Panel Session 8:
Technical Language Processing (TLP) and Large Language Models (LLMs)

November 13, 2024

Nashville, TN.

16th Annual Conference of the Prognostics and Health

Management Society

Technical Language Processing (TLP) and Large Language Models (LLMs)

Technical Language Processing (TLP) is a set of tools, techniques, and methodologies designed to adapt Natural Language Processing (NLP) tools to expert-driven, text-based data within engineering and other specialized domains. Adapting Large Language Models (LLMs), which are trained on vast amounts of text for processing, to specialized applications is an emerging area of innovation. As the intersection of industry, artificial intelligence, and human-centered intelligent automation accelerates, mastering communications and language processing is essential for maintaining a competitive edge. This panel will explore key aspects of TLP and LLMs, including: state-of-the-practice and best practices, risk awareness, needs and gaps, sector-specific applications, metrics and measurement, and the evolving role of humans as observers in optimizing the interaction between operators and monitoring systems.

- **Moderator:**
- Sarah Lukens, LMI
-
Panelists:
- Michael Sharp, NIST
- Neil Eklund, Oak Grove Analytics
- Hao Huang, GE Vernova

How this panel is going to shake out…

- Context & Motivation – TLP background
- Introduce the panelists
- Canned questions: topics from the past few days
- Audience questions – what else?
- Get out of here at 4:30 to get on the first bus!

Bayesian Belief Bourbon XXXX

What came out of the LLM panel last year?

"Gen AI: Where is the win?" Panel Fall 2023

Generative **AI** in PHM applications

Where is the **W**in?

Panel Session #4
13th Annual PHM Society Conference
Moderator: Asma Ali
Panelists: Kai Goebel, Karl M. Reichard, Olympia Brikis, Sarah Lukens, Mark Roboff

https://phm2023.phmsociety.org/wp-content/uploads/sites/13/2023/11/Panel04-GenAI-Panel-at-PHM-10-31-2023.pdf

"Let's not throw out our best practices in engineering just because we have a shiny new toy"

