Novotek

Introduction to industrial data

A guide to planning, management and strategy

Data is the backbone of the modern industrial revolution happening around us. Despite its importance, a recent McKinsey survey identified that businesses still approach data with a blinkered view or short-term initiative that fails to make effective use of it.

In this whitepaper, we will guide you through how to make the most of the precious data being generated every second in your industrial plant. It begins at the field level, with proper consideration of system specification and capabilities. But beyond that, effective industrial data management is as much about mindset and culture as it is technology — and, as we explore in this paper, businesses must make that central to their data strategies.

Taking a grounded approach

Talk of Industry 4.0 has been useful in helping businesses understand the value of industrial automation and connectivity. But a lot of it is just hype, and many of the technologies are not being adopted as broadly or effectively as they could be. Here, we explain why it's not only vital to work with a systems integrator to get the most from industrial connectivity, but to also take a grounded approach.

In the three decades years since Chuck D's group, Public Enemy, rapped "Don't Believe The Hype" in their same-titled single, it seems that we've become no less susceptible to hyperbole. After years of Industry 4.0 talk, it's unbelievable to think that the idea of "Industry 5.0" is now being floated in some corners, that humans and machines will collaborate more fully to supply more bespoke products and services.

To be fair, as with Industry 4.0, there is truth behind the "5.0" concept. Novotek's customers in the food and beverage industries, in particular, are seeing the markets stagnate to some degree. This is pushing manufacturers towards smaller batch-sizes, more bespoke and customised products — otherwise known as discreet manufacturing — and greater collaboration between workers of varying skillsets and automation technologies, such as cobots.

"A grounded approach means having specific goals that relate to your own company's objectives.

However, whether we're talking about Industry 4.0 or 5.0, it's my view that businesses must see beyond the buzzwords. To illustrate my point, at Novotek we've been working with Manufacturing Execution Systems (MESs) since the term was first coined in the early 1990s. Yet, even today, the technologies are not being adopted as broadly or effectively as they could be. This similarly applies to the wider Industry 4.0 concept.

It's been suggested that businesses are not achieving lofty enough goals for Industry 4.0 to be considered a revolution. At Novotek, we recommend firstly taking a grounded approach. Secondly, being mindful of how all these systems are interconnected and connect back to your business.

A grounded approach means having specific goals that relate to your own company's objectives. For example, instead of specifying a system to their current needs, businesses should be mindful that it must last over the next five, or even 10, years. This where the expertise of a systems integrator comes into play. We understand the pros and cons of systems available on the market.

Next, connecting these systems back to your business is dependent on how effectively you manage data. Historian software, like GE Digital's Historian, can prove critical in this regard by gathering, archiving and compressing data together. Furthermore, the data can then be analysed and distributed — in real-time — among everybody within an organisation. That's how big data can be matched to your goals.

Going back to the food and beverage sector, companies can achieve more through personalised data from customers, digital forecasting of changeable markets and real-time monitoring of processes and information.

While the hype of Industry 4.0 has helped light the spark that's driving the world towards industrial connectivity, a grounded and truly integrated approach will be essential to getting the most value out of things like remote maintenance, predictive analysis and digital twins.

Historian software is just one way in which connectivity can be mapped to tangible business results, in the long-term — to ensure that Chuck D's intonations of "Don't Believe The Hype" might ring true for the next 30 years.

