COGNITE | AkerBP

Open Industrial Data

Inspiring innovation
& fueling collaboration
in the oil & gas ecosystem

Open Industrial Data

Inspiring innovation & fueling collaboration in the oil & gas ecosystem

About Cognite

Cognite is a global industrial SaaS company that supports the full-scale digital transformation of asset-heavy industries around the world. Our core Industrial DataOps platform, **Cognite Data Fusion®**, enables data and domain users to collaborate to quickly and safely develop, operationalize, and scale industrial AI solutions and applications.

Cognite Data Fusion® codifies industrial domain knowledge into software that fits into your existing ecosystem and enables scale from proofs of concepts to truly data-driven operations to deliver both profitability and sustainability.

Table of contents

Executive summary

In a digitalized future, it's difficult to predict what the landscape of the Oil & Gas sector will look like. Metamorphosis is certain, but what kind? And what can we do to position ourselves to remain relevant and resilient?

For too long, Oil & Gas companies have taken a passive posture toward digitalization, reacting to each new technology as it comes. But this halting, piecemeal approach is inefficient and leaves companies vulnerable.

Though the industry has become expert at collecting and aggregating data, it has yet to find a way to leverage the data as an asset. The crucial next step will be a fullscale embrace of technology at every level of operation, transforming business models and activities through the strategic use of data.

The vast amount of industrial data collected every day contains hidden insights that can improve company workflows, optimize operations, and reduce costs. Oil & Gas companies can (and must!) find a way to extract these valuable revelations.

And they don't have to do it alone. Data science vendors, research communities, and students are eager to contribute to this progress. Unfortunately, these groups face a number of obstacles. Often they are working with outdated data sets, or are forced to spend considerable effort getting their hands on relevant data. Potential collaboration partners within data science have traditionally been made to jump through several hoops just to get started: show-and-tell, commercial contracts, data sharing agreements, NDAs, you name it.

The Open Industrial Data project is a crucial first step toward a more open and collaborative industry, where data sharing is the norm rather than the exception.

By sharing this live stream of industrial data freely, Aker BP and Cognite hope to accelerate innovation within data-heavy fields, such as predictive maintenance, condition monitoring, and advanced visualization techniques, as well as other new, unexpected applications. Advancement in these areas will directly benefit Aker BP's operations and will also improve the health and outlook of the industrial ecosystem on the Norwegian Continental Shelf.

That's why Aker BP and Cognite also want to inspire other industrial companies to join the Data Liberation Front. Imagine the speed of innovation and improvement in a world where hundreds or thousands of industrial companies share live data this way.

We're not sure how this venture will play out. Who will be interested in this data? How will they use it? Which specific problems will they attempt to solve? What solutions will be developed to solve them? We have some ideas, of course, but that's not the point of the project.

This is about open industrial data. Injecting real, live data into the realm of innovation and invention. Then stepping back to see what happens, to be surprised.

So, whether you are an ML startup focused on the oil and gas industry, a graduate student looking for an interesting data set for your thesis, or just curious about what real live data from a running oil platform looks like – here it is! Take it and learn with it. Show us what you find. We're excited to see what is possible.

With your help, the days of intentionally concealing data in industry will come to a necessary close. Openness will be the standard of the future. And that future will be bright.

Introduction

The Open Industrial Data project presents the world's first live industrial data stream. The data originates from a single compressor on the Valhall oil platform in the North Sea. The project is the result of a collaboration between Aker BP, one of Europe's largest independent oil companies, and Cognite, a fastgrowing tech startup headquartered in Norway. Aker BP uses the Cognite Data Fusion® to organize the data stream and create a digital representation of the Valhall compressor's industrial reality.

With the Open Industrial Data project, Aker BP has positioned itself has a pioneer of industrial data sharing. This openness is a radical proposition within the Oil & Gas industry. Most companies still consider such historical and live data to be strictly confidential information. However, Aker BP and Cognite believe data liberation is essential to human innovation. Every journey begins with a single step; this stream of compressor data is only the beginning.

The data stream will be provided through the Open Industrial Data website, open to the public on a subscription basis free of charge. Registered users will receive access to Cognite's Operation Intelligence application, a visualization tool for analyzing, monitoring and planning data, in addition to an API-key, which gives access to all relevant data in the platform.

