# Leveraging AI in the Nuclear Sector

## Data Driven Decision Making

Robert Austin
Senior Program Manager

World Institute for Nuclear Security (WINS)
December 10 -11, 2024

# Overview

- EPRI Overview
- EPRI 3DM program overall
- Some specific applications
  - Large Language Models
  - “Old School” AI for Non-Destructive Examination
- Upcoming engagements

EPRI

# Vision

To be a world leader in advancing science and technology solutions for a clean energy future

# Mission

Advancing safe, reliable, affordable, and clean energy for society through global collaboration, science and technology innovation, and applied research.

*Together…Shaping the Future of Energy®*

**Explore EPRI's research across the Nuclear, Generation, and Energy Delivery and Customer Solutions sectors ranging from decarbonization to grid modernization to low carbon resources.**

## COLLABORATION

EPRI's collaborative platform is unrivaled. Our R&D:

- Leverages your research dollars
- Connects you to a global network of peers
- Accelerates deployment of technology
- Mitigates the risk and uncertainty of going it alone
- Positions you as a leader in addressing industrywide challenges

## CREDIBILITY

EPRI's independent research is guided by our mission to benefit the public. We offer:

- Objective solutions
- A proven track record
- Scientifically based research you can trust


## EXPERTISE

For more than 50 years, EPRI has been applying R&D to help solve real challenges. With EPRI, you can:

- Reduce expenses and increase productivity
- Be more resilient today and better prepared for tomorrow
- Access an industry repository of collective experiences, technical expertise, and training resources
- Extend your staff and make your teams more robust and more confident
- Benchmark, learn and share best practices
- Increase your awareness of challenges that others are facing and alternate solutions to challenges you might be facing
- Save time and money troubleshooting problems EPRI and its stakeholders have seen before

## Who We Are

Founded in 1972, the Electric Power Research Institute (EPRI) is the world's preeminent independent, non-profit energy research and development organization, with offices around the world.

## Our Experts

EPRI's trusted experts collaborate with more than 450 companies in 45 countries, driving innovation to ensure the public has clean, safe, reliable, affordable, and equitable access to electricity across the globe.