– Karl Reichard
Associate Research Professor
The Pennsylvania State University
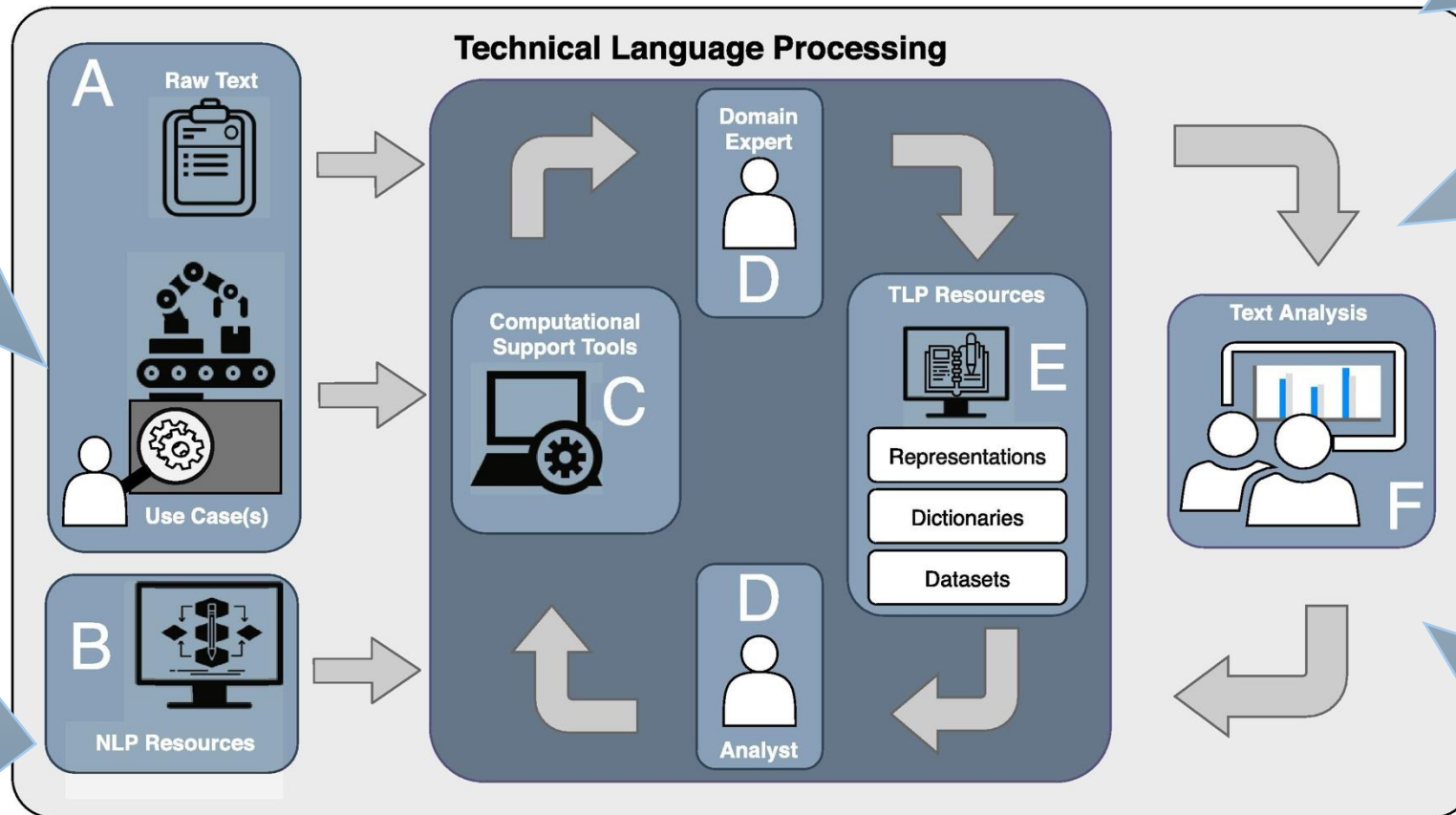(Generative AI Panel @PHM 2023)

Approach it like engineers:

- Where are the *strengths*?
- What are the *risks*?
- What am I *assuming* when I use these tools?

What is TLP?

A. Use case is explicitly considered as an input

B. Use NLP resources where it makes sense

C. Use computer tools to alleviate burden on domain expert

D. Collaboration between analyst and domain expert

E. TLP resources such as **standards** or **technical manuals**

Technical Language Processing

A — Raw Text — Use Case(s)

B — NLP Resources

C — Computational Support Tools

Domain Expert — D

TLP Resources — E — Representations, Dictionaries, Datasets

Text Analysis — F

Analyst — D

Technical language processing: Unlocking maintenance knowledge; Brundage, Sexton, Hodkiewicz, Dima & Lukens (2021). *Manufacturing Letters*, *27*, 42-46.

TLP Content and Context at PHM

Sept 24-25 NIST TLP Community of Interest Virtual Workshop. Organized and led by **Michael Sharp**, NIST.

Nov 11 – Keynote: Dr. Shauna Sweet "Action, Not Reaction"

Nov 13 – Today! At PHM:

Time	What	Title	Speakers
9:00-10:30 AM	Tutorial	LLMs & Multimodal AI for PHM;	**Neil Eklund**
10:45-12:15 PM	Paper session 6A	• "Physics & Data Collaborative Root Cause Analysis: Integrating Pretrained LLMs and data-driven AI for Trustworthy Asset Health Management" • Adapting LLMs to practical copilots in PHM: Requirements identification and performance evaluation • Towards a Fault Management Analysis Tool for Model Centric Systems Engineering	• **Hao Huang** • Sarah Lukens • Ksenia Kolcio
	Paper session 6B	• Diagnostics LLaVa: A visual LLM for domain specific diagnostics of equipment	• Aman Kumar
3:30-5:00 PM	Panel 8 – TLP and LLM's	Us! Here and now!	

