UPISC UNIVERSITY OF PITTSBURGH INFRASTRUCTURE SENSING

COLLABORATION WORKSHOP

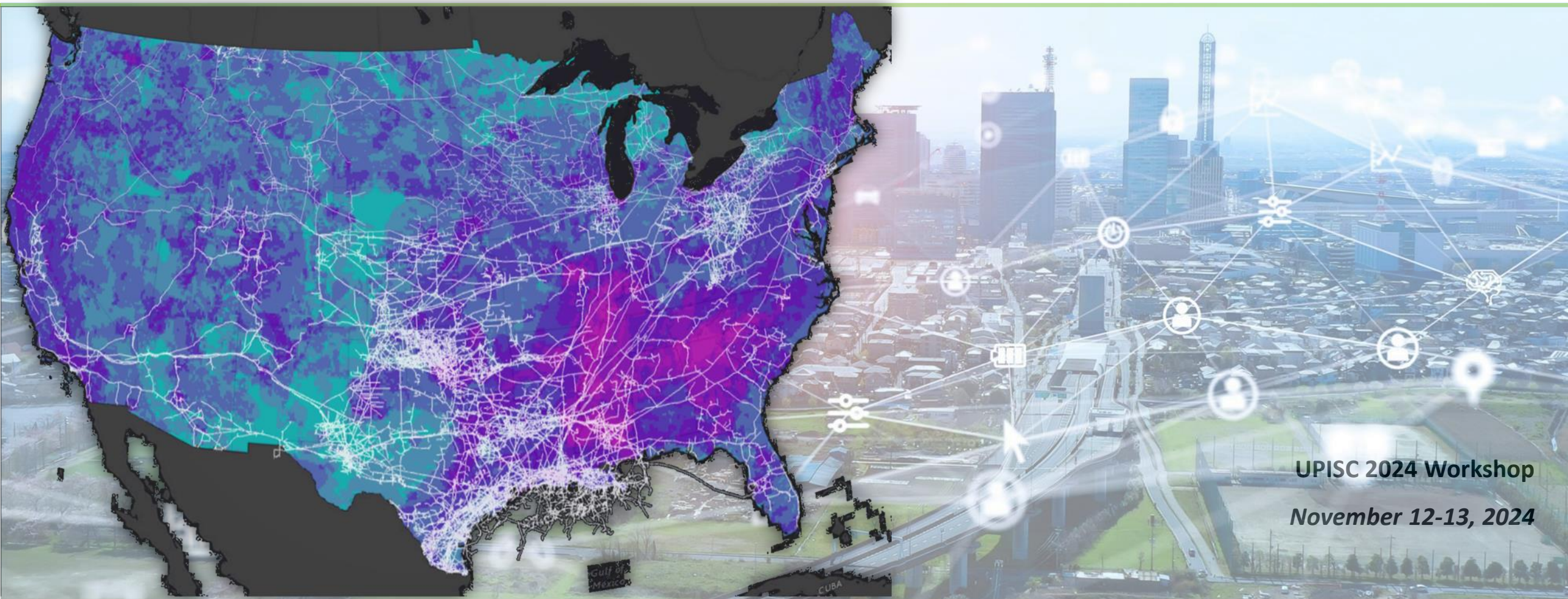
NATIONAL ENERGY TECHNOLOGY LABORATORY

Jennifer Bauer - Geo-Data Scientist
Department of Energy's National Energy Technology Laboratory (NETL)

Jennifer's work applies data science, machine learning, programming, spatio-temporal statistics, and GIS to support several research projects aligned to onshore and offshore risk reduction and prevention, characterizing energy infrastructure integrity, informing infrastructure reuse strategies and energy remediation efforts, identifying energy communities, and supporting energy transitions. Products of her work include the public release of terabytes of data, several custom-built tools and applications, and publications. Bauer serves as the Chair for DOE's Geospatial Science Steering Committee (GSSC) and a technical advisor on NETL's Science-based AI/ML Institute (SAMI).

Informing infrastructure resiliency and reuse strategies with Artificial Intelligence (AI), digital twins, and data-driven solutions

Jennifer Bauer
Geo-data Scientist



UPISC 2024 Workshop

November 12-13, 2024

Disclaimer

This project was funded by the Department of Energy, National Energy Technology Laboratory an agency of the United States Government, through a support contract.

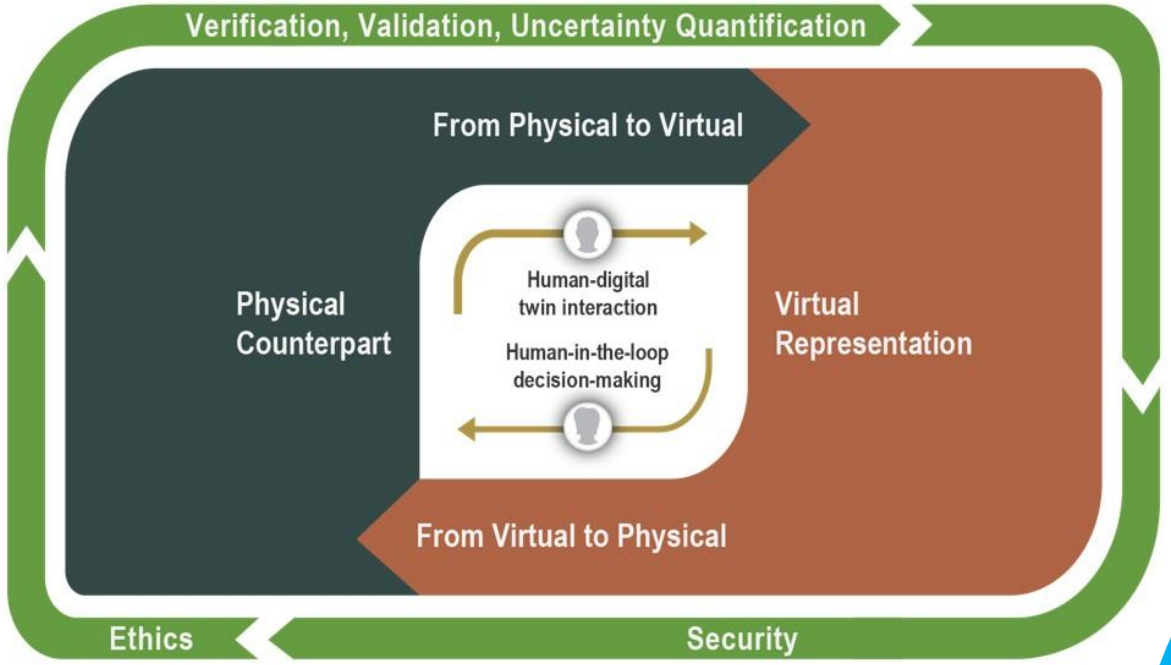
Increase in Infrastructure Investments

For energy infrastructure, investments largely designed to support resiliency & scalability

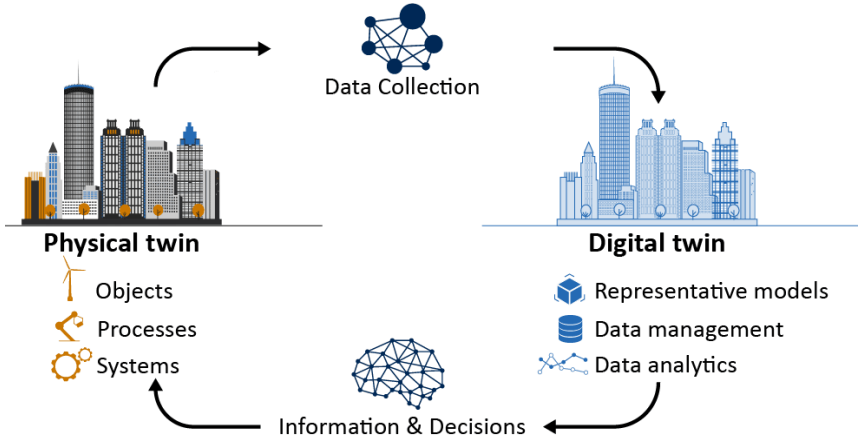
Incorporation of more sensors, more connectivity, and analytics

