

Data Integration - Cathodic Protection Survey

Cognitive Integrity Management

August 2020

Introduction

At OneBridge Solutions, data is imperative to our mission as a data science and machine learning SaaS software company. Lots of data.

As we evolve our enterprise level integrity management solution, Cognitive Integrity Management (CIM), we need to understand the datasets available in the industry. Discussing this with our customers, prospects, and potential partners, two common concerns seem to dominate: how much data is there, and in what format is it?

Before we can start to apply machine learning models to gather insights from any new dataset, we first have to spend quite a few development cycles on ingesting, normalizing, standardizing, QA/QC validating, and correlating new and existing data. We know this from experience: we've spent the last couple of years creating a very solid foundation around in-line inspection (ILI), pipe property (GIS / PODS), and repair data (in the ditch), and are now extracting important insights that are creating business value and saving pipeline operators money. (As a side note, we are also starting to see encouraging results using [Microsoft Cognitive Services - Text Analytics](#) to extract data from PDF documents.)

In the latest release of CIM, we have begun the development cycle around data related to cathodic protection (CP) surveys.

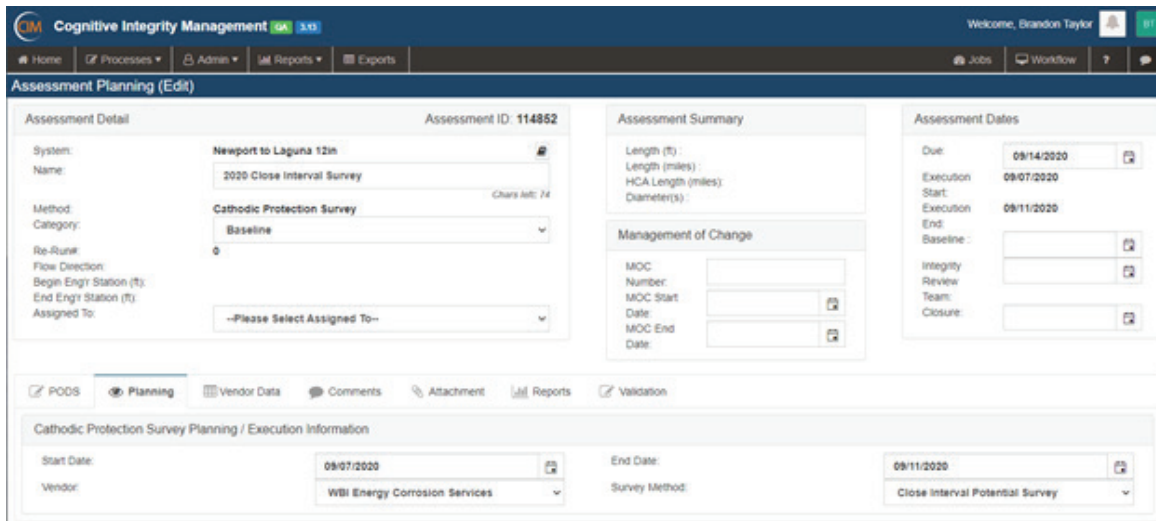
This white paper describes how close interval survey (CIS) data gathered as part of CP Survey will work in CIM, both during the CP Survey preview launch and as it evolves. We believe that by articulating our full vision for CP Survey, operators will better understand how to leverage CIM most effectively within their organizations. OneBridge is working hard and plans to include new features around this exciting new dataset as part of our normal releases.

Assessment Planning

CIM provides the metadata pipeline operators need to plan, schedule, and execute CP Survey assessments.

The following survey methods and types of CP Surveys are supported during the initial preview launch.

- Alternating current voltage gradient (ACVG)
- Alternating current current attenuation (ACCA)
- Direct current voltage gradient (DCVG)
- Close Interval Survey (CIS)
 - On only
 - On/off
 - Depolarized (native / static)



We also leverage existing CIM functionality such as comments, attachments, reports, and QA/QC validation routines that run during the ingestion of the data.

As is the case with each major new feature/function release, our customers will guide the evolution of CIM. We expect updates around this new feature to come out every few weeks as part of our standard SaaS release cycle.

Data Ingestion

When we launched the first version of CIM in the fall of 2017, there was no “truth” ILI data (fields, feature types, etc.). Since then, we built a user interface in the solution that allowed integrity engineer SMEs to categorize and build up enough truth data for us to create our first ILI data classification ingestion algorithm. That algorithm has now been trained on more than 5,000 ILI files.