Technical Language Processing Community of Interest

UPDATES

NIST to Host Technical Language Processing Virtual Workshop on 24-25 September 2024

July 1, 2024

NIST to Host Virtual Workshop on Technical Language Processing on 24-25 September 2024

Credit: CTL

TECHNICAL CONTACT

Michael Sharp
michael.sharp@nist.gov
(301) 975-0476

ORGANIZATIONS

Communications Technology Laboratory
Smart Connected Systems Division
Smart Connected Manufacturing Systems Group

NIST will hold a virtual workshop on Technical Language Processing on 24-25 September 2024. Online registration is $33.00 USD and is available via the workshop webpage. The two-day virtual event brings together industry, government, and academic stakeholders through the NIST Technical Language Processing Community of Interest (TLP COI), and it offers a valuable platform to connect, share insights, and explore the evolving landscape of TLP.

SIGN UP FOR UPDATES FROM NIST

Enter Email Address

TLP COI meeting and workshop covered:
- Risk Awareness.
- Accessibility
- Longevity: Additional Key Introductory TLP
- Current TLP Landscape:
- Data Sharing and Resources: Needs, Gaps, and Applications
- Metrics and Measurement.
- Exploring LLMs
- Humans as Observers

https://www.nist.gov/el/technical-language-processing-community-interest

Interested in more? email tlp-coi@nist.gov

About this panel

Sarah Lukens, Ph.D.
Moderator

TLP Community of interest

Michael Sharp, Ph.D.
NIST

LLM tutorial

Neil Eklund, Ph.D.
Oak Grove Analytics

Technical paper session – industry research areas

Hao Huang, Ph.D.
GE Vernova

Michael Sharp

- **Dr. Michael E. Sharp** is a Reliability Engineer at the National Institute of Standards and Technology (NIST) located in Gaithersburg, MD.
- He received a B.S (2007), M.S. (2009), and Ph.D. (2012) in Nuclear Engineering from the University of Tennessee, Knoxville, TN, USA
- Research interests:
    - signal analytics,
    - machine learning,
    - artificial intelligence,
    - optimization
    - NLP.
- Michael has worked on a wide array of projects including image processing for elemental material recognition, navel reliability monitoring, and manufacturing robotics diagnostic monitoring.
- He currently works as a project lead on Industrial Artificial Intelligence Management and Metrology (IAIMM) at NIST in the Communications Technology Laboratory.

Smart Connected Manufacturing Systems Group
michael.sharp@nist.gov
(301) 975-0476

Neil Eklund

- Dr. Neil Eklund, FPHMS, is a leading expert in Asset Health Management and a seasoned technologist with over 25 years of experience in data science, industrial artificial intelligence, and machine learning.

- Has led high-impact projects across industries including aerospace, energy, healthcare, and oil & gas, resulting in 16 patents and over 70 technical publications.

- Dr. Eklund is one of the co-founders the PHM Society and serves on its board of directors.

- Roles at General Electric Research, Xerox PARC, and Schlumberger, where he was the Chief Data Scientist. At Schlumberger, he led the development of the first successful deployed Internet of Things (IoT) application in the oil industry, generating over $20 million in its first three months of operation.

- Collaboration with organizations like DARPA, NASA, the DoD, Lockheed Martin, ExxonMobil, Ford Motor Company, and Boeing highlights his impact on both commercial and government sectors.

- Taught graduate-level courses in machine learning and classes through the PHM Society. He is also a former contributor to the International Standards Organization (ISO) in the area of diagnostics and prognostics for complex machinery.

Who is this guy?

2024 Tutorial: "LLMs and Multimodal AI for PHM: The Future of Maintenance Intelligence"

https://phm2024.phmsociety.org/tutorials/

Hao Huang

Senior AI Scientist

GE Vernova Advanced Research

hao.huang1@ge.com

Dr. Hao Huang is a Senior AI Scientist at the Advanced Research Center of GE Vernova. He earned his Ph.D. and Master's degrees from Stony Brook University and holds a second Master's degree as well as a Bachelor's degree from Sun Yat-sen University, China.

