Data collaboration means visualizing data from all of your different data sources, and getting this data to the right people, in the right format, in time to use it in making effective decisions.

Today's operations are swamped with data from multiple sources from every layer of the business. Of critical concern is data collaboration -- finding the right data, getting it to the people that need it to make their decisions, and making sure it gets to them in a timely manner. By using monitoring and analysis tools to pull together all of your different data sources and visualizing those results in an intuitive way, personnel can make an immediate difference in responding to conditions, resulting in reliable and agile operations.

So, how do we facilitate this? Let's run down a wish list of what a data collaboration and decision support system should look like.

- We want a web-based, thin-client platform - something anybody can easily access to get at the information they need.
- We need the system to make use of and add value to existing systems and data infrastructure.
- We want read/write access to any data source, through a single consistent environment.
- We want to have intuitive analysis and visualization capabilities that turn data into knowledge through trending, reports and KPI dashboards.
- We want to be able to share that knowledge in collaboration, through automated reports and alerts, so proper action can be taken quickly.

This paper will run down some of the data management issues facing industry today, and sketch out the requirements for a system that will resolve those issues and recover the real value hidden in the mass of information flowing through your plant and process.
**Table of Contents**

The Disparate Data Challenge ................................................................. 4
The Data Collaboration Solution ............................................................ 4
Visualizing the Data .............................................................................. 5
Data Trending, Analysis, Monitoring and Reporting .................................. 6
Put Your Data to Work! .......................................................................... 7
Table of Figures

Figure 1 – Data collaboration solution ...........................................................................................................5
Figure 2 – Treemap screenshot .......................................................................................................................5
Figure 3 – Digital Dashboard Screenshot ........................................................................................................6
The Disparate Data Challenge

The foundation of our system has to be data access. We need to get connected to any data source across our sites – ideally, we’ll be connected to clients, customers, and suppliers, as well. We want to consolidate that data in our system. We don’t actually want to move the data to a giant data warehouse or anything; we want to leave it where it’s at and just provide the ability to access it from a consistent environment.

The challenge in making this happen is the problem of disparate data. At the base of the data structure, there’s the real-time layer: instrumentation, sensors, PLCs, etc., all pumping out streams of information. At the top level, there’s the business layer where we find things like financial applications.

What we have in between is the operations layer. We have all these point solutions. So, for example, we might have truck hauling applications. We’ve got operator log books, reliability management, maintenance management, simulation models, etc. All these applications require data from other point solutions. Your LIMS data, for example, is required in your shipping and receiving. This causes a lot of complexity in the field, a lot of cost. Special drivers need to be developed for the applications to talk between each other, or there’s limited access. Somebody has to manually enter data into a spreadsheet and pass that information up to another department who can then correlate that information with their information… and so on. The end result of this is “Excel® Hell,” hundreds or even thousands of spreadsheets at a single site.

In this situation, there’s no useful correlation. It’s point solutions. They’re very good at what they do but they’re not providing a bigger picture, so you’re kind of running blind. Because you need to use the single-point solution tools to get at that information, you might find yourself locked in to proprietary applications. So, it’s very difficult to bring this data together and to get a look at the bigger picture. And when people use Excel or an equivalent engineering tool to manually pull all this information and use their own calculations to get their numbers, you end up arguing over these results -- because who knows whose calculation is correct? So, you end up with something called, I like to say, “Excel Hell.”

To be effective, we need to get rid of that. We need to bring everyone’s data together into a single, shared version of the truth.

The Data Collaboration Solution

The solution to the disparate data problem is the implementation of what’s called a collaborative plant management layer. Powered by open data standards such as OPC, this layer correlates plant data from point sources, puts it into context, pulls it together into graphics, reports and trends and then makes those available to the people who need it.