The importance of context in data collection

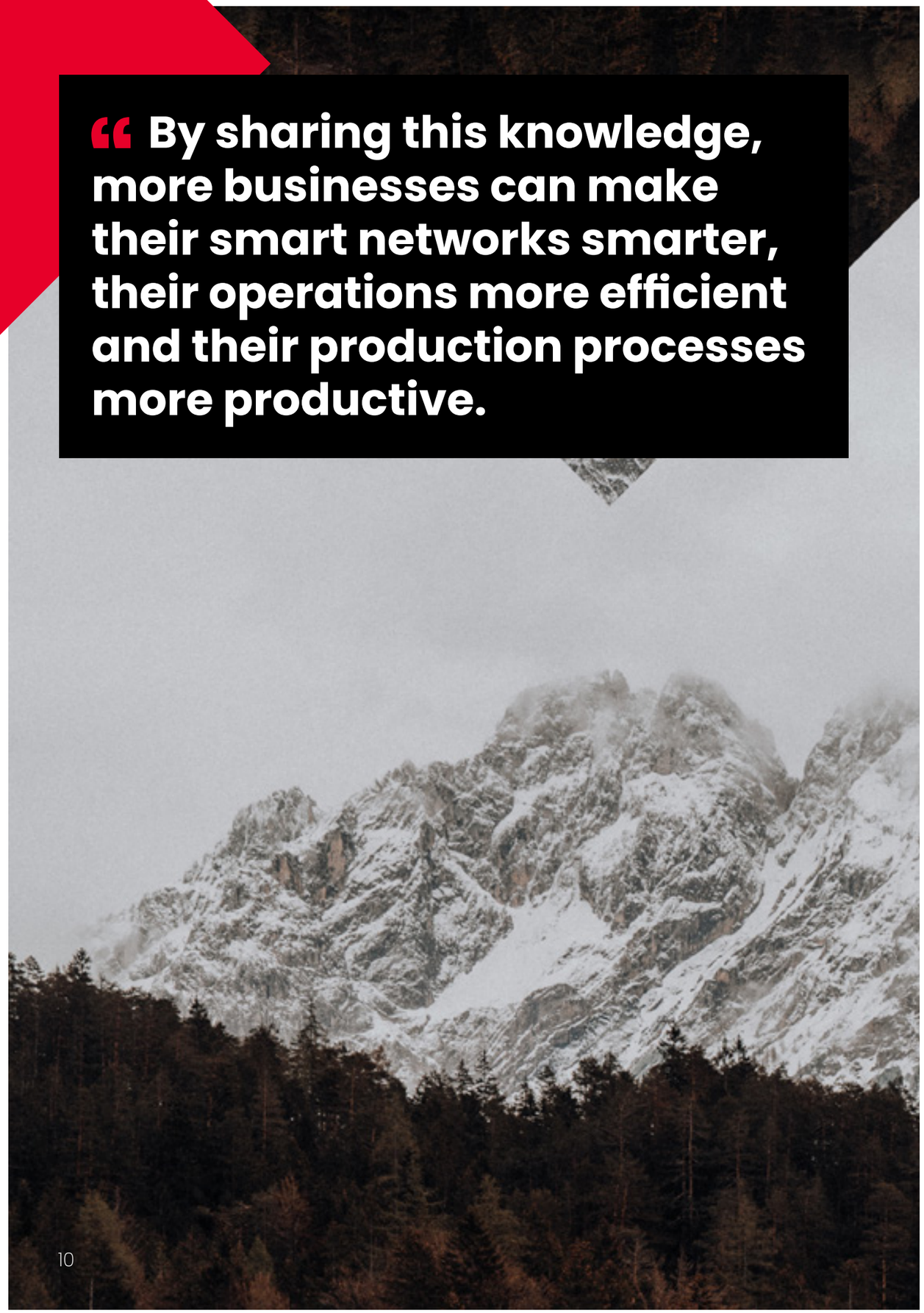
It's no secret that data collection is important for generating valuable insight that helps plant managers improve efficiency in industrial systems. However, what seems unknown to many in the industry is the importance of understanding the context of the data being analysed.

Tunnel vision is a problem across a variety of fields, from sports to industry. It is easy to get fixated on a single goal and to forget to take a step back for a wider view of a situation. Yet, taking this step back can provide invaluable insight and understanding of the reality of a situation.

Let's take a hypothetical scenario. Imagine that a maintenance manager has a machine that is bending wires. These wires have to bend to a very accurate angle to properly work. However, the machine isn't bending them properly at certain times of the day, causing the business to produce faulty goods that have to be discarded.

To fix this issue, the maintenance manager brings up the data collected by the device's onboard software. The manager then analyses the data with a digital twinning platform. Looking at the analysis, it becomes apparent that the machine is vibrating anomalously at certain hours of the day. The manager dismantles the machine, reassembles it and even runs it in an isolated scenario, but is still unable to find the source of the problem.

In this instance, if the manager had taken a moment to take a step back and looked for context, they may have realised that the anomalous vibrations coincided with the activation period of a nearby piece of heavy machinery. There is nothing wrong with the machine in question, but its surrounding context reveals the cause of the error. This is why context in data collection is vital.



> **“ By sharing this knowledge, more businesses can make their smart networks smarter, their operations more efficient and their production processes more productive.**

With an array of smart sensors and devices, paired with a digital twinning system like GE Digital’s Predix platform, across the whole production line, the manager could have clearly seen the correlation by viewing the data in context. Another method of achieving contextual understanding would have been by comparing the machine with other similar ones in different plants.