Given the limited depth of CIS data, we plan to follow a slightly different path. Initially, and to promote consistency in the data, a defined import template will be used until we have a sufficient volume of data to allow an ingestion algorithm to be created. Users who have data in PODS can use our CIM / PODS synchronization agent to push data directly from PODS to CIM and will not be required to use the template.

That said, we do expect to facilitate the following:

- **Simple part of business process.** Software solutions help standardize, streamline, and create significant efficiencies in the business process. We have ensured that performing CIS analyses in CIM can be done in the same fashion that CIM is already used by more than 250 integrity engineers.
- **Normalized QA/QC (validation).** Customers tell us that data is mandatory for critical engineering. We take it further: Valid data is mandatory. CIM performs data validation checks on CIS data before it gets ingested into the solution. This ensures that all downstream processes are accurate, consistent, and efficient.

- **Scale.** CIM enables loading of both current and historical data for the entire pipeline system after pre-ingestion data validation has been run in our onboarding tool. We expect to provide the same capabilities for all CIS data.

Alignment to In-line Inspection, Repair, Pipe Property Data

The core algorithm in CIM is our alignment algorithm. More than 22 releases of this algorithm have resulted in some very significant learnings that have facilitated alignment at the system level, that is, all ILI assessments ever done on a system. We leveraged our algorithm as the baseline for the ILI-to-GIS (PODS) alignment within CIM, which has been extended to include mutually visible assets (valves, casings, AGMs, etc.) that provide spatial-(latitude, longitude, elevation) to-engineering stationing (begin/end, m-value) alignment.

As the data captured from CP surveys is limited in depth (on, off, dB μ V, %IR, depth, notes, latitude, longitude, elevation, and a few others), there are only a few options to align it to existing ILI and/or GIS (PODS) data. Today, CIM utilizes the spatial latitude and longitude coordinates to snap it to the nearest ILI location. As CIS measurements are typically taken every 5-10 feet and ACVG/DCVG indications occur as coating anomalies are identified, this process will overlay multiple CIS readings and possibly multiple ACVG/DCVG indications onto a single joint of pipe. We then correlate each anomaly (corrosion, dent, crack) that has already had pit-to-pit alignment across all ILI assessments done to the closest CP survey results point. CIM can then overlay rate of corrosion growth and anomalous CP patterns.

CIM now supports and manages data from ILI, GIS (PODS), Repair, and CP surveys that have been aligned at the anomaly level where there is sufficient data across an operator's entire pipeline system.

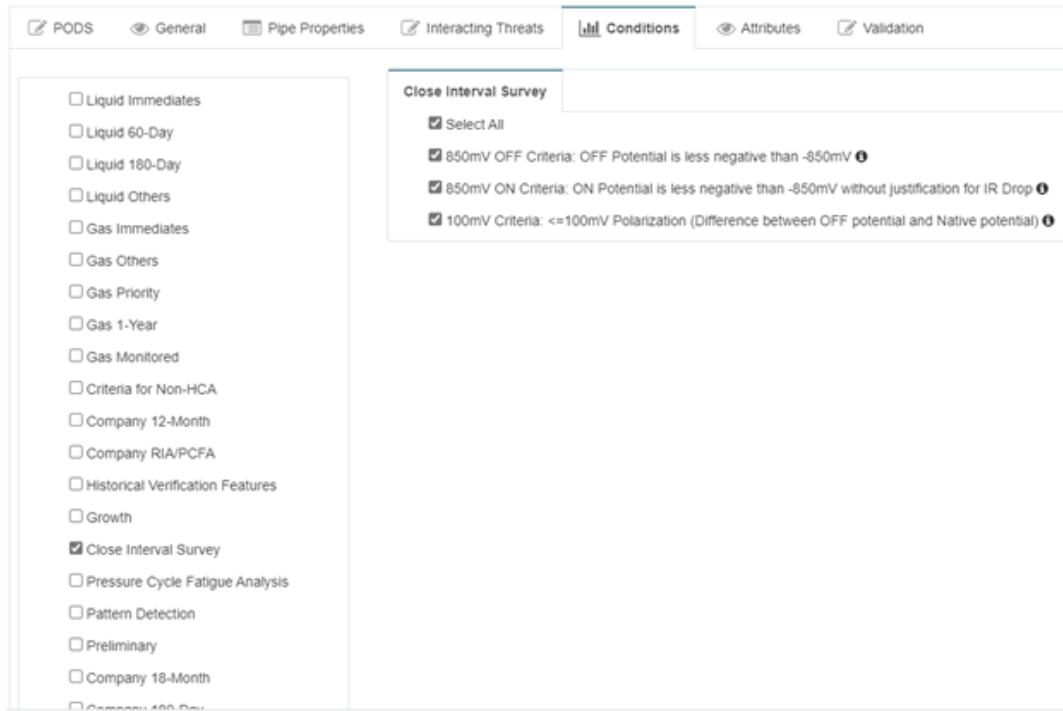
Data Science & Integrity Compliance

There are significant benefits and very real business value to managing your data in a modern database, but as one of our customers likes to say, "Data is mandatory for critical engineering." Our data science team agrees. They can now run all kinds of machine learning models and statistical methods against new data to determine if new patterns emerge.

