PENSPEN

‘The Tide is Turning: The Role of Digitalisation in the Pipeline Integrity Industry’ Q&A

Thank you to everyone who registered and joined us for our webinar entitled ‘*The Tide is Turning: The Role of Digitalisation in the Pipeline Integrity Industry*’.

Our presenters, Nigel Curson and Aidan Charlton, have answered your questions that were submitted during the webinar.

If you do have anything further that you would like to ask our presenters, please contact them directly.

Nigel Curson, Executive Vice President for Technical Excellence

Email: n.curson@penspen.com

Aidan Charlton, Technical Manager

Email: a.charlton@penspen.com

Questions and Answers

Q: How do you integrate different ‘information format/input’ (e.g. Excel, DB, Word, printed/scanned data)?

A: Data objects such as pipeline tallies/listings containing In-Line Inspection (ILI) or similar results are commonly transferred inter-company via spreadsheet, nominally to a common format such as PODS. Formats such as PODS leave enough discretion to data originators that the data is typically not machine readable or useable out of the box. THEIA is trained to recognise many common formats from various vendors, under a variety of formats, and to detect and correct for many of the vagaries in inherent with a looser data format and the choices made by data originators in applying data formats. Where this is not possible, THEIA has been trained to use heuristic data recognition

routines, and to detect low certainty data identifications for subsequent detailing in the automated report.

Similar routines are available for the upload of other file formats, such as .JSON files or .DB files for bulk upload.

Penspen directly control the data recognition routines so are also able to provide custom data recognition routines to suit client requirements.

Q: How can the veracity of the data be validated?

A: THEIA has been programmed to use heuristic data recognition routines and to detect and correct data formatting errors. Low certainty data is identified for subsequent detailing in the automated report. In all cases, a conservative assumption is made for data that cannot be positively identified by the automated alignment algorithms.

When developing functional code, each code operation is subject to a specific function operations test to ensure proper and robust operation. Similarly, the overall function of the code is verified against known solutions from previous SME led calculations. This includes verification that proper conservative assumptions have been made where the veracity of input data is questionable.

Q: Who are the leaders in applying machine learning/predictive analytics to integrity management?

A: We are not aware of any criteria for measuring this but do know of a few companies working in this area including us; however, consider the following:

It is entirely possible to use current machine learning algorithms to predict the safe working pressure of a metal loss features. If technology is applied correctly, the results can align to a range of established methodologies, such as ASME ModB31G or Kastner, with a very high degree of accuracy. But is applying Machine Learning in this way really a good idea?

In this example, the existing equations are quick to perform using standard computing; no appreciable performance gain is expected, so there would be little or no performance benefit in using machine learning. The bigger consideration really must be, at what point are we as engineers happy to allow critical engineering decisions to be driven by the results of a black box process? Yes, machine learning and artificial intelligence can improve the performance of some applications, but careful consideration must be given to the balance of potential improvements and potential risks introduced by new technology.

A simpler criterion for now is to work on the basis that the process and processing behind all safety critical calculations should be recorded and be human reproducible. This does not necessarily preclude the use of machine learning and artificial intelligence from all tasks, just that these technologies must be applied intelligently. Where an alternative direct mathematical model can be developed, then the use of the mathematical model is likely to be more appropriate than the use of machine learning or artificial intelligence in most areas of integrity management.

Q: Will PHSMA eventually mandate full digitisation for integrity management? Anything in the Mega Rule.

A: Very unlikely, since this would infer a standard approach, which is hard to envisage. It would also be burdensome on some operators.

Q: Is it feasible to repurpose offshore oil and gas pipelines to transport HP hydrogen back to shore?

A: Yes, but there are potential issues with hydrogen embrittlement. Stresses must be limited. Existing H_2 pipelines only use low strength steel and at low stress levels. In practice, this means that existing pipelines fabricated from lower grades of steel may be more suitable for conversion to H_2 transportation than pipelines fabricated from higher grades of steel. This is, of course case, specific.

