LEADING PIPELINE RESEARCH

Pipeline Research Council International

The Data Institute

What it is, what's in it, and updates

PRCI LEADING PIPELINE RESEARCH

1

What is the Data Institute?

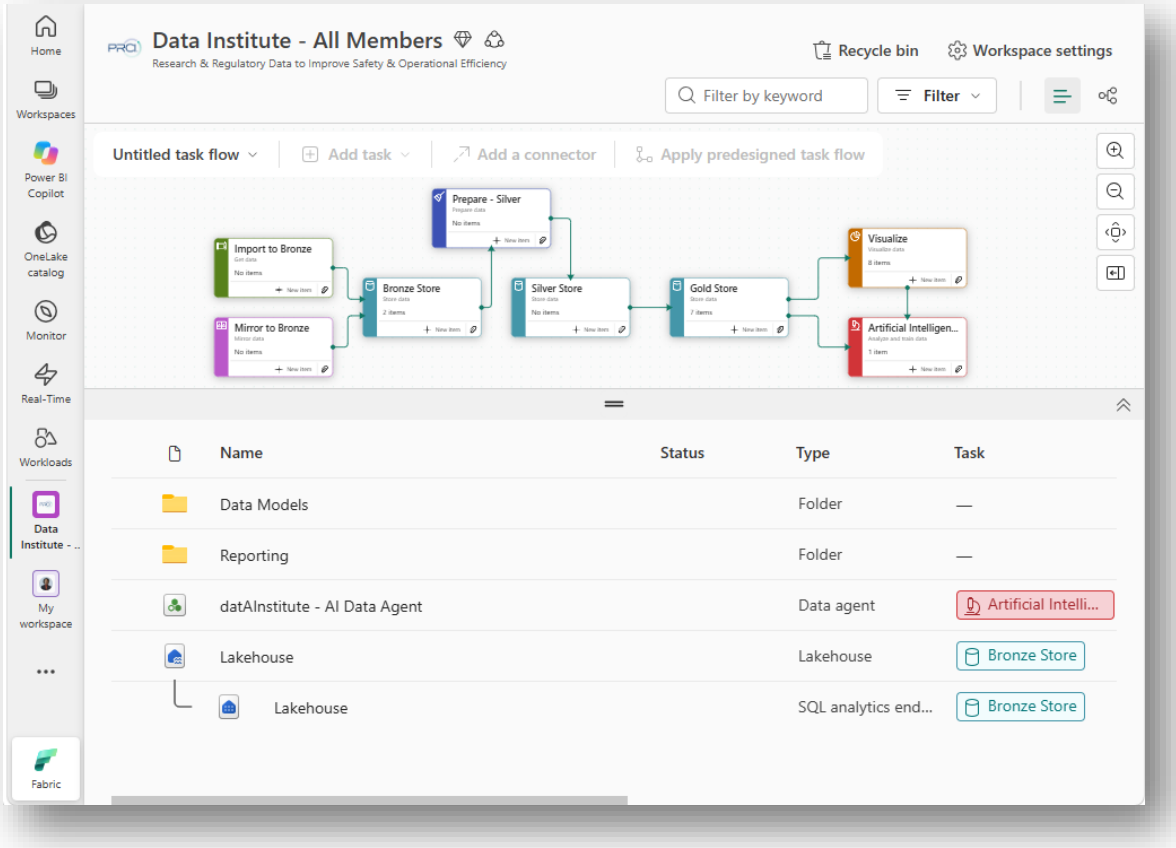
The Data Institute transforms PRCI's research, pipeline, & incident data into information you can act on



© 2026 Pipeline Research Council International

What is The Data Institute?

- A single platform for research, source data, sharing, and analytics
  - Consolidates all completed PRCI Research, the VTDC, agency data, public analysis
  - Provides all past data and best-in-breed toolset to active research
- Built in Microsoft Fabric
  - Picked after interviewing Technical Committee & DI Advisory Committee members
  - Key deciding factors:
    - ➢ Members already have/use Power BI and Fabric
    - ➢ Gets as much into one place as possible as opposed to a more fragmented cloud solution



© 2026 Pipeline Research Council International

PRCI LEADING PIPELINE RESEARCH

2

Why did PRCI create DI?