Integrity Compliance

There are more than 200 conditions in the library for liquid and gas integrity concerns, with new ones added each time a customer is onboarded. The library includes CFR 192 & 195 regulations, industry best practices, CIM data science and machine learning patterns, and company specific conditions. All customers have access to the entire library and can conduct what-if scenarios using these conditions. We call this shared or concurrent learning.

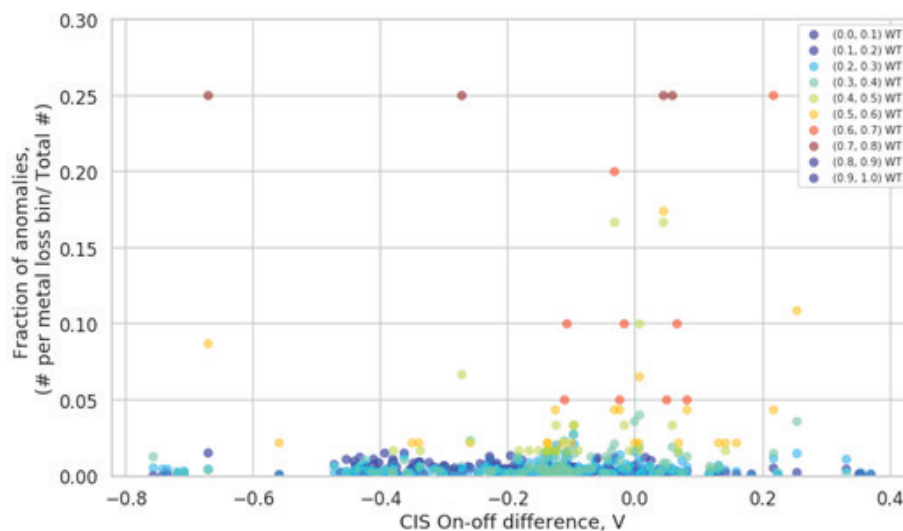
With the addition of new CP Survey data, this library will continue to grow. Initially, the conditions will incorporate known industry best practices like those listed in the screen shot below. However, as customers start ingesting and working through the CP Survey process, we expect new requests to start flowing into the development queue. For example, one that has already been discussed is corrosion growth where on/off separation approaches zero.



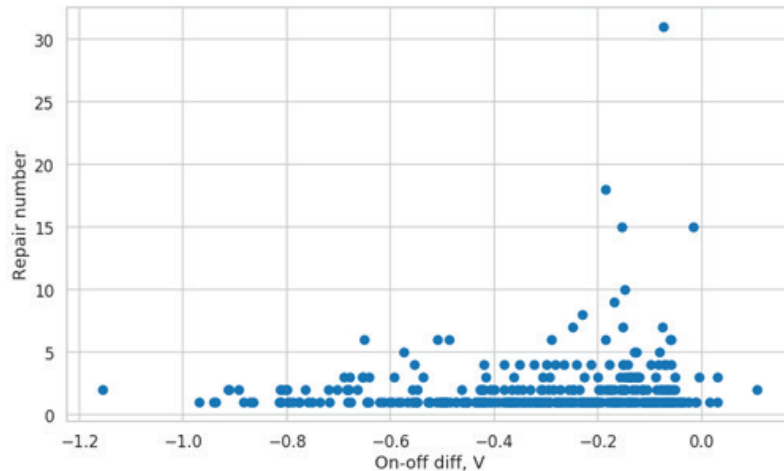
Data Science Patterns

Now that the CIM platform can ingest and align CP survey data, our data science team has been hard at work developing statistically proven insights and patterns that operators can use in their daily decision-making. Although based on only a limited set of pipeline systems, there are already promising developments:

- With 40-foot bins there is a higher density of deeper corrosion within a region the on/off voltage difference approaches zero and/or there is a reversal.

