Putting Your Operations Data to Work:

Understanding and Taking Advantage of *Data in Motion*

VisualKPI

Transpara™
NOW YOU KNOW.

Content

Introduction

All operationally-driven businesses have staggering amounts of data coming at them from all directions, internal and external, that must somehow be put to effective use. Given the staggering rise of IoT and IIoT devices in industrial and other applications, even more is on the way.

It probably comes as no surprise that various data analysis and presentation tools have been created to meet this demand for collecting and visualizing information. The problem is understanding how they differ, and what makes a particular solution the best fit to make sense of operational data and ideal for your specific use cases.

This is where those looking to get the most out of their operational data run into problems. It's not just about features and functionality – it's much bigger than that. It's about the approach.

Most tools and projects are either built to generate static reports or suffer from a "experts-only" mentality (read: a big, bespoke project with extensive training required and no ability for most users to get value). These same tools focus on periodic (and therefore historic) reports, standalone displays and slow-moving data, mostly from relational databases (SQL) and Excel spreadsheets. They then visualize this data using traditional reporting tools such as Power BI, Tableau, or Excel itself. Piles of data on the screen to read, but usually missing the simple answer to "how's it going?" or "what needs my attention?"

Moving To Real-time Decision Making

Many of the world's most successful enterprises – from the C-Suite to operations, maintenance, and field staff – demand the ability to see what's happening as it unfolds, including the visualization of current and leading indicators and being alerted when they deviate from expected behavior. These companies view operations the way a stock trader watches the market. Imagine if they only got a daily or weekly report of how the market performed? Disaster.

Despite the existence of in-motion data flowing freely in various underlying systems and between various assets, most applications don't know how to handle this unlocked potential and, specifically, how to surface issues to the right people at the right time for optimum performance. Worse, most applications focus on data without context or meaning by just putting piles of data and charts on a screen, forcing users to do all of the cognitive work themselves. They aren't focused on obvious insights and actionable KPIs. In the rare case where this is possible in an application or a tool, the custom configuration requirements call for specialists to handle the setup and demand extensive employee training for them to develop the skills needed to get value out of the software. The process ends up being expensive and time consuming – and is still fragile and subject to human error in the end.

Here's where things get interesting and traditional thinking needs to be thrown out the window. Large amounts of data, especially when moving in real-time and across many systems, are seen as chaotic and messy (and indeed they are).

Traditional thinking goes something like this:

"Let's move it all into one central location, get it all organized, and build the dream application that lets us do anything with it. The data scientists will be perfectly happy, all the latest buzzwords (cloud, artificial intelligence, machine learning, predictive analytics, forecasting, observability, and more) will be in there, and decision making will become perfect. Simple!"

Now, let's examine what really happens:

- The company spends at least a year evaluating consulting companies and technology platforms before deciding to get started.
- Millions are spent building the ultimate plan that will answer all the critical questions, and then work begins.
- Politics and roadblocks reign, deadlines are missed, and time passes.
- As time goes by, technology platforms advance, various teams across the enterprise adopt even more data sources and applications (that were never considered in the master plan), and the big company most likely buys or merges with other companies, further confusing everything and adding even more forms of data to the mix.
- Finally, the day comes (often years later) where it's time to "go live" with the big, shiny new system. Two problems: the questions you once cared about have changed, and the system is slow and already out of date, with no well-defined user experience. Even if you do get to the finish line - and most never do - millions have been poured down the drain and separate and competing projects have started across the company that further dilute the original effort. Sound familiar?

Let's return to a statement we made at the beginning of this example: Data is chaotic and messy, which gives you two options:

- Try to control it by eliminating the chaos (Good luck!)
- Assume that there will always be chaos, embrace it, and put it to work to your advantage

In this guide, we discuss embracing this chaos to deliver immediate and long-term value. We will highlight some of the most common challenges with surfacing and analyzing operational and asset performance data and the benefits of having an organized and future-proof system without going on a fool's errand. We will also cover what companies can do to start benefiting immediately from the data already flowing through their existing systems.

The Chaos Of Operations Data

Operations-intensive organizations face significant challenges with data collection, quality, visualization and analysis. While existing business intelligence tools that report on and analyze past performance are enough for many departments and functions, those with time-sensitive processes face unique obstacles.

At its foundation, the challenge stems from a myriad of data sources not being put to use effectively (or at all) and operational data not being delivered to teams at the right time and in a way that makes it easy to understand and act on.

