PNNL-SA-157397

Novel Signatures from Deployed Transmission Infrastructure Sensors

Project Number 72954

Kayte Denslow

Non-destructive Evaluation Team

Pacific Northwest National Laboratory

U.S. Department of Energy

National Energy Technology Laboratory

Oil & Natural Gas

2020 Integrated Review Webinar

Outline
(required contents)

1. Program Overview
   - Funding
   - Period of performance
   - Participants
   - Overall project objectives
2. Technology Background
   - Envisioned use
   - Fundamental science basis
   - Development prior to project
   - Technical/economic advantages and disadvantages
3. Technical Approach/Project Scope
   - Experimental design
   - Work plan
   - Project schedule
   - Project success criteria
   - Project risks and mitigation strategies
4. Progress and Current Status
   - Description of technology
   - Significant accomplishments, tied to technology challenges
   - Performance levels achieved so far, compared to project goals and tied to economic and technical advantages
5. Plans for future testing/development/commercialization
   - In this project
   - After this project
   - Scale-up potential
6. Summary
   - Project summary
   - Key findings, lessons learned, and future plans

Appendix
- Organization chart
- Gantt chart

Funding, POP, Participants

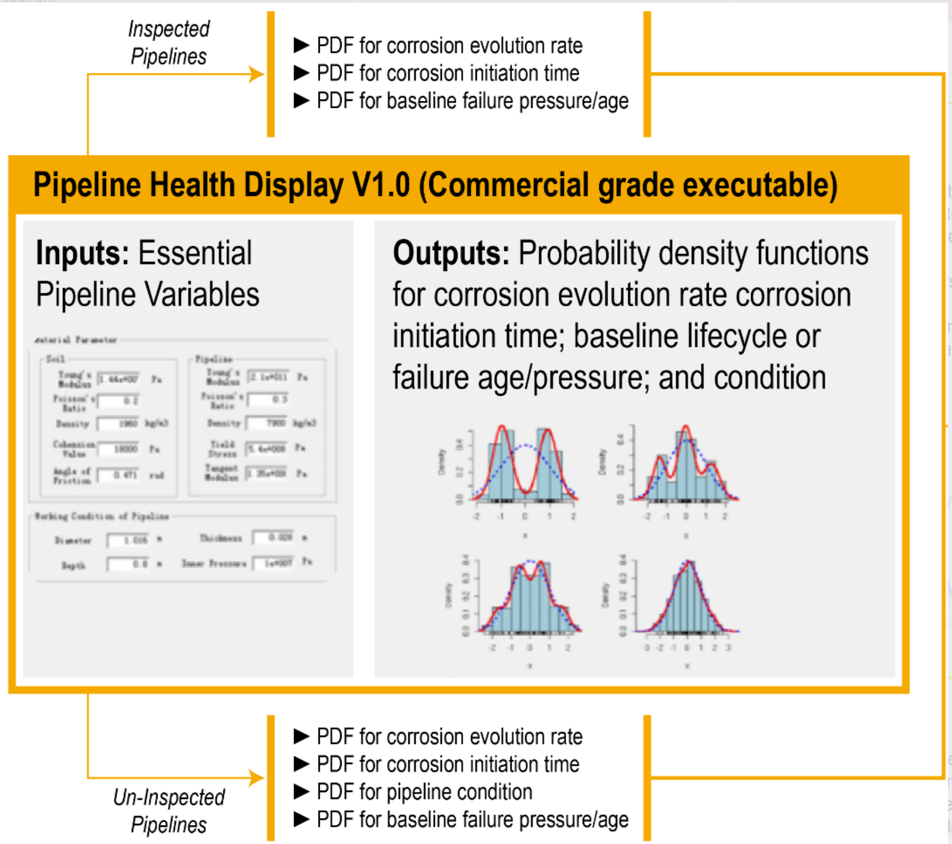
- 10/2019 - 10/2022
- $1.5M, DOE
- "Data Share" Partners: Natural gas transmission (NGT) pipeline industry
  - Open to NGT pipeline operators
    - 1 partner so far (unnamed)
  - Open to ILI companies, ILI research labs/consortia
    - 1 ILI partner so far (unnamed)
    - 1 consortia member (PRCI)

Overall Project Objectives

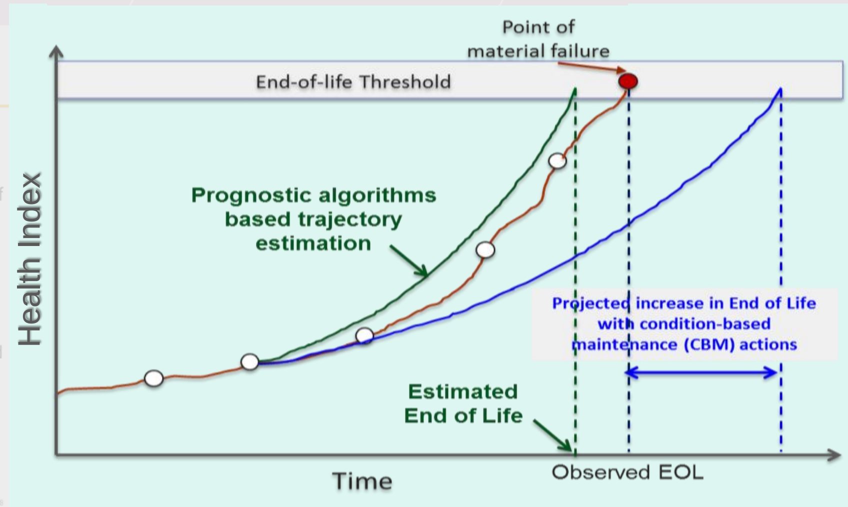
1. **Team with NGT pipeline industry and apply ML to historical NGT pipeline data sets**
   - ILI data (ILI tool signal data or flaw sizes listed in ILI reports)
   - Pipeline attributes (material, environmental conditions, construction history, etc.)
2. **Uncover "novel signatures" in data sets to gain new insights on current & future pipeline condition**
   - Non-obvious ILI signal features used to increase flaw detection probability, resolution & accuracy of flaw size (MFL or UT)
   - Non-obvious relationships between pipeline corrosion initiation time, corrosion rate, and pipeline properties/attributes
3. **Use novel signatures to build model**
   - Diagnostic model for assessing current pipeline condition
   - Hybrid physics-based, data-driven prognostic model for predicting future pipeline condition
4. **Generate algorithms with model, transfer to industry**

