

Integrating Environmental Systems - Why?



EHS Data Ltd - 2018

www.ehsdata.com

Why?

The majority of environmental monitoring is done because it's a legal requirement, usually through the issuing of a permit to a site. For many or even most larger scale operations, permits require many different media to be monitored (e.g. air, soil, surface or ground water and noise). Often this needs to be combined with supplementary data (e.g. production or weather) and meta information (e.g. flows, height, location). Different frequencies apply to different data streams, difference compliance levels, calculations and so on.

Often monitoring equipment comes with its own software but this rarely seeks to combine its data with data from all of the other equipment on the plant, field readings and lab reports etc. - nor manage many of the other permit requirement such as managing monitoring schedules, analysis of collated data sets or reporting. These information silos of data are often different from one another with varying interfaces and capabilities.

This is where integration comes in. Data integration involves combining data residing in different sources and providing users with a unified view. An integrated system solves many of the problems inherent with information silos. They can;

- Hold and manage all key parts of the permit pertaining to monitoring.
- Manage the whole monitoring plan.
- Harvest, validate and alert issues.
- Combine, Collate, aggregate and calculate different data sets.
- Provide a single, uniform platform for analysing and reporting all monitoring data.
- Permit sharing of information across an organisation in a controlled, secure way.
- Prevent many instances of human error through copying and pasting from one system to another and into, for example, spreadsheets.

How?

With so many different datasets – how is it possible to combine into one single interface? This needs a lot of skill and experience to do, however with the right tools, it can be done with virtually any type of monitoring data. Automation is also vital. Ideally, a system should minimise as much as possible the need for users to continually populate it with data and fix import problems. Some of the things to consider are;

Delivery Method

- How is the data coming into your system?
- Is it being emailed? Can it be automatically processed, and imported, along with any ancillary reports (e.g. lab reports).
- If needed can your system connect to other site systems and databases to harvest data?
- Can it keep an eye on folders to 'pick up' files of data that land in them? Will it synchronise with field entry devices?

Layout

Files coming from multiple different silos will inevitably have almost as many different layouts as there are data sources, and it is usually simply not practical or possible to seek any consistency between them.

Things like header information, column names, column order, file orientation - is every column a different parameter in a cross-tabulated layout, or maybe sample point headers, of the file could be a long list. So your importer needs to be intelligent and flexible. It must



be able to read most files with no user input, to interpret column order and orientation and so forth. Sometimes this will not be sufficient, so user configuration for special cases should be provided, along with the ability to insert custom scripts or macros to deal with the trickiest of layouts.

Special Cases

Many data streams have special cases or rules which need applying. Some equipment might provide US dates. It is imperative that the importer knows this and how to handle it. Is 07/08/2018 August or July? There are many varying date formats – 01/02/2018 2018-01-02 01-Feb-2018 and so on.

Many .csv files have comma column separators, but other delineators can be found such as semi-colon or tab. The importer needs to know which to expect. Other files may be missing columns. For example, a flow meter knows which meter it is and might not include its name in the output file. This would have to be specified to the importer, or 'all files arriving in the is folder is from x laboratory', etc.

With laboratory readings, < symbols are common (e.g. <0.01). They simply mean 'below detection limit'. This doesn't mean zero, but different regulators have different requirements for their handling. Is it the full amount up to the detection limit, half of it or zero? Your importer needs to know which in order that it can interpret not just as text, and post the correct number in the database for use in interpretation, but also as text for reports. Comma decimal separators (e.g. 0,01) are common especially in continental Europe or Latin American countries. The importer needs to understand that this is not a column separator but a decimal 'point'. So on setup, the user or administrator must have choices and flexibility on how to set up any data stream.

Consistency

Consistency in naming can be a huge problem if not handled correctly. Different equipment might call the same location something different, e.g. 'channel 1' might identify the same location as 'stack 1', or over time people manually recording a name differently, e.g. BH1 BH01, Borehole 1. These must all be combined, and the imported has to know the variants. A good importer will learn these aliases over time and hold the mappings in order that it only needs to ask once.

Format

Can your system handle multiple file formats, for example csv files, excel spreadsheets, text files and so on.

Calculations

Automatic aggregation and calculation is often not as simple as it might seem. Take one seemingly simple example of a mg/m³ compliance limit. There might be a sensor measuring mg/l and a flow meter measuring m³/minute, both at different intervals. The calculator must know where to search back in time to find the nearest applicable reading to use. Or rules on how to handle a meter going round and back to zero. Once the result is obtained, there may now be a new data set which now needs a compliance check (and alert if appropriate). There are many complications which need handling.

Conclusion

So, in conclusion, there are many reasons for the need for integrating environmental systems, however, some of the few examples above demonstrate some of the difficulties. It is not simple to build and implement, specialist tools and skills are required.