Here are a few of the most critical challenges specific to operations data:

1. Embracing Chaos, Part I: Organizing Assets & Information

Organizations with extensive operational data are often running large and expensive projects. To manage significant amounts of operational data, companies often rely on establishing a user hierarchy with strictly defined and divided roles, including process architects, business analysts, operations managers, operations engineers, and more.

There is a significant divide between dedicated specialists who are trained to set up and read the results from reporting systems and the people on the ground who manage the day-to-day operations.

This discrepancy in skills means that only small groups of people are capable of making the most of the technology, and they inevitably become the gatekeepers of that information. The only way to truly scale great decision-making is to put power in the hands of more users without distracting them with technology.

2. Embracing Chaos, Part II: No Migration Required

A traditional approach to data visualization is to focus first on getting all of your data centralized before you can put it to use (data lakes, big data, centralized BI, etc.). This usually takes years, is incredibly expensive, and is never complete, as systems come and go during the effort. Data centralization also introduces significant delays and overhead to everything you want to accomplish.

While this approach is proven to work for some forms of data, such as sales and marketing data or slower-moving finance data, it doesn't work for operational data, where things can change quickly and time is of the essence – it can even lead to dangerous and even catastrophic surprises.

Users are left with static reports about past periods, often missing important context or the ability to fix problems as they occur or even prevent problems.

3. Embracing Chaos, Part III: Empower ALL of your users

An influx of data coupled with the absence of solutions that provide clear context leads to at least one major problem: dependency on highly skilled specialists who can make sense of the troves of data. Anyone outside of this core group (executives, management, field staff, etc.) is required to submit a report request to these specialists, which also distracts them from their more valuable analysis work.

Analyzing streams of raw data may be fine for data scientists, but it forces everyone else to rely on these expensive, highly trained specialists for even the smallest question. Even if users are able to spot data that suggests a certain issue or opportunity for improvement, they are often unable to explore further on their own.

This reliance leads to several outcomes:

- It takes a comparatively long time to obtain the requested data – by the time it arrives, it might too late or irrelevant
- It ties up specialists' time with processing menial data and report generation
- It hampers innovation, as employees can't work freely and independently
- It's an all-around misuse of company resources

Data scientists and other highly trained teams should be focusing on solving hard problems rather than being constantly distracted with requests for building or adjusting reports and dashboards that users should easily create and modify themselves. More importantly, the power of curiosity in the hands of more users scales insights and leads to better decisions.

4. Embracing Chaos, Part IV: Big Project = High Cost, No Results

Earlier, we highlighted the many challenges of large data and analytics projects, but it bears repeating. The cost and speed of implementing traditional data applications are ripe for disruption, along with the time it takes to get real value from them.

One key driver is the process of migrating data from the various sources across an enterprise, which is a prerequisite for most of today's business intelligence technologies and other applications. This is often a massive process that spans months or years, requiring major programming and data science skills, as well as the reformatting and cleansing of data and adjusting data sources to make them suitable for implementation into the new platform.

We believe there is a better way, and we have used this method with hundreds of the most successful enterprises for decades to prove it. It involves several key approaches:

- Get value immediately before moving anything and before embarking on a huge project – it will often change your priorities
- <u>Read</u> data where possible without moving it, especially real-time or streaming data
- Copy data only when it makes sense (e.g. the "analytics cache")
- Migrate only when absolutely necessary, but still get value from it before you do

Missed Opportunities

The traditional approach to data visualization and analysis, which is rooted in traditional reporting and business intelligence techniques, misses some of the most valuable rewards you should be getting from your data:

- Your teams don't have timely access to data they need to do their best work.
- Decision making is limited to only those with access to the reports and tools, which doesn't scale and wastes the intelligence of a large swath of your best people.
- It is usually only concerned with what *happened*, rather than paying attention to what *is happening* and what *is about to happen*.
- It forces people to "look around" these reports and try to uncover issues or opportunities instead of automatically surfacing metrics (KPIs with obvious color and status) that are outside of expectation and need attention.
- Should users have to look at all? With core metrics turned into KPIs, they can alert users the moment something unexpected happens and lead users right to the situation, often removing the need for reports entirely.
- It misses what happens between reports. Regardless of how fast your data moves, you should know when things aren't going as planned and shouldn't wait for a predetermined period to address the problem. The faster you catch deviations, the better chance you have of fixing them before they become bigger or more expensive issues.

Another downside of reporting is that bad news doesn't travel fast enough and won't get to the right people in time to fix it. Any indications of performance issues on traditional reports don't come with alerts and often land with management instead of with users who can fix the problems directly and immediately.