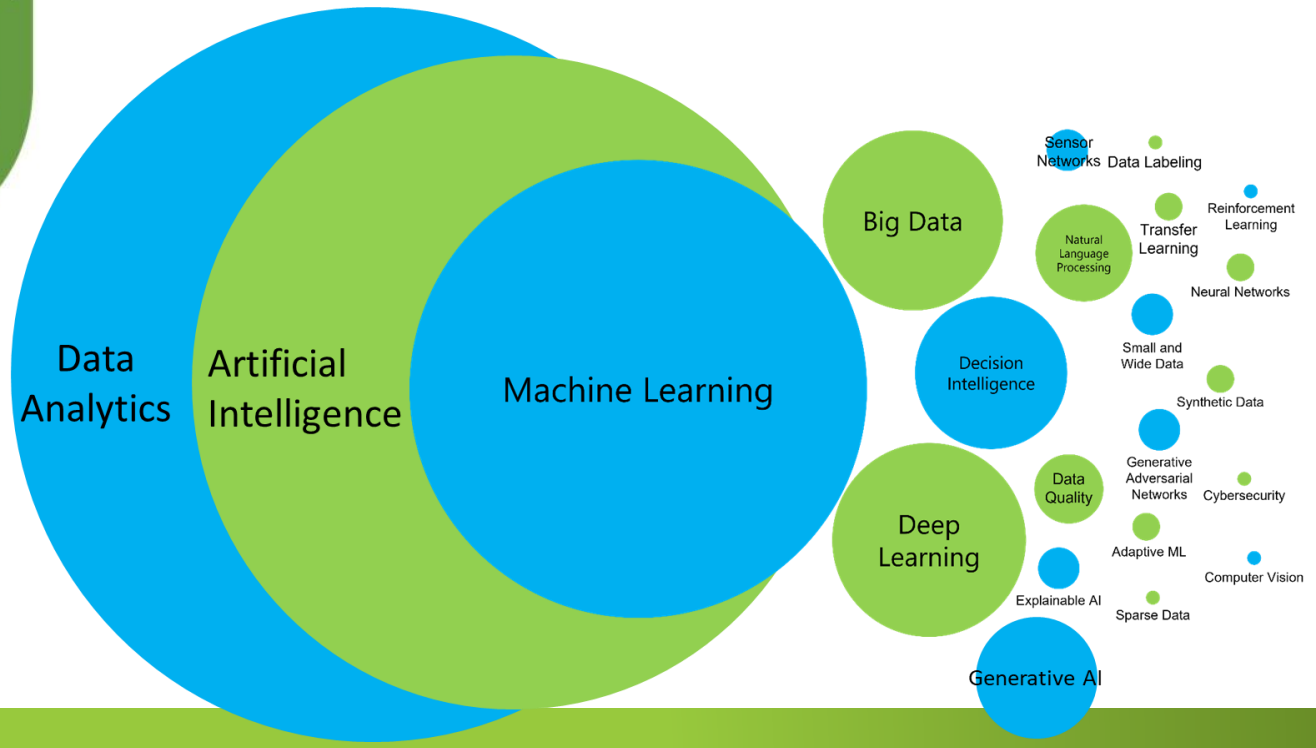
Great opportunities for safely repurposing and reinvestment into existing infrastructure assets

The opportunity for AI & Digital Twins



"A digital twin is a set of virtual information constructs that mimics the structure, context, and behavior of a natural, engineered, or social system (or system-of-systems), is dynamically updated with data from its physical twin, has a predictive capability, and informs decisions that realize value. The bidirectional interaction between the virtual and the physical is central to the digital twin" – [NASEM, 2023]





Source: https://www.gao.gov/products/gao-23-106453

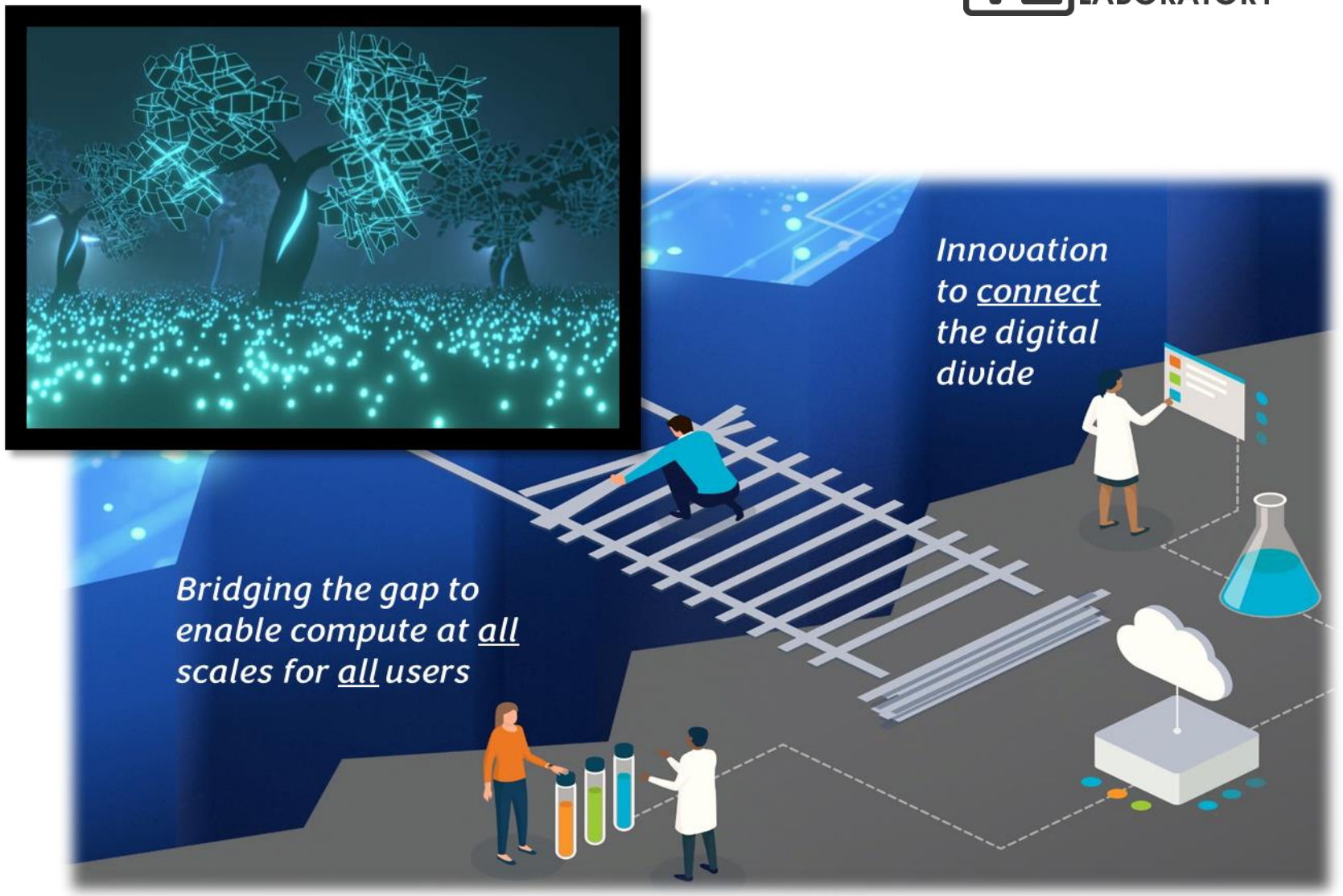
Data are a keystone for innovation

- **Both AI and Digital Twins are heavy data consumers**
- **Data-driven research projects are starved for information**