This will include time series data,maintenance history, and Process & Instrumentation Diagrams (P&IDs) for Valhall's first stage compressor and associated process equipment. Cognite's extensive and welldocumented APIs and user-friendly toolkits (in multiple languages) will allow users to build machine learning models and applications on top of the liberated industrial data. To mitigate the risks concerning sensitive information from the Valhall's compressor (e.g., equipment downtime), the data stream will be delayed by one week.

The Open Industrial Data project will provide the most exciting, publically available industrial data set in the world. Free, straightforward access to a continuous live stream of real data is a groundbreaking opportunity for academic researchers and startups alike. Real, reliable, live data is the best way to evaluate state-of-the art models and algorithms, as well as to develop graphical user interface applications on top of industrial data.

Project background

Aker BP's intention behind the Open Industrial Data project is multifaceted. They want to offer the live data stream to

1. Establish Aker BP as the leading example of data sharing in the O&G industry;

2. Test and benchmark potential vendors within machine learning, predictive maintenance and other data science fields through defined problem statements related to the shared data;

3. Attract interdisciplinary talent, expertise, and backgrounds to the company;

4. Test potential new job candidates within the data science discipline;

5. Contribute to worldwide research by providing real data for evaluating state-of-the art models and algorithms;

6. Innovate on methods of sharing by using The Cognite Data Fusion®, via a non-commercial application;

The first step of a journey isn't always easy. With this project, Aker BP wants to break trail on behalf of the Oil & Gas industry. By being the first to share a live data set, **Aker BP can go on to leverage and share what they learn along the way, enabling more data sharing by more players, more assets, and even more data sets made live to the public.**

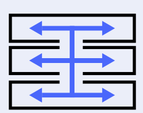
Project vision

Aker BP and Cognite hope to inspire more companies to join the project in order to increase the amount of shared data from the industry. Because no industrial player is an island. Each company exists as part of a larger ecosystem made up of suppliers, vendors, and other external stakeholders. The individual pieces impact one another in many different ways, including via innovation and collaboration. The Open Industrial Data project is a step towards increasing this kind of positive impact by promoting an openness that benefits everyone in this ecosystem.

It's true that increased data openness assumes a corresponding increase in business risk. But the corresponding benefits, even the ones we can't yet name, are likely numerous too. For Aker BP and Cognite, the vision for the Open Industrial Data project began with the as-yet unfathomed depth of value that comes with revolutionary openness.

Organized and clean, easily accessible and easily tested, live data is the perfect stimulant for rapid, sustainable innovation. Without trying, we'll never know what openness might empower us to do.

It's time to:



Explore the upside(s) of sharing data

Take the initiative in this field on behalf of the industry

Start a conversation with evidence in hand

How will the data be shared?

Cognite Data Fusion®

The Cognite Data Fusion® is a horizontal platform for all types of production and operations data. It breaks down existing data silos to liberate data across the industrial value chain. This dramatically reduces the cost of integration and maintenance while enabling scalability, speed of development, and data openness throughout the organization. Internal and external experts are able to apply algorithms to visualize and solve critical business problems.

The Cognite Data Fusion® is currently deployed across all five of Aker BP's operational assets. Here, the Cognite Data Fusion® handles live and historical data for close to 200,000 sensors, with a peak transfer of 800,000 data points per second. A range of third-party applications and data scientists use the 1+ trillion data points in the platform to create value and support Aker BP's strategy for day-to-day operations.

The initial focus has been to aggregate and process data from sensors and contextualize it, structuring it in relation to process diagrams, production information, 3Dmodels, and event data (maintenance, incidents). Everything linked in the real world is also linked in the platform. This enables machine learning applications for optimization and automatization, as well as human-facing applications, such as advanced visualizations and apps for the digital field worker.

The Open Industrial Data project website provides access to the live data stream on a subscription basis, free of charge. Once they register using a Google ID, users will receive an API-key, which gives access to all data in the platform. Cognite also provides access to their Operation Intelligence application, a visualization tool to improve data comprehension.

