WHITEPAPER

Intelligent upstream: Data management strategies for oil and gas production

Authored by:

-

Russell Herbert

Global Oil & Gas Industry Principal, AVEVA

Executive summary:

By refocusing data management strategies and taking a holistic approach to build the right data foundation, upstream oil and gas companies can use new tools to optimize production, increase safety, and improve bottom-line results.

A critical foundation: Why operational data management holds the key to digital transformation success

From digital twins to machine learning, predictive maintenance to advanced analytics, oil and gas companies are increasingly focusing on deploying new digital technologies to improve operations. And for very good reason. The slew of new applications promises to improve asset health, sustainability, and the ability to boost production with less expense. In fact, the World Economic Forum projects that digital transformation efforts have the power to unlock around $1Tn in value for oil and gas companies. In an industry impacted by fluctuating oil prices, ever-changing regulations, and uncertain economic conditions, increasing revenue and efficiencies is the key to bolstering the bottom line.

Less than a decade ago, a barrel of oil could garner $100. If a company wanted to increase revenue, teams simply drilled more wells. As oil prices have plummeted, these same organizations must now squeeze every bit of value out of the oil fields. This requires a keen focus on increased production, optimization, and operational efficiencies, which includes extending asset life, reducing maintenance costs, mitigating downtime, and increasing crew safety. The pressure is high to do more with less, making every business decision critical.

While deploying new tools and digital technologies is imperative to achieving future goals and ensuring a safe, agile operation, the industry is increasingly aware that introducing new technology often brings significant complexities. Companies are recognizing that the real challenge of digital transformation is not adding new systems or analytics platforms. The challenge lies in actually integrating everything together so that data properly flows, analyses produce accurate insights, and the important information gets to the right people when they need it. Every new solution must be properly managed in order to fully realize the potential benefits.

Although there are early signs of change, many organizations are unfortunately still focusing on user-facing tools and applications rather than the digital foundations that underpin everything. This failure to prioritize improving underlying data management mixed with sometimes confusing and conflicting guidance from technology vendors means many companies are embarking on digital journeys on somewhat shaky foundations. Unless this issue is recognized and addressed, it can quickly lead to increased future workload and costs, decreased value, and, ultimately, it can jeopardize future success.

Companies that prioritize proper operational data management in conjunction with advanced systems integration strategies will be the organizations that reap the most benefits of new technology investments.

An operational data management platform is the foundation for advanced oil and gas operations

Operational data is nothing new to upstream oil and gas companies. In fact, most operators have collected and stored asset data in historians for over 20 years. This data is typically used to monitor production and underpin the oilfield tools and digital models that companies rely on every single day. As data has become more valuable, organizations have recognized the opportunity to expand their use of digital technologies. As a result, data collection and volume have grown. It is not uncommon to find companies managing well over a million data points, compared to hundreds of thousands just a few years ago.

This increase in data volume is opening doors to new insights around production, maintenance, safety, energy management, and drilling. Remote operations are also becoming a reality, where real-time operational data can provide specialist teams accurate, up-to-date asset information in safe, centralized remote operations facilities.

To get the most value out of operational data, the upstream oil and gas industry is deploying predictive analytics, digital twins, and more on top of traditional historians. While connecting these new tools and technologies to the data historian is quite straightforward, integrations have presented challenges. These challenges call into question whether legacy historians are the right foundation for advanced operations. Data volume is not the key hurdle with data historians; the problem lies in how that data is represented and catalogued to enable users and systems to quickly access and use relevant information.

The demand is great for readily available and meaningful operational information to support new digital objectives. In response, the industry's leading real-time data technologies, such as AVEVA's PI System™, have evolved from simple data historians to far more powerful operational data management platforms. These platforms can be used not only for data collection and storage, but also analysis, event-based surveillance, quality management, and to facilitate advanced data connectivity to the wider enterprise. From there, users can respond to insights in real time and extract maximum value from operational data.