# Poll – Your Data Science / AI Experience?

1. Actual degree or certificate (Cool!  You can answer hard questions.)

2. Some coursework & coding (this is me)

3. Some reading & study

4. Played with ChatGPT, Copilot, etc. a little

5. None

# Data-Driven Decision Making (3DM) at EPRI

Leverage data science for the Nuclear Power industry

Launch & support activities across the Nuclear Sector

General application areas

*Insights*: learning from the past

*Prognostics*: anticipating the future

*Automation*: increasing reliability

*Optimization*: increasing efficiency

More details on 3DM program page

Projects

Results

**Applying Data Science in the Nuclear Power Industry**

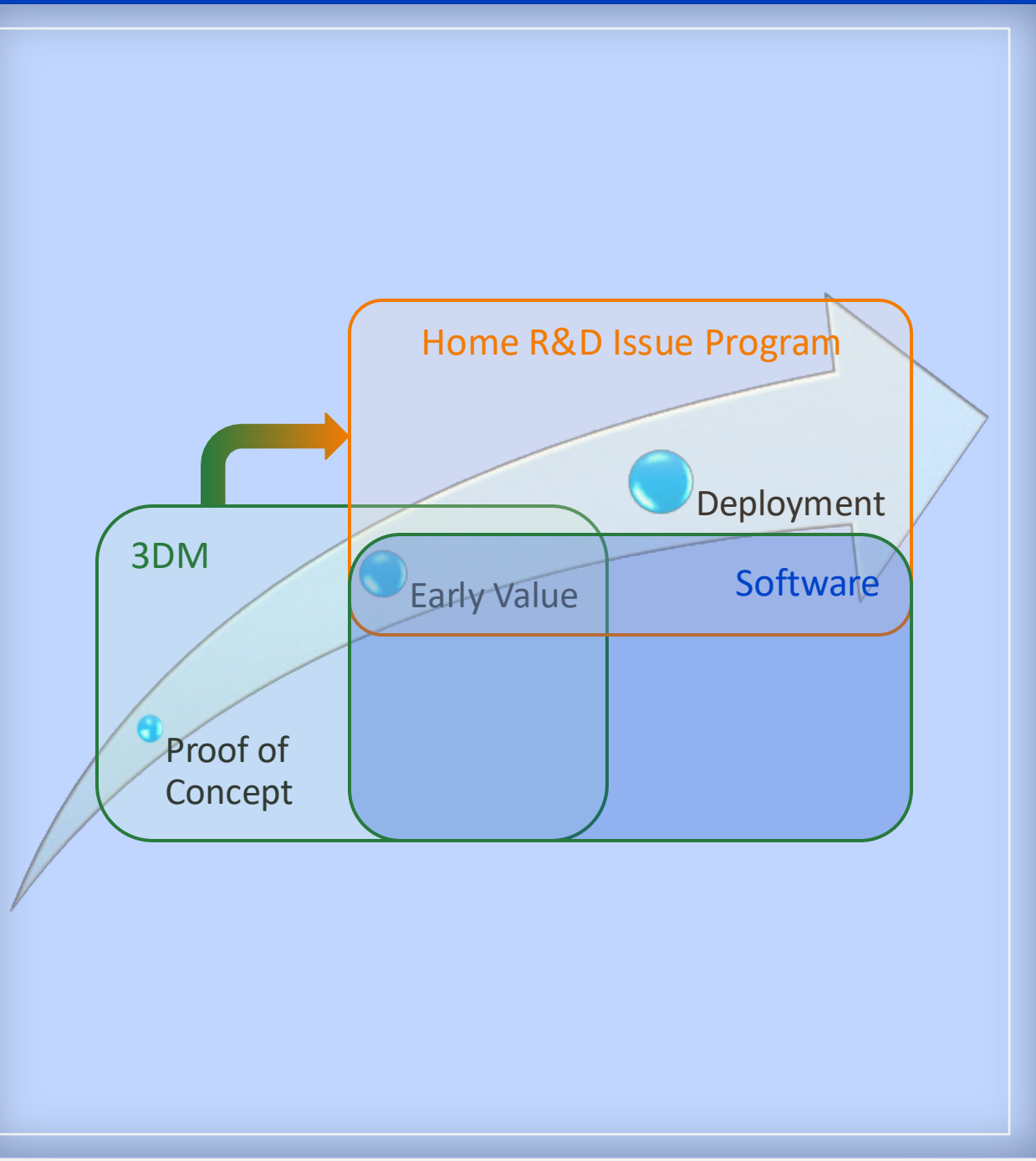
# Overall Approach

**Launch**

- Initiate key proof of concepts
- Sponsor early implementation & tech transfer
- Develop use guides

**Support**

- AI projects across the industry
- Strategic initiatives



# Large Language Models

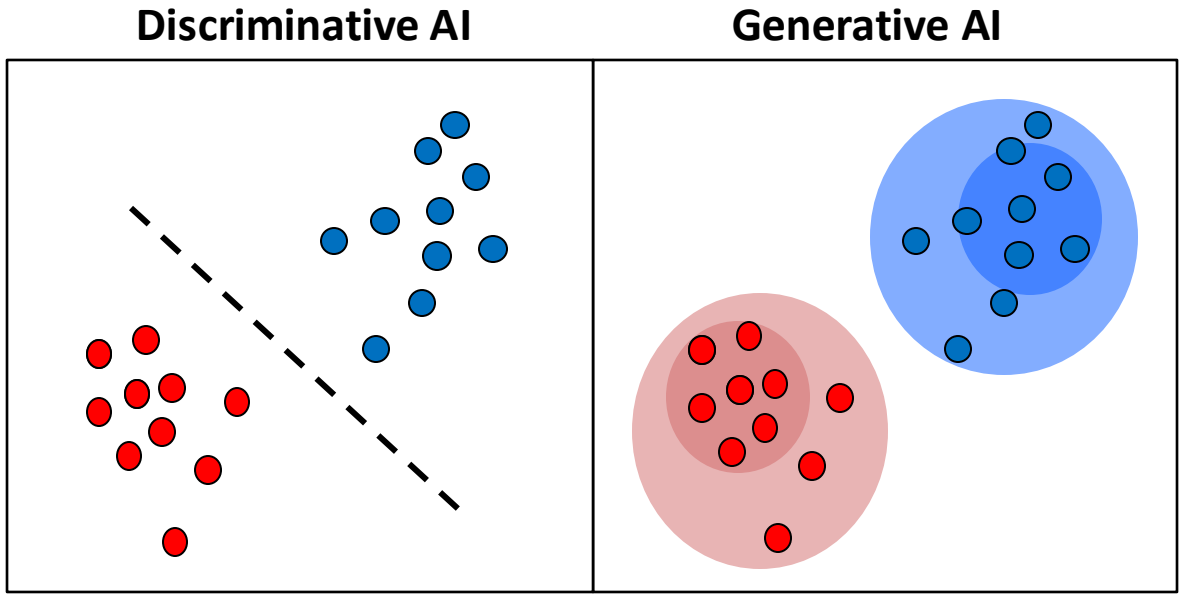
# LLM, e.g. ChatGPT, Breakdown

**GPT** stands for **G**enerative **P**re-Trained **T**ransformer, one type of Large Language Model (LLM)

*Generative* means that it is creating something, in this case text or “chat”

*Pre-Trained* means that the probabilistic models that predict what word gets generated next are based on large data sets

*Transformer* refers to the type of model




Discrete boundaries
Linear regression
Decision trees
Statistical NLP

Probabilities in space
Recurrent neural networks
Diffusion models
Large language models

# Selection (Dated) of Available Large Language Models

| Model | On-prem | Number of parameters | Architecture | What type of data is it trained on? | Use case |
|---|---|---|---|---|---|
| GPT-3.5/ChatGPT (OpenAI) | no | 175B | decoder | General, web pages, books, more | Conversational, content generation |
| Llama 2 (Meta) | yes | 7-65B | decoder | web pages, books, other written materials | Same as GPT family, but more efficient/faster |
| GPT-J 6B (EleutherAI) | yes | 6B | decoder | C4 dataset | Text generation, fine-tuning, English only |
| T5 11B (Google research team) | yes | 11B | encoder-decoder | Web pages, Wikipedia, etc. | Summarization, content generation |
| Dolly 2.0 (Databricks) | yes | 12B | decoder | instruction prompts | Conversational |
| Flan-T5 | yes | 11B | encoder-decoder | T5 is based on C4 dataset, Flan-T5 was trained on prompts | Primary use is research on zero and few shot learning |
| RoBERTa | yes | 125M | encoder | BookCorpus, Wikipedia, CC-News, OpenWebText, Stories | Decision making, classification |
| mT5 | yes | 13B | encoder-decoder | mC4, covers 101 languages | Translation |

# A Concern - Hallucinations

- When model says something that ***sounds*** fluent, but factually incorrect.
  - Example: EPRI is headquartered in London.
- This usually happens when the model does not have the answer or the correct context so it “guesses” based on probabilities.
- A lot of work is being done to limit these hallucinations.