  Crowd Flower study estimates 80% of researcher project time is spent searching for relevant, existing data

  https://visit.figure-eight.com/2015-data-scientist-report

- **Improve access to data resources makes it easier to integrate info from different sensors, disciplines, and resources to improve our understanding of the whole**



An example of science-based AI + data for infrastructure

Offshore data, models, & tools supporting…

- Risk prevention & decision-support planning
- Encompasses natural-engineered system
- Measures potential social & environmental impacts

AIIM Advanced Infrastructure Integrity Modeling

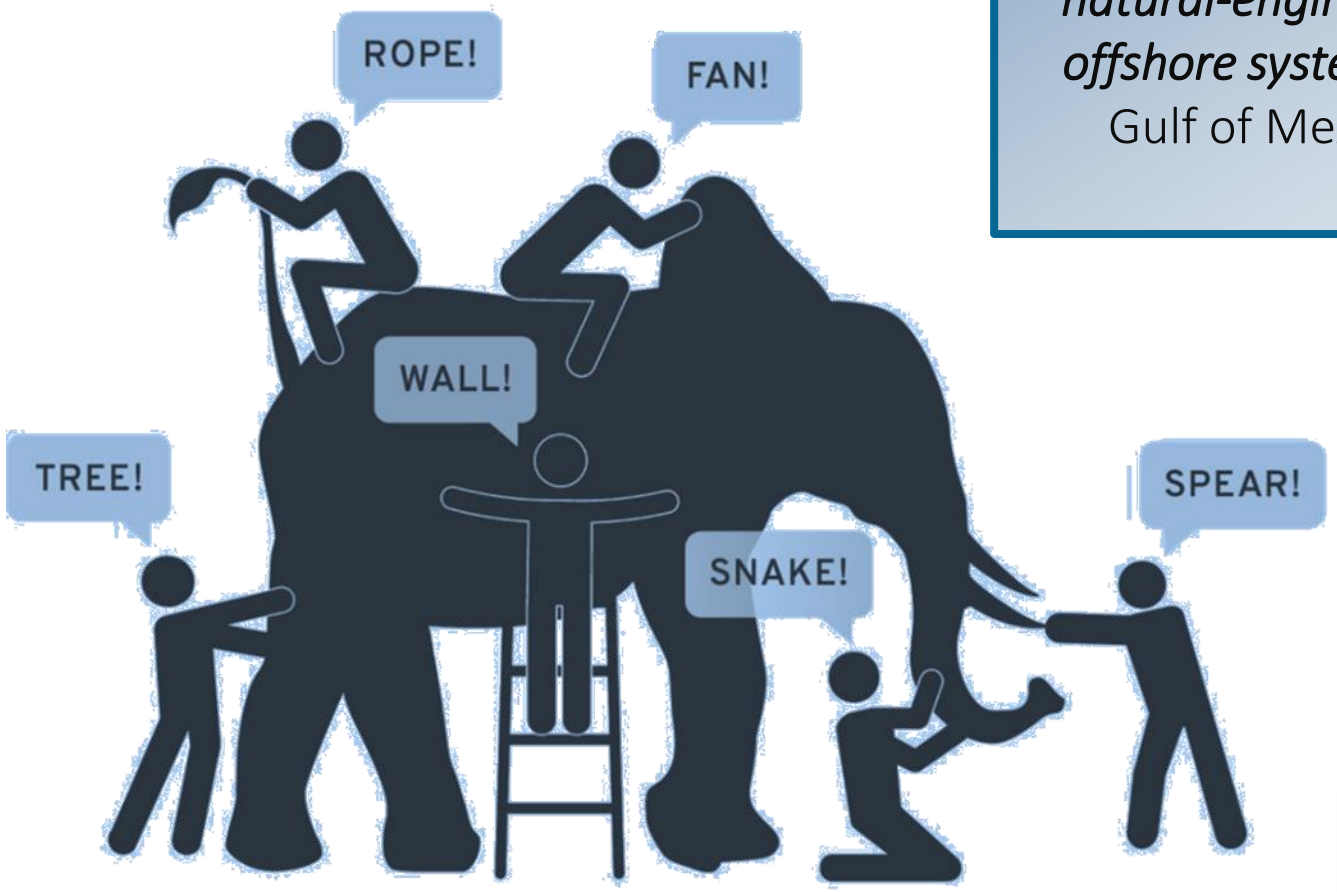
- Energy infrastructure-focused
- Published ML/analytical models on evaluation of **infrastructure integrity to inform reuse strategies**

1. **Remaining lifespan**
2. **Risk likelihood**

TechConnect WORLD INNOVATION CONFERENCE & EXPO

Big Data & Big Data Computing-Driven Insights

Using the Whole to Inform Local Trends & Predictions

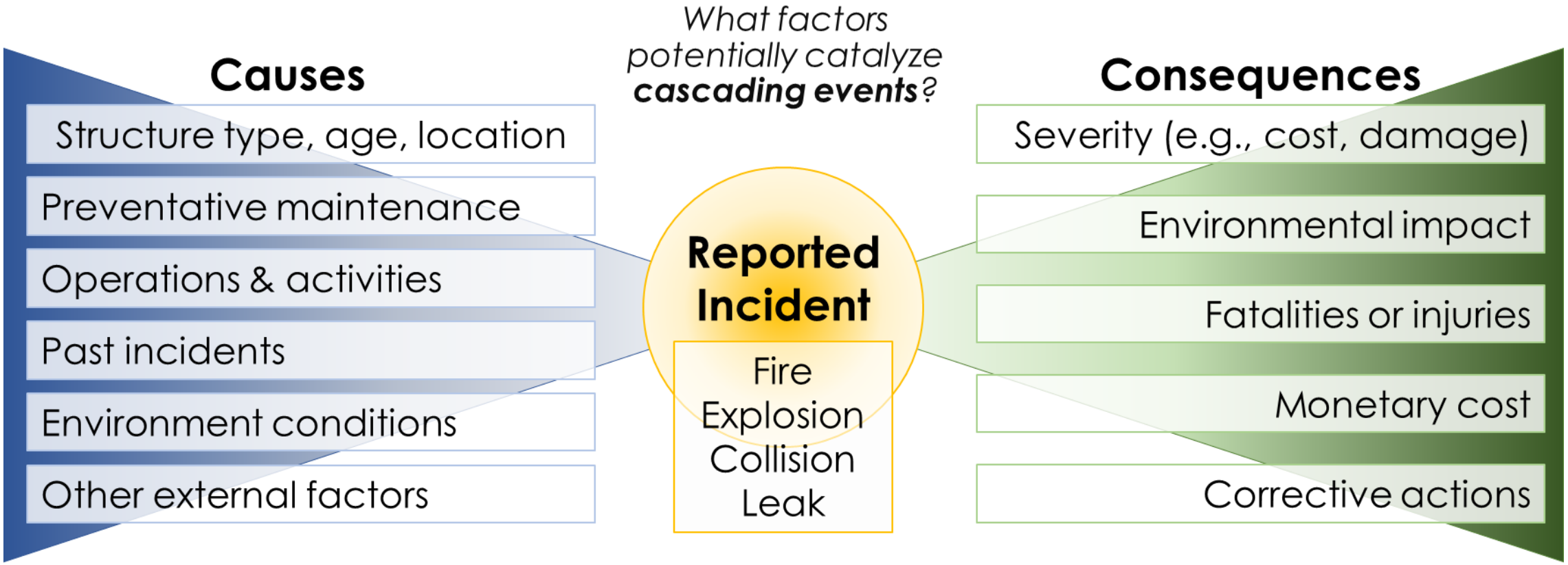


Data representing the ***natural-engineered offshore system*** for Gulf of Mexico