Immaturity of condition monitoring anomaly detection algorithms, etc., is due in large part to the fact that there hasn't been a lot of data available. The Open Industrial Data project presents a steady stream of stable data, stored and relatively clean, made possible by the Cognite Data Fusion®.

Once data has been liberated and contextualized within the Cognite Data Fusion®, it can be made openly and securely accessible to approved third parties. Aker BP can grant this permission to strategic partners and vendors via user-friendly Application Program Interfaces (APIs). Or, as in the case of the Open Industrial Data project, they can make a specific data set publically accessible in a secure, functional environment, to inspire innovation and equip innovators.

The supplied APIs for the Open Industrial Data project allow for flexible parameters to empower quick aggregations. For example, users can call for values every five minutes or for the average values each hour. This allows the user to aggregate targeted data quickly and zoom into specific points in time at high speed. Cognite also provides an open source graphing library to help users visualize their data.

Live data

Convenient, continuous access to a live stream of reliable industrial data is a game changer for researchers and entrepreneurs dealing with machine learning (ML).

One of the most challenging and time consuming parts of developing an algorithm or application is often tracking down data sets to use during development. These products need to be tuned and tested. Previously, most testing has been done using historical data. A researcher might develop an algorithm using two-year-old data, then test it on last year's data. This practice has been generally accepted because it was the only practical option, but it has shortcomings.

Historical data is dead data. In the case of historical sensor data on an oil platform, for example, an old data set leaves developers vulnerable to blind spots such as the interim replacement of upgraded

equipment or the deployment of new technology. These simple, common occurrences could render the an ML algorithm skewed or and application obsolete before it's even complete.

Live data, on the other hand, opens exponential opportunities, not only for product development, but for the disruption and revitalization of research in general.

A living data set will change. What is good one year won't necessarily continue to be good. Environmental circumstances can change. It's like pressure testing. Algorithms and applications developed on live data will naturally grow stronger, more adaptable, prepared to respond to the the curve balls of reality.

These dynamic qualities of living data may even be pulled forward into the research published on it. The status quo for research dictates that any published finding is trapped in time. Others may pull out the thesis and respond to it with tests of their own, but these new theses, too, will likely be based only on similar data and will be tied to the date of their publication. Technology and time march on.

With a live data stream, the concept of publication becomes more flexible. Released from the constraints of historical data, research findings will continue to live and grow along with the data it is based upon, carried forward and continuously tested.

Universities and employers can also use this source of living data to their advantage in evaluating the performance of students and employees. Competitions can be conducted over time, comparing teams; if something unexpected happens, evaluators can see who picked it up and who didn't. During a hiring process, developers can showcase their work to the potential employer based on current circumstances in the real world.