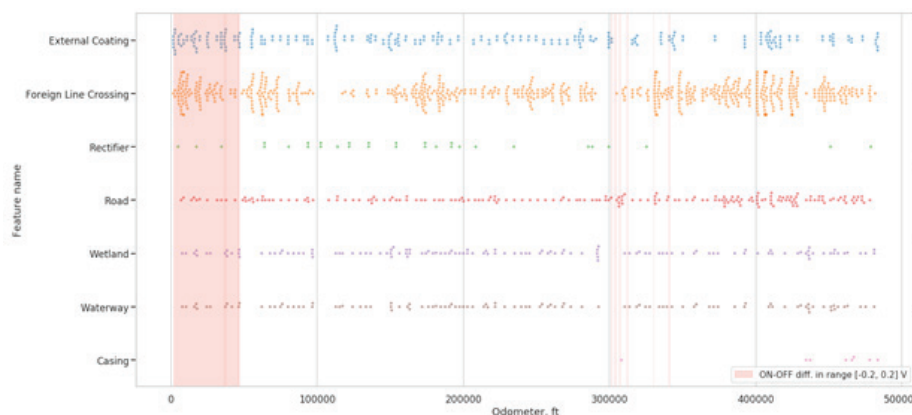


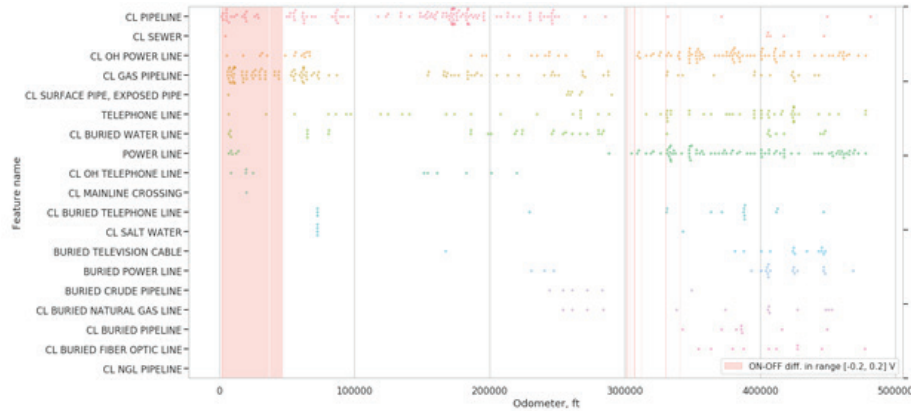
- There is early evidence that more repairs have been done when on/off voltage difference approaches zero.



- With 500-foot bins, the patterns based on the note field within the CIS files show an on/off jump when encountering roads, powerline crossings, etc. We expect this work may further help refine the confidence interval of our alignment.
- Different signals within CIS against various GIS (PODS) attributes show differing clustering depending on the attribute
- We expect that over time we will be able to train new classifiers to be used for other standard machine learning methodologies.

These are some of the plots coming out of the data science experiments. They show how GIS (PODS) data correlates to CIS data and provide additional details where specific foreign line crossings overlap with regions of diminished CP protection.

