data

The analytic power of the desktop spreadsheet is derived from the ease with which users can create, replicate, and combine mathematical formulas. When introduced, the spreadsheet was a significant improvement over the calculator. (In both cases, users created the data they were analyzing by typing in numbers.) The desktop spreadsheet remains an extremely useful tool when the user originates data. But much has changed in the last 25 years. Today, the data for many spreadsheet applications originate in enterprise databases. These databases are subjected to administrative disciplines very different from what is routine or even possible on users’ desktops.

“Spatial abstraction” is the term used to describe the renaming (abstraction) of data based on the spreadsheet cell location (spatial context). For example, a measure like “Unit Sales”, as stored in the database, becomes “B2” when copied into that cell of a user’s spreadsheet. Similarly, “Unit Price” becomes “C2”. Obviously, “B2” and “C2” are meaningless without the spatial context of the particular spreadsheet. The user can add another layer of abstraction by placing a label in an

adjoining cell. There is nothing linking this label to the formula, other than the spatial location of the formula and label. The formula "B2*C2" might be labeled "Revenue" by one user but "Gross Income" by another. Of course, many layers of abstraction can be created by using the results of one formula as an input to other formulas. While we all understand and even rely on this spatial abstraction of data, it is a source of errors and data inconsistencies. These errors often go undetected, surfacing only when the magnitude of the error poses a serious problem.

Because spatial abstraction links application logic to cell location, spreadsheet applications lack the data transparency that we expect from applications associated with a database. The common process of drilling down to lower levels of data in a database dimension is not easily supported in desktop spreadsheets.

Of course, if it is not necessary to link analytic and reporting applications to a database, then spatial abstraction is not an issue. But the reality is that more and more applications are driven from enterprise databases, especially data warehouses.

Most Data Warehouses Are For User Applications

The simple goal of data warehousing is to consolidate and organize information from multiple sources as an accurate time-variant data repository. In other words, the data warehouse is updated periodically to provide one source of reliable up-to-date performance information along with a complete history of performance measures.

Users rely on the data warehouse for reports, and in many cases, on the data that is copied into their desktop spreadsheets for analysis. In fact, many vendors of Business Intelligence applications provide capabilities to export data from online reports to users' desktops and spreadsheet applications. To some, this may be a useful feature; however, it is an acknowledgement that Business Intelligence has not addressed a very basic analytic requirement. The acronym OLAP, often associated with Business Intelligence, was derived from **Online** Analysis Processing, yet desktop spreadsheets are **Offline** Analysis Processing.

As databases continue to grow, desktop spreadsheets do not have the capacity to support analysis of large databases. The current release of Excel®, has capacity to manage one million rows of data on the desktop. While this is a significant amount of data, the real question is "Why manage this amount of data on the desktop?" Databases are updated routinely and therefore there is a need to continuously synchronize analysis with the underlying database.

True Online Analysis Processing

In a perfect world, the analytic flexibility of desktop spreadsheets would be available online, so that applications could be seamlessly integrated into collaborative workflows. This **online** spreadsheet would mimic the traditional row-and-column display and would match the flexibility of desktop spreadsheets, but would also provide greater data transparency and more online analysis capabilities. With these improvements, the online spreadsheet would be a true analytic interface to the data warehouse.

The ten critical requirements for online analysis are as follows:

1. Support multi-dimensional and multi-hierarchical databases that reflect the business model of the enterprise. Provide users with data transparency, allowing them to easily navigate the database from the spreadsheet interface.
2. Provide a library of calculations that is centrally managed and verified. Eliminate the dependence on spatial abstraction of data.

3. Synchronize security in both the database and application. Avoid downloading more data to users' desktops than is "displayed on the glass."
4. Allow users to work independently to model different scenarios, but provide an audit trail for all changes that users make to shared workspaces. Support information sharing and collaboration throughout business process workflows.
5. Facilitate the consolidation of data inputs and analysis performed by multiple users.
6. Allow users to invoke agent processes to highlight exceptions or generate alerts based on exceeding performance thresholds. These agent processes should also be used to trigger alerts when data entries are out of an acceptable range. Support bi-directional alerts that allow users to communicate actions that were taken as a result of an alert — in other words, closed-loop communications.
7. Provide cell locking, both within the spreadsheet, to model scenarios within constraints, and in a shared collaborative workspace, to enforce the locks during collaboration.
8. Allow users to create and compare multiple scenarios — "What if?" models — used to assess different assumptions in identifying the best solution.
9. Provide an interface for incorporating advanced procedural analytic models for statistical forecasting, optimization, goal seeking, and other complex formulas and procedures used to model business drivers.
10. Provide reporting from the online spreadsheet to produce presentation-quality, printer-friendly reports.

More Than a Concept

eiVia meets these ten critical requirements for online analysis by providing a comprehensive application development platform: Predictive Intelligence. It is designed for delivering applications that are data-intensive and must be integrated into collaborative business process workflows.

The CEO of a major consumer products company stated his goal succinctly: "I will no longer tolerate learning that we will miss our plan when it is too late to do anything about it." His desire was to improve performance visibility and gain greater decision-making transparency. Today, this company uses eiVia's Predictive Intelligence to forecast sales volume for 3,000 key customers and 550 products. The process is repeated weekly and involves 300 users who are active contributors of key performance metrics.

eiVia replaced a cumbersome desktop spreadsheet and email application with an application that provides continuous visibility, including the latest estimate of sales and the predicted financial implications. Over time, the application has evolved and adapted to new business challenges. Today, the organization not only has improved revenue visibility, but also visibility into marketing spending, costs and margin performance. Some users have access to more than 90 volume and financial variables for their analysis and modeling. Having this level of detail means that performance visibility is extended to operational levels to support tactical decision-making.

If this application could be replicated in a single desktop spreadsheet, the application would translate into 140 million rows of data. This is 140 times the capacity of the current release of Excel®.

Conclusion

Despite the “fear and loathing” engendered by desktop spreadsheets, these spreadsheets will not go away anytime soon. In fact, few companies have a goal of replacing spreadsheets, despite growing audit costs to meet regulatory compliance. The emphasis, today, is on finding ways to better manage spreadsheets.

But organizations want to improve business performance and recognize that, in many cases, spreadsheets are simply not up to the business-critical task of supporting collaborative workflows and ensuring data integrity at each step in the analysis process. This is especially true as the ability to generate a torrent of data outstrips the capabilities of desktop applications.

Business performance management relies on secure data and validated business rules used in the analysis of this data. Applications must facilitate information sharing, data consolidation and collaboration as organizations develop their “best practices” for performance management. It is important that the convenience and flexibility of desktop spreadsheets is not lost, while increasing online analysis capabilities.

eiVia provides Predictive Intelligence, the combination of forward-looking predictive models and user-initiated, interactive analysis. Predictive Intelligence is essential for developing applications to better manage business performance. Performance visibility, with speed and precision, is now a reality.

eiVia and the eiVia logo are registered trademarks of eiVia, Inc. Excel is a registered trademark of Microsoft Corporation in the United States and other countries. eiVia is not affiliated with Microsoft.

Actionable Business Intelligence

For more information visit: www.eiVia.com