- **11k+** platform records
- **26k+** miles of pipelines
- **55k+** well records
- **51k+** environmental layers
- Geohazard layers
  - Landslide prediction surface from **NETL's Ocean and Geohazard Analysis** smart tool
- **46GB+** biochemical data
- **Spatio-temporal production data** at the well, platform, and lease block level
- **70+** years of platform incidents
- **30+** years of pipeline incidents
- **50.6GB** of monthly ship trackline data

Expanding Analytics on Reported Incidents

Examining Risk Cause and Effect



Answer questions such as,

- *At what age are platforms most prone to incidents?*
- *Where are infrastructure more vulnerable to incident?*

The AIIM Approach:

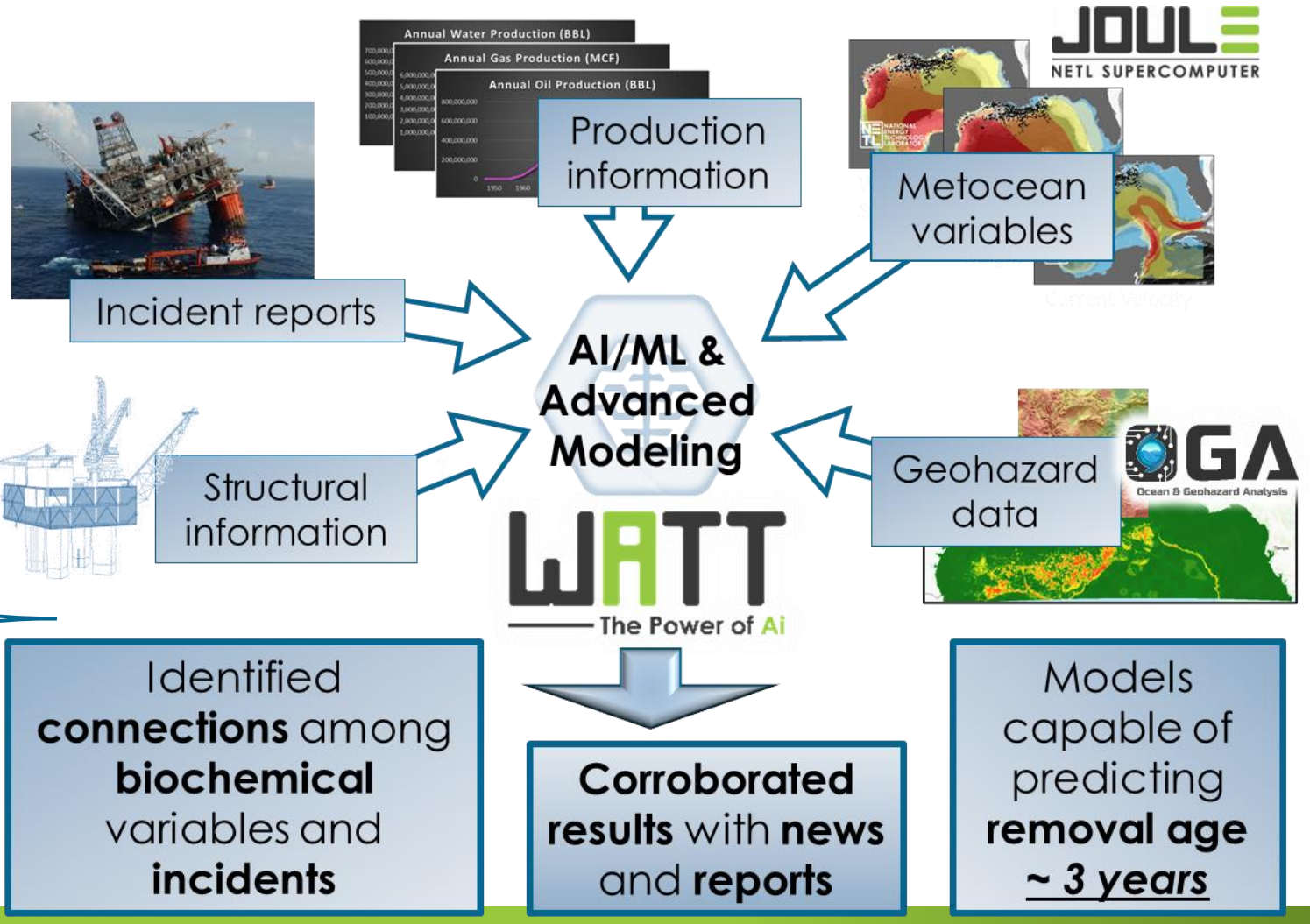
Evaluating Multi-ML Model & Advanced Analytical Insights

Machine Learning Models
(Dyer et al. 2022)

- Gradient Boosted Decision Trees (GBDT) (2 models)
- Artificial Neural Network (ANN) (2 models)
- Bayesian Network

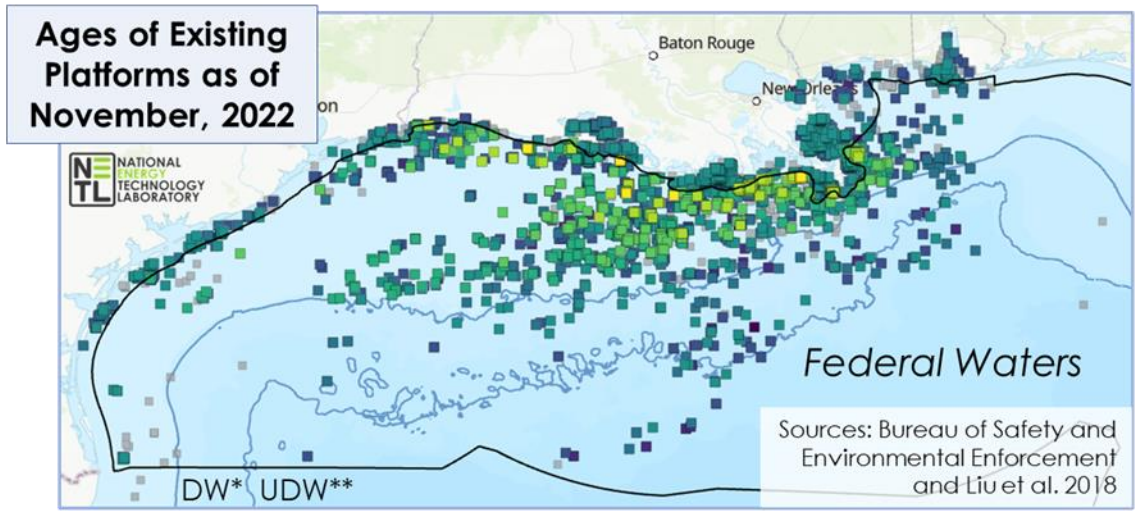
Advanced Analytics

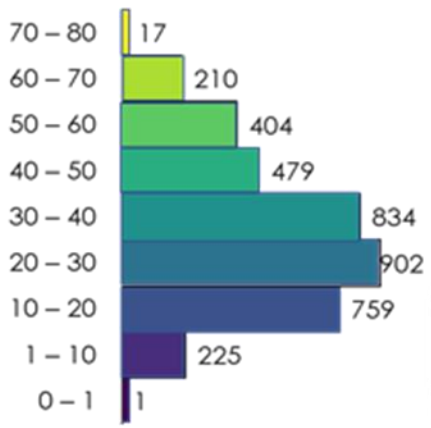
- Geographically Weighted Regression (Nelson et al. 2021)
- Causality/Time Series Analytics