At the very least, delayed alerting will result in unexpected surprises in operational outcomes. At the other end of the spectrum, the consequences can be considerably more dire. For example, the critical malfunctioning of equipment can cause a significant environmental hazard or even safety issues.

Static reporting is a direct cause of management's biggest fear: surprises.

Eliminating Surprises

Unwanted surprises can come in the form of asset failures, downtime, performance issues, and other critical events that – even at their most benign – can negatively impact the bottom line and damage the company's reputation.

Operations and management cannot afford to wait a week to learn about operational issues from static reporting. Unfortunately, this is way too common. While this might be acceptable in areas such as sales or financial reporting, it is rarely optimal for operations and can even be dangerous. Ideally, you would learn about changes to leading indicators before they become problems so that you rectify them in a standard way instead of through an emergency response (and the associated cost).

For operations, not building an early warning system based on monitoring and alerting on leading indicators is a missed opportunity to avoid a plethora of profit-leeching errors or even catastrophes. These systems make it possible to perform proactive and predictive monitoring and preventative maintenance, as opposed to emergency maintenance when the situation has already deteriorated.

Fostering A Real-time Culture

Another important aspect is the effect that real-time KPIs and data visualization can have when shared and made accessible at every level of the organization.

When users are given access to live data and are empowered to explore it, they gain a heightened sense of responsibility for company-wide goals and performance. They feel more engaged to creatively and proactively investigate and solve problems, thus making fuller use of the available talent in the organization.

Providing employees with more agency in the workplace is likely to benefit the bottom line. A recent study by Gallup found that employee empowerment leads to a 21% increase in profitability, a 41% reduction in absenteeism, and 59% less turnover.

Getting There: Creating A "System-of-systems" For Operations

Different types of users and use cases can have very different needs. For that reason, the "one size fits all" approach doesn't work. Rather, there are two questions you should ask yourself when evaluating any new data-centric solution:

1. Who are the users (personas), and what are their specific challenges?
2. What is the default time horizon (past and future) that aligns with your critical operating decisions?

Then, follow these three steps to understand what needs to be visually represented:

Step 1: Identify the problem(s) you're trying to solve, the words you use to describe or categorize your assets and information (attributes), and an initial set of key metrics that matter at each level of the business.

Step 2: Decide what KPIs best represent success for each user and their job, and how they might be organized (hierarchy) for rolling up performance (e.g., by geography, department, site, asset type, attribute, etc.).

Step 3: Use the information above to uncover which sources of data you need to interface with in order to obtain the respective information or to find gaps in your data.

Most existing solutions aren't focused on the needs of operations teams, which include operations management, field staff, maintenance, and others. As a result, they can't provide users with the most relevant information to do their jobs. When implementing a data visualization and analytics platform, you would ideally think of the operations staff and choose a solution that's designed specifically with their needs in mind.

Data that operations staff require for improved decision making include:

- Direct access and the ability to natively connect and work with multiple sources of data in a single application, without replacing existing systems
- Natively integrated internal data, including both slower-moving and real-time industrial systems (sensor data, control systems, historians, time-series, etc.)
- Flexibility to integrate external data such as weather, commodity pricing, supplier information, and other sources

Does that mean that you have to replace existing analytics solutions?

No – they're usually good at what they're designed to do and are often very expensive to replace. These two approaches can work in tandem, with business intelligence working with static data and real-time tools working with operational and real-time data. Together, they complement one another to best support users at all levels.

Required Capabilities

Historically, when thinking about data visualization, we think about traditional business intelligence tools like Microsoft Power BI, Tableau, SAP Business Objects, and hundreds of others. These were specifically designed and built to interact with relational data (SQL), spreadsheets, and OLAP data, and they're great at that.

But when it comes to operational and streaming data, the requirements and use cases change dramatically. These BI tools are not purpose-built to handle streaming or real-time operations data and the unique types of analysis and alerting required to excel in these scenarios. If you've ever tried to use moving data or scroll back in time across time ranges in a BI tool, you've seen a lot of spinning circles and done a lot of waiting.

Operations and streaming data is stored in multiple sources and formats, and many are not supported by traditional business intelligence tools. Even when they have integrations, they are not first-class citizens – the process can often involve making copies of data, migrating it, or batching it up (not streaming).