Q: If my customers does not have sensors nor specialists to handle it, can you deal with data collecting with instruments as well?

A: Yes, but we would typically tend to partner with somebody that has the geographic coverage to do the physical work and somebody to provide the sensors.

Q: Would you be able to perform freespan analysis?

A: The development of freespan analysis is nearing completion and is aligned with commonly used methodologies. Following a successful internal peer review, this analysis class will be made available to users of THEIA in the very near future.

Q: Should I have few ROV reports, would you be able to correlate the reports and predict the freespan behaviour?

A: Extracting span lengths from video data automatically using artificial intelligence is something we are working on. In common with other vendors, developing access to suitable example datasets is a priority. Penspen is working with industry to assist where possible.

Existing feature correlation routines in THEIA are perfectly capable of being applied to the correlation of existing data detailing features, such as freespan.

Prediction of future span behaviour, such as change in location, dimensions and calculated values is being considered and is likely to form part of the THEIA development roadmap.

Q: How to use cloud computing without internet connection? Even get data from sensors... Onshore pipelines crosses full countries, so many internet blank spots...

A: If it is field work, we would propose/implement a synchronising database, which updates when a connection is obtained. This is quite common practice under these circumstances.

Even for permanently attached sensor networks, we are tending to see satellite connectivity reducing in price to a point that can be considered cost effective. This can be further optimised to reduce data throughput and improve latency by the intelligent application of edge computing. Integration of data derived from edge computation is something that is built into the fabric of THEIA and would seem to offer a lot of advantages for connectivity blank spots.

Q: Are you going to sell it by modules? I don't need to buy the whole software, but just the parts I need.

A: Yes, we are doing exactly that.

Q: Is there any offline version to be used in locations where no-internet connection exists?

A: We are not planning an offline version of the whole system, but for elements of this, yes. Particularly those focused on field work.

Q: Considering the lifetime of a pipeline, is it really worth it to invest this much into a pipeline, which is in most cases a temporary construction? Even lasting for decades it is still a temporary construction.

A: The investment is mandatory in a sense, because we have the responsibility to manage risk, comply with the law and protect people, the environment etc. THEIA is designed to help with this.

Q: Do you need historical failure data to actually predict future failures? If so, what is the time span needed for data collection?

A: This really depends on the style of analysis being conducted.

Simple deterministic assessments, such as ASME ModB31G and Kastner, are widely available and accepted based on existing historical test data.

Probabilistic methods similarly have examples of previously developed and accepted methodology. Commonly, these are promoted for acceptance via bodies such as EPRG and PRCI and can be based on previously collected datasets.

Where additional training data can be required is in the development of artificial intelligence and machine learning models. In many cases, historical datasets can be used to jump-start the learning process, and in many cases, there is sufficient data already collected. This is, of course, dependant on sufficient data being available to be appropriate with the actual frequency of event under analysis.

Several working groups have been established in the industry to attempt to pool data for more difficult training problems.

Q: As the oil and gas industry is moving towards digitisation, are the new skill sets required in the pipeline integrity industry?

A: The digitalisation of the oil and gas industry in relation to skills and competency has two distinct sides.

The approaching primary skills and competencies gap that the pipeline industry is expected to encounter with increasing severity in the coming years is well understood. Part of the solution will be better management of the transfer of knowledge from more experienced engineers to less experienced engineers, in addition to better recruitment and retention within the industry.

There are recognised methods of addressing the primary skills gap, however, technology and an industry 4.0 philosophy can offer some respite from the problems of maintaining the required size of competent workforce.

Using pipeline integrity as an example, the capacity of a company or department to perform pipeline integrity tasks in a timely manner is directly related to the number of engineers with appropriate competence who can be maintained within that company or department. Where industry 4.0 technology is properly implemented with appropriately controlled data flows and algorithms are designed by subject matter experts, it becomes possible to provide detailed integrity services on an immediate basis, with a much-reduced requirement for difficult to find pipeline engineers.