You asked for a go-to place where a question takes clicks, not a new research project

From the 2025 strategy interviews with operators:

> One structured, trusted, secure source for industry data

> An evergreen capability, that lowers the cost of future projects

> Members fund research for: a shared pool bigger than any one company, and analysis of that pool

© 2026 Pipeline Research Council International

Why PRCI created it

Research data has often been fragmented, hard to discover, and costly to recreate



Data Institute:
preserve → govern → connect → reuse

© 2026 Pipeline Research Council International

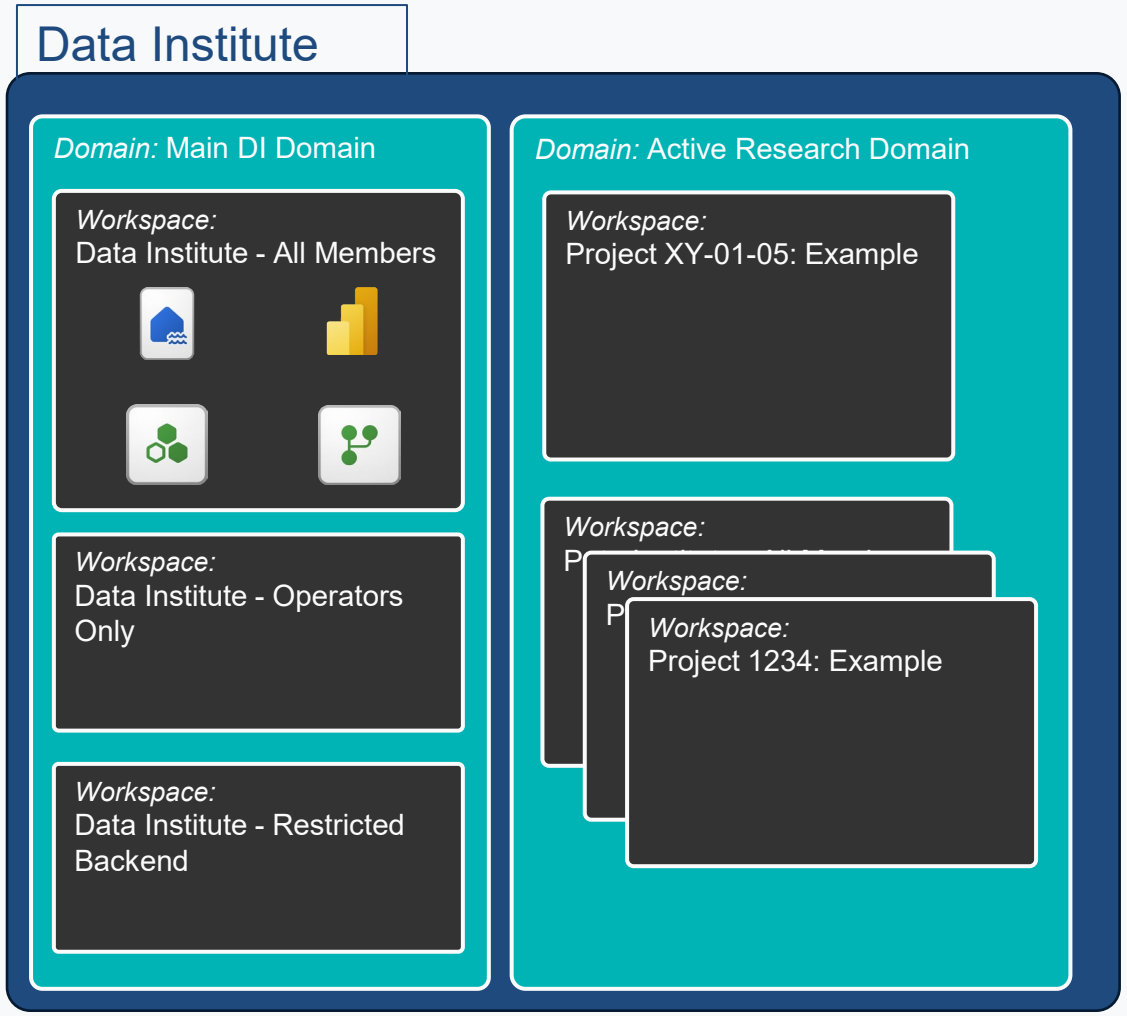
PRCI LEADING PIPELINE RESEARCH

3

How is the Data Institute Structured?

Microsoft Fabric is the digital foundation

Power BI is the interactive reporting and visualization layer



The basic building block of Fabric is an individual **asset**.

Assets are grouped together into **workspaces**. Workspaces have standard permissions.

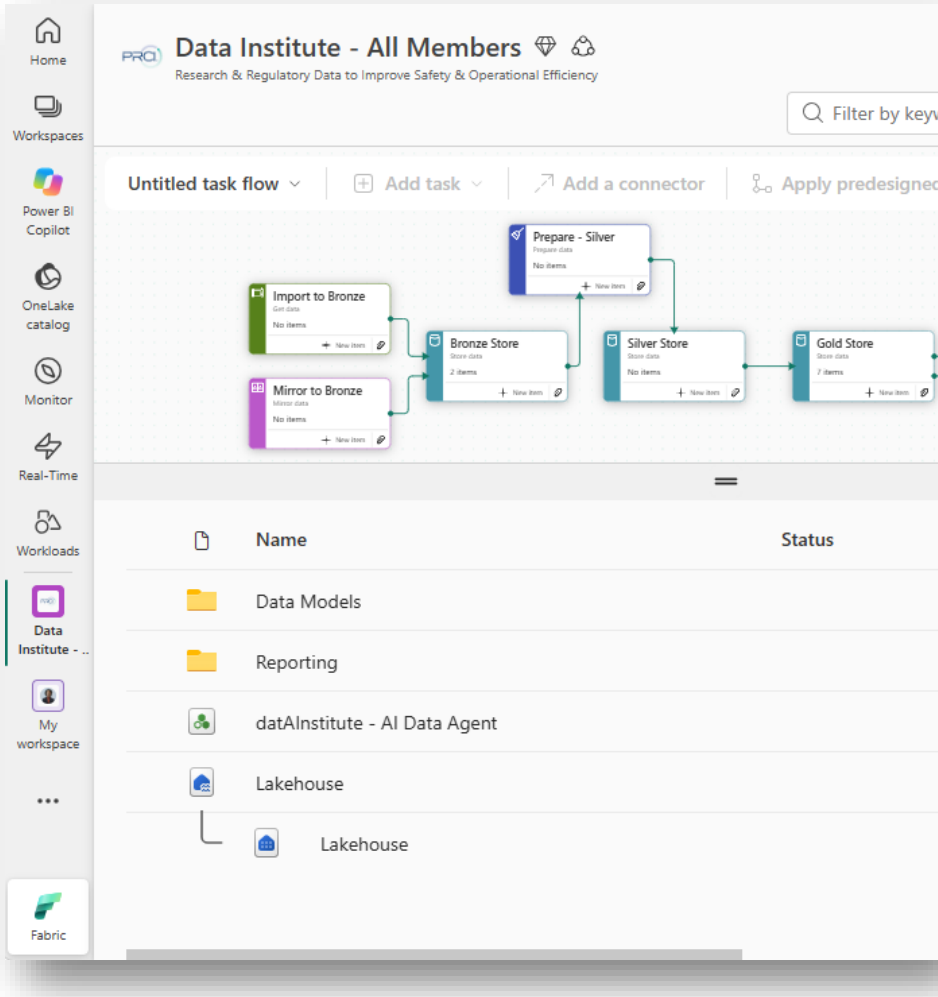
Workspaces can be organized into **Domains**.

