BREAKING DOWN DIGITAL SILOS FOR CENTRALISED REPORTING AND IMPROVED OPERATIONAL INSIGHTS

Quentin Griffiths[1]

1. Lutra Limited, Wellington, New Zealand

KEYWORDS

Operational insight, data silo's, automated reporting, performance improvement.

ABSTRACT

Municipals are data rich environments, with field teams, SCADA, external laboratories as well as a wider range of IoT devices and external data providers, all acting as sources. With these rich data sources comes the challenge of data silos, with each source often held within a firewalled system. The following paper looks at the process of aggregating these data sources into a single platform, Infrastructure Data (ID), and the data-driven outcomes that Lutra's clients have achieved by gaining this operational insight across their assets, allowing for improved efficiency and compliance management.

INTRODUCTION

The operation of water and wastewater treatment plants is by its nature, a complex task, requiring trained personnel who have a good situation awareness of what is happening at the plant at any given time. Supported by office-based compliance, asset management, planning and management teams, it is natural that a range of systems and processes have been developed to ensure the supply of safe drinking water whilst also safeguarding against environmental harm from discharges.

Because of this range of teams, with their dedicated roles, it is common that each will have a different set of tools and systems that they use to aid in their day-to-day work. Access to these systems is often limited to the team directly responsible for the task, who are then trained appropriately. Whilst this creates efficiency in the aspects of training and licensing costs, it also develops silos between teams and databases.

A common tool used to combat these data silos is Microsoft Excel, with manual extracts being taken from the different systems of record, manually manipulated, and then used to generate an output. This requires time and effort to complete and needs each team to be responsive when sharing the datasets. Whilst this helps to bridge a gap, it takes time to manually prepare, can be error-prone, has no version control and is outdated by the time it is ready to be used causing disconnects to occur.

The following paper, therefore, examines a range of short case studies that highlight the process of aggregating real-time data into a single location, works through how the FAIR (Findable, Access, Interoperable, Reusable) data principles come into play, and the outcomes of rolling out an aggregation platform like Infrastructure Data to the day-to-day operations of a municipal.

IMPLEMENTATION PROCESS

Data aggregation is not a new concept, with the creation of data lakes being commonplace in enterprise organisations. The driver behind the implementation of a data lake is to create a central repository for all the organisation's data, in a single, trusted location.

Whilst this method is commonly used in enterprise organisations, the effort required for a data lake to be implemented, and more importantly, maintained is often unattainable for many utilities.

To address this challenge, Lutra work with water utilities to assist with aggregating operational and compliance data into a single platform; Infrastructure Data (ID).

Before kicking off a project, the first step is to define the outcomes the utility is looking for – whether it is a compliance-based focus or an operational insights driver.

Taking the example of compliance-based reporting, the first step is to understand the treatment processes within individual plants and to define the standards that they need to comply with. This gives a clear indication of the types of datasets which are required, refining the focus of the data sets, whilst also building knowledge between the Lutra team and the municipality.

From this point, the source of each dataset is defined, starting with SCADA for the process variables. The data points from the historian are identified, and any appropriate data cleansing and adjustments are made to the frequency and reported accuracy of the variables.

Focusing on a smaller subset of the SCADA data creates a specific dataset that is easily found, identified, understood and reusable within the desired reporting frameworks, closely aligning with the FAIR data principles.

Following the integration with the SCADA historian, the focus shifts to external and internal laboratories. Taking a similar approach, a complete list of sampling points is provided, with the most appropriate subset of data being selected for the required reports. A connection is established, and the data transfer begins – connecting the two systems and taking lab results from a static PDF document to a live dataset. Depending on the external integration capabilities, these connections range from automated CSV export to an external server via FTP, or through a direct API connection.

The final step in the process, when referencing a compliance-based example, is the field measurements or observations that are undertaken by the operational teams. Within many utilities, the data collected from the field is either entered into a paper-based logbook, or an Excel spreadsheet.

With ID paper-based manual entry forms are replaced with digital versions, allowing for field measurements, calibrations and the likes to be captured in a digital format from the field.