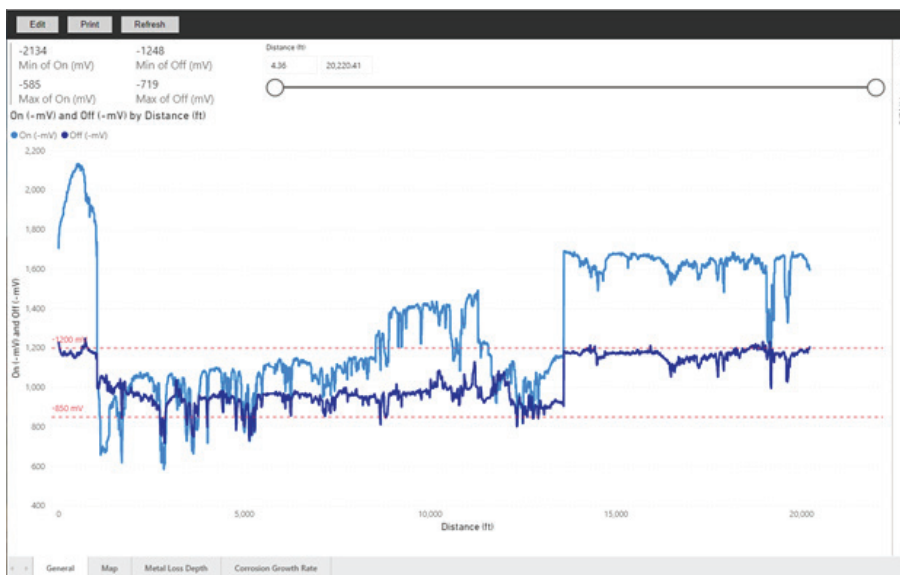


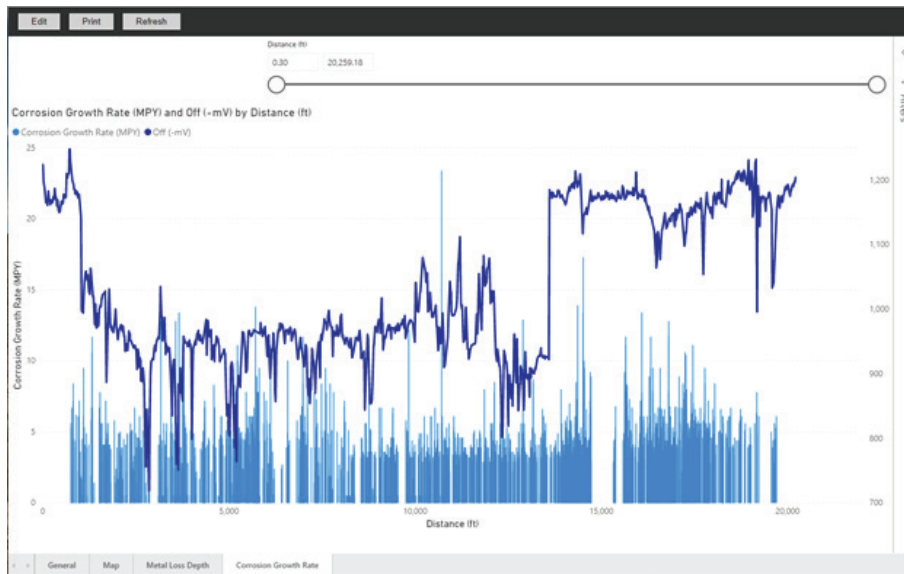


As stated in the introduction, lots of data is imperative to our work and the addition of CP surveys expands the learnings beyond ILI, GIS (PODS), and repair data. We can now leverage the information provided by operators' SMEs to guide us in our data science roadmap. This will help ensure that the new features/ functions we incorporate into CIM support SMEs and their daily work.

Business Intelligence

The ultimate value of organizing and consolidating data into CIM is the enterprise-wide reporting that is made available to an operator. CIM utilizes [Microsoft PowerBI Embedded](#) as the tool to remove the noise and surface the learnings and intelligence needed to make informed decisions. Initially, CIM will ship with the following two dashboards. Additional dashboards will quickly follow as SMEs start digging into the results.





Conclusion

Hopefully, this white paper provides enough detail around the process and value CIM could provide to your organization. We expect this new CP Survey feature/functionality to evolve quite rapidly as our current customers start loading data, processing integrity analyses, and closing the loop with actual excavation results.

As part of our continuing evolution toward PRCI PR-331, our team has created user stories within our roadmap to utilize components of this work. Additional NACE SP0502 concepts and calculations will ultimately be incorporated into the overall framework of the solution. This cadence will largely depend on the priority assigned by users of CIM and the development effort required.

Future Research / Development

CP survey data is the first of many planned new data sets being evaluated for inclusion into CIM. A process similar to the one outlined in this white paper will be used as the framework and pattern for such data as Soil, Depth of Cover, etc. We will apply significant scrutiny to each new data set in line with minimum requirements regarding data usefulness in true data science and machine learning.

About Cognitive Integrity Management

Cognitive Integrity Management Cognitive Integrity Management™ is a platform that addresses regulatory compliance and internal operational goals. It is a fully integrated enterprise level solution for a complete overview of asset health in the highly secure Microsoft cloud.

OneBridge Solutions Inc.

OneBridge Solutions Inc., develops and markets revolutionary new SaaS solutions that use advanced Data Sciences and Machine Learning to analyze big data using predictive analytics to assist Oil & Gas pipeline operators to predict pipeline failures and thereby save lives, protect the environment, reduce operational costs and address regulatory compliance requirements.

Visit <http://www.onebridgesolutions.com> for more information.

Contact Us:

OneBridge Solutions, Inc.

#4227, 10230 Jasper Avenue
Edmonton, AB T5J 4P6
Canada

sales@onebridgesolutions.com

1.877.261.7045

US Operations:

1775 W. State Street, #141
Boise, ID 83702, USA