The Data Institute has many elements, but is predominantly one workspace:
"DATA INSTITUTE – ALL MEMBERS"

© 2026 Pipeline Research Council International

All Members Workspace = Home

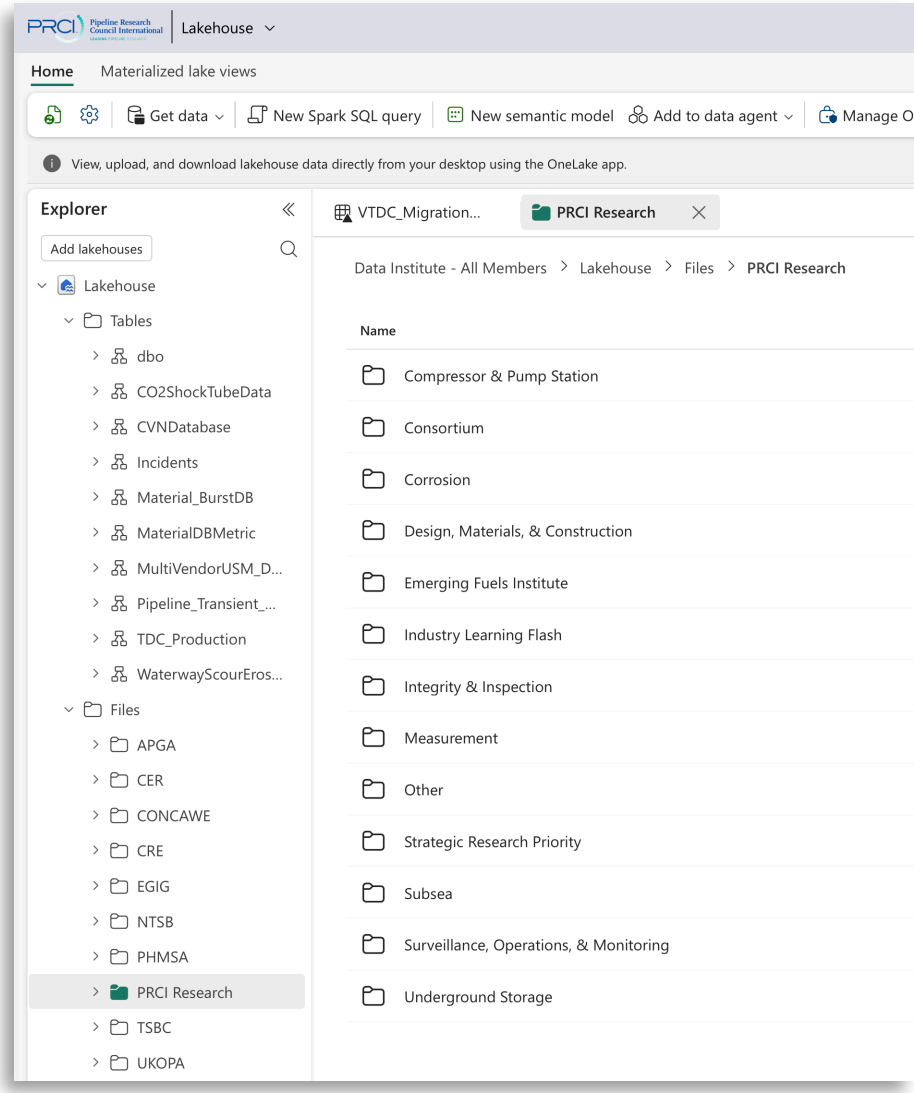
- **“Data Institute – All Members”** is the primary **workspace** of the DI
  - All completed PRCI research reports are located here
  - Projects can now generate Power BI reports as data products
  - Gradually adding the underlying research datasets
  - Agency(PHMSA, EGIG..) datasets are centralized here
- **“Data Institute Main Lakehouse”** is the most important asset.
  - Most of the things mentioned above are housed there



© 2026, Pipeline Research Council International

Main Lakehouse

- Tables
  - All of the VTDC tables have been migrated
  - Incident records are put into SQL tables
- Files
  - All historical PRCI research reports, organized by Technical Committee
  - A growing number of agencies' datasets:
    - USA: PHMSA, NTSB
    - Canada: CER, TSBC
    - Europe: EGIG, UKOPA, CONCAWE
    - Australia & New Zealand: APGA



© 2026, Pipeline Research Council International

PRCI LEADING PIPELINE RESEARCH

4

What is in there today?