This shows why digital twinning is such a powerful tool. Being able to recreate an entire plant in a digital model breaks silo mentality. This allows managers to have holistic insight, which reveals issues that were previously not apparent. Whatever the issue, it is evident that context in data collection matters. Being able to analyse systems is now a reality and should be used to the full extent of its potential.

With potential to reduce waste and unnecessary expenditure, digital twins will allow stronger operations. With context being easily achievable in the modern industrial arena, it no longer needs to be a trade secret among industrial businesses. And by sharing this knowledge, more businesses can make their smart networks smarter, their operations more efficient and their production processes more productive.

Keeping ahead of the data curve

The amount of data that is generated, communicated and manipulated in modern industries is truly staggering. For instance, over a terabyte of data is saved each day in power production alone, and this is quickly increasing as time marches on. Handling this volume of data requires a certain calibre of system — here is what goes into these multi-terabyte data handling systems.

It's almost hard to believe that merely 30 years ago, a single terabyte would have filled nearly half a million floppy disks. As a species we've become quite desensitised to these dizzying figures. A terabyte just isn't all that much in the modern world — you can even get it on a fingernail-sized SD card.

Aside from the modern internet of on-demand video and conference calls, industrial data systems have been encouraging this trend upwards for decades. From 1960s punch-card computers, to 1980s magnetic tape reels, to modern SSDs; industrial data collection has always taken advantage of the latest developments to push the limits.

This theme of increased data collection has become fossilised by modern automation, as well as health and safety standards, both of which rely on complete data handling to operate robotics and keep workers safe respectively.

The net result of industry's data reliance has produced some interesting statistics. According to predictions recently published by industry analyst IDC, sectors such as semiconductor manufacturing and automotive manufacturing are producing upwards of two terabytes of data every day — as much as the combined water, energy and gas utilities sector. Both these figures are expected to skyrocket in the next five years, with semiconductor manufacturing expected to quadruple.

Data generated and stored by industry (per day)
2019 figures and 2025 projection
(Source: IDC)

Automotive OEM:
2019: 1.9 TB
2025: 5.84 TB

Food and beverage:
2019: 0.85 TB
2025: 1.53 TB

Power:
2019: 1.15 TB
2025: 4.52 TB

CPG:
2019: 0.74 TB
2025: 1.34 TB

Semiconductor:
2019: 1.73 TB
2025: 13.54 TB

Pharmaceutical:
2019: 1.27 TB
2025: 2.29 TB

Treading data

Many organisations, from a wide range of industries, collect data from their processes on these huge scales. The question is, what do they then do with it?

The simplest, traditional and most obvious way is to generate a regular report to be examined by plant managers and engineers. For instance, a bottling plant might create a report concerning pasteurisation temperatures and conveyor motor speeds, among other parameters. This way problems can be addressed if anything reports out of expected values.

While reporting is easy, it has pitfalls. Namely, even with daily or even hourly reports, the system cannot predict where or when the next fault will appear, as a report approach is entirely and inextricably reactive. Furthermore, in a plant with potentially tens of thousands of sensors, how do engineers decide what to include and exclude from the report? It just takes one oversight or mistake to wreak havoc.

What computers are best at

Luckily, industrial data handling software has been keeping a good pace with industrial data growth.

Modern SCADA and industrial historian software, such as those supplied by Novotek UK and Ireland, is well equipped to crunch the terabytes of data created by production lines around the world today. Instead of collecting data, generating a regular report and expecting the engineer on the day to parse it, these systems keep a watchful eye on all connected devices.

With integrated machine-learning algorithms, the systems can get a feel for the normal, smooth operation of the plant. If anything perturbs the algorithm, be that a hot-running motor, increased pump cycle-time or any other aberrant parameter, the system alerts engineers that an issue is developing.

This way problems can be identified and corrected before they start posing issues. For instance, a SCADA system monitoring a municipal sewerage plant discovered and notified engineers of a potentially catastrophic blockage, over twelve hours before it occurred. Under a report-based system the subtle operational effects that this blockage building up caused may have been missed, or the entire fault could have occurred between reports. Either way, the result is at best a more stressful maintenance experience, or at worst a horrendous mess.

So, while terabytes might make modern teenagers yawn, when that amount of data is flying at you uncompressed, in real time, it can quickly become overwhelming. It's best to hand the parsing off to those who are far better equipped for these logic-intensive— computers with industrial data handling software.

Developing an effective data strategy

The amount of data that is generated, communicated and manipulated in modern industries is truly staggering. For instance, over a terabyte of data is saved each day in power production alone, and this is quickly increasing as time marches on. Handling this volume of data requires a certain calibre of system — here is what goes into these multi-terabyte data handling systems.