Advanced data strategies net deeper upstream insights

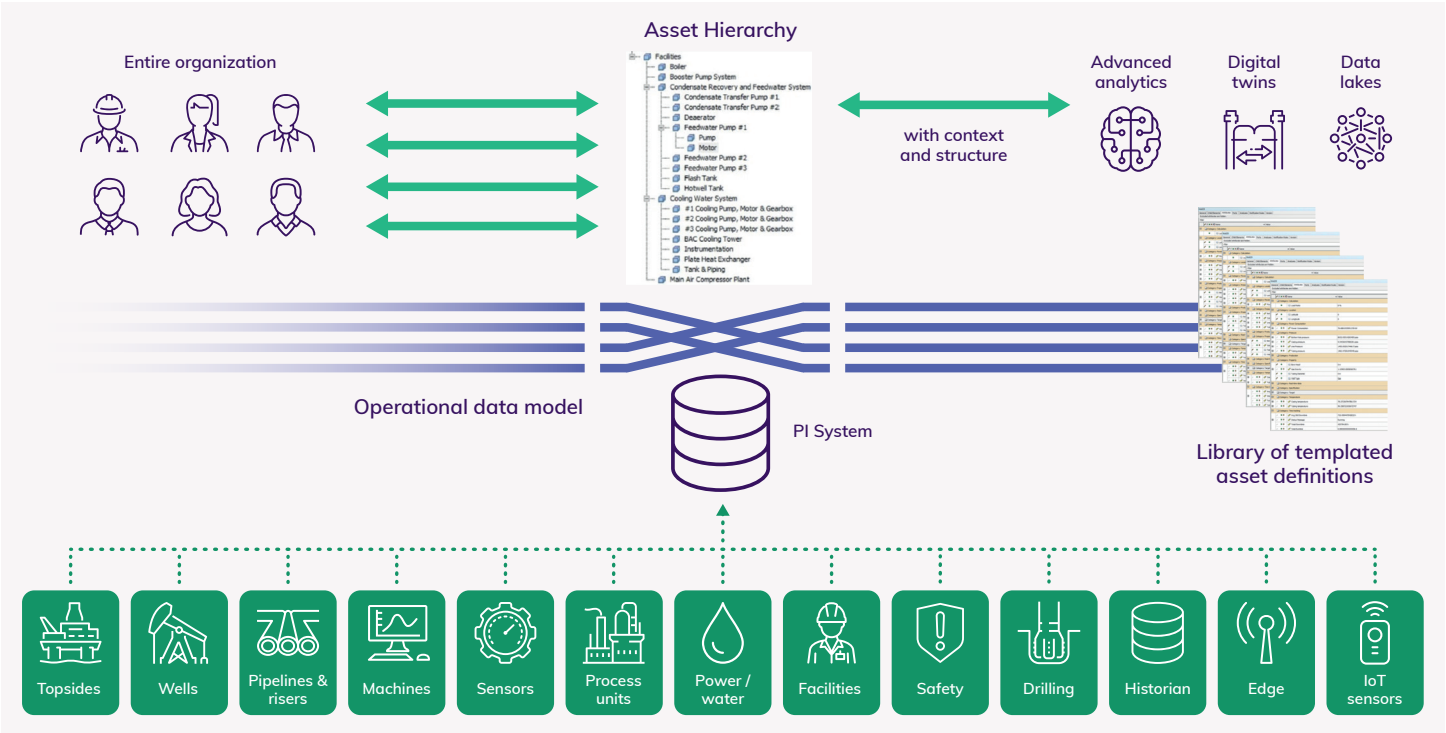
Successful operational data management strategies must:

1. Allow teams to bring context and structure to sensor data in real time.
2. Enable teams to enrich data with standard time-series calculations and analytics.
3. Provide mechanisms to automatically identify, analyze, and respond to patterns and operational events in the data.
4. Support analytics and data processing to improve data quality and reliability.
5. Facilitate advanced data integration and synchronization with the broader enterprise, including advanced analytics and digital twins.
6. Provide mechanisms to bring sensor-driven predictive insights and forecasts back into operations.
7. Work seamlessly from the edge to the cloud.
8. Ensure maximum security for process networks and data.

It’s time to evolve from traditional, tag-based systems

For many years, data historians have provided high-performance access to time-series data. Although these historians continue to deliver huge value, these tag-based systems heavily depend on team member knowledge and are not always intuitive for people who aren’t familiar with tag naming conventions. That’s because, unlike other operational systems, such as maintenance and finance applications, data historians typically lack metadata or a descriptive data model within the software to add much-needed context.