Industry 4.0 therefore has the possibility to transform many pipeline operations and integrity functions from being limited by the number of man-hours available, to an on-demand SME basis where the technology never sleeps.

The second skills gap relates to the pipeline engineers of the future, who will be required to have the digital skills to be able to develop and interact with data and uses of data as required by modern technology in order to develop the coding behind digital applications. Therefore, in the future at least, some pipeline engineers will require all of the skills and competencies of existing mechanical based engineers but will also be expected to be competent with data science techniques and computer coding. The second skills gap can be succinctly defined as the people who have the competence to write efficient code and who also possess specialist domain knowledge of the pipeline engineering industry.

Essentially, the pace of future technological development within the pipeline engineering industry will be directly related to the capacity of the industry to solve the second skills gap.

Q: How is it different from PODS?

A: PODs is a data based standard/data structure. THEIA is a data management, visualisation and analysis environment.

Q: Can you talk more about how THEIA's development schedule is relative to some other service providers?

A: We, of course, would like to know the answer to this too, but in practice we are only vaguely aware of what others are doing. We are aware of what already exists in the market.

Q: How long before THEIA is better than human experts?

A: This really depends on the definition of better. We are building in as many insights as we can based on the knowledge base developed by Penspen over many years in the pipeline industry. In this case, the output is only as good as the person building the system. The real benefit of THEIA is the ability to automate and correlate multiple processes that a human may typically undertake but in a fraction of the time. In that sense, the machine is already better than the human.

Where machine learning and artificial intelligence are applied appropriately to specific cases, particularly pattern recognition, the machine can undertake tasks that are simply not possible to achieve via human action.

Q: How is field assessment related to the In-Line Inspection (ILI) reports alignment done?

A: Locational data can be imported, exported, and converted to suit client requirements. This includes typical references, such as chainage, girth weld number, WGS84 positional data or data based on local datums. THEIA is EPSG compatible, so any typical EPSG datum can be used. Additionally, custom datums using normal input variables, such as reference ellipsoid, meridian and geoid, can be defined to suit client requirements.

As all data correlation is automated, THEIA can output the correlated locational data from multiple data sources simultaneously, in your choice of locational data type. This can come in useful where a single data source has drifted out of alignment.

Q: Can THEIA be customised to different types of transport systems? E.g. mining industry slurry, water or tailings transport systems.

A: Certainly, we have a background with pipelines under a variety of operating conditions, including slurry, water and tailings transportation so are well placed to understand the differences and develop and implement appropriate methodologies.

Q: What can sensor manufacturers do to make it easy to connect to the THEIA platform?

A: This really depends on the route taken to deliver data to the THEIA platform, if the sensor data is already being uploaded to a client database and is formatted at least reasonably coherently, then data can be requested and integrated into assessments via a simple API call.

Direct connection to THEIA is feasible, however any direct data connection would have to be audited by our technology partners to ensure that client data remains fully protected at all times.

An alternate solution that may be of interest to those with cost of transmission, bandwidth or latency limitations - such those sometimes experienced using satellite communications - would be edge computing. This is essentially a small part of THEIA in a box to which sensors can be connected for use in the field.

If sensor manufacturers or operators are considering solutions such as direct data connection or edge computing, the most appropriate recommendation is to contact the THEIA team for a discussion.

Q: How does it work for facility integrity management? Do you have any case studies you can share?

A: THEIA is focused on pipelines right now. Facilities are on the roadmap, but currently are some way off.

Q: Who are your major clients who have signed up to use THEIA?

A: We would have to seek their permission before disclosing this. We hope to include some use cases on our website in the future.

Q: How are you overcoming the lack of interoperability between IT systems? This seems to be one of the largest issues.