What is available today

A growing collection of raw data, prepared data, research content, and analytical products

	Regulatory and industry data	PRCI research content	Insights and tools
Main DI Domain	Incident data: PHMSA incident &reports CER – Canadian Incident data APGA – Aus and NZ data NTSB, TSBC, EGIG, CONCAWE, UKOPA reports	Project deliverables: Final research reports Structured project data VTDC datasets and shared data Completed project data: Underground storage data	Incident Analysis reports Interactive Power BI views Semantic models Research-oriented analytics
Active Research Domain		Ongoing pilot project data: 7 ongoing pilot workspaces	Interactive Power BI views Semantic models Research-oriented analytics

© 2026 Pipeline Research Council International

PRCI LEADING PIPELINE RESEARCH

5

Research Facilitation

Why PRCI Standardizes Research Data Onboarding

- Establish a consistent, repeatable approach for onboarding research projects with data
- Ensure research data is hosted, governed, and secured within the Data Institute
- Set clear expectations for researchers, contractors, and PRCI project teams
- Enable collaboration and consistency across projects
- Enable PRCI to scale research activities efficiently over time

LEADING PIPELINE RESEARCH

Who This Process Is For

PRCI Research Leadership

Sets expectations that research data is hosted and governed through the Data Institute

PRCI Project Managers

Coordinate intake, kickoff, and delivery using a consistent process

Research Teams and Contractors

Execute research work within a provisioned environment with clear standards

Data Institute / Technical Support Team

Provides infrastructure, access management, governance, and optional support

© 2026, Pipeline Research Council International

LEADING PIPELINE RESEARCH

What This Process Answers

- Where does my research data go?
- Who has access to the data, and when?
- When does the Data Institute get involved?
- What happens while research is in progress?
- What happens when research is complete?
- What is included by default?
- What requires additional scope or approval?
- How does this process scale across many research projects?

© 2026, Pipeline Research Council International

How research project data enters the Institute

A seven-stage onboarding process connects project execution to long-term reuse



Standard process

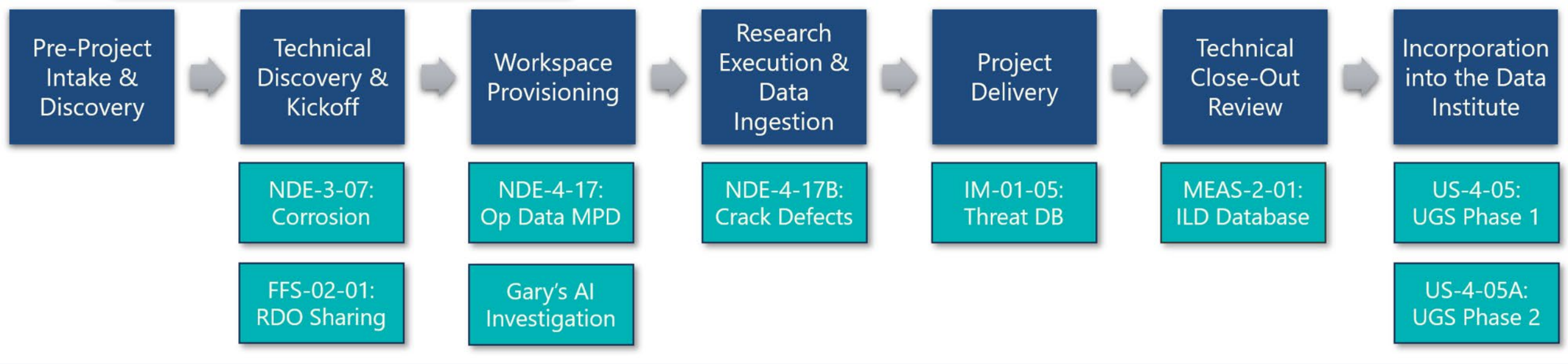
Intake · kickoff · workspace · access · light oversight · close-out · placement