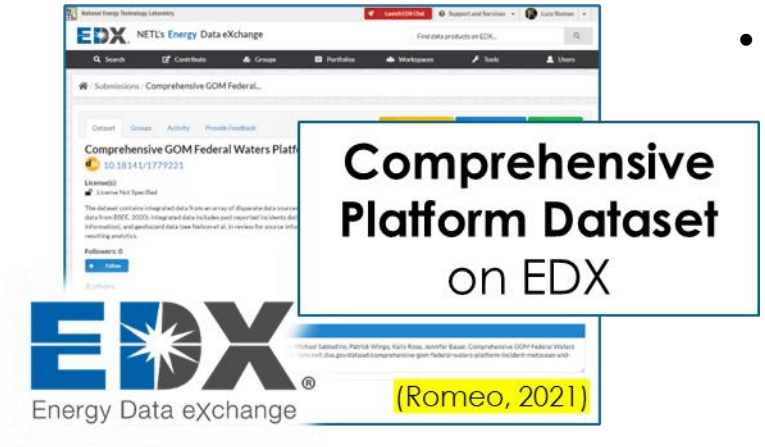


AIIM: Research Products & Insights

Evaluating Platform Infrastructure







- Integrated ***70+ years*** of data
  - **11k+ structure records**
  - **1,700+ features**
  - ***Updated incidents & metocean statistics***
- **Updated ML models** Capable of **predicting removal age within ~3 years**
- **Key stressors** included:
  - Storm occurrences (wind gust, minimum central pressure)
  - Corrosive factors (biochemical variables)
  - Number of well slots
  - Water depth
  - Reported incidents

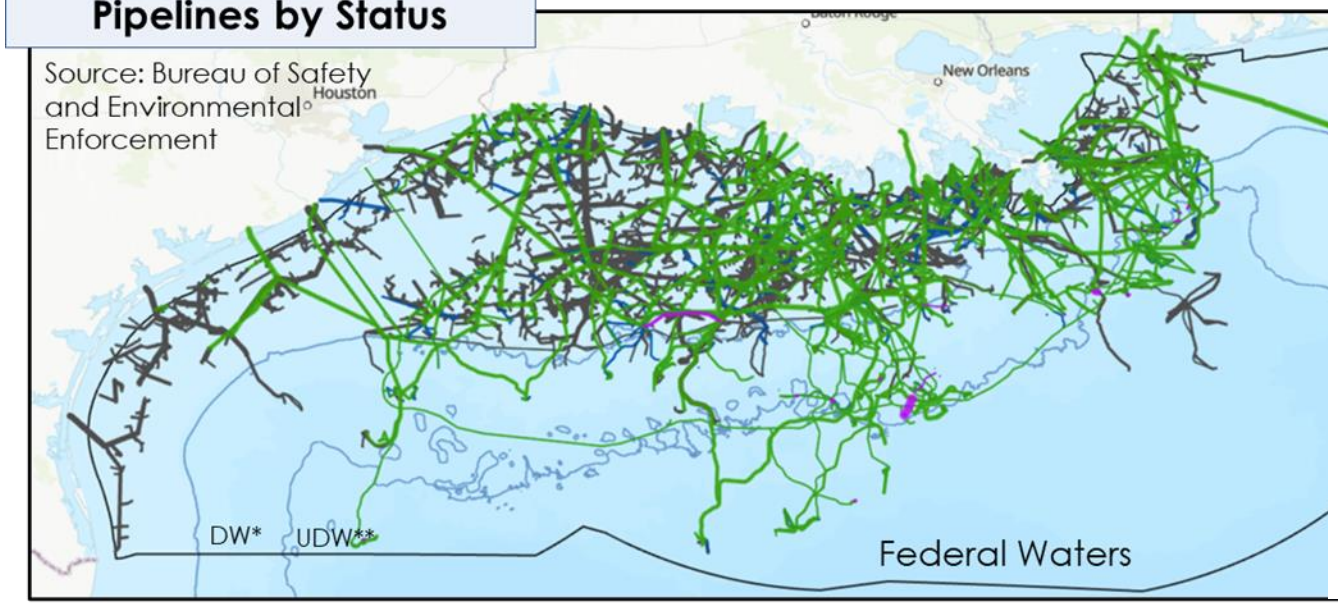
Model	Root Mean Square Error (RMSE)
ANN	**5.3 years**
GBDT (XGBoost)	**3.4 years**
GBDT (CatBoost)	**3.1 years**

AIIM: Research Products & Insights

Evaluating Pipeline Infrastructure

- 26k+ miles of pipelines
  - 100k+ locations (every 1km + end points)
  - 400+ features
- Added **seafloor metocean data** (floor current, bottom pressure)

Pipelines by Status



- **Preliminary model runs** capable of **predicting <u>abandonment</u> age within ~3 years**
- **Key stressors** included:
  - Biochemical variables
  - Installation date/proxy date
  - Cathodic code
  - Status code
  - Facility operator

Status	Miles
Abandoned or Removed	23,838
Active	19,731
Cancelled, Proposed Abandon or Remove	3,914
Proposed	109

Model	Root Mean Square Error (RMSE)
ANN	**3.3 years**
GBDT (XGBoost)	**3.1 years**
GBDT (CatBoost)	***0.7 years*** ***0.99 R^2***

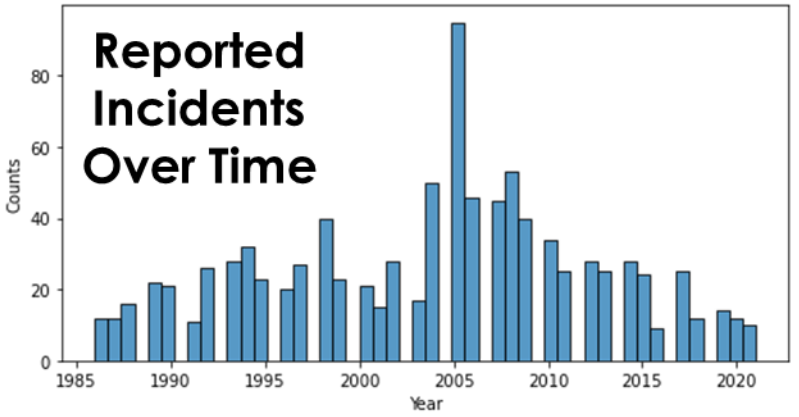
Database Publication
(Pfander et al., 2024)