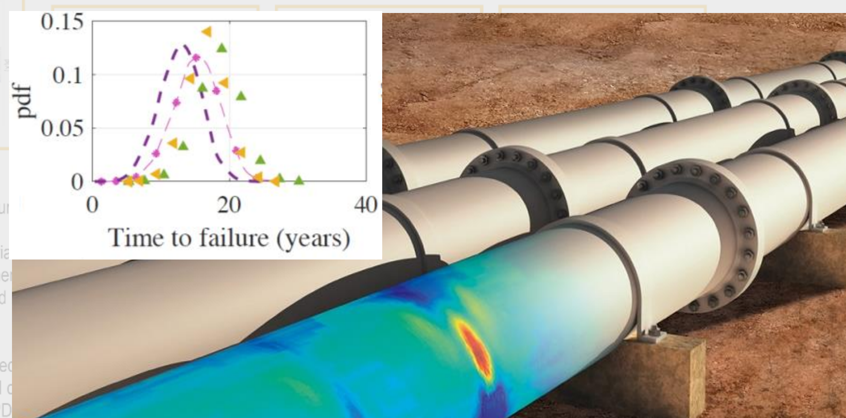
Envisioned Use

Research (Us)



Operations (Industry)

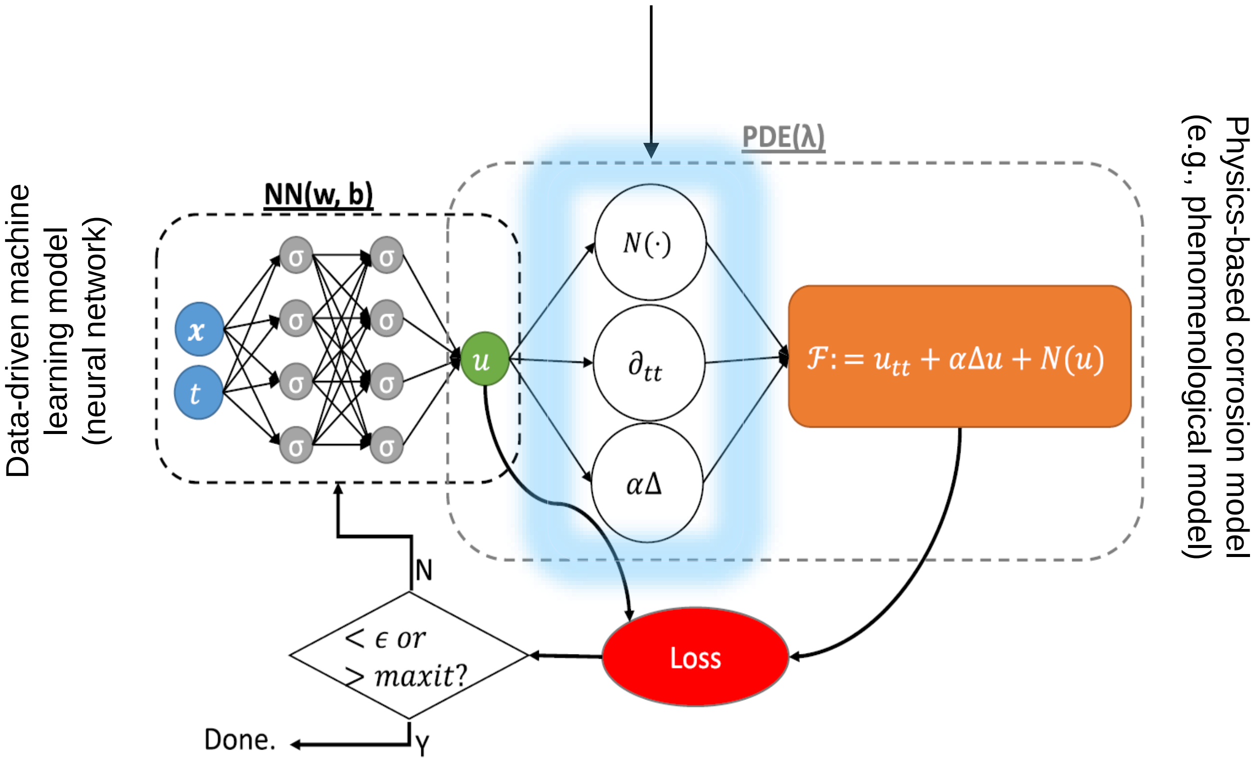




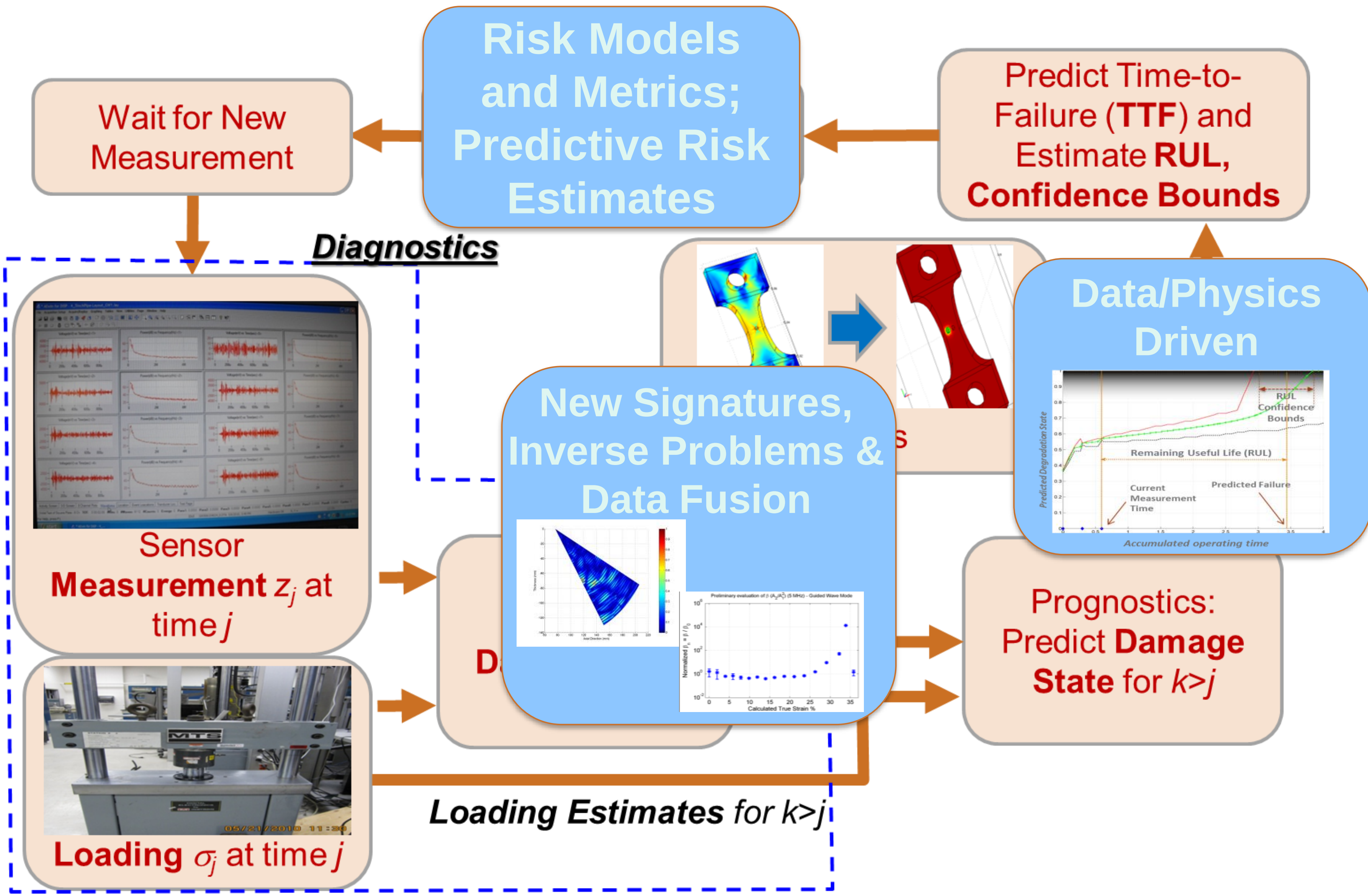
Amaya-Gómez R, M Sánchez-Silva, E Bastidas-Arteaga, F Schoefs, and F Munoz. 2019. "Reliability assessments of corroded pipelines based on internal pressure – A review." *Engineering Failure Analysis*.