Research Interests:
- LLM-Assisted Anomaly Detection and Root Cause Analysis
- Large-Scale and Streaming Industrial Data Analysis
- Green AI
- AI-Based Design Optimization

Recent Publication:
- Physics and Data Collaborative Root Cause Analysis: Integrating Pretrained Large Language Models and Data-Driven AI for Trustworthy Asset Health Management, PHM 2024
- Deep Root Cause Analysis: Unveiling Anomalies and Enhancing Fault Detection in Industrial Time Series, IJCNN 2024
- Energy Efficient Streaming Time Series Classification with Attentive Power Iteration, AAAI 2024

Physics and Data Collaborative Root Cause Analysis: Integrating Pretrained Large Language Models and Data-Driven AI for Trustworthy Asset Health Management

Hao Huang, Tapan Shah, John Karigiannis, Scott Evans

GE Vernova Advanced Research, Niskayuna, NY, USA

hao.huang1@ge.com, tapan.shah@ge.com, John.Karigiannis@ge.com, evans@ge.com

http://www.papers.phmsociety.org/index.php/phmconf/article/view/3881
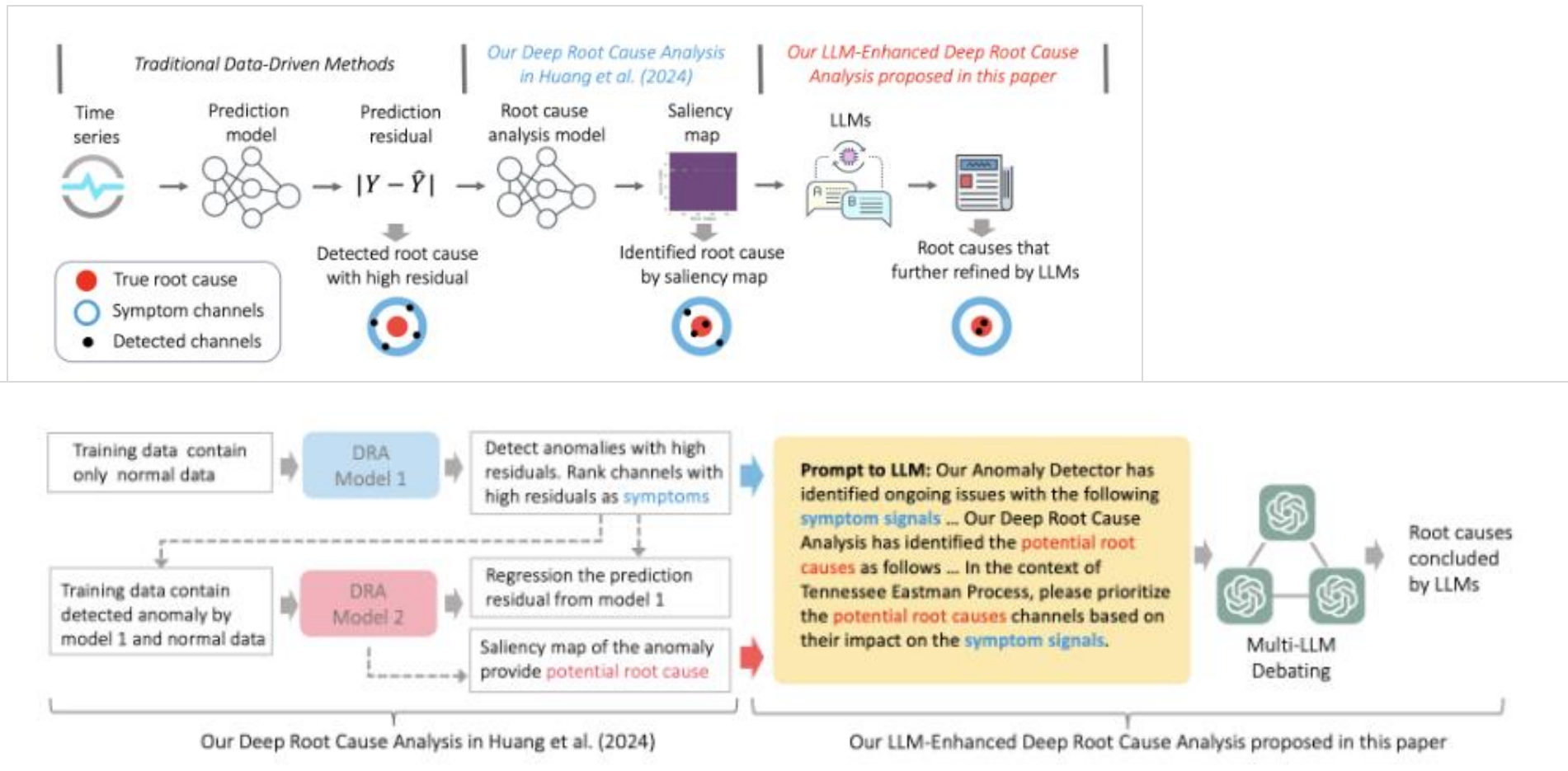
Discussion from TLP COI Meeting