The result is often a quick and reliable operational data store that lacks the foundation to fully support advanced technology, such as predictive analytics and digital twins. Extracting value from these tools requires a real-time feed of contextualized, relevant information to yield best analyses and enable data integrations – and this is why an operational data management platform is critical. This feed ultimately enables the analysis and integrations necessary to properly connect and synchronize all operational information, and this, in turn, yields optimal results.



At the heart of operational data management platforms is a customizable operational data model, which is not present in a simple data historian. This model allows users to define the context and structure of sensor data and apply it in real time as new information arrives. Beyond defining that structure, the model also enables operations teams to further enrich the measured data with additional metadata and standard real-time analysis. Having this critical context immediately available in the data layer eliminates the need for teams to know exactly which tags to query and allows operations personnel to rapidly extract the insights they need.

Beyond enhancing data consistency and availability, the operational data model also provides external applications and systems with a consistent set of contextualized operational information. This mitigates complex and costly duplication efforts across the application portfolio and simplifies any necessary changes. Centralizing and standardizing operational data management accelerates data science initiatives and lays the groundwork for successful, enterprise-wide integration.

Asset Framework (AF): The operational data model foundation

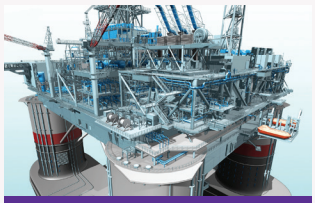
AF is the operational data model within the PI System. Users can define an asset model, enrich measured information with additional metadata and standard real-time analysis, and establish relationships between data and assets.

Why is an operational data model so important?

Engineers and operators

Enterprise integration

Advanced analytics



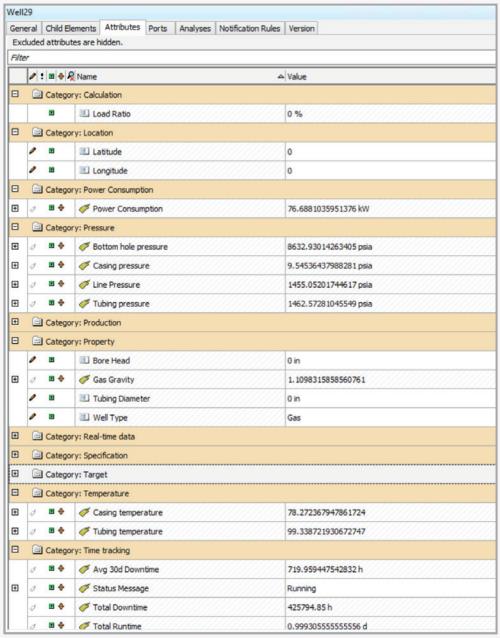
Digital twin and unified operations

- Operational data becomes meaningful, consistent and quickly useable by an entire organization
- It ensures all operational information relating to a physical piece of equipment is grouped together and can be found in one space, an important single point of truth
- It allows us to ensure all processing of data and analytics is applied consistently and correctly to a specific type of asset, event across different sites or operating entities
- It empowers engineers and operations teams to do a lot more with their data
- Helps scale deployment and avoid missing data or making mistakes
- It provides a mechanism for integration and synchronization with other enterprise systems and digital twins

Defining the operational data model unlocks future value

Before upstream companies can reap the benefits of an operational data model, teams must first define the information, data processing, and analytics that will be stored. While every organization will have different requirements, the data model will ultimately connect all time-series data sources and other pertinent information into one centralized location. The model must represent how the operations team wants to organize information and is typically comprised of measured sensor data, standard KPIs and calculations, descriptive metadata, and the results of third-party data analysis and future predictions.

What information should be stored in the data model?



Asset definition template

Time series data

- Process control systems
- DCS, SCADA, PLCs
- Historians, cloud and other DBs
- Laboratory data
- IOT and EDGE data

Metadata

- Engineering design data
- Descriptive and geospatial information
- Maintenance information
- Performance curves
- Operating windows

KPIs and calculations

- Performance metrics
- Asset status
- Summaries, averages, highs and lows, rate of change

Analytic results

- 3rd party insights
- Modeling results
- Operational recommendations
- Cleansed data

Future predictions

- Production estimates
- Expected equipment performance
- Predicted drilling performance