Underpinning Science

Hybrid Physics-based/Data-driven Model of
Corrosion Initiation and Evolution Rate



Jagtap AD, K Kawaguchi, and GE Karniadakis. 2020. “Adaptive activation functions accelerate convergence in deep and physics-informed neural networks.” *Journal of Computational Physics* 404.



Risk Models and Metrics; Predictive Risk Estimates
Wait for New Measurement
Predict Time-to-Failure (TTF) and Estimate RUL, Confidence Bounds
Diagnostics
Data/Physics Driven
RUL Confidence Bounds
Remaining Useful Life (RUL)
Current Measurement Time
Predicted Failure
Predicted Degradation State
Accumulated operating time
New Signatures, Inverse Problems & Data Fusion
Sensor Measurement z_j at time j
Loading σ_j at time j
Prognostics: Predict Damage State for $k>j$
Loading Estimates for $k>j$

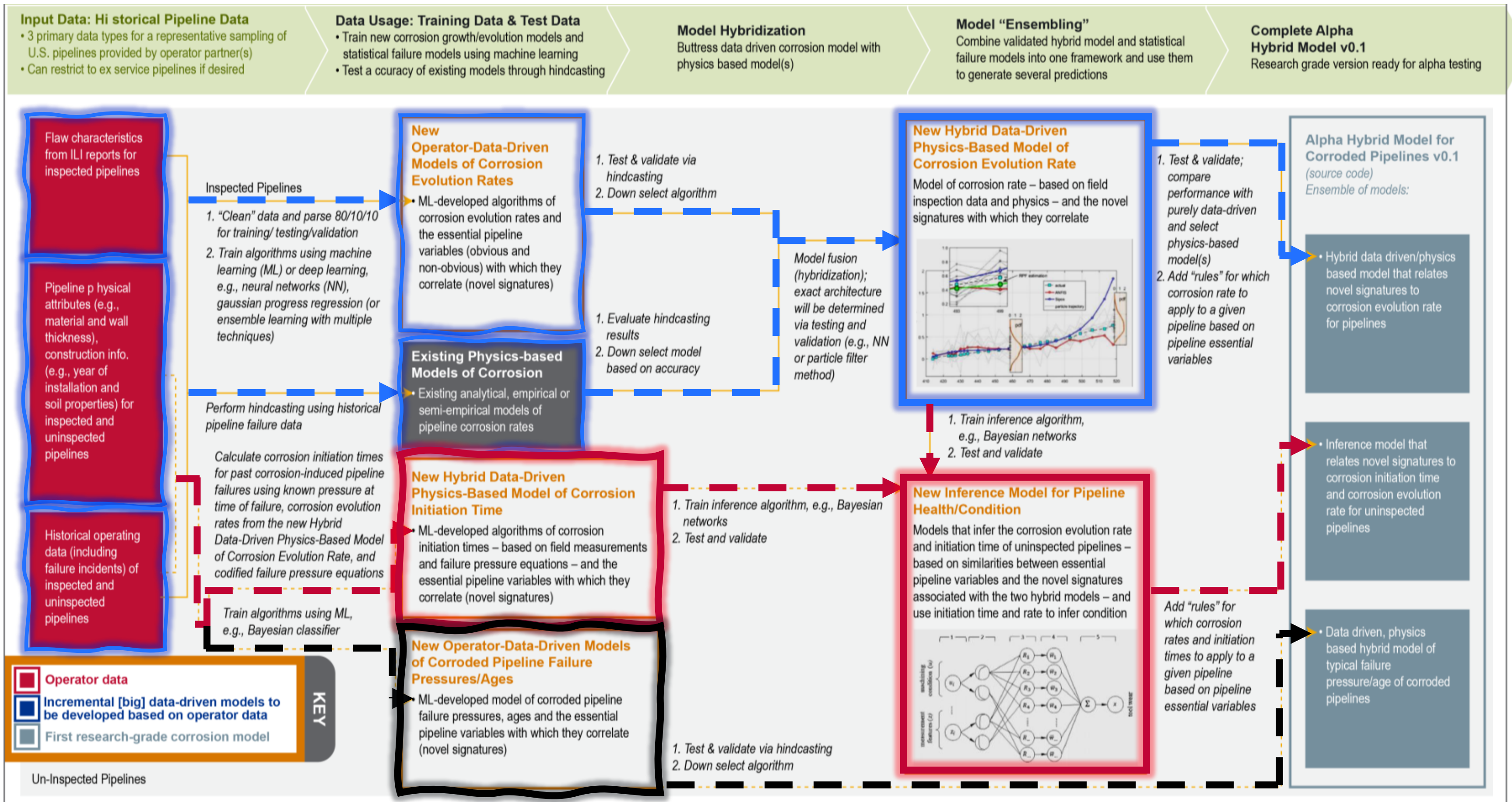
Advantages/Challenges

- Technical/economic advantages:
  - Data sets needed to train data-driven/ML aspect of model already exist (data collection is expensive)
  - Will leverage machine learning algorithms developed at PNNL and by others, so not starting from scratch
  - Hybrid model approach allows data-driven model to be imperfect and physics-based model to be imperfect
  - Hybrid model should result in increased certainty of pipeline condition and degradation/corrosion rate
    - Benefit to stakeholders: If industry has higher certainty of pipeline reliability and remaining useful life, then it drives focused/efficient predictive maintenance
    - Results in more reliable “gas grid” and more efficient use of integrity management resources
  - Cost is concentrated in the initial development and testing phase; scale-up and technology transfer of software over time will cost less
- Technical/economic challenges:
  - Need large data sets to train data-driven aspect of model
  - Reliant on data access provided by others – takes time to build trust/credibility with partners

Pacific Northwest
NATIONAL LABORATORY

Experimental Design

Model Development



Work Plan & Schedule

#	Milestones and Deliverables	Current Due Dates	Status
1	Year 1 Summary report	N/A	*Complete*
2	Diagnostic Algorithm Evaluation Report	5/31/2021	Not Yet Due
3	Prognostic Algorithm Evaluation Report	9/30/2021	Not Yet Due