# Another Concern - Data Security

- Training models can require uploading data to the model
- This data then potentially can inform the model, resulting in loss of data security
  - "The complaint by the New York Times states that OpenAI took millions of copyrighted news articles, in-depth investigations, opinion pieces, reviews, how-to guides and more in an attempt to "free ride on the Times's massive investment in its journalism."

    (https://theconversation.com/how-a-new-york-times-copyright-lawsuit-against-openai-could-potentially-transform-how-ai-and-copyright-work-221059)

# Model Improvement and Hallucination Mitigation

## Prompt Engineering

- ***$$ work and resource intense method***
- ***Needs SME input***
- ***Few-shot prompting***
  - Few examples to guide the model's understanding
  - Improves on zero-shot prompting
- ***Chain-of-thought prompting***
  - Series of prompts that build upon each other, forming a coherent chain of thought
  - Maintains context, connectivity in responses, and logic
  - Especially useful for complex problems that require reasoning

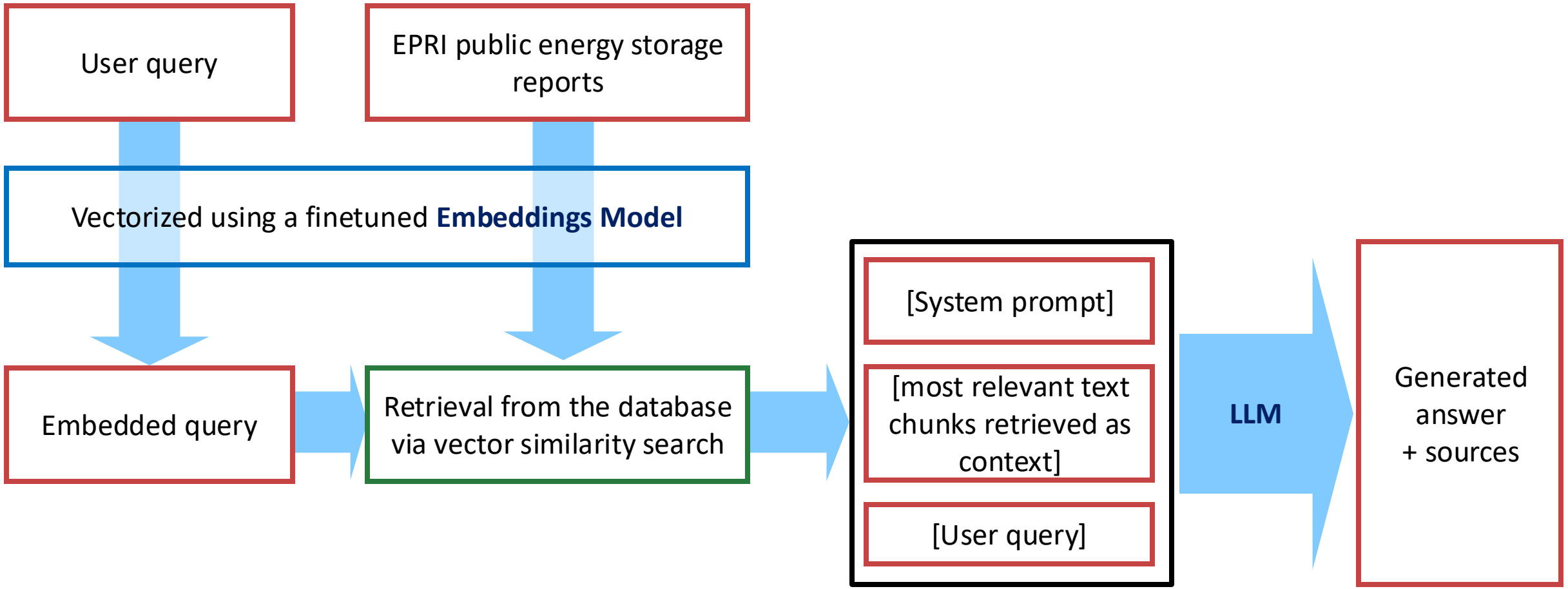
## Fine-Tuning

- ***Supervised Fine-Tuning ($$$)***
  - Train pre-trained model with labeled examples, adapting its parameters to achieve task-specific performance
  - Ideal for complex tasks with well-defined labels
- ***Unsupervised Fine-Tuning ($$$$)***
  - Adapts a pre-trained model for specific tasks without explicit labeled data
  - Ideal when a task is less well-defined or when capturing general language understanding is the primary goal
- ***Reinforcement Learning from Human Feedback (RLHF) ($$)***
  - Involves refining a language model's behavior by receiving feedback from human evaluators

## Grounding

- ***$ Easy to set up and maintain***
- AKA Retrieval Augmentation (RAG)
- Information can be added at any point in time
- Vector database for information retrieval
- Connects the language model to external knowledge or real-world context
- Incorporates information from external sources to improve model performance for specific tasks