Moving to this digital format, as with laboratory data, allows an otherwise static dataset to be immediately addressable and usable within reporting or trending applications, helping to close the loop from what is happening in the field, to what is being visualised within a head office.

Once the datasets are identified, and the staff are onboarded to the platform, Lutra work with the operations and compliance-based teams to develop the appropriate regulatory reports, as well as the supporting operational dashboards to aid in managing compliance. During this process, both the municipal and Lutra get a deeper understanding of the data, where to access this and the typical naming conventions used.

This process also helps to show the value of data aggregation to the municipality, which in turn aids in wider organisation buy-in and understanding.

This understanding then helps to inform the second stage of the process, whereby additional datasets, such as weather, energy or chemical consumption can be added. Taking a similar approach to the compliance-based example, these are added as an application is identified, ensuring that the data being added is useful, well-understood and applicable. Examples of these applications include water scarcity modelling, chemical usage, demand forecasting, and energy consumption, all of which provide a wider view of operational efficiency, asset performance and environmental reporting.

CASE STUDIES

The results of implementing a platform such as Infrastructure Data, will depend on the desired outcomes of the individual project. A key outcome experienced across all projects is that when data silos are removed and data aggregated into a single source of truth, the gap between teams is bridged, improving trust, and creating cross-pollination between operations and office-based teams.

Case study 1 – Compliance Alignment

The first example is presented in Figure 1, whereby the combination of SCADA data with a digital field form, records actions undertaken by the operations team at the physical treatment plant. In the example shown in Figure 1, a large turbidity spike occurs on the online meter. Referencing the field data, it can be clearly seen by an office-based team, that a verification process was undertaken on the equipment at the time of the spike, highlighting the likely cause of the jump.

The office-based team can dive further into the dataset to understand the actions that have been undertaken at the site, without needing to physically visit the plant, nor interrupt the operations team to understand the cause of the issue.

With this data being captured in real-time, even if the spike was not picked until the monthly reporting period was due, then there is still an online record of the events that have happened on-site during the month. This negates the need to trawl back through paper logbooks to understand the routine maintenance events that happened at the beginning of the month. The outcome of combining these datasets is a time saving, whilst also giving insights as to the routine maintenance that is occurring on site.

Applying the maintenance records against the live SCADA data also helps to pick up changes in the equipment's behaviour following a manual intervention, aiding in long-term trending and troubleshooting in the event of regression, or similar errors.

Case Study 2 – Pump Efficiency

Understanding the performance of a treatment plant is paramount to improving the overall efficiency of the plant.

Within a treatment plant or reticulation network, a portion of the ongoing operational costs comes from operating pumps.

In many cases, pumps are installed, tested and then left in operation until either a failure occurs, or a time-based asset replacement trigger is reached.

Whilst time-based maintenance and replacement is an easy metric to track and plan for, it gives no indication as to the rate of degradation of the pump, nor the increased operational costs related poor performance.

In Figure 2 the efficiency of a pump is modelled out, using the available parameters such as RPM, flows and pressures. Combining these data points allows for the efficiency of the pump to be modelled out in near real-time, with a direct comparison to the expected efficiency curve, as documented in the manufacturer's datasheets.

An energy cost is then attributed to the operation of this pump, allowing asset management teams to use data-driven insights and costs to deem when an asset is due for renewal, rather than just using a time-bound replacement scheme.

In November 2023 a New Zealand municipal used this approach to model out the performance of a duty-standby pump arrangement. At the time, they were experiencing what was initially thought to be a large leak, which was causing a reservoir level to drop at an unexpected rate.

Figure 3 represents the visualisation that was used by the municipality, which identified that a recent duty change had been made to pump 2, limiting its control to 40 Hz, rather than 50 Hz. The result was that the treatment plant was unable to meet the demand, resulting in the reservoir level dropping. The office-based team alerted the site operations which made the appropriate adjustments. The outcome was that the municipal did not have to deploy a leak detection team, saving both time and money.