Additional scope when needed

Custom dashboards · transformation · advanced analytics · AI · cross-project integration

© 2026 Pipeline Research Council International

R&D Projects 2026



© 2026 Pipeline Research Council International

PRCI LEADING PIPELINE RESEARCH

6

What to expect in 2027 and beyond?

2027 Data Institute Objectives

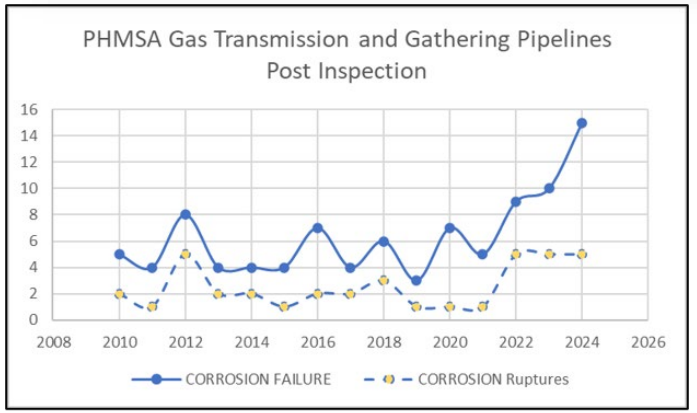


1. Research Project Facilitation

- Pilot 15 projects total to cover all types of projects
  - Provide structured storage, analytics and reporting
  - Provision secure and blinded spaces
  - Increase insight, useability and re-useability with cross project sharing
- Record data requirements for each type of project facilitation

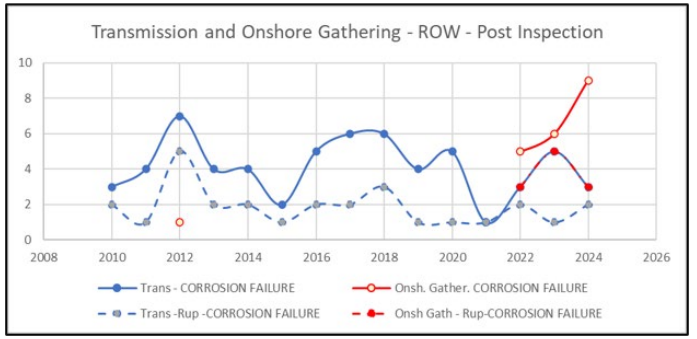
2. Improve Safety Performance and Operational Efficiency

- Collated data from PHMSA, CER, AUS+NZ - that we can interrogate
- Annual Incident Trend Report –trends by failure cause, sectors, type, etc.
- Enable company performance comparison by members
- Provide trustworthy and seamless secure workspaces for incident data sharing



3. Develop Advanced Data Analytics Initiatives

- Deliver Research-based Decision Optimizer by end of 2027
- For each ILI and related info shared - optimized scenarios for most economical safe options



• 1. Research Project Facilitation

2026 to 2027:

- 7 Pilot projects being tested – based on previous DI initiative
- Develop Data standards for typical project categories piloting 15 projects

2028 and beyond:

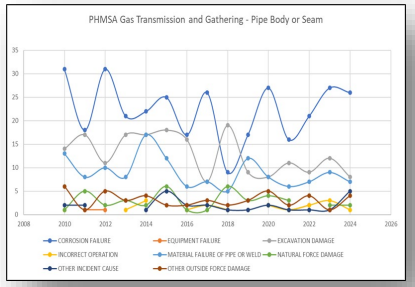


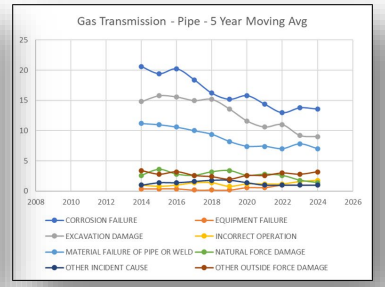
- Support data needs as shown in process and implement learning from pilots
- Develop annual enhancements process