4	Pipeline Reliability and Lifecyle Health Management System Design Report	9/30/2022	Not Yet Due

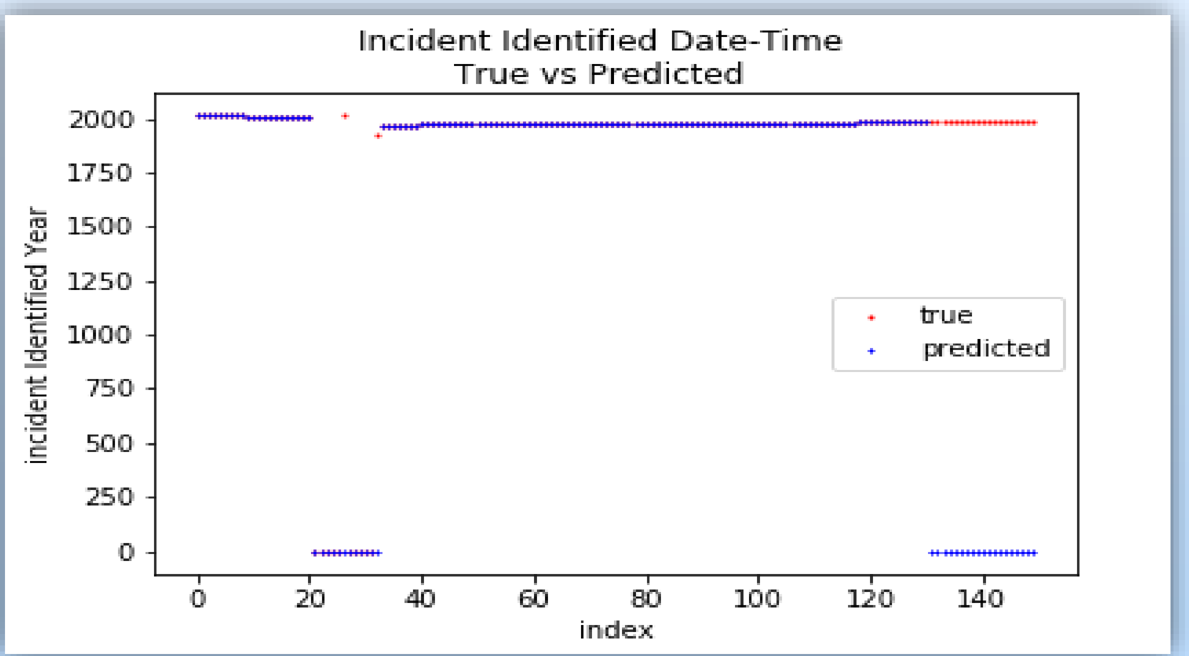
Success Criteria & Risk Mitigation

- Project success criteria:
  - Engaged "data-share" partners from NGT pipeline industry (go/no-go)
  - Model predictions accurately reflect ground-truth answers (e.g., 90+%)

Project Risks	Risk Mitigation
No MFL/UT signal data <u>with</u> corresponding ground-truth data are provided. -or- No ILI report data and meta-data (associated pipeline attribute data, construction info., environmental conditions, operating history, etc.) are provided.	MFL/UT signal data without ground-truth data is not useful for model training. The model framework is flexible enough that it can make use of MFL/UT signal data with ground-truth data if the data sets are available (for diagnostic model development); however, it can forego diagnostic model development and just focus on prognostic model development that uses tabulated flaw sizes, such as those found in ILI reports. If pipeline ILI report data are provided, but no corresponding meta-data, then the ILI report data will not be useful for model training. If meta-data are provided, but no ILI report data, then the data can be used to train a prognostic model, but it would be based on time in service before failure, instead of based on pipeline condition.
MFL/UT signal and ground-truth data sets are provided by partner(s), but only for a modest quantity of pipelines.	Modest signal/ground truth data sets are acceptable because inspections are performed on long runs of pipeline and each pipeline typically contains thousands of features that are examples on which the model can be trained.
ILI report and meta-data are provided by partner(s), but only for a modest quantity of pipelines (e.g., narrow range of materials and/or geographical regions).	Modest ILI/meta-data sets are acceptable, but not preferred, because the range of pipelines to which the model can be applied in the future will be limited by the diversity of the pipeline data used to train the model. A modest quantity of pipeline data sets just means the model will need to be applied to a narrower range of pipelines. However, the model can be updated as more data sets are provided to expand applicability.