Documentation links

- Design documentation
- Engineering schematics

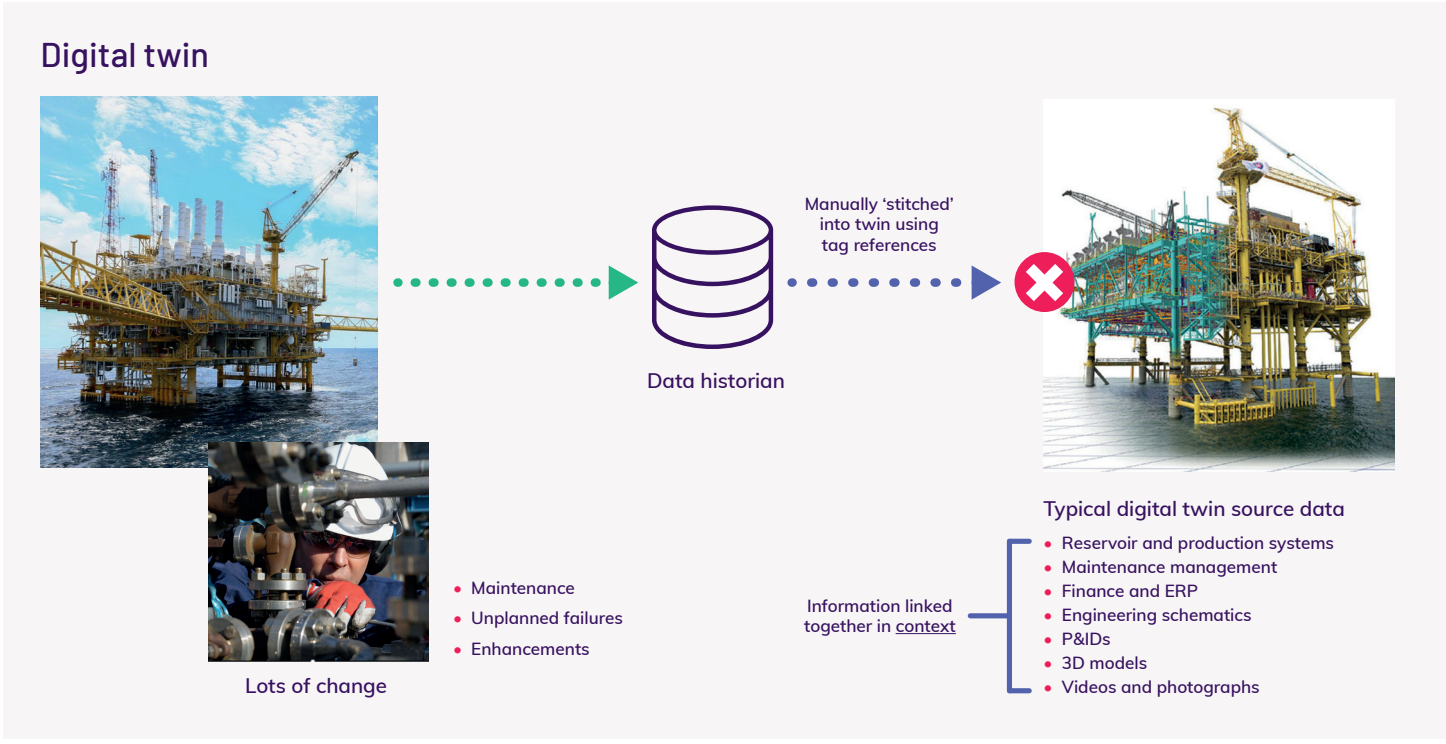
System references

Once this groundwork is laid, the data model transforms all raw, sensor-based data into a standardized and meaningful single source of truth. Every person within the organization can then utilize the model to extract relevant and contextualized information without searching for specific tags. Data availability coupled with this context enables this system of record – and holds the key to unlocking the value of the data.

An operational data model ensures accurate, meaningful digital twins

Digital twins are just one example of how an operational data model supports advanced operations. Digital twins are virtual representations of physical objects. These often three-dimensional replicas are created using a contextualized set of linked asset information. Digital twins typically include data sets from maintenance, finance, and engineering systems. To use these digital twins in operations, companies will often look to integrate sensor-based data into the model. Including operational data is what enables users to leverage the twin for real-time monitoring, optimization, and predictive analysis.