Case Study 3 – Chemical Predictions

Chemical usage on a treatment plant can provide insights to the overall efficiency of the plant, whilst also helping to highlight changes in the plant's behaviour – due to maintenance issues or changing conditions to the raw water intakes.

Referring to Figure 4, the chemical consumption for each process within the plant is modelled out and tracked. A prediction of the assumed chemical consumption is provided, based on the previous behaviour of the plant, as well as the expected performance related to the raw water quality.

This chemical tracking is useful from a OPEX standpoint, providing a deep understanding of where the costs lie within the treatment process, with the costs being updated regularly to match ever-increasing costs. Using this detail, municipals can target costly processes and invest in experts, such as Lutra, to optimise the process.

Outside of the pure cost analysis, chemical consumption can trigger an alert when a large variation is observed – be it instantaneous or over a long period of time. This provides operations teams with a proactive alert to investigate what has changed, either in the plant or within the water source – allowing for contamination or equipment degradation to be identified sooner, before a major issue occurs.

Case Study 4 – Summer Demand

Water scarcity is becoming a common challenge across the globe, with summer periods becoming longer and potentially drier. Understanding the demand during the summer periods provides insights as to when water restrictions need to be put in place, whilst also informing scheduled plant shutdowns, ensuring that there is enough headroom in the network.

Referring to Figure 5, weather data, historic consumption patterns and production data from across the network, ID accurately models (~<10% variance) the demand for the coming 5 days.

This display is actively used by a large New Zealand municipality to inform their water restriction levels, whilst also providing information on their available product headroom for the coming weeks.

Case Study 5 – Drinking Water Reporting

In the context of compliance-based reporting, New Zealand-based municipals have reduced their monthly compliance reporting down from 5 working days to ½ of a working day, allowing their compliance teams to focus on improving their outcomes, rather than wrangling spreadsheets.

This time saving is driven by ID's ability to collate laboratory data, SCADA data and field measurements into automated reports. Powered by a rules engine, each report confirms the level of compliance against the appropriate drinking water standards, proving a pass or fail result. Failures can be commented on and have a direct connection to the New Zealand water regulator, Taumata Arowai.

Table 1 highlights an example of lab data reporting, tracking algae counts within waterways. Active alerts are generated automatically when the presence of algae is detected within the sampled zones. This removes the need for manual review of laboratory PDFs and provides immediate feedback on potential issues.

CONCLUSION

In conclusion, municipals are data-rich environments that often struggle with data siloing. When combined into a platform such as Infrastructure Data, they can gain insights that help to improve their operations and highlight issues ahead of time.

Working with over thirty-five municipals across New Zealand and Australia, our clients have experienced time savings of up to 4 working days a month for compliance reporting, as well as aiding operations teams in diagnosing issues with their treatment plants before they take a plant offline. Data is a powerful asset, that when properly combined can change the way that a municipal operates.

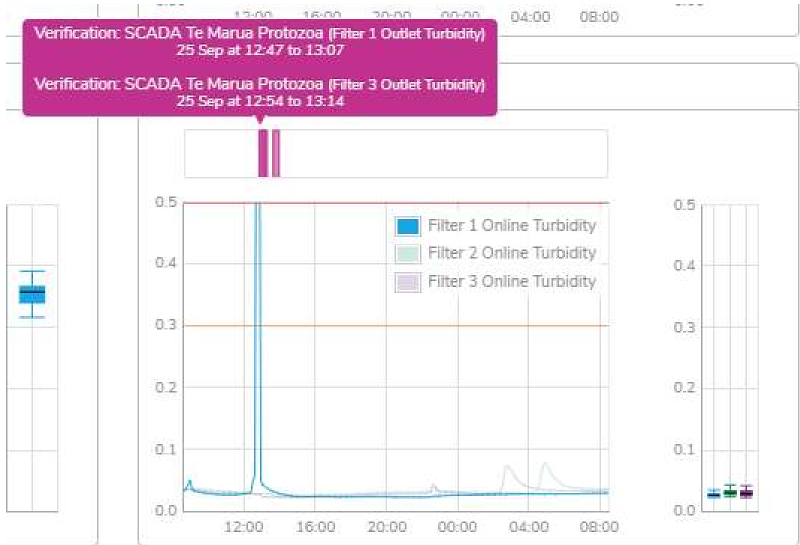

Figure 1: Combining field data with SCADA data provides explanations to large data spikes for filter turbidity.