Accomplishments

Feature Variables		Target Variable
Column Name	# Possible States	Column Name
IYEAR	49	INCIDENT_IDENTIFIED_DATETIME
ON_OFF_SHORE	3	
ITEM_INVOLVED	3	
PIPE_TYPE	6	
PIPE_DIAMETER	83	
PIPE_MANUFACTURER	61	
PIPE_MANUFACTURE_YEAR	56	

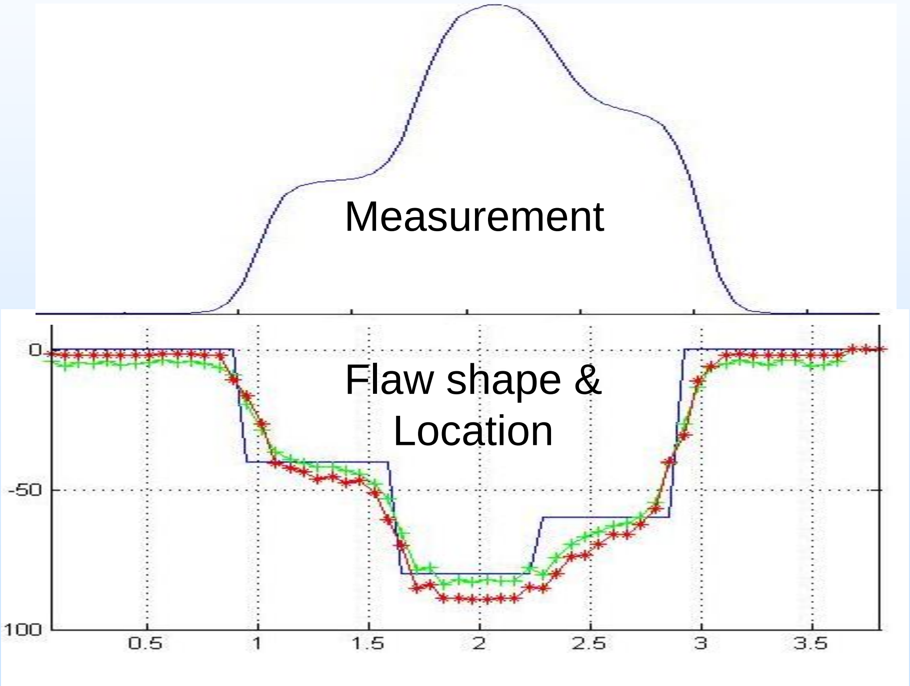


- Used past NGT pipeline incident data to show DNN applications
  - Limited to pipeline failures due to corrosion, about 1500 samples (80/10/10 split)
  - Model predicted incident year with ~86% accuracy i.e., model correctly predicted the incident year ~86% of the time
- Presented at PRCI REX2020 in March and invited members to participate on project
- **“Data-share” partners added to project team**
  - 1 pipeline operator partner
  - 1 ILI company partner
- PRCI hosted webinar where PNNL presented project goals
  - To keep opportunity open
  - Attended by ~100 members
  - Follow-up from 9 companies

Current Status

Diagnostic component: Predict flaw characteristics from signal data

- Data received (thanks to ILI partner):
  - 2-D MFL scans (axial, radial, circumferential) from ILI
  - Ground truth measurements and scans of defect location and surface
  - Processed features of 2-D MFL scans (measurements of signal peak, peak-to-peak, footprint, etc.)

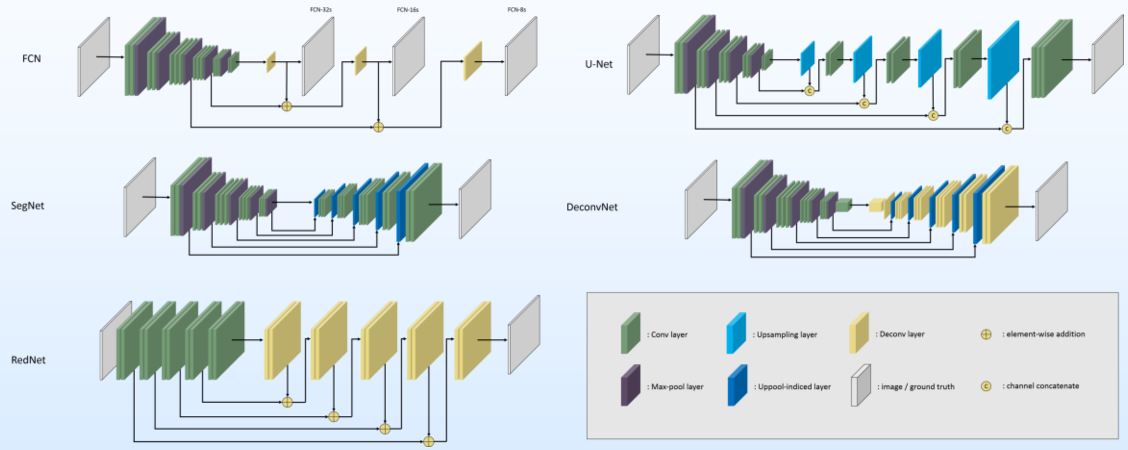


- Model input/output structure:
  - Data inputs: (raw) MLF data or (processed) its geometric features
  - Prediction target: infer properties of flaws (length/width/depth or depth map)
- Key research questions:
  - Is the signal local or global (interference between defect signals)?
  - Do candidate flaws need to be located or picked (automated vs human detection of “bounding box”)?
  - Can a model understand raw data, or does it signal properties or other metrics (rescaling can highlight subtle features)?
  - Can a model reconstruct the defect surface, or just geometric properties?

Current Status

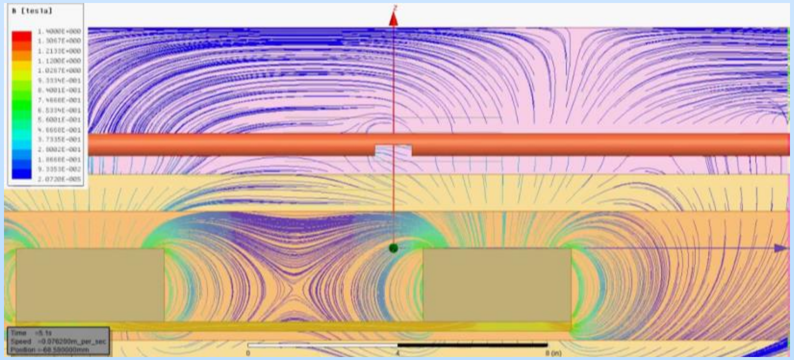
Diagnostic component development: Defect detection

- "CenterNet" is a centered-point object detection approach
- Objective: Use CenterNet to detect flaws from MFL data, classify based on length, width, depth
- Inputs: 3D matrices of MFL data, labeled using the COCO data format
- Outputs: Defect location and depth, and rough defect length/width