While manually linking sensor data from a data historian via tag references into the twin is possible and often done, this process is time-consuming and error-prone. Maintenance activities, unplanned failures, enhancements, and modifications are just a few of the areas where oil and gas assets experience rapid change. For the digital twin to remain both accurate and useful, all these changes must be continually updated and synchronized. In the busy operational environment, tag-based models quickly become outdated because asset changes are often not properly reflected back into the twin. As references break, predictive analyses miss important new information, and engineers begin to lose faith in the accuracy or the usefulness of the digital twin.



The addition of an operational data model dramatically improves the effectiveness of the digital twin. The model gives the digital twin a point of reference to automatically query and synchronize contextualized information whenever needed. Physical asset changes can be quickly queried and updated so that new sensor data replaces referential links from equipment that is no longer being used.

Overall, the operational data model facilitates a robust link between the sensor data and the digital twin, allowing the twin to serve as a continual source of insight that engineers and operations teams can trust.

Event-based surveillance

Outside of contextualization, upstream oil and gas companies must have systems in place to analyze and record operational events relating to asset behavior as they occur.

These events could represent planned operational activities, erroneous conditions, predefined data patterns that might suggest equipment failures, or even conditions relating to data quality. For example, event alerts may show if a well has tripped, or highlight integrity issues or indicators of potential future problems. Identifying these events in real time within the data platform enables event-based surveillance so users can prioritize activities, flag anomalies for further review, manage alerts, and take action.

Not only does event-based surveillance enable better equipment management and safer, more productive operations, it requires fewer people to monitor an operation while simultaneously removing the risk of missing performance anomalies. Event-based surveillance ultimately allows teams to focus on other tasks while still ensuring an appropriate response to any issues that might arise.

In addition to the immediate operational benefits, event-based surveillance also creates the contextualized slices of data that are critical for advanced analytics.

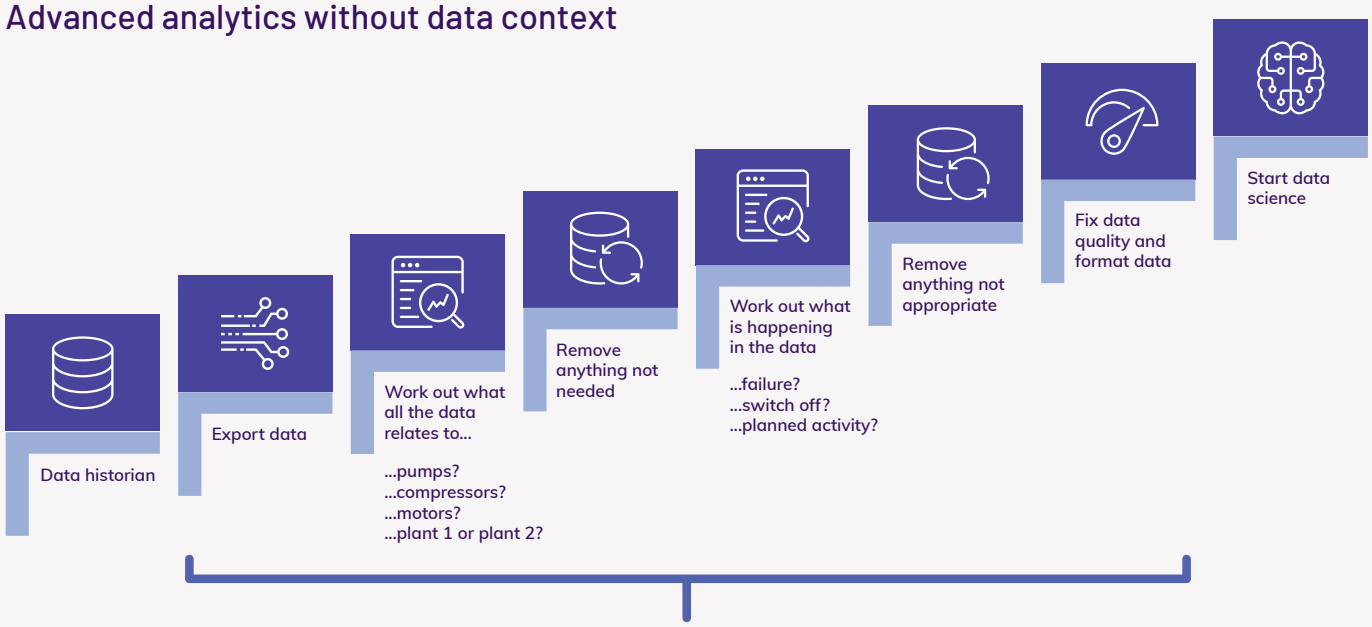
Advanced analytics requires contextualized data

Advanced and predictive analytics allow upstream oil and gas companies to make data-driven decisions, optimize maintenance, and maximize production. Before machine learning models can be successfully trained and used, raw data must first be cleansed, processed, and contextualized to ensure that only relevant, high quality, event-based information is used in analyses. Training machine learning models using sets of data that specifically relate to operational events allows data scientists to isolate specific problems. In turn, they can produce better models that net more accurate predictions.