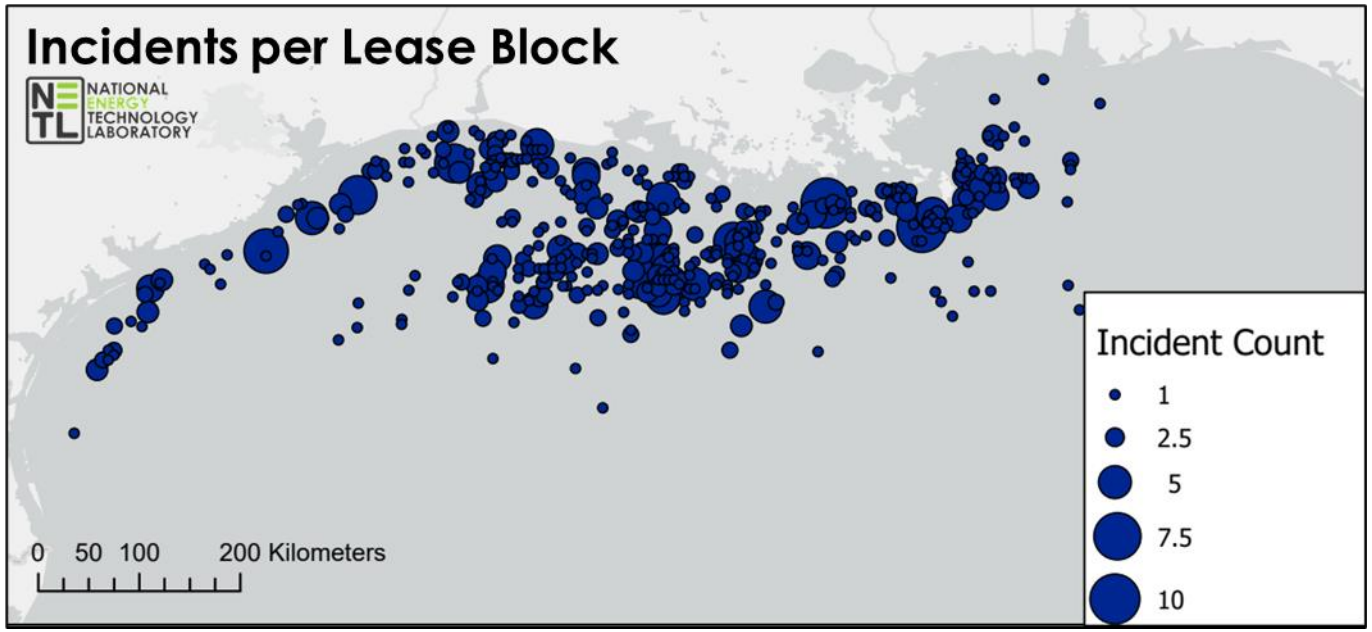
EDX Energy Data eXchange

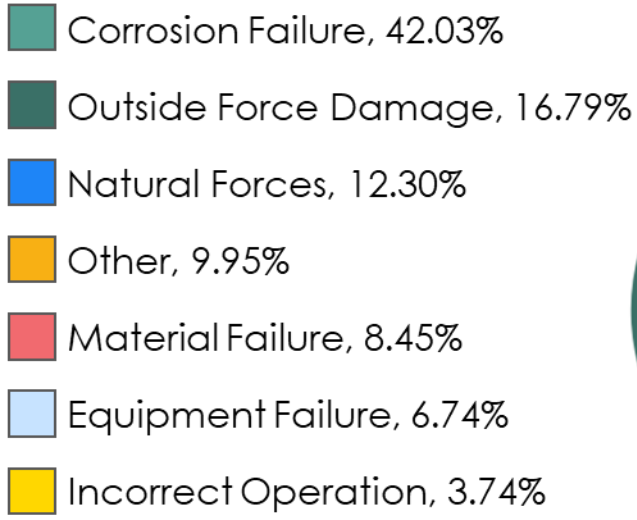
Pipeline Incident Processing

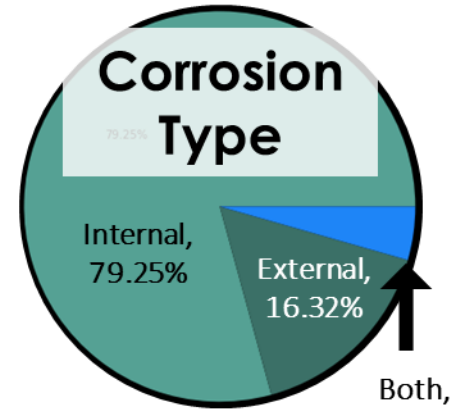
Understanding 30+ Years of Incidents

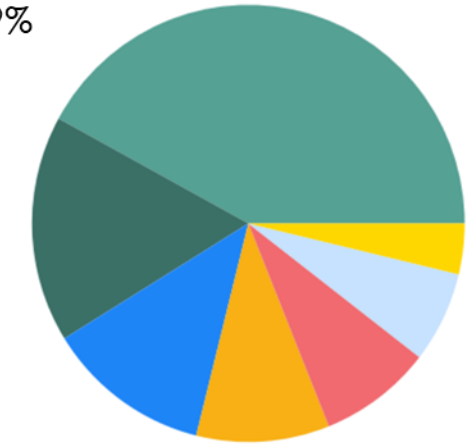
- Compiled, cleaned, and mapped **970 incidents**
  (Pipeline and Hazardous Materials Safety Administration)
- **30+ years** of incidents (1986 – 2021)
- **Mapped** more than **80%** to lease blocks
- Calculated **impact-based severity**