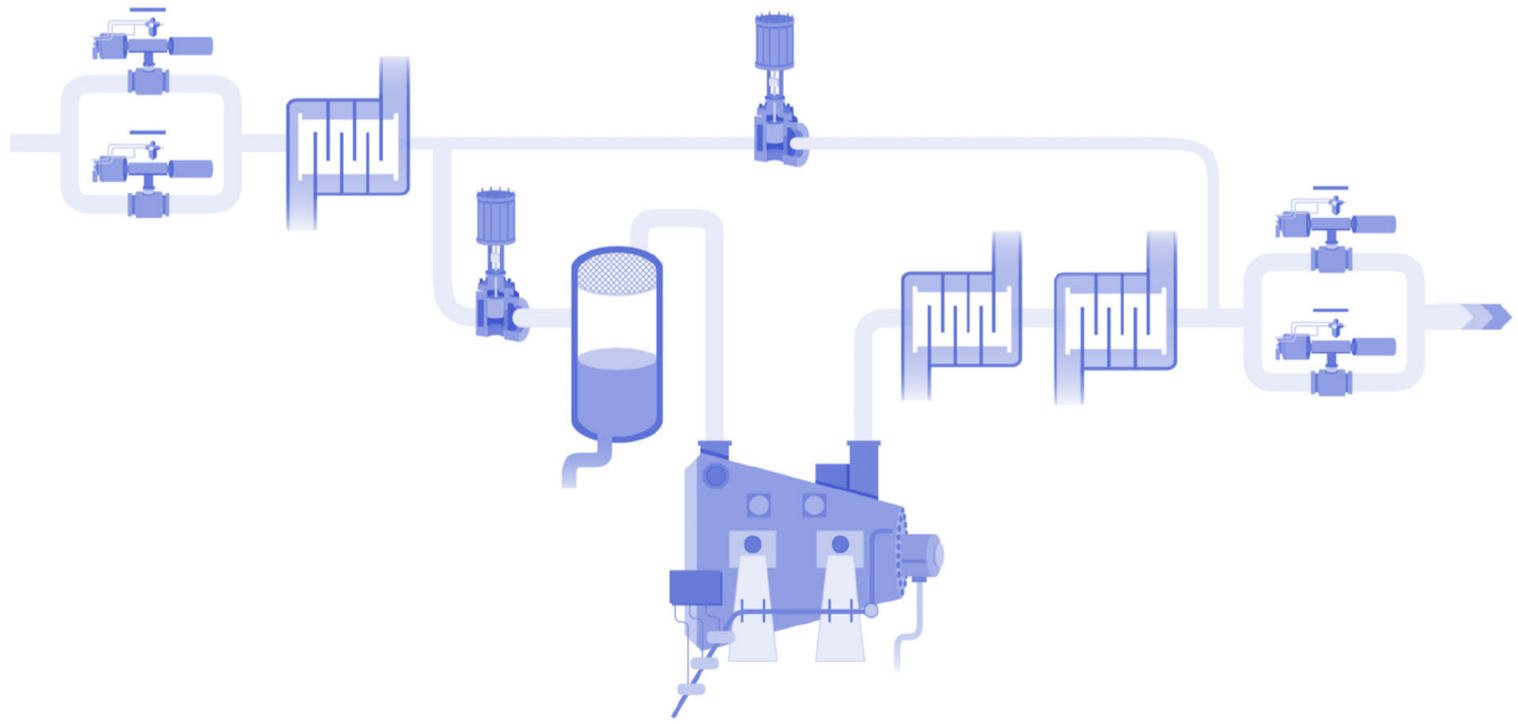
What data will be shared?

High level system description

The data set represents the first of four stages for compression of natural gas on the Valhall PH platform. The purpose of the gas train is to compress and treat the gas to meet the required export pressure and specification, with a total capacity of 4,06 MSm3/d (143 MMscf/d).

Aker BP selected the first stage compressor on the Valhall because it is a subsystem with clearly defined boundaries, rich in time series and maintenance data. The data set available in the Cognite Data Fusion® will include time series data, maintenance history, and Process & Instrumentation Diagrams (P&IDs) for Valhall's first stage compressor and associated process equipment: first stage suction cooler, first stage suction scrubber, first stage compressor and first stage discharge coolers.In addition, data from the compressor's lubrication system, dry gas seal system and condition monitoring system (temperature and vibration) will be available.

The first stage compressor (23-KA-9101) is an electrically-driven, fixed-speed centrifugal compressor which receives gas from the separators at approximately 3 barg pressure. Before reaching the compressor, the gas is cooled in the first stage suction cooler (23-HA-9103), which is a shell-and-tube heat exchanger. The cooled gas flows into the



first stage suction scrubber (23- VG-9101), which removes liquid droplets before the gas enters the compressor. The gas is compressed to approximately 12 barg and flows through a pair of discharge coolers (23- HA-9114/9115) before entering the second stage of compression.

Live data

Automatic control of the first stage compressor and associated process equipment is achieved by the following control loops:

1. The temperature at the suction cooler outlet is controlled by regulating the amount of cooling medium flowing through the shell side of the suction cooler.

2. The level in the first stage suction scrubber is controlled by regulating the amount of liquid flowing out of the scrubber. Excess liquids are sent back to the 2nd stage separator.

3. The pressure in the first stage separator is controlled by regulating the suction throttle valve (STV) upstream of the first stage suction

scrubber, which controls the amount of gas sent to the compressor.

4. Anti-surge protection is achieved by recycling gas from downstream the discharge coolers back to the inlet. The amount of gas to be recycled is controlled by regulating the antisurge valve (ASV). Since the recycle line is connected just upstream of the STV, the pressure control and anti-surge control loops have strong interactions.