An operational data platform automatically contextualizes all new data and identifies events as they arrive from the field. By associating operational data with equipment and processes, users can quickly understand which data should (or should not) be included.

In a simple historian without automatic contextualization, users are left with no option but to extract raw data and manually weed through it to establish relevant and complete information. Performing this type of data cleansing away from operations and after events have occurred is time-consuming, costly, and risks inaccurate results.

Advanced analytics without data context



Properly managing and preparing operational data for advanced analytics significantly shortens time to value and increases the success of advanced analytics projects. As a result, many upstream oil and gas companies are contextualizing all operational data in a trusted data layer prior to using it in data lakes or other analytics tools.

Data quality

The growing importance of modeling and advanced analytics makes data quality and reliability more important than ever. An operational data management platform can analyze data quality in real time by looking for events that indicate potential problems. Users can define rules that look for gaps in the data, data that is out of range, noisy data, or even stale sensors that have stopped working.

Once those events are recognized, operations teams can restrict the use of certain analytics or use third-party tools to provide better alternatives through the operational data management platform. Users can even mark the data to be ignored when training future machine learning models.

It's clear that managing operational data better is key to realizing value from all digital initiatives. Establishing data management best practices creates the foundation that will fully support current and future operational needs. Across the industry, important strategic decisions will need to be made by every operator regarding where and how to address data management needs.

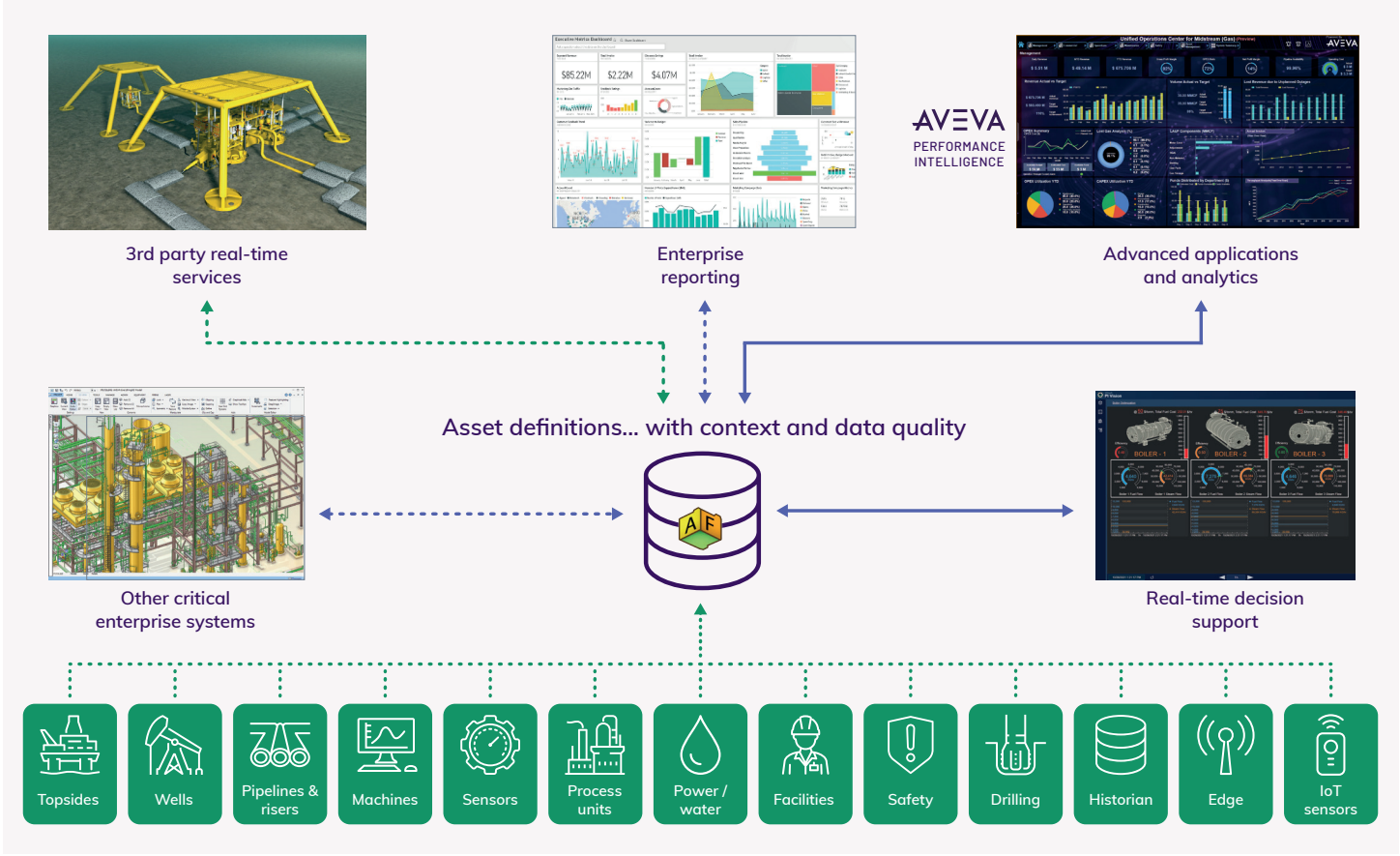
Why do this in an operational data layer instead of a data lake or analytics platform?