# Retrieval Augmented Generation (RAG): Basics



# Some Examples of Grounded Answers

| Question | SME answer (real answer) | Llama-2 13B with RAG (grounded) | GPT-3.5 (no RAG, not grounded) |
|---|---|---|---|
| What is the energy density of molten salt energy storage? | 400-800 MWh/acre | Based on the information provided in the table A-7, the energy density of molten salt energy storage is approximately 400-800 MWh/acre. ✓ | As of my last knowledge update in January 2022, typical energy density values for molten salt systems used in CSP are in the range of 50 to 150 megajoules per liter (MJ/L). ✗ |
| What is the energy density of lithium-ion battery storage? | 100-300 MWh/acre | Based on the reference [2], the energy density of lithium-ion battery storage is approximately $300/kWh for a 4â€“8 hour duration system ✗ | approximately 100 to 265 Wh/kg. ✗ |
| What is the energy density of lithium-ion battery storage in *MWh/acre*? | 100-300 MWh/acre | 100-300 *MWh/acre* ✓ | The energy density of lithium-ion batteries is typically expressed in terms of watt-hours per liter (Wh/L) or watt-hours per kilogram (Wh/kg). However, expressing it in terms of MWh/acre involves converting these units and considering the area. ✗ |

# On-prem Efforts with the DGX-2

DGX-2 specs

- 16X NVIDIA® Tesla® V100 GPUs, for a total GPU memory of 512GB
- 81,920 CUDA® cores, 10,240 Tensor cores

Comparison

- Dell laptop: 16GB memory, 4 cores
- iPhone X: 3GB memory, 6 cores

- All processed data, including the documents and prompts, are completely private as everything happens locally

- This could accommodate even some of the largest open-source LLMs, such as LLama-2 65B (65 billion parameters)
- Successfully deployed both **LLama-2 7B, LLama-2 13B** (7, 13 billion parameter models), **Mistral 7B** & **Mistral 8x7B** running on all 16 GPUs

# Revised On-prem Results

- Total of 34 questions, answered by Energy Storage SMEs at EPRI
- For grounding of information retrieval model, public reports were used
- Even small prompt changes improved the on-prem results

**ChatGPT (GPT-3.5) in Azure Playground**

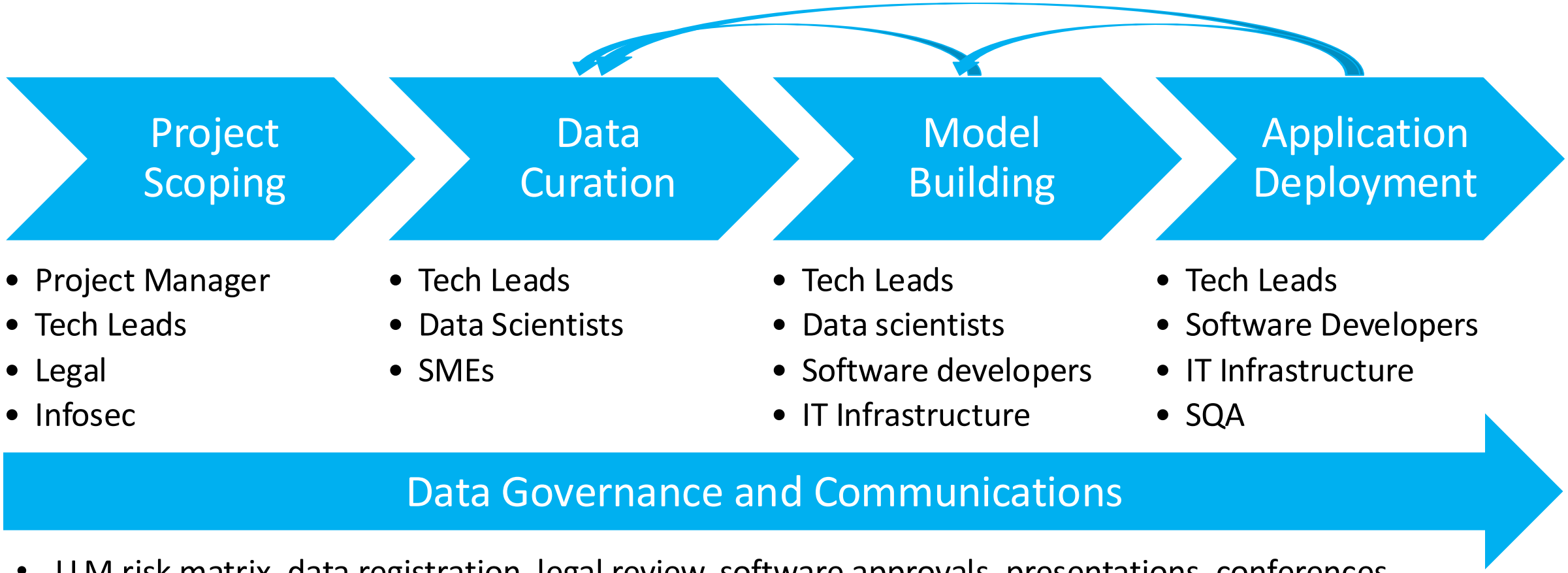
- Speed test: ~3.6 seconds/question
- Word matching: 46%
- Semantic matching: 57%
- Human verified answers: 29%

**LLama-2 7B on the DGX-2**

- Round 1: With grounding
  - Speed test: ~6 s/question
  - Word matching: 39%
  - Semantic matching: 43%
- Round 2: With grounding + prompting (Aug 2023)
  - Speed test: ~7 s/question
  - Word matching: 50%
  - Semantic matching: 56%
- Round 3: With revised grounding + prompting (Dec 2023)
  - Speed test: ~7 seconds/question
  - Word matching: 89%
  - Semantic matching: 64%
  - Human verified answers: 82%