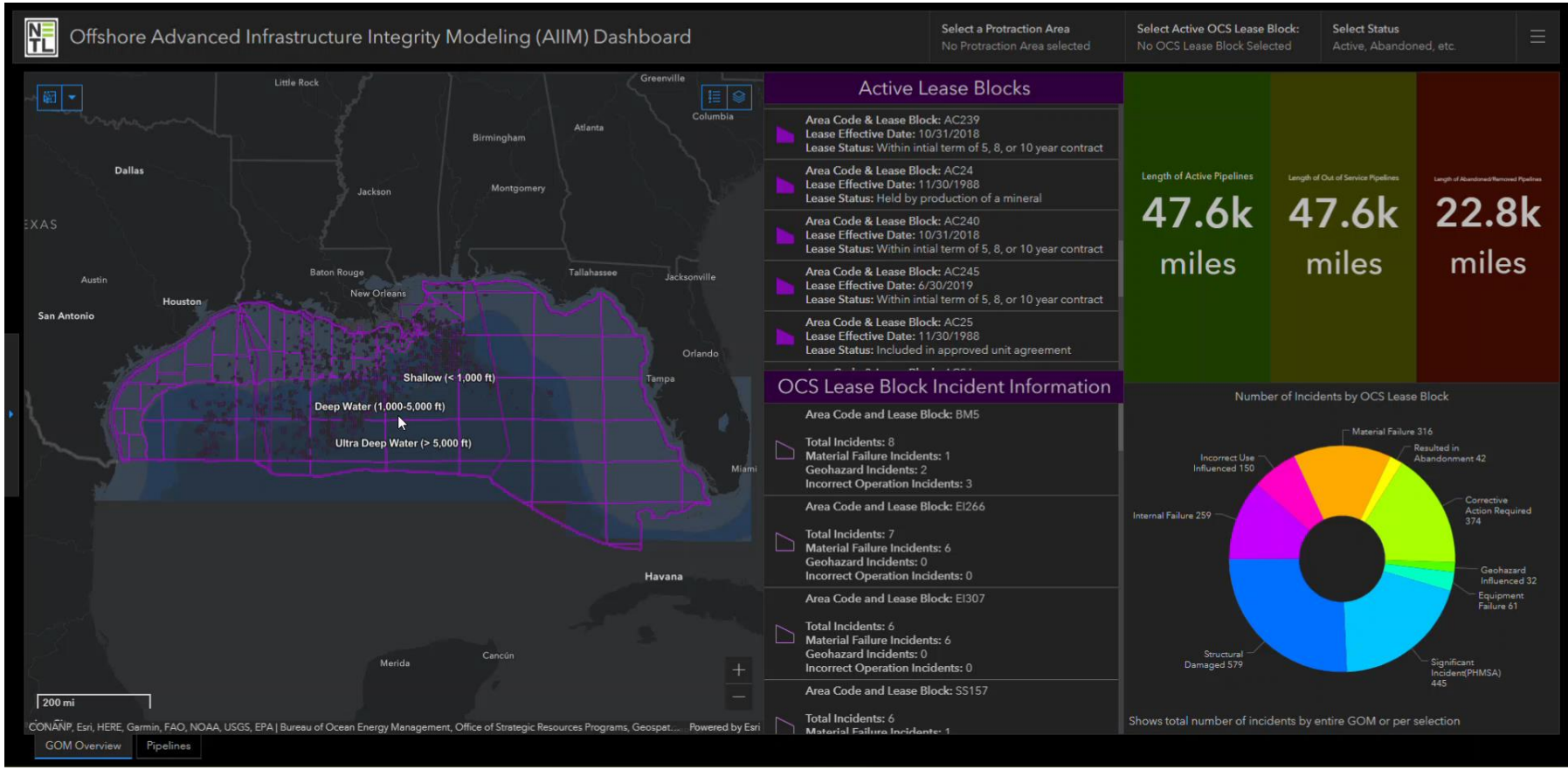




AIIM Dashboard

Delivering Data- and Model-Driven Insights

- Developed as an **ESRI Dashboard** to support data visualization & interrogation
- Contains **pipeline data** and model results
- Adding in **updated platform data** and model results
- Enable spatially querying by **areas of interest** (lease blocks, protraction areas, etc.)

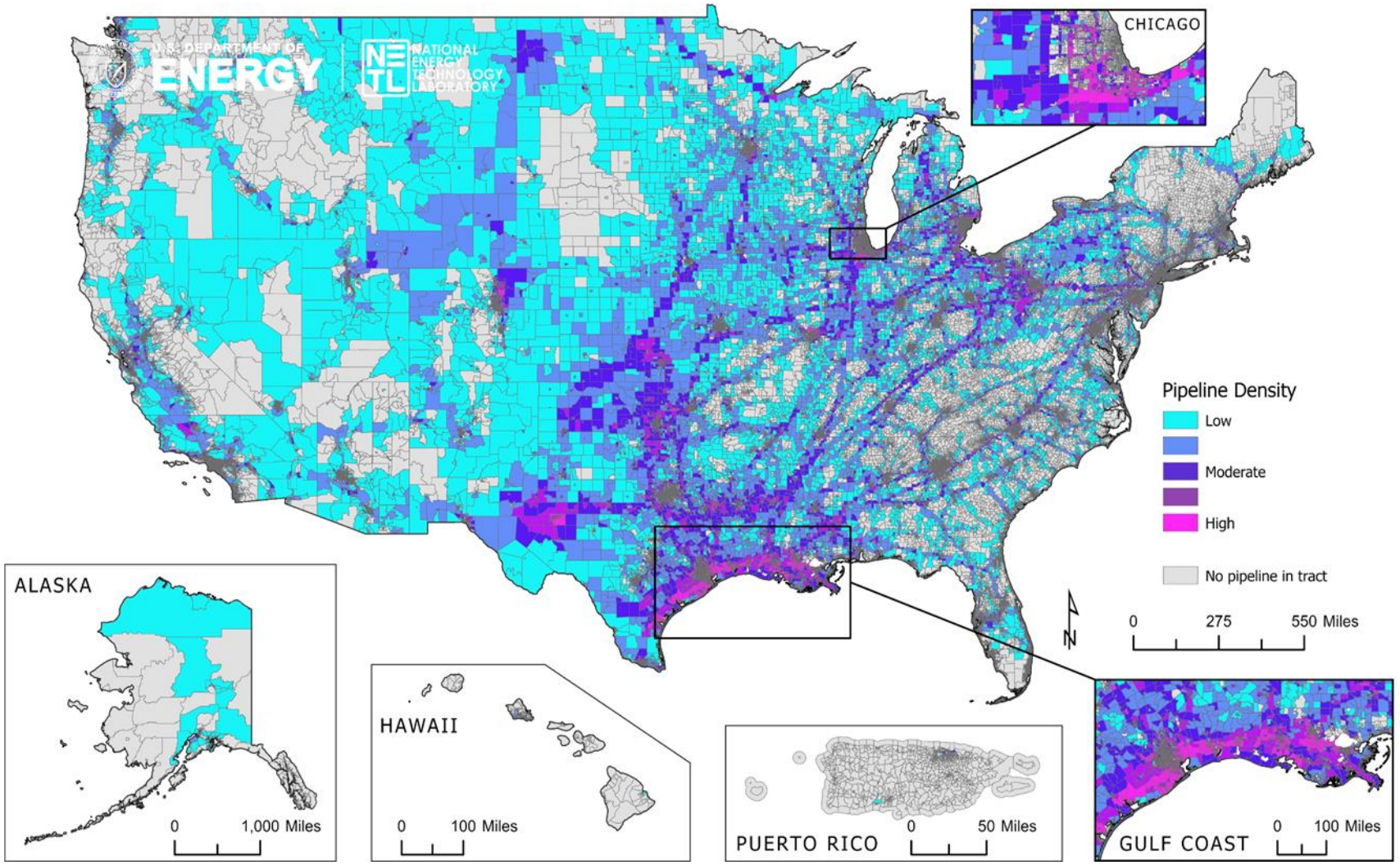


EDX Energy Data eXchange

Database & web tool available

Catherine Schooley, Lucy Romeo, Dakota Zaengle, Isabelle Pfander, Rodrigo Duran, Jennifer Bauer, Kelly Rose, (2024) Advanced Infrastructure Integrity Modeling Dashboard, https://edx.netl.doe.gov/dataset/offshore-aiim-dashboard