In Herman Melville's Moby Dick, we find one of the earliest recorded references of a multitool: a "Sheffield contrivance", containing numerous tools in the exterior of a pocketknife. In the story, the item serves as a metaphor for the multitalented carpenter responsible for maintaining the ship in its pursuit of the titular white whale.

The metaphor is equally comparable to a good industrial data system, which plays a critical role in the pursuit of an effective data strategy — industry's own white whale.

In its recent global survey, McKinsey Global survey found more businesses agree that data and analytics are changing their industries in increasingly significant ways. Despite this, McKinsey also noted that many companies are "responding to these shifts with ad hoc initiatives and one-off actions, rather than through long-term strategic adjustments".

Part of the challenge is that each department of many companies, from procurement and C-suite to field service engineers, still gravitate towards a silo mentality. Yet in order to be effective, the data strategy should run throughout every level of a business, with the capacity for each area to have their own view of the data. If this isn't considered, the strategy will never truly work.

However, before we determine how to improve a data strategy, we should define what an effective data strategy is.

Defining good data strategy

An effective strategy is one where collection and aggregation of data is supported for every process, practice and piece of machinery that is pertinent to operations. From this data, a company can perform suitable analysis to scale granular information in such a way that it provides macro insights. These can be viewed by, and reported to, any stakeholders with an interest in an area of operation.

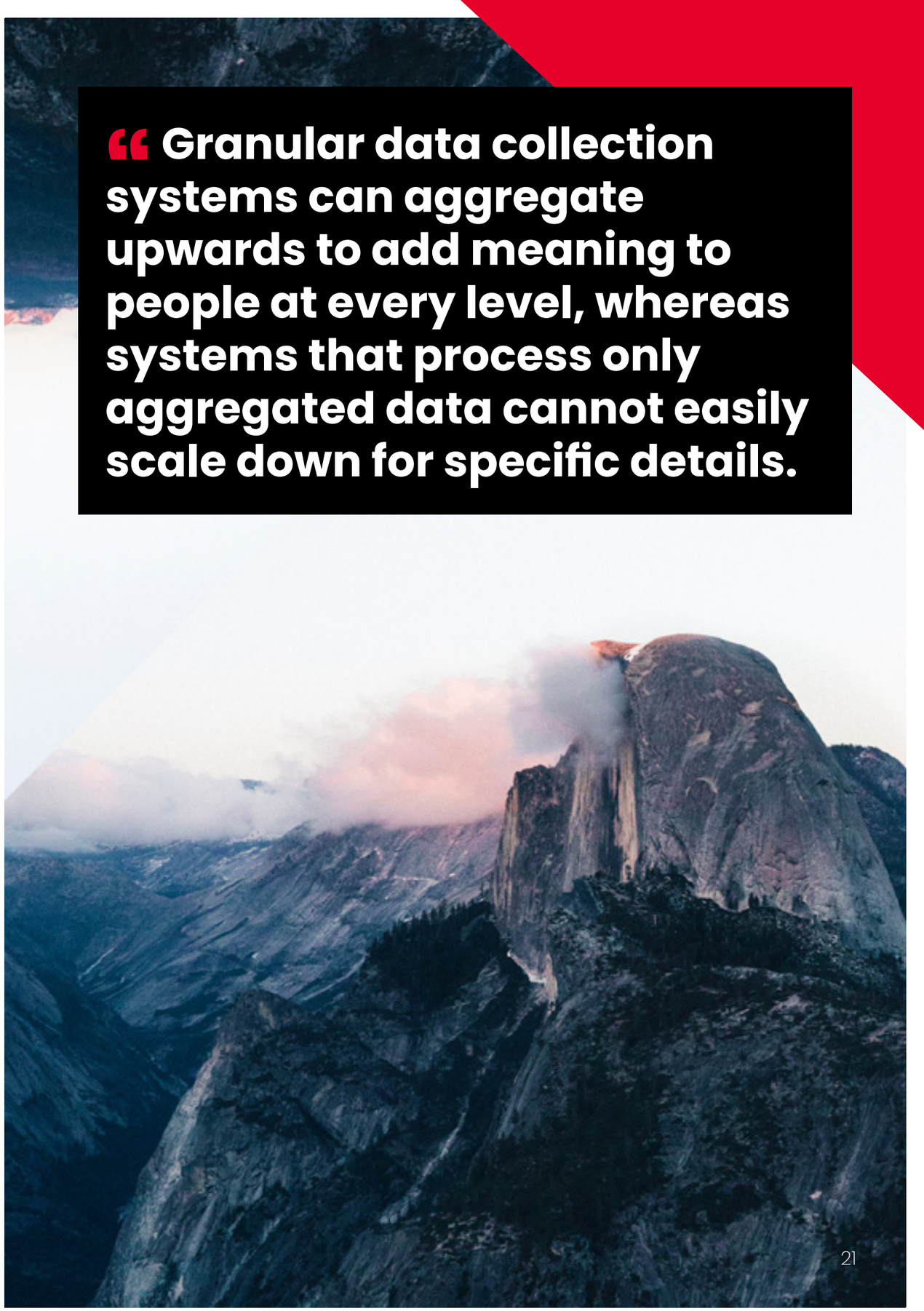
As an example, a chief executive for a multinational company might have a top-level view of plant performance, with a simple traffic light system to indicate normal, caution, or warning levels of overall plant health. This is enough for the executive to monitor overall operations and satisfy their interests.