Example data sources include:

- High-speed, proprietary time-series databases and industrial historians like AVEVA PI System (and AF), GE Proficy, AspenTech InfoPlus.21, Rockwell FactoryTalk, Honeywell, Yokogawa, and others

- Open-source time-series databases like InfluxDB, TimescaleDB, etc.

- Cloud time-series databases like Microsoft Azure Time-Series Insights (TSI), AWS Timestream, and more

- Control systems like DCSs and PLCs (via standards like OPC and Modbus)

- Engineering tools

- IoT/IIoT systems and sensors (via MQTT and others)

- Vertical industry applications like LIMS

- External data like real-time and forecasted weather, energy, or commodity pricing, and more

So, what is necessary to get the most value out of your operations data? We'll dive into these key capabilities next.

Data From Multiple Sources

The goal is to integrate data from a multitude of sources and formats into one single view – also known as a "single pane of glass." The ideal solution must work with multiple sources of data in real-time (or near real-time), connecting directly or via API, and it should do this without requiring a major project.

Many data sources have their own data visualization tools for use with that specific source. However, when multiple sources are combined, having multiple user experiences is no longer an option. They create confusion, work in different ways, require unique training for each, and also depend on users checking them all individually.

Lastly, data sources and integrations shouldn't have to be permanent decisions. Sources and the way they communicate change all the time, and any good solution should be built with those changes in mind.

Simplify By Adding Context

Streams of data are useless until context is added. This contextual information, often called metadata or configuration data, is where all the power is. However, the approach matters. Many data sources also have metadata in them, but they all store it in different and inconsistent ways. The trick – and real power – is to store this information outside of the data sources themselves so that it remains consistent and lets you work with many sources as if they are all one single application.

Without this critical layer, there would be no way to do any of the following:

- Create KPIs that combine multiple sources of data with obvious and consistent status, color, and meaning, including limits, targets, and the ability to provide alerts on various conditions

- Generate visualizations automatically: With consistent patterns and attributes, there is no need to handcraft each visualization; in fact, you can have hundreds of visualizations already running without building anything and customize on top of this existing base with ease

- Organize your information and assets: Piles of data are difficult to work with, especially when coming from multiple sources, if they are not organized into a hierarchy or grouped based on various attributes or characteristics. This is critical to making information obvious to users and eliminating the need for deep user training – more importantly, it allows you to automatically rollup performance at any level with no custom work and guide users to drill into deeper levels by visually seeing where issues exist, even if there are millions of metrics

- Perform analytics without coding or a big project: Adding attributes to your infor-mation allows you to see which types of assets are having issues, without building complex OLAP cubes or other bespoke work. Imagine easily viewing your information and KPIs by status, equipment type, brand, location, capacity, class, site, or any other attributes you care to apply – or stack ranking them by volatility and other means

- Provide alerts on deviations from expected behavior: Without consistently defined KPIs with limits and thresholds, you would have to build every alerting scenario in a custom way – instead, these core patterns allow for powerful alert templates and scenarios that can be applied easily across any number of metrics

- Generate insights: Imagine a system where, even with data coming from many places, you can surface insights about information and assets that have attributes in common – this is common in relational databases and BI tools but has been a huge challenge with time-series databases, especially when these sources are streaming data in real-time, mixed and/or used at scale

Most importantly, the right application should be so simple and obvious that anyone can understand the performance of key metrics at a glance, compare statistics, summarize performance, and more – all without the need for any coding or advanced training.