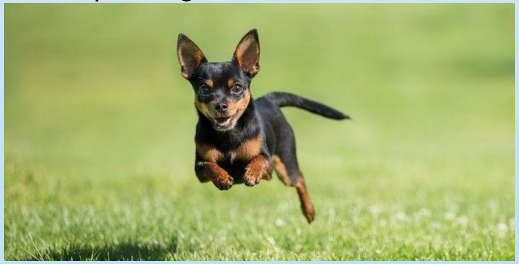
CenterNet: Can use different models as backends depending on the task (e.g. object detection, object segmentation, etc.)
https://medium.com/@sunnerli/simple-introduction-about-hourglass-like-model-11ee7c30138

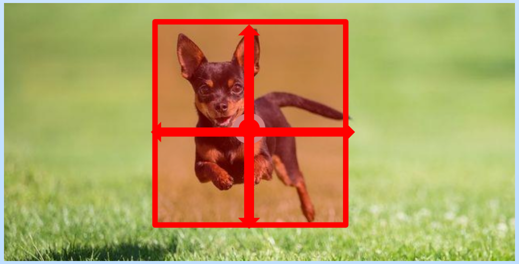
MFL Principle: Magnetic field will leak in areas where there is a defect



https://www.youtube.com/watch?v=b_v957tnCek

CenterNet: Learns to find the object's center and the length and width of the encompassing box.

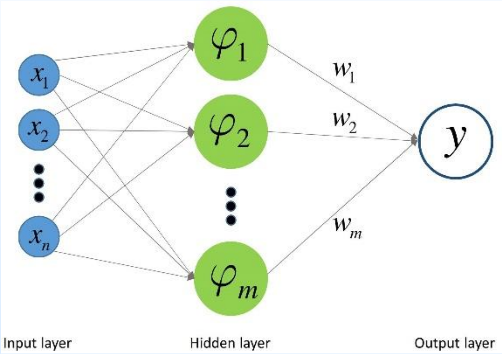




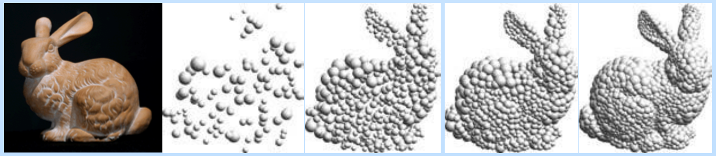
Current Status

Diagnostic component development: MFL to defect reconstruction

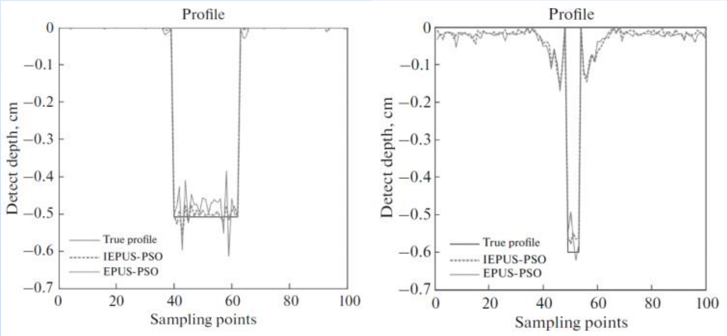
- Radial basis functions (RBFs) are used to represent complex functions
  - E.g. the sum of several RBFs can be used to learn detailed 3-D surfaces, and multi-dimensional signals like MFL
- RBF networks are used as a general function approximator
  - Have been used to learn maps from MFL to defect surfaces
- RBF networks are deep neural networks with one hidden layer
  - Traditionally, higher accuracy is achieved with more RBFs.
  - However, "deeper" networks may provide higher accuracy.



Traditional RBF NN: Source: Kandroodi et. al. (2017) IEEE Trans. Magn.



Sources: (left) http://graphics.stanford.edu/data/3Dscanrep/; (right four) Ohtake et. al (2004), http://www.cs.jhu.edu/~misha/Fall05/Papers/ohtake04.pdf



Reconstruction of defect profiles from MFL signals: (left) simulated MFL data, (right) real MFL signal from engineered defect. Source: Han et al. (2017) Russ. J. Nde. Test.

Future Development/Tests

Current project

Stage 1: Alpha Testing
Apply hybrid model to in service pipelines that have ILI data and to in service pipelines that have no ILI data to demonstrate accuracy and versatility of the model

Stage 2: Beta Testing
Apply hybrid model to in service pipelines having ILI data and in service pipelines that have no ILI data to further demonstrate accuracy and versatility of the model and to qualify it for official technology transfer

Testing Complete
Model is reliable and ready to advance to commercial grade

Alpha Hybrid Model for Corroded Pipelines v0.1 (research grade prototype software)

Alpha Software for ILI Signature Screening v0.1 (research grade prototype software)

Perform Alpha Testing - *researcher/operator team applies the model to pipelines that:*

- had failed in the past or had been preemptively removed from service,
- whose "true state" are known and can be compared with model predictions, and
- whose data were not part of training/testing/ validating the model during the development stage.

Alpha Testing Data Examples:

- Pipelines that have failed in the past (reportable or non reportable), the data for which are made available by the operator partner(s) or are available in publicly available historical failure databases
- Pipelines that are part of PRCI's sample library at the Technology D evelopment C enter (TDC) in Houston, TX
- Pipelines that are part of EPRI's sample library in Charlotte, NC.

Beta Hybrid Model for Corro ded Pipelines v0.1 (engineering prototype software)

Beta Software for ILI Signature Screening v0.1 (engineering prototype software)

Perform Beta Testing - *researcher/operator team applies model to pipelines that:*

- had failed in the past or had been preemptively removed from service,
- whose "true state" are known and can be compared with model predictions, and
- whose data were not part of training/testing/ validating the model during the development stage.