A: Historically, technology companies have tended to operate on a basis of data being very easy to import but very difficult to export – e.g. try exporting your music collection from certain brands to a format that is recognisable universally such as MP3.

To have a controlled but otherwise free flow of data between potentially competing specialist technology and services providers will require a shift in philosophy from a typical technology provider's viewpoint, to one more compatible with the pipeline industries requirements.

THEIA tries to make the collation of data from disparate sources as simple as possible. Multiple file formats are recognised, collated and aligned automatically using a combination of heuristic recognition and defined formats. Custom formats are comparatively trivial to add into the code for additional expansion capability. Similarly, data export is designed in.

Ultimately, the application of industry 4.0 to pipelines will be made easier by addressing the fundamental flaws in the data structures currently being transmitted between companies, as well as by addressing the lack of a live and controlled data-sharing framework, which retains robust master data management probably owned by the pipeline operators themselves.

This shift is something that will have to be driven by operator demand and action. Penspen would be happy to work with other stakeholders within industry to drive this forward and already are doing so with the likes of the PDEF JIP (Pipeline Data Exchange Format).

Q: Is there any application/idea of machine learning in flow lines that mostly are not piggable?

A: Unpiggable pipelines pose a unique challenge to operators. The Penspen-THEIA team are developing algorithms in this field and would expect to make a further announcement in the near future.

Q: How do you align data for corrosion assessment?

A: For any assessment, relevant data can be obtained from a variety of sources, which can also be based on different reference systems. For example, historic construction data may be based on chainage, ILI data may be based on counted girth welds, external surveillance may be based on Northings and Eastings, and excavation and direct assessment in-field data may be based on latitude and longitude.

THIEA selects and applies the most appropriate algorithms to align all data to a common internal reference system for further analysis, such as for corrosion features. All alignment algorithms have been custom developed in-house by the Penspen THEIA team.

Penspen follow the ethos of only using artificial intelligence and machine learning where a better, fully traceable and justifiable mathematical model cannot be developed to complete the task in an appropriate timescale. In the case of data alignment, the Penspen THEIA team has developed better direct mathematical models, and therefore the use of artificial intelligence and machine learning has been rejected, as it is inappropriate.

As responsible engineers, we must remember that the pipelines we work with daily can have a high consequence of failure. For that reason, we must only use techniques such as artificial intelligence and machine learning where it is absolutely required and where it is fully appropriate to do so.

Q: Can THEIA be used when the In-Line Inspections (ILI) are from different vendors?

A: We can accommodate all of the commonly received formats.

Q: Is the software trained to make engineering judgement for FFS? How is that made possible?

A: No, FFS is based on facts and standard methodology.

Q: If digitisation is taking place, don't you think people will be jobless and training must be provided to operate the software by each and every company?

A: That is a couple of different questions; irrespective of what happens with the digitalisation of industry, the need for suitably trained and competent engineers is immutable, at least for the foreseeable future.

A black-box scenario cannot be permitted to develop as a result of any new technology. Ultimately, this and the skills gap caused by a naturally reducing workforce will require that both existing engineers and new entrants to pipeline engineering will be expected to develop competence more quickly than previous cohorts.

There will, of course, be an obligation on employers to assist in knowledge transfer and competency building to assist the development of engineers. THEIA has been designed to assist in knowledge transfer and competency building. One of the three fundamental principles underpinning THEIA is the open access to Penspen training material for clients of THEIA.

In terms of training for specific software, yes, users must be competent to use any software that they are expected to operate.

Penspen believe that software does not have to be difficult to use. A poorly designed software or the use of any software by a user who is not fully competent would be undesirable. THEIA is designed to be coherent and logical to pipeline engineers with the minimum of specialised training, and to offer appropriate affordances so that the potential for user error is minimised.

Q: Can this be integrated with different existing pipeline data models, such as ESRI APDM, PODS and other existing data source

A: Yes, we can map to and from any of these.