1. An operational data layer caters to the operations side of the business, which empowers teams to extract more value out of operational data.
2. It is far easier and more reliable to build a trusted data layer in tools specifically designed for managing operational data. When users are more comfortable, organizations see an increased likelihood of adoption.
3. Sometimes the data alone can't tell the full story, which is why operations teams who are familiar with assets and processes must be closely involved with data management activities.
4. The further a user is from the source of data, the more likely errors will occur and performance will be an issue.
5. Introducing context to operational data in applications and analytics platforms located deeper into the architecture can create data silos and introduce significant challenges.
6. The key is to centralize data management strategies – and do it once. Then that data can be used everywhere.

The operational data management platform is the key enabler of digital transformation

An operational data management platform is critical for the oil and gas industry because it serves as the foundation for solid data management strategies. The operational data model transforms tag-based sensor data into a consistent, understandable, and actionable set of master information.

With enriched operational data sets, event-based contextualization, and improved data quality, upstream oil and gas companies have the critical building blocks in place to accelerate digital transformation efforts and extract value from new applications. A solid operational data management strategy empowers organizations to quickly extract operational insights for better bottom-line results.



The path to performance intelligence

After years of building technology ecosystems on top of traditional data historians, the oil and gas industry is rethinking data management. Companies are now recognizing the importance of a holistic approach that includes applications and analytics underpinned by robust data management strategies. As a result of its acquisition of OSIsoft, AVEVA's end-to-end approach is now enabling operators to lay the groundwork for success. By focusing on data management from collection to analysis, connecting that data to advanced tools, and delivering those insights to operations teams, AVEVA is helping oil and gas companies achieve even greater levels of performance.

Case studies: Embracing operational data management

A solid data management strategy starts with the PI System. Below are some examples of how oil and gas companies are standardizing operational data strategies within the PI System to optimize production and mitigate risk.

BP: Global data standards improve maintenance strategies

For years, BP has successfully used AVEVA's PI System as a simple historian to monitor and optimize its upstream operations. Around 2014, BP began looking to deploy advanced tools and data lakes on top of its operational data. Connecting historians to the data lakes was relatively simple, but the raw, tag-based data lacked the structure and context to make it useful. This meant users had to manually perform all contextualization whenever it was needed. This process was complex and time-consuming, resulting in latent analyses and delays. BP quickly determined it needed to adopt better operational data management strategies. Around 2017, the company realized that advances in the PI System and the addition of AF could support its new requirements.

The team used AF to create a federal data structure inside of the PI System, enabling users to contextualize operational data before it was sent to the data lake, all while making that same information available to the entire BP organization. Once BP's operational data standards were in place and implemented through AF and the PI System, engineers and data scientists around the world all had access to the same set of useful information. This shifted the way BP viewed the PI System, transforming it from a source of data into a valuable system of insights.