5. The temperature at the discharge cooler outlet is controlled by regulating the amount of cooling medium flowing through the shell side of the discharge cooler(s).

Utility systems

The following utility systems are also included in the data set:

- The lubrication system for the first stage compressor. The system is comprised of pumps, coolers, filters, heating elements, and associated monitoring equipment, the purpose of which is to reduce friction and mechanical wear and to prevent overheating and corrosion in the compressor internals.

- The dry gas seal (DGS) system, the purpose of which is to prevent the process gas inside the compressor from escaping to the atmosphere.

- The condition monitoring system (CMS), which monitors temperatures and vibrations in the motor, gearbox and compressor.

Strategic thinking, struggles, & precautions

Publishing any important data means confronting a variety of risks. There are reasons why this kind of data has long been classified as confidential. Before releasing the live stream of compressor data to the public, Aker BP measured the list of possible risks against the potential value of the venture. The exercise helped them understand what precautions should be taken to mitigate the risks, but also highlighted other obstacles between them and a golden age of industrial data sharing.

Public visibility of equipment downtime

Transparency could be a business risk. For example, what might the consequences be for letting everyone know when equipment is down? Oil is a global commodity, which means increased scrutiny on the companies involved in its extraction and production. There's a possibility that public visibility of real-time equipment breakdown could impact the stock exchange. After all, consumers don't necessarily have the context, history, and/or sophistication to appropriately gauge the seriousness of an event. Their reaction could be premature.

Aker BP opted to instate a delayed release of the data to defend against this particular risk. While the data stream will be continuous and live, the one week delay between the creation of data and the release of data will give the company time to resolve downtime issues or deploy a necessary public relations strategy.

Third-party access & security

Opening previously confidential data to public access seems like it could also increase the company's vulnerability to a data breach.

The Open Industrial Data (OID) project is set up within the Cognite Data Fusion® as its own secure, separate project or "tenant". The live data stream inside the tenant is a duplication of the stream from the Valhall compressor. No navigable connection exists between Aker BP's own environment and the OID's public tenant, making it impossible for anyone to access and tamper with the original data. In addition, registered users receive read-only security credentials that are limited to the public tenant environment and authenticated at each log in.

While the Cognite Data Fusion® allows for infinite scalability and can run countless separate projects/tenants simultaneously, these environments are kept separate and secure. For practical purposes, what the user of any one of these projects can see and work with might as well be on its own server/computer.

GDPR & protected personal information

In the post GDPR-era, there is also a heightened awareness of the potential exposure of data that includes personal information. Aker BP needed to make sure any individually identifiable information was removed or redacted from the records.

For example, while maintenance logs (Events) are part of the Open Industrial Data stream, the names of the personnel who serviced the equipment could not. This is achieved by removing all text fields which may contain names.

The legal framework

For each production license on the NCS, the license partners are required to enter into a Joint Operating Agreement (JOA), a standard document issued by the Norwegian Ministry of Petroleum and Energy (OED). The agreement includes formal restrictions on the sharing of data. While the operator is entitled to share necessary data with a third party to achieve some business goal, public sharing of data is generally restricted unless permission is granted by all license partners.

Approaching the Open Industrial Data project, Aker BP secured this permission from the Valhall license partners.

However, Aker BP intends for the Open Industrial Data project to be a first step toward an industry-wide shift toward increasingly proactive data sharing. More complex assets involving multiple parties on a single JOA could become major obstacles to this goal.

Extracting & leveraging the value of big data

Oil & Gas companies already have an incredible volume of data, and it's expanding all the time. But data is only as valuable as it is accessible, readable, and open for connections. Dropped into a data lake on a distant server, companies cannot leverage its value.

To make data openness an industry standard, JOAs must change, not least because any supplement, amendment, or exception to the current JOA is subject to approval from the Ministry. A good first step would be to remove these kinds of legacy restrictions, which stem from a time when the data landscape was entirely different. Formal restrictions on data sharing in this foundational contract predispose whole ventures toward a mindset of exclusivity and protectionism. This negative stance cannot be healthy for a joint venture.

Every industrial player on the NCS needs to reconsider the benefits of data sharing, the gains to be won simply by injecting more clean, reliable data into the ecosystem. Opening our minds on a company level can and will encourage the authorities to do the same.