"If you aren't seeing your core issues in seconds, you aren't doing it right."
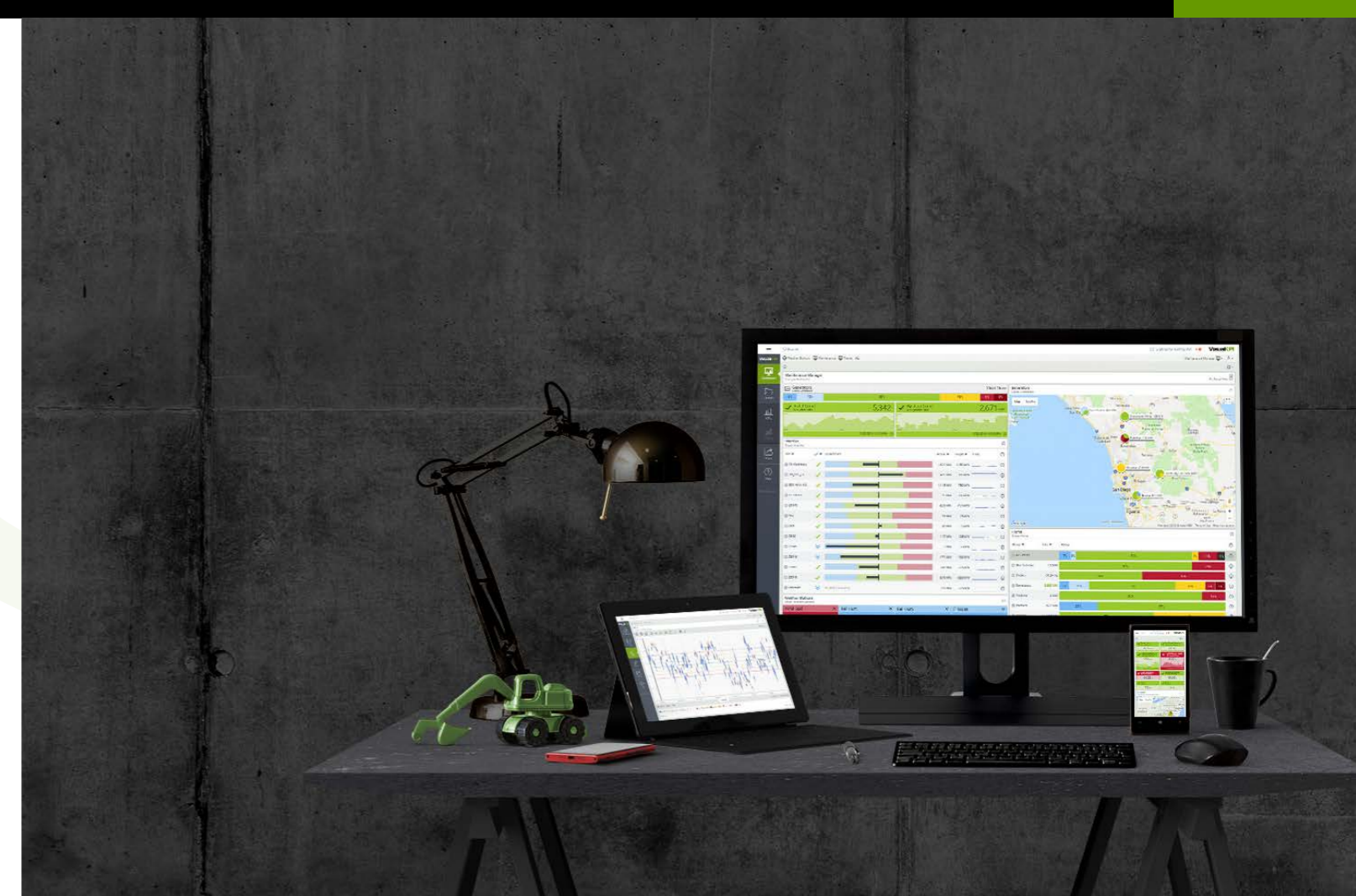
Access From Any Device or Location

While many vendors and applications have mobile device support, these claims should be examined more closely. For example, with a tool where visualizations are custom built by users or with process graphics, you have no choice but to either a) shrink these down for smaller devices to the point where you can't read them, or b) force users to create separate displays for each device or screen size.

Alternatively, when you combine data and configuration information (metadata, attributes) to generate automatic visualizations, the system can easily support all features on any device without any custom configuration. Even custom dashboards can scale down to phones or embedded devices and up to large screens, with zero custom work.

To get this right and provide the most value to users in various locations, all information must:

- Be accessible to any user from any device, including switching devices at any time

- Provide the same context, organization, performance roll ups and drill down

- Have a similar user experience, regardless of device

- Make every feature available everywhere

Most applications that "support mobile" require customization for every device or screen size, and they often require apps that differ from their browser-based versions. Look for systems where mobile is not an afterthought and where every feature is supported on every device without extra work.

Automatically Generated Visualizations: No Customization Required

You should never have to start with a blank screen and build everything yourself. Unfortunately, this is how most data visualization applications work. They are a series of displays and dashboards that stand alone and know nothing about each other, instead of a system with patterns that do the work for you.

Let's look at both approaches:

- Display/dashboard-centric: This usually starts with creating a display, then either adding data or a specific visualization, customizing it, and publishing the display or dashboard for users. Each one is considered an individual item, like a document or file.
- System-centric: In this approach, you define the system first. This means interfacing with data sources, creating metrics and KPIs, and organizing them into groups or a hierarchy. Once this is complete, you now have a system with many visualizations already running and the ability to drill down based on performance or see information rolled up and summarized at every level. Once that is running, you have the ability to rapidly create personalized views and dashboards that belong to an organized system and can even be automated. Oh, and changing your mind is easy.