If we assume that one plant in the UK indicates caution, the executive can speak directly with the plant manager. The manager's view of data will identify specific areas where there is a performance risk such as a specific engine on a particular production line, and have a quick view of relevant KPIs such as availability and uptime as a percentage.

Maintenance staff can view the specific technical issue using the same data at a more granular level, such as engine load information or oil temperature figures, and address it accordingly. If a replacement part or engine is needed, procurement and finance teams can view the make and model and process an order in their system.

This is the white whale, where the flow of data allows different stakeholders and teams to have different views and systems for interacting with one central pool of data. It's something that has been made theoretically possible by the advent of the Industrial Internet of Things (IIoT), but also hindered by many systems that have been dubbed "IIoT solutions".

“ Granular data collection systems can aggregate upwards to add meaning to people at every level, whereas systems that process only aggregated data cannot easily scale down for specific details.



Collecting and scaling

Many IIoT solutions encourage engineers to embrace a simplified approach to data collection, such as only state change data or certain parameter deviations. With this, it is impossible to collect the richness of data required to support several different improvement processes, from fixing the issue to understanding why an engine broke and predicting future breakages.

The best choice is a modern Historian software that can collect various types of data such as state change, alarms, quality system information and process parameters. This allows for very granular insights into specific issues, but it can easily integrate into other software, such as a manufacturing execution system (MES), and run calculations on data sets to provide higher level insight.

It's here that an effective data strategy makes an immediate difference. One issue that we at Novotek encounter all too often is that businesses employ several systems to collect the same data without realising, and are reluctant to invest in multiple analysis and visualisation tools, such as an MES alongside a Historian software or an IIoT solution. Fundamentally, businesses are spending money in the wrong places and are subsequently not getting the return on the investment that they want.

The best advice is to think of data like the "Sheffield contrivance" multitool alluded to in Moby Dick, or the modern Swiss army knife. It's most practical and efficient to use one system with various tools that are applicable to other people, teams and situations. Granular data collection systems can aggregate upwards to add meaning to people at every level, whereas systems that process only aggregated data cannot easily scale down for specific details.

The good news for industrial businesses is that with this in mind, and the white whale in view, it becomes a much less complicated process to capture it. You should set your industrial strategy as one that, for the first steps, endeavours to streamline the number of systems collecting the same data. Instead of a wide array of IIoT solutions, look instead to one store-all system like Historian software.

Then, look at ways to analyse and present the micro-level, granular data at a macro level to serve different business areas. For example, consider the calculations run to present information differently and yield alternate insights. If you are not thinking about who will use the data or have an interest in the data, then your strategy will not be right. Put more than one tool in place if necessary but look at how to share data between systems.

As an industrial data specialist active across the manufacturing, processing, power and utilities sectors, Novotek has extensive experience guiding plant managers and executives alike in the right direction to improve their data strategies. Beyond noting how data can present different insights, the most frequent advice we also give is to not think of system investment as an afterthought of strategy. Incorporating these options into the strategy means you can identify where multiple systems are complementary and beneficial, or where they are a wasted investment.

Data, and the value it possesses, can provide numerous short-term benefits but the biggest advantage comes from a long-term, scalable strategy. With the right approach, underpinned by the right software and systems for collection and analysis, industry can at last catch and conquer the white whale.

Are you starting your industrial digital transformation journey and want some assistance with developing an effective, future-proof industrial data strategy?

+44 (0)141 332 1551

Novotek