**New on-prem results surpass OpenAI**

# LLM Workflow: Team Roles




- LLM risk matrix, data registration, legal review, software approvals, presentations, conferences
- Legal, Infosec, Project Manager, Tech Leads, Stakeholders

# Summary

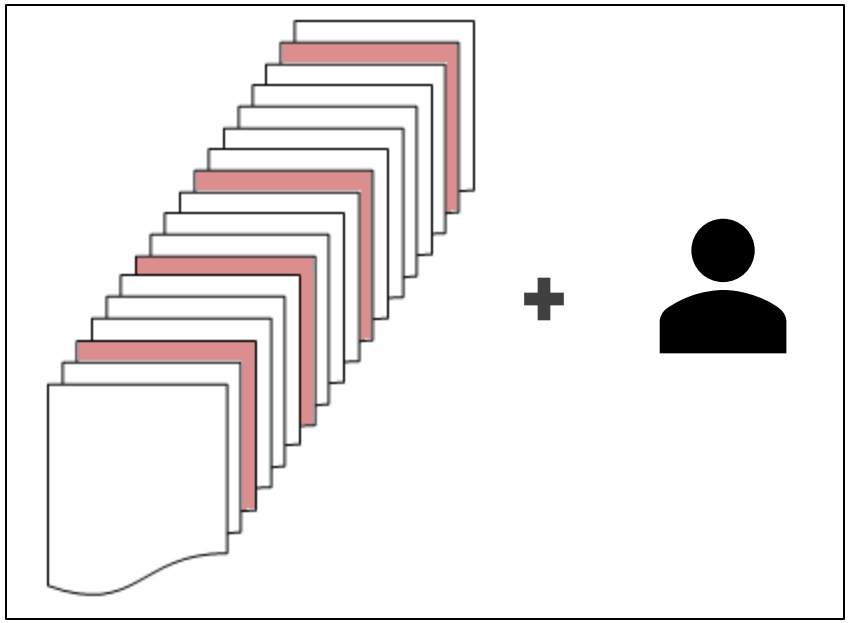
- Large language models offer potential for productivity improvements, but still a new technology with concerns with “hallucination” and data security
- Good results can be obtained with smaller, local models
  - Mitigates IP loss concerns
  - Runs on less (i.e. cheaper) hardware
  - Might give BETTER results (but again, new technology)
- Requires a multi-disciplined team

**Will be an area of active R&D for some time**

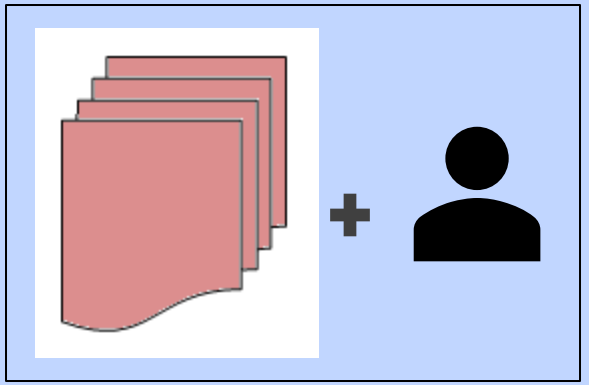
# AI-Assisted Ultrasonic Testing (UT) Inspections

# How Would AI Assist in UT Inspections?

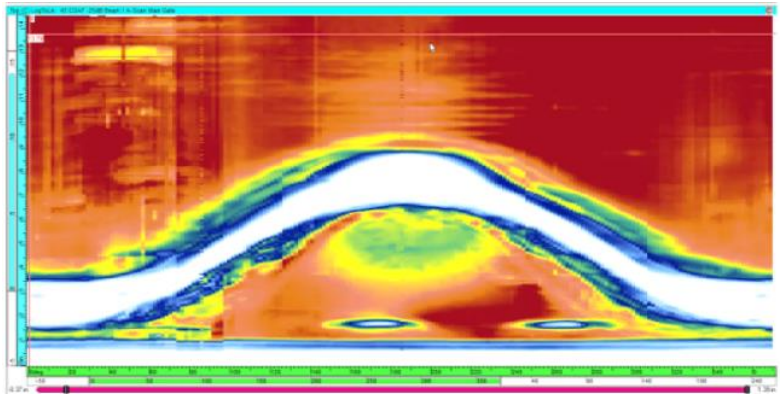
- Current Inspection
  - Examiners distribute their energy across a high volume of (mostly benign) data