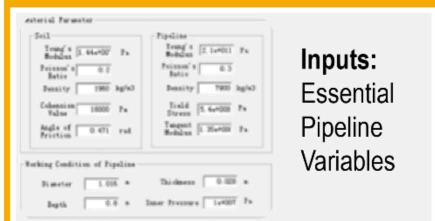
Model Complete PNNL

Technology Deployment & Outreach Office prepares Technology Transfer Plan

Optional: engage an industry partner to commercialize software

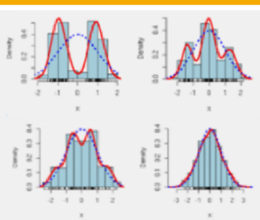
Add graphical user interface (GUI) to Beta Hybrid Model for Corrosion v0.1 to facilitate entry/uploading of pipeline essential variables and flaw characteristics from ILI

Pipeline Health Display V1.0
(Commercial grade executable)



Inputs:
Essential Pipeline Variables

Outputs:
Probability density function s for corrosion evolution rate corrosion initiation time; baseline lifecycle or failure age/pressure; and condition



Verify and validate GUI software

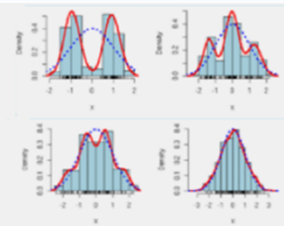
GUI Software Complete Advance to Technology Transfer

Future Tech Transfer

Pipeline Health Display V1.0
(Commercial grade executable)

Inputs:
Essential Pipeline Variables

Outputs:
Probability density function s for corrosion evolution rate corrosion initiation time; baseline lifecycle or failure age/pressure; and condition



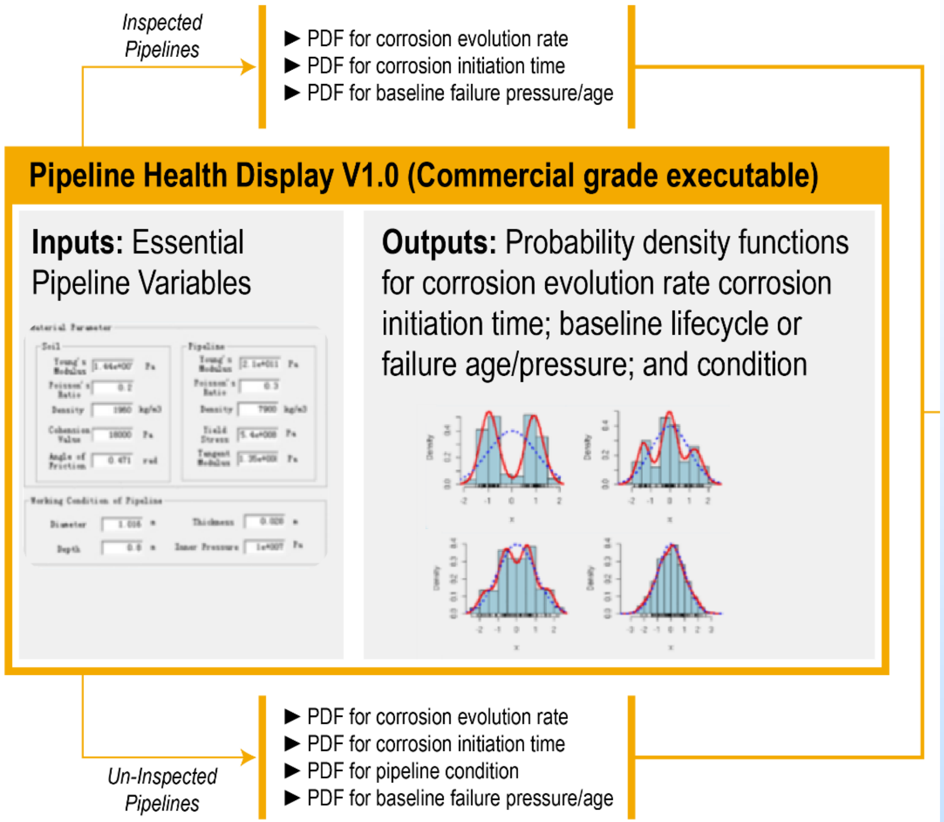
Execute PNNL Technology Transfer Plan

Examples of Technology Transfer Options

The Hybrid Model will be made available to data providers and partners at no-cost under a royalty free software site license (for executable code) for internal use only. The Hybrid Model will also be marketed to software service support provider(s)/vendor(s) to ensure a long-term commercial offering and support services. All data providers/partners will receive a discounted subscription to the commercial code offering from the vendor(s).

Implementation by Partners

Scale-up



A Living Model:
Sustained support by periodically issuing updated versions of the PHD commercial grade software following updates to the hybrid model source code, which entails:

- ➔ enrichment of the training data set as additional pipeline data are made available;
- ➔ updating the algorithms; verifying and validating the updated algorithms; and
- ➔ updating the executable for the commercial-grade PHD grade software.

Enrich and scale-up training data set to increase scope of application

Summary

- Project is in “Year 2” in terms of progress, Year 3 in terms of time
- Finally building momentum on data receipt and model building
- Key findings:
  - ~86% accuracy of predicting year of pipeline failure (based on 1500 examples, limited to failure cases only)
  - Engaged “data-share” partners to expand training data set to include representative sampling of pipelines
  - Receiving batches of ILI signal data (MFL) and developing diagnostic aspect of model
- Lessons learned
  - Took extra year to build data-share partner relationships
  - Industry open to partnering on this project; some have attempted ML themselves or with others and are open/optimistic about continuing
- Future plans
  - Continue developing prognostic aspect of model
  - Begin developing prognostic aspect of model
  - Follow up with interested companies from PRCI webinar

Appendix

Organization Chart

Casie Davidson PNNL FE Sector Manager

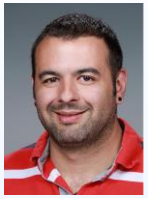
Angela Dalton, PM

Janie Vickerman, Project Coordinator

Kayte Denslow, PI NDE, Sensors

Naveen Karri, Task Lead (Mech Eng)