Features to look for include:

- Automatic visualization generation – This should cover 90% of the user or administrator's effort but should not eliminate their ability to make custom charts or other visualizations.
- No custom development required – Every view must be able to be set up just by configuration and should be available for both end users and power users.

Having these core capabilities out-of-box significantly increases speed to value. As a bonus, it allows companies to evaluate the application quickly, without an extensive project.

Future-proof: Easily Adding Or Changing Data Sources

The ability to easily add or change data sources (or just change your mind) is a key aspect that ensures monitoring, visualizing, and alerting on operations data is efficient, cost-effective, and future-proof.

Long data migration projects should be a thing of the past. You should now be looking for functionality that allows for:

- Rapid addition of new data sources – Adding them or changing them should be swift and easy, while keeping the "single pane of glass" mentality
- Robust, secure, portable, and scalable – Look for industrial-strength security, with the option to combine on-premise and cloud sources, real-time and slow-moving data, internal and external data, and various formats (time-series, relational, etc.)
- No requirement to migrate or copy data – The tool should have the flexibility to "read" data from sources without the need for agents or any customization on the data source side. This is especially important when integrating with sources that you don't directly control or that have political ramifications, like your ERP system
- No compromise on scale or data source performance – Data speed and frequency should never have to be sacrificed to meet platform interoperability requirements

As technology has advanced, it is now possible to ask for more from your operational and business data to drive better decision making and keep your finger on the pulse.

Value In Days, Not Months

When it comes to managing and monitoring performance, time is literally money – not only when it comes to implementation, setup, and configuration, but also in terms of usability and the amount of time new users need to get acquainted with the interface.

Any great application should provide immediate value and use without training, both for basic users and power users. You should be able to start quickly – in just days – and have a system that can grow, change, and evolve over time.

Not only does this save valuable time and resources, but it also allows companies to evaluate the solution in days and weeks instead of months. The most successful companies follow a "succeed or fail quickly" mentality, which applies not only to the initial decision and implementation, but to every change along the journey.

Judging Success

Just like with large-scale projects, a critical aspect of choosing a software platform and its implementation is defining evaluation criteria. There must be a clear finish line and a set of concrete objectives in place to judge success (or failure).

Here's a list of criteria to consider when evaluating the success of your operational data analysis tool:

- Data – Does it easily integrate with the core data sources, and is it easy to add new ones? Does data flow easily, without taxing the data sources and at the frequency expected?

- Personas – Do the core user types understand it, and does it provide the answers they need without requiring deep training?

- Self-motivated usage – Will core users actually use it, both at the beginning and in the longer term, without management demanding it?

- Increase in "self-service" – Has it reduced the burden on teams that previously created displays, reports, and dashboards by putting more power in the hands of users? Are teams like business intelligence and data science free to use their skills for critical analysis and determination of the right metrics, thresholds, algorithms, and insights?

- Speed of problem identification and resolution – Has the amount of time needed to surface and solve a problem been reduced or eliminated?

- Elimination of surprises – At a management level, has there been a reduction or elimination of unwanted surprises related to operations?

Additional aspects should also be considered:

- Has business performance, visibility, and decision making generally improved?

- Do more employees have access to information that allows them to do their job better?

- Do employees have more headroom for innovation?

- Has employee satisfaction increased? This is particularly relevant, as many employees (especially younger ones) dislike working with older technologies.

Conclusion

Companies that rely on managing complex operations should strive to provide employees at all levels with obvious and actionable information across all relevant data sources and from any device.

While this has been the promise of many data analysis and visualization tools for decades, most applications still fall short of meeting the specific requirements of real-time decision making, or require expensive, time-consuming projects that rarely succeed or stand the test of time.

The information and characteristics defined in this guide are not new – Transpara's software is used every day by hundreds of the world's most successful operating companies. Transpara and its Visual KPI real-time platform embody these capabilities and drive value for thousands of users around the world in industries like energy and utilities, oil and gas, manufacturing, telecommunications, mining, IT and data centers, pharmaceuticals and biotech, core infrastructure, and more.

You can learn more at www.transpara.com, including scheduling a "no sales pitch" live demo, requesting a free trial for use with your own data, digging into our free video training, or just reaching out for advice. As we like to say… "Call for any or no reason."