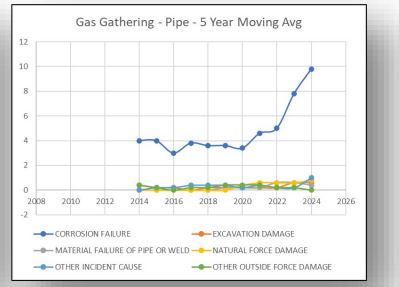
© 2026, Pipeline Research Council International

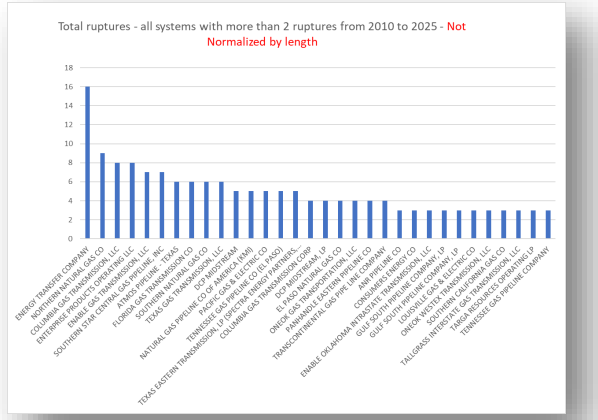
2. Improve Safety Performance & Operational Efficiency

- **There are three kinds of lies: Lies, Damned Lies, and Statistics ~ Twain/Disraeli**
- Enable user defined slicing and dicing for meaningful data analytics



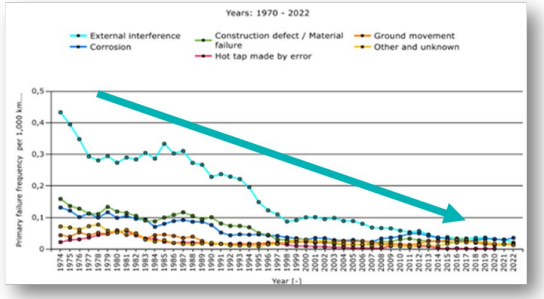




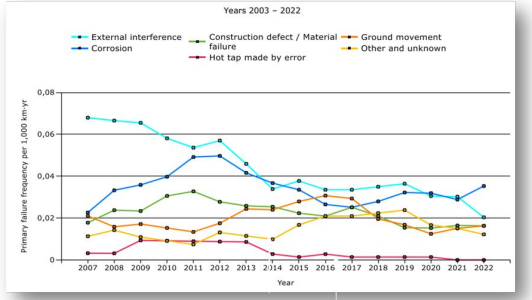


- Enable company performance data analytics by members
- See International trends and learn early and drive R&D – Annual Report