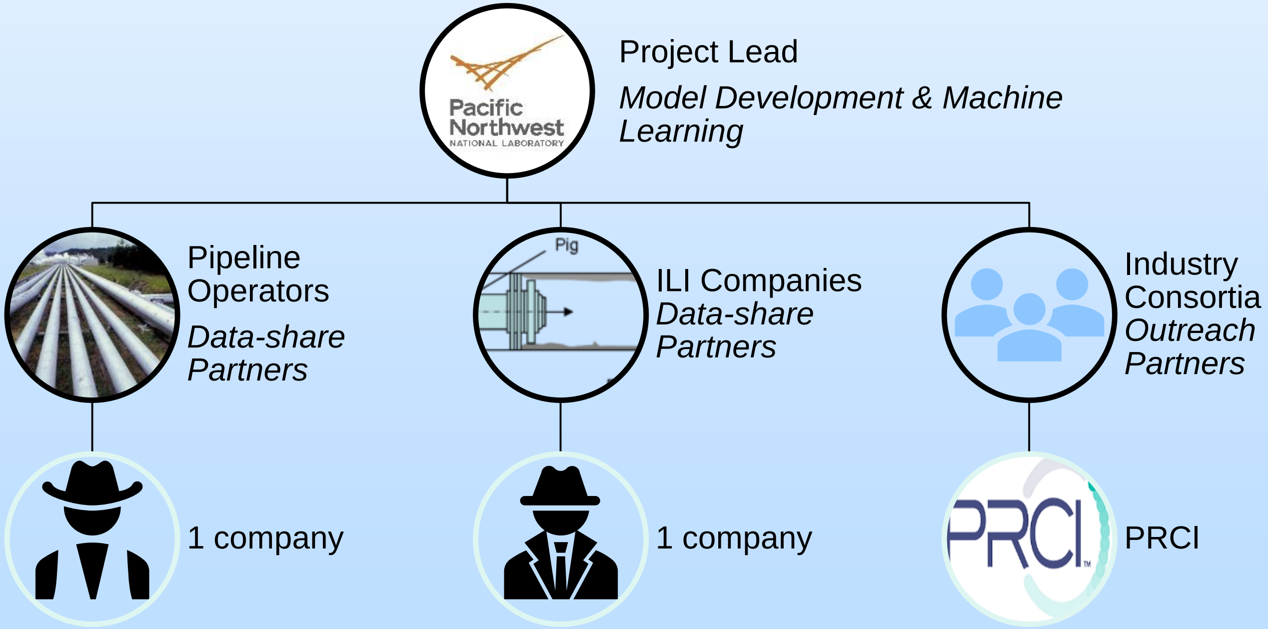
Steven Rosenthal, Task Lead (Math/ML)



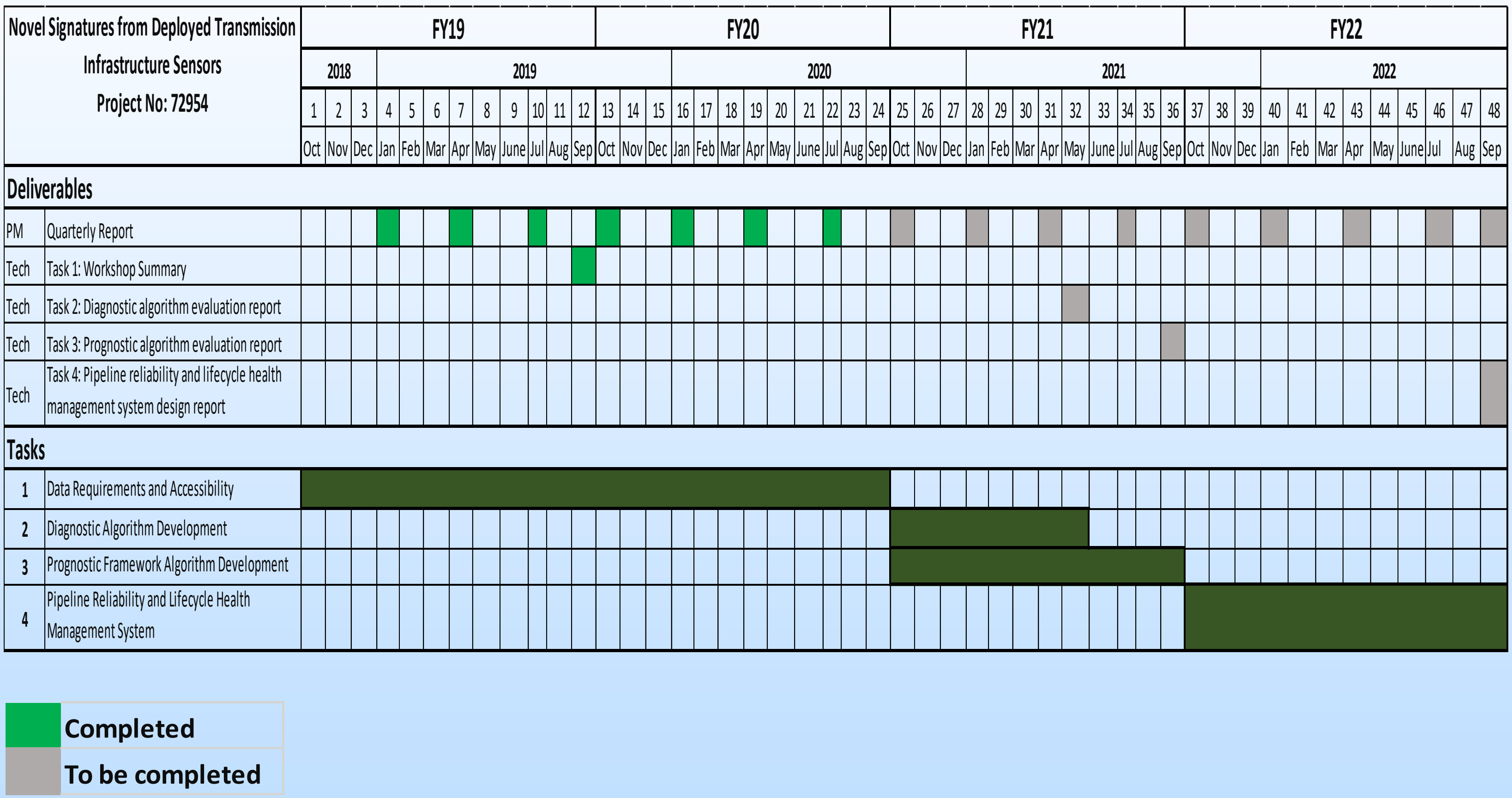
Juan Brandi-Lozano Math/ML/ Industry

Arun Veeramany, Risk & Reliability

Ken Johnson, Technical advisor



Gantt Chart


Novel Signatures from Deployed Transmission Infrastructure Sensors
Project No: 72954
FY19
FY20
FY21
FY22
2018
2019
2020
2021
2022
Deliverables
PM Quarterly Report
Tech Task 1: Workshop Summary
Tech Task 2: Diagnostic algorithm evaluation report
Tech Task 3: Prognostic algorithm evaluation report
Tech Task 4: Pipeline reliability and lifecycle health management system design report
Tasks
1 Data Requirements and Accessibility
2 Diagnostic Algorithm Development
3 Prognostic Framework Algorithm Development
4 Pipeline Reliability and Lifecycle Health Management System
Completed
To be completed