Engineers on BP's oil platforms began using AF Analytics and PI Vision to build self-service analytics tools and gain access to relevant insights. For example, in just three days, engineers used contextualized PI System data to build PI Vision analytics on one of their North Sea assets, improving production by more than 20,000 barrels per day.

Eni: Real-time insights feed simulation models

Eni extracts oil and gas across 79 countries. The company's success hinges on optimizing its vast network of oil and gas assets. Several years ago, Eni built a digital oil field, or eDOF, to monitor and optimize assets in real time. The company selected the PI System to collect and manage all underlying operational data within eDOF. The PI System captures real-time sensor data from the field and feeds this information into a variety of specialist oil field applications and asset models. These asset models provide production engineers and operations teams with valuable insights to optimize their oilfields.

After selection, the eDOF team used AF to contextualize operational data, standardize calculations, and monitor data quality. This foundation prepared all operational data for advanced analytics because data scientists now had access to the same meaningful information in the PI System as operations teams in the field. This enabled data scientists to leverage operations expertise without the need to fully understand the underlying data.

Eni's Digital Energy Analytics (e-DEA) is part of the eDOF solution. Data scientists from the e-DEA group used PI System data to monitor and forecast energy efficiency at an operation in southern Italy. This enabled technicians to detect anomalies and suggest corrective actions. By using contextualized PI System data, teams accelerated data science projects by more than 300% and saved up to 5% on energy usage.

Petronas: Real-time, centralized monitoring for rotating equipment

For years, Petronas used the PI System to collect and analyze operational data. When it became clear that Petronas needed to enable real-time condition monitoring, particularly on its high-value assets, the company had to decide between purchasing a commercial system or developing an in-house solution. The commercial options were expensive and Petronas had the development expertise and tools internally. The team quickly realized the foundation for a real-time condition monitoring solution was already in place thanks to their existing PI System.

In 2016, the team began development on the Protean solution with the goal of enabling advanced monitoring and predictive maintenance on its fleet of rotating equipment and gas turbines.

The team focused on data management strategies from the beginning, using AF to build standard equipment templates, performance metrics, and real-time asset analytics for all turbines, regardless of manufacturer. Thanks to the data foundation in the PI System, Petronas built its custom solution at a fraction of the cost and quickly deployed it across many operating assets.

In the first two years of operation, the Protean solution was rapidly rolled out and delivered significant value in a very short period of time. Protean was deployed across 45 different gas turbines where it generated 225 alerts, saved 135 days of lost production, and helped avoid 15 catastrophic failures.

About AVEVA

The world's most essential and complex industries rely on AVEVA to manage the lifeblood of the industrial enterprise: operations data.

AVEVA is a pioneer and global leader in operations data management software, with over 40 years of experience helping industrial organizations meet next-generation demands for efficiency, reliability, security, sustainability, and resilience.

AVEVA's market-leading PI System is the proven system of record for operations data in essential sectors such as power generation and utilities, water, oil & gas, mining, metals, manufacturing, pharmaceutical, facilities, transportation, food and beverage, and more. Every day, industrial professionals in 146 countries rely on the PI System to improve operational performance, protect health and safety, keep the lights on, and make the world run more smoothly.

Learn why two-thirds of Fortune 500 industrial organizations choose PI System at **www.aveva.com**

The PI System manages more than two billion sensor-based data streams that enable better operations management and outcomes. For example, plant operators can spot problems with a remote pump before it fails. Environmental scientists can predict the accumulation of harmful elements in city water before it is tainted, while process engineers can fine tune production variables to increase profitability. With the PI System's high-quality curated data, data scientists can rapidly construct smarter AI algorithms, and executives can review dashboards that inform better business decisions.

About the author

Russell Herbert, Global Oil & Gas Industry Principal for AVEVA, brings extensive experience following his time working in a range of development and operations roles within the oil & gas industry, including time at BG Group, Schlumberger and Tullow Oil.

Much of Russell's work within AVEVA involves providing industry insight and strategic guidance regarding how the latest developments in Operational Data Management can significantly improve operations whilst also enabling important emerging technology trends such as Digital Twin, Predictive Analytics, Machine Learning, Artificial Intelligence, IIoT and Integrated Operations. Recent engagements have included projects with Shell, BP, ENI, Total, Petronas, Saudi Aramco, ADNOC, SBM Offshore and Repsol.