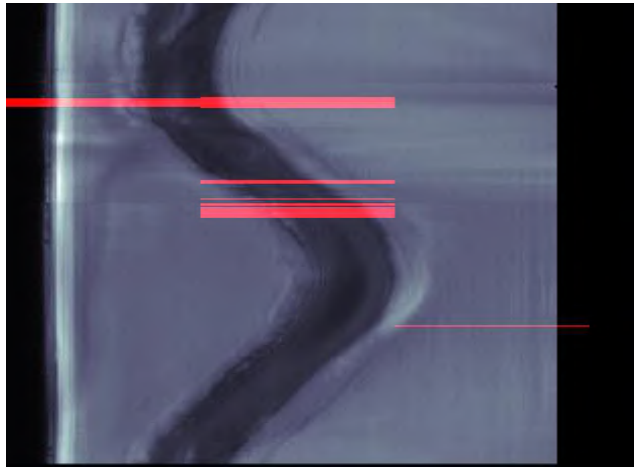


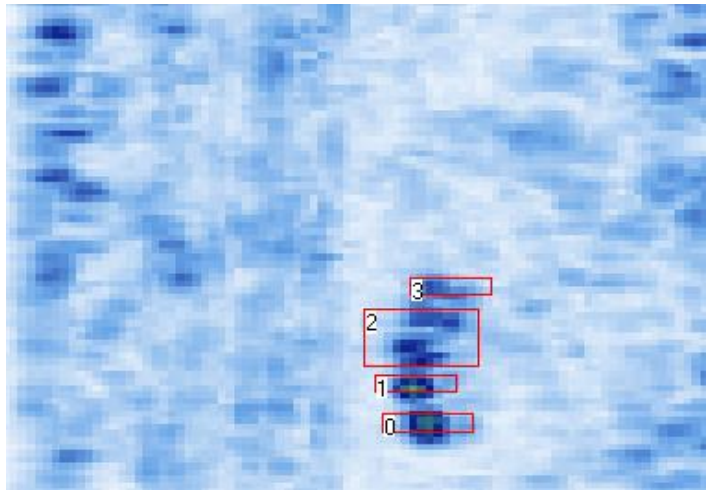
- AI Assisted Inspection
  - Examiners focus their energy on the regions that require more careful review, while AI takes care of the more monotonous portion



# Deployment








**1** UT data is saved in ML Box through local connection

**2** Box automatically detects and processes UT data…

**3** … and generates output for review

**All data remains local; no SW needed**

# Status

## Complete

### Reactor Vessel Upper Head (RVUH) Penetrations

READY NOW!!!

- Two successful field trials (2022-23)
- Positive qualification assessment
- ➢ European utility currently seeking qualification

## Work in Progress

### Dissimilar Metal Welds

- One field trial completed (Sep 2023)
- Addressing lessons learned now
- ➢ Looking for additional host volunteers

## Coming Soon

### Core Barrel + Core Shroud

- Starting in 2024
- Looking to add automated length-sizing capabilities

# RVUH 2022 Field Trial

Watch the video

EPRI

SUCCESS STORY

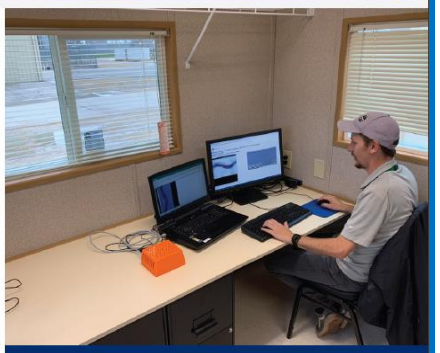
AI TOOL DEVELOPED BY EPRI SIGNIFICANTLY CUTS ANALYSIS TIME IN U.S. NUCLEAR PLANT FIELD TRIAL

**Data analysis takes four hours compared to four days without artificial intelligence**

A field trial at a U.S. nuclear power plant of an automated data analysis application being developed by EPRI has shown that the innovation offers not only substantial time savings but also a possible path to easing future staffing issues.

**BACKGROUND: MILES OF DATA**

During scheduled outages at nuclear plants, critical components are inspected for service-induced degradation. As part of the first-of-its-kind field trial at a U.S. plant in April 2022, ultrasonic inspection data from non-destructive evaluation (NDE) of 78 reactor pressure



*EPRI staff performing AI-assisted analysis during the field trial. The orange box performs the analysis locally and produces the report for review.*

***Success Story: 3002025510***

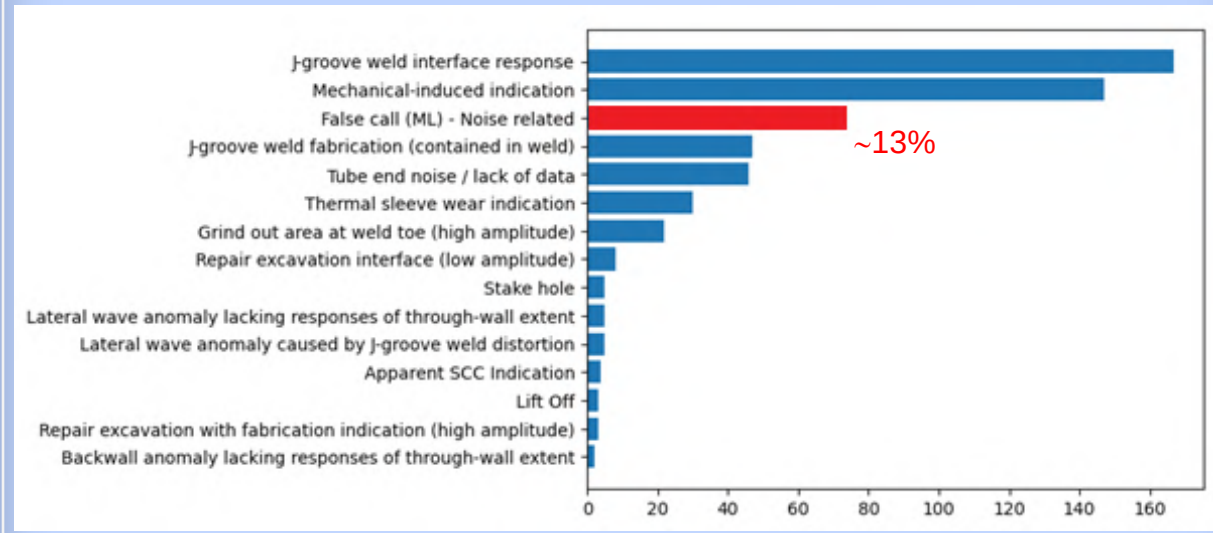
*AI-assisted analysis*:

- Flagged **2%** of the volume for review (after on-site retraining)
- Flagged **all SCC** locations
- Yielded **same results** as the one from the qualified manual technique

# Amount of Data Requiring Review

| Pre-AI | | Post-AI | |
|---|---|---|---|
| 4.4 | miles | 463 | feet |
| 7.0 | km | 141 | meters |

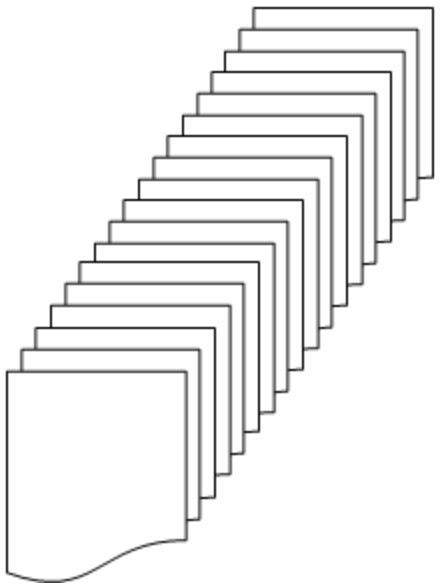
***All values are approximate***




*AI-Assisted Analysis of Ultrasonic Inspections* (3002023718)

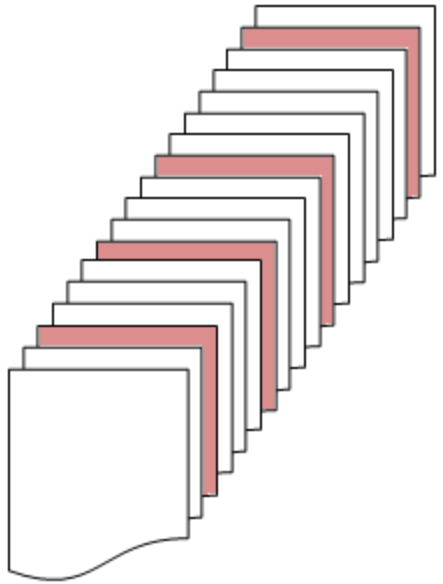
# Combine AI with Inspector

- Procedure would be updated to include a defined process for the AI evaluation
- AI algorithms for qualification would be developed and provided to the qualification body: model is **frozen** at this stage



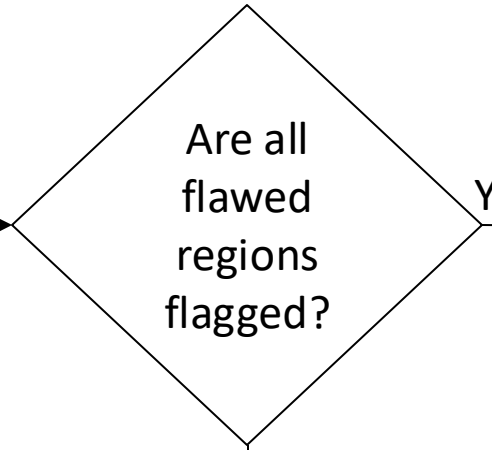
*Pre-recorded data*

*AI model is applied*



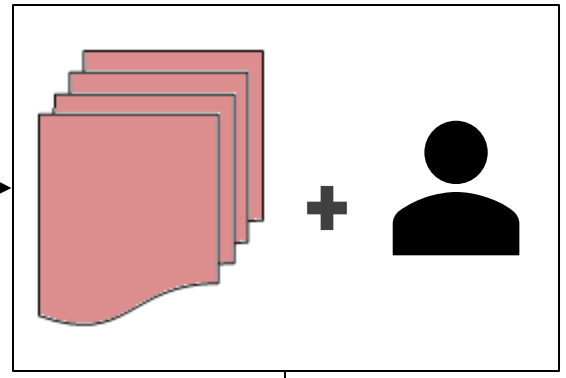
*AI flags areas for review*

*Qualification Body Determination*




Fail

*Only flagged regions are presented to candidate, who reviews and makes calls*



Pass/Fail

*Based on typical detection + false call criteria*

All observations point to AI being qualifiable when paired with human analyst

# Summary

- AI / ML can improve the efficiency and efficacy of inspections when paired with a human inspector