Looking at 50 years of Data



More Recent Data - 20 years



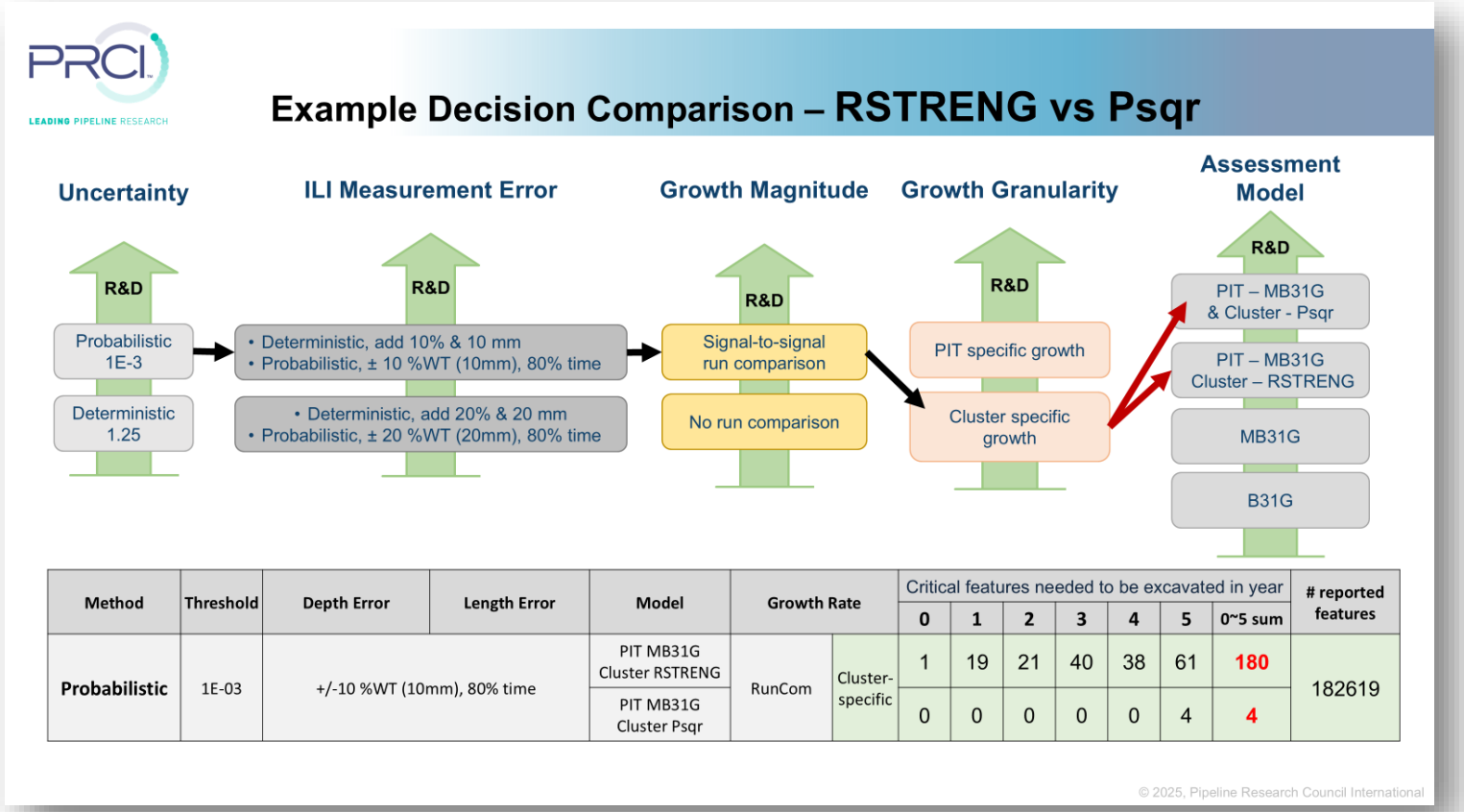
Know your stats so that no one can fool you!

© 2026, Pipeline Research Council International

3. Develop Advanced Data Analytics Initiatives

- Deliver Research-based Decision Optimizer by end of 2027
- For each ILI and related info shared - optimized scenarios for most economical safe options
- Moving PRCI Research to tried and tested value driven implementation

Supporting Research-in-Action!



Method	Threshold	Depth Error	Length Error	Model	Growth Rate		Critical features needed to be excavated in year							# reported features
							0	1	2	3	4	5	0~5 sum	
Probabilistic	1E-03	+/-10 %WT (10mm), 80% time		PIT MB31G Cluster RSTRENG	RunCom	Cluster-specific	1	19	21	40	38	61	180	182619
				PIT MB31G Cluster Psqr			0	0	0	0	0	4	4	

© 2025, Pipeline Research Council International

© 2026, Pipeline Research Council International

Thank you!

Questions?

© 2026 Pipeline Research Council International