Aker BP responds to these issues internally by using the Cognite Data Fusion® to collect, clean, and contextualize their data. Now, with the Open Industrial Data project, they are using the imperative fourth functionality of the Cognite Data Fusion® to maximize data value: facilitating collaboration and equipping innovation by easily sharing selected data.

Potential rewards outweigh the risk

Somehow, a fully digitalized future in Oil & Gas seems both distant and inevitable.

Technologically speaking, operators cannot accurately plan their next one to five years. No one knows who the best vendors will be, because so many new companies are entering the market. No one knows what the interplay between vendors and customers will be, as technology will impact those interactions and new partnerships are sure to emerge. And no one knows the pain points for customers and vendors in this new world, because only direct experience with these problems can lead to solid insights on how to solve them.

Rather than stand overwhelmed by the impending metamorphosis, Aker BP has determined to go on the offensive. Daring to invest in philosophical shifts, even or especially in the early stages. Considering the possibilities, the risks, the rewards, and picking a side.

Openness should be the new standard. All industrial data should be open by default.

This is a striking deviation from historical norms. Immediate, total transition probably isn't possible, nor would it be advisable. Instead, Oil & Gas companies need to take on this initiative in phases, deploying the conceptual shift in practice on a small scale in order to experience trial and error. The Open Industrial Data project does exactly that. With it, Aker BP wants to demonstrate the benefits of making live data easily available to students, startups, and other interested innovators. It's also a chance to troubleshoot the process and set best practices in a secure environment. With the Cognite Data Fusion®, it will be possible to scale this project quickly to realize benefits on more sensitive, more critical data, once stakeholders can be assured of the value and reassured regarding the risks.

Call to action

Given the demand for live data, the Open Industrial Data project could easily become the most referenced public data set in academic publications and the most utilized public data set in tech startups innovating on ML.

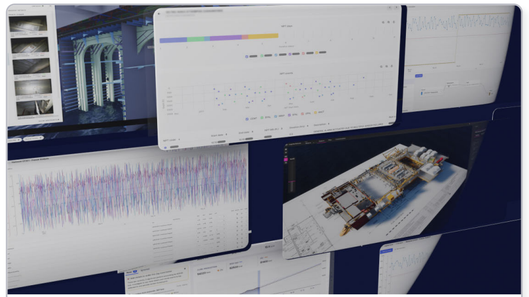
Are you ready to test an API or algorithm? Want to see how the data works or share ML insights with a like-minded community of developers?

REGISTER FOR OPEN INDUSTRIAL DATA ON COGNITE HUB →

No matter what the digitalized landscape of Oil & Gas looks like in the future, Aker BP and Cognite want to do more than be a part of it. They want to shape it. They want to break it wide open.

Want to know more about our product?

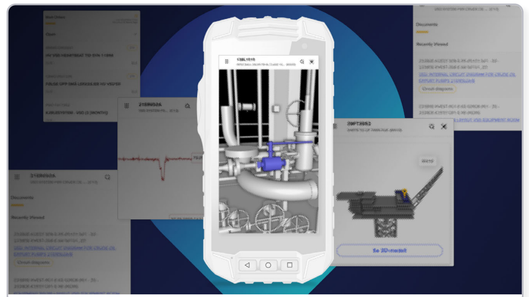
Explore more insights from Cognite



PRODUCT TOUR

Learn from Cognite customers and product managers how Cognite Data Fusion® simplifies and streamlines the data experience of a subject matter expert.

WATCH NOW →



CUSTOMER STORIES

Discover how Cognite Data Fusion® makes data more accessible and meaningful, driving insights that unlock opportunities in real-time, reduce costs, and improve the integrity and sustainability of your operations.

GO TO STORIES →

ANALYST REPORT

Customer interviews and financial analysis reveal an ROI of 400% and total benefits of $21.56M over three years for the Cognite Data Fusion® platform.

READ THE REPORT →

BLOG

Discover our rich catalog of industry insights and technology deep dives.

READ OUR NEWEST BLOGS →

WHY COGNITE DATA FUSION® → BENEFITS FOR YOUR TEAMS → CONTACT SALES →

COGNITE | AkerBP

COGNITE.COM →