- Accuracy, especially no false negatives or misses, is paramount

**Ready for deployment!**

# Upcoming Plant Modernization & DX Related Meetings

**Plant Modernization Committee Update Meeting**, November TBD, 2024

- Virtual
- Registration will be available soon

**AI & DX In Electric Power Summit**, January 7-9, 2025

- Details on next slide
- Register here

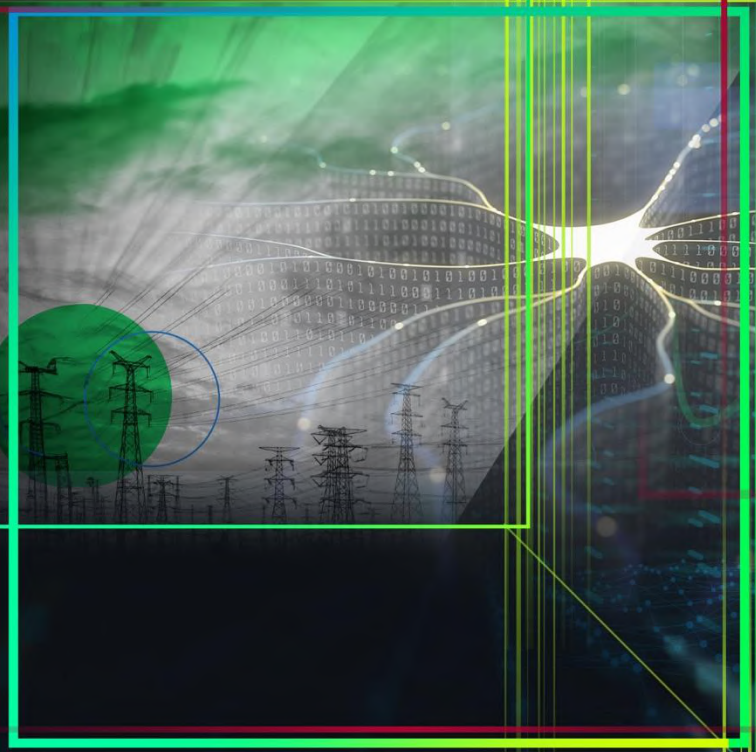
**PMC/3DM Meeting as part of Nuclear Power Council**, Feb 2, 2025

- In-Person
- Orlando, Florida
- Registration will be available soon

EPRI

Stanford | ENERGY
Bits & Watts Initiative

# ARTIFICIAL INTELLIGENCE (AI) AND DIGITAL TRANSFORMATION (DX) Electric Power Summit



January 7-9, 2025
Crowne Plaza Palo Alto

# Applications to Nuclear Security

## Discussion


www.epri.com

# APPENDIX
# Additional Content

# Data Science Training

## Computer-Based Training (CBT)

| Course | Product/Availability |
|---|---|
| GridEd Short Course: AI in the Electric Industry: an Introduction for Managers | 3002022336 |
| GridEd Short Course: Load Forecasting | 3002022501 |
| GridEd Short Course: Labeling Images for Machine Vision Applications | 3002024816 |
| Enabling Digital Transformation: Application Integration Leveraging Common Information Model (CIM) | 3002026121 |

**And more to come…**

## Live/Online Training

## https://grided.epri.com/courses.html

| Course | Product/Availability |
|---|---|
| GridEd Short Course: Machine Vision in the Electric Industry: a Hands-On Training for Practitioners | 3002021065 |
| GridEd Short Course: Machine Learning and Big Data Analytics in Smart Grid | 3002019825 |
| Machine Learning Applications for Time Series Data in Power Systems – a Hands-On Training for Practitioners | 3002026234 |
| Machine Learning Applications for Smart Meter Data – a Hands-On Training for Practitioners | Target first offering: 2024 |
| Machine Learning Applications for PMU Data – a Hands-On Training for Practitioners | Target first offering: 2024 |

# Reference & Upcoming EPRI Products

## ChatGPT and the Power Sector: What's Hype? What's Possible?

- White paper with initial considerations and potential use cases
- Published April 2023, 3002026468

## Large Language Models at EPRI

- EPRI Insight with revised considerations and proof-of-concept use cases
- Expected to be published Q4 2024

## Guidance on LLMs for Data Science Practitioners

- Slide deck for LLM learnings on models and metrics
- Expected to be published Q2 2025

## Guidance on LLMs for IT/OT Practitioners

- White paper for LLM learnings on deployment considerations
- Expected to be published Q4 2024 with DXRI Supplemental

## Storage Technology Enhanced Large Language Application (STELLA)

- Energy Storage chatbot
- Beta testing for proof-of-concept completed

# Current R&D LLM Internal PoC Projects

- Energy Storage Chatbot (STELLA) – Internal tool completed and tested with SMEs
- Nuclear Decommissioning Knowledge Management – Kickoff meeting and legal/infosec approval in progress
- Operating Experience (OE) Consolidation and Part 21 – Data curation, legal/infosec approval in progress
- Hydrogen Safety Bot – Internal tool completed and tested with SMEs
- Voice-to-Form for Field Workers – Pending kickoff, initial vendor meetings
- NMAC Troubleshooting Tool – Kickoff meeting and legal/infosec approval in progress
- Relief Request – In progress
- PRA Storyteller – Pending kickoff

# AI in NDE - References

- Machine Learning to Improve ID Pit Sizing in Balance-of-Plant (BOP) Heat Exchanger Tubes: Target—Type 304 Stainless Steel Tubing (3002021048)
- A Case Study to Improve Flaw Depth Sizing in Heat Exchanger Tubes with Machine Learning Techniques (3002021042)
- AI Tool Developed by EPRI Significantly Cuts Analysis Time in U.S. Nuclear Plant Field Trial: Data Analysis Takes Four Hours Compared to Four Days Without Artificial Intelligence (3002025510)
- AI-Assisted Analysis of Ultrasonic Inspections (3002023718)
- Using AI to Streamline Data Analysis

  (https://www.youtube.com/watch?v=4KNkE-Wfg9M)
- PRE-SW: Assisted Analysis of Remote Visual Examination (AARVE) v0.1 Beta (3002026528)
- PRE-SW: Assisted Data Analysis – Eddy Current: Bobbin (ADA-EC) v0.2 Beta (3002026529)

**EPRI is VERY active in this area**