Beyond offering easier access to and higher confidence in data, collaborative plant management has the virtue of being many times faster than manually crunching numbers or poring over spreadsheets. Instead of spending all of Monday and all day Tuesday pulling in your numbers, formatting them in a little spreadsheet, and e-mailing them out and then finding out at the end of the day that it’s already three days too late to make a decision on that information, you’re getting solid operational information in near-real time. Obviously, this can be of great benefit to the business.
There’s also a cost savings involved when everybody’s using the same tool to access multiple types of data. You have the training cost savings. You have maintenance cost savings. Now, you have one place to manage all of your tools and applications from your server and people are just using Microsoft® Internet Explorer® to get at this, so now you’re not managing 50-60 work stations and having to go out and do updates and support them.

Finally, a web-based system that works in conjunction with your existing infrastructure becomes infinitely scalable. It can provide visibility into all your operating facilities, worldwide, from anywhere. For example, your people can instantly look at meeting customer requirements in a certain region by viewing all the producing plants in that region and seeing that, say, tomorrow or a week down the road one plant is going to be ramping down. They can then put the process in motion to ramp up another plant to compensate. So, you can start to make some educated decisions very quickly, regardless of distance.

**Visualizing the Data**

So, let’s talk about how we present this information -- our dashboards. It’s important that your collaboration tool offers the ability to create graphical displays that are representing real-time conditions in a useful, intuitive format. You need to be able to look at various operating areas and see, at a glance, what’s going on. What are the current problems? What things are going well? Do I need to dispatch maintenance? ‘Treemap’ visualization technology, for example, is an excellent approach in this area.
On a related note, these dashboards need to be easily customizable by individual users, for individual situations. We want universal data access, but every user is going to have different needs or preferences at any given moment. The ability to create custom dashboards on the fly, without programming, will add a lot of value to your data collaboration system.

Data Trending, Analysis, Monitoring and Reporting

Beyond at-a-glance visualization, we’re going to want tools for deeper analysis. At the most fundamental level, we want to be able to perform real-time trending on any tag-based data sources. We also need historical trends. We want to be able to look at, for example, pressures and temperatures throughout history. We want to be able to see these trends across different time ranges, on multiple scales: how are things running this week as compared to maybe a year ago? A system like the one we’re describing can give you these kinds of answers with a click.

We also want to be able to extract data from our relational data sources. So, this is another analysis tool. For example, we might want to look at Maximo, our outstanding work orders or anything waiting approval. Ideally, we want to be able to create SQL queries visually so users don’t need to know the SQL language to start querying these databases. What we want is a drag-and-drop environment, an ad hoc tool for pulling back the information we need, saving that, and even publishing it to the rest of the user community. What we want to end up with are formatted web reports that can be saved, shared and exported to other formats.

The data collaboration system can also perform some alarming functions, monitoring correlations between tags. This is in no way meant to replace DCS alarming; that’s very good at saying ‘Okay, this tag has exceeded its engineered value. An operator needs to take action.’ What this system can do is provide a layer on top of that: ‘When this pressure exceeds this, and this temperature is at greater than this, and tank volume is lower than this, there could be a condition that I need to pay attention to.’ This report can then be quickly and easily e-mailed out to operators. With the scripting engine behind this, you can get at pretty much any system event by interacting with databases. You can even write back to the DCS or other control system.

Finally, the system should have functions for scheduled reporting. Once you’ve got your created reports built up and you’re starting to use them on a daily basis, you’ll want them delivered automatically. So, for example, at 6:00 a.m., the system will run your production reporting for the past 24 hours and e-mail that out to the ET decision makers to decide what they need to be doing that day.

Because every situation will have different needs, this reporting must be able to be done in a variety of formats, from basic PDFs and text emails to trends and queries within the system’s own modules. The key is to get people making decisions instead of looking for information.
Put Your Data to Work!

There are three things you want to do with your data: consolidate, visualize, and collaborate. Bring information together to reduce the amount of time spent looking for data; take that data, apply business logic to it, and turn it into actionable information; and present it to the people that need it, when they need it. You want to be able to leverage your existing infrastructure, and to take advantage of the information currently hidden in spreadsheets.

That’s data collaboration -- putting your information assets to work for you. There really is some valuable information in there; you’ve put these systems in place for a reason. Meaningful access to all that data will turn it into actual actionable information, enabling better, faster decisions.

Microsoft, Excel and Internet Explorer are registered trademarks of Microsoft Corporation in the United States and other countries.