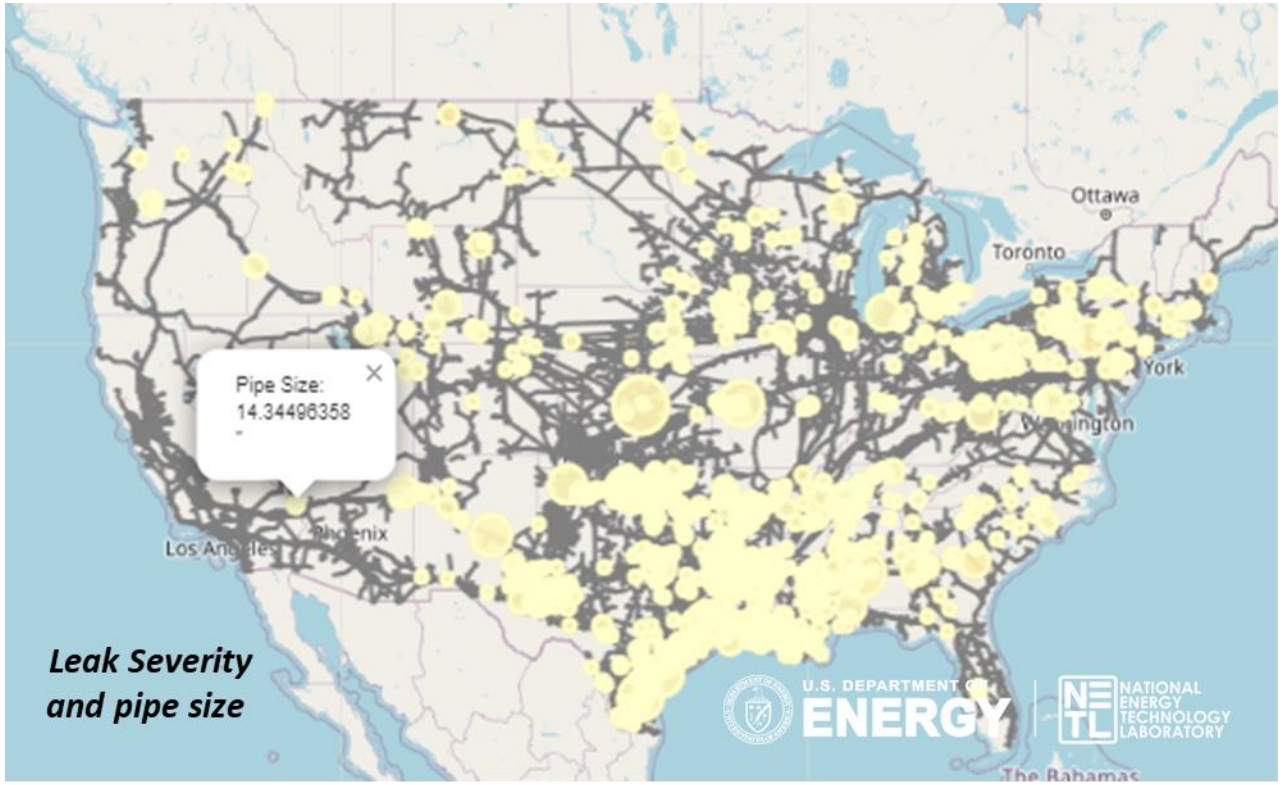
Adapting AI/AIIM for onshore pipelines

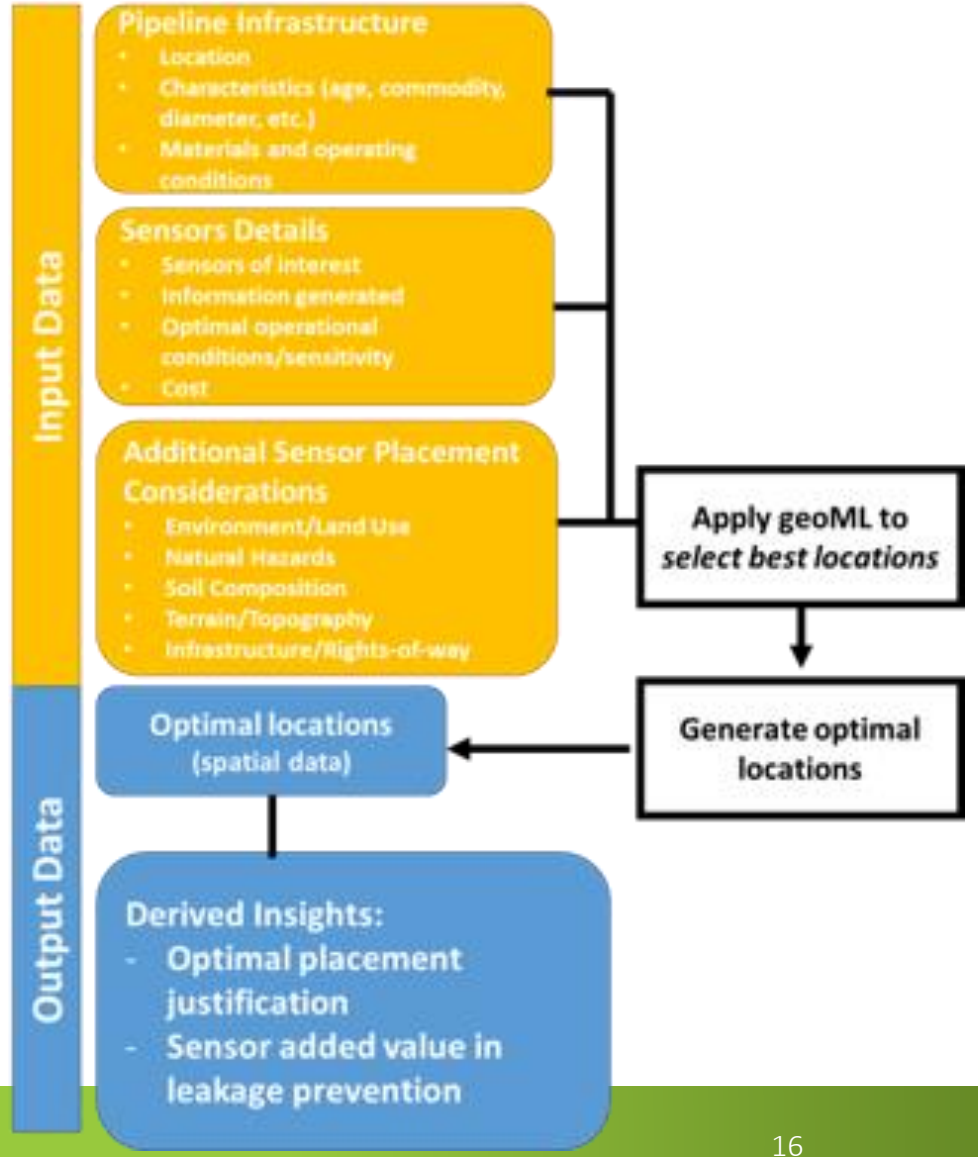


- Looking at existing pipeline network with decades of incident histories to offer insights into future or repurposed pipe
- On-going work; currently refining model variables and starting applying models

Using AI insights to support sensor placement

- Leverage much of the same data and exploratory analytical insights to begin developing a geoAI model to select optimal placement for different sensors
  - Limitations/challenges due to data availability on both pipes & sensors





Repurposing data for Infrastructure Planning

NATIONAL ENERGY TECHNOLOGY LABORATORY

Geospatial Database

Process Data

Collect Data

Create Metadata

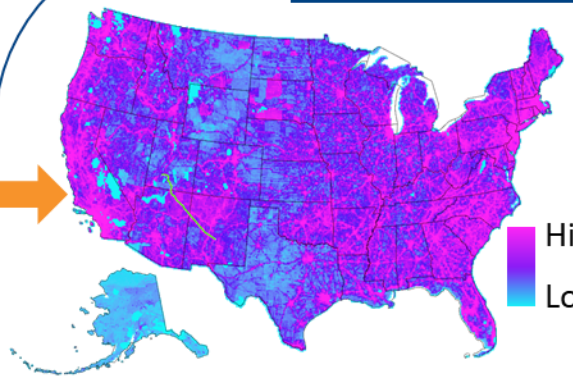
Review Resources

Apply Preliminary Weights

Publish Database

Smart CO2 Pipeline Route Planning Database - Submissions - EDX

Smart CO2 Transport-Route Planning Tool



Build Multivariate, Weighted Grid

User Inputs

Identify Route → Select Start & End Location

or

Evaluate Route → Upload Area or Line of Interest

Model-based Reinforcement Learning using Monte Carlo Tree Search (MCTS)

Outputs

Identified Route Layer

Evaluation Report

Spatially Assess Intersects to Database by 10km area & Census Tract

Output

Evaluation Report

Smart CO2 Transport-Route Planning Tool - Submissions - EDX

Enduring Challenges & Opportunities

Emphasis on data interoperability & sharing

Build confidence in AI and Digital Twins

Leverage data from other systems (telecom, civil, ecological, etc.) in secure manner

NETL RESOURCES

VISIT US AT: **www.NETL.DOE.gov**

@NETL_DOE

@NETL_DOE

@NationalEnergyTechnologyLaboratory

Jennifer Bauer
jennifer.bauer@netl.doe.gov

https://edx.netl.doe.gov