Figure 2: The efficiency of a pump is modelled out vs the manufacturers efficiency curve, and a cost to run is attributed through the resulting performance.

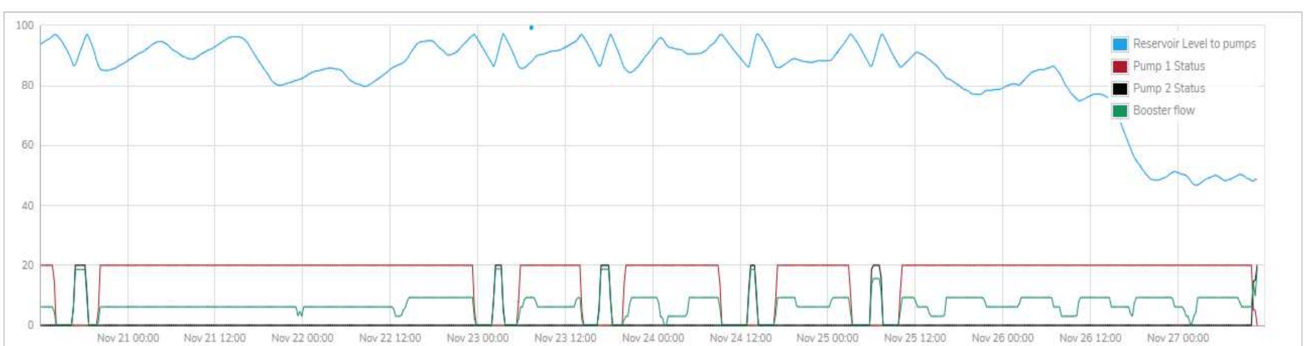

Figure 3: Duty-standby pumps are modelled out vs reservoir level and used to track down a rate limiting issue.

Figure 4: Tracking chemical usage against expected or predicted values offers insights into reduce plant performance.

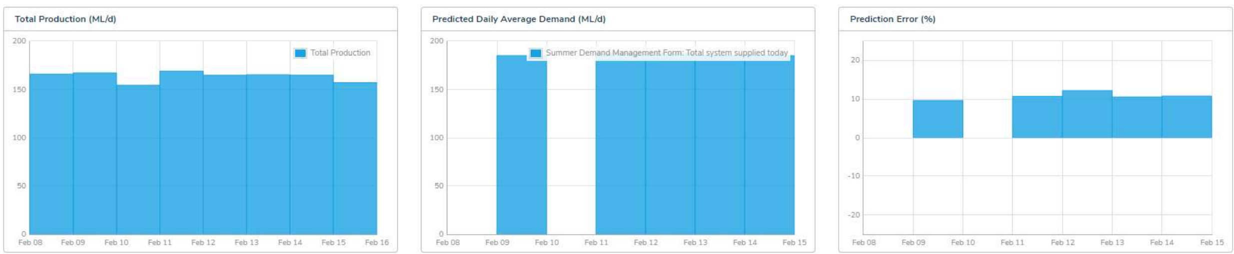

Figure 5: Predicted summer demand for the coming week, allowing for water restrictions planning and plant maintenance.

Table 1: Highlights the presence of algae within the waterways, raising alerts when algae is present in laboratory samples.

FAILED - SM1S1

	Presence	Not Detected	Possible Positive	Not Tested	Number of Samples
mcyE	0	21	0	6	27
anaC	8	19	0	0	27
sxtA	1	21	0	5	27
Cyr	1	21	0	5	27

FAILED - SM1S3

	Presence	Not Detected	Possible Positive	Not Tested	Number of Samples
mcyE	0	20	0	14	34
anaC	17	17	0	0	34
sxtA	0	20	0	14	34
Cyr	0	20	0	14	34