Let's discuss our commonalities,
across....

Problems

Solutions

Practices

Still from So I Married an Axe Murderer (1993), directed by Thomas Schlamme, TriStar Pictures.

From Michael Sharp

Topics of interest

Canned Questions...

From the keynote Monday Nov 11...

Dr. Shauna Sweet had a message:

Quantitative evidence that pre-trained LLM's severely overfit: what is the implications on generalizability? On extrapolation?

"It's sometimes the unusual things that matter the most"

Question: From an engineering perspective, do you have any experience, recommendations or best practices for how to explore, poke at or harness unusual behavior from language models?
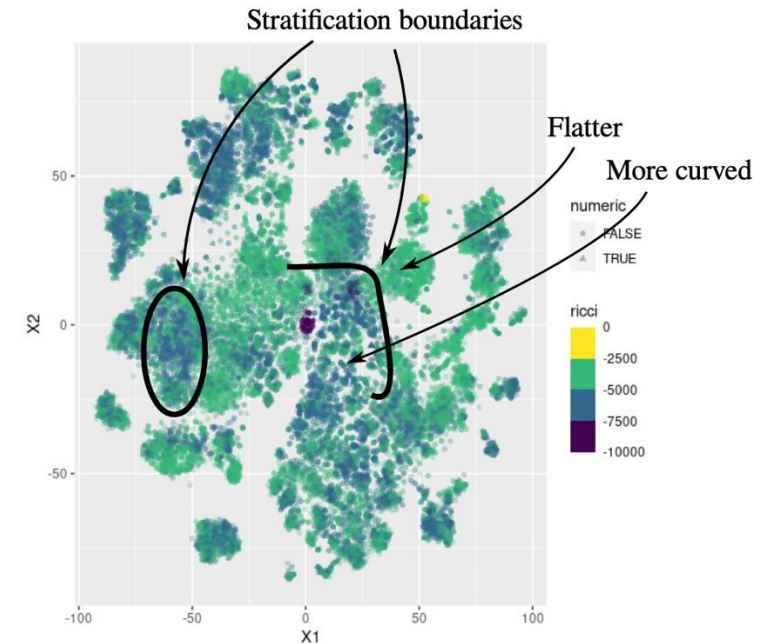

Figure 15: Estimated Ricci scalar curvature for MISTRAL7B, plotted using tSNE coordinates. Curvature is indicated by the depth of color: darker points have higher (more negative) curvature.

https://arxiv.org/pdf/2410.08993

Hallucinations

ated Image: "A Ball Bearing Bearing Down the Bering Strait"

From Michael Sharp

In the tutorial, Neil made the comment about how it feels like LLMs **hallucinate** more when you know more about a topic

"I was fixing the wheels on the generator…"

In the keynote, Dr. Sweet observed that numbers are on the same manifold in the embedding space, and the model does not understand the individuality of tokens across a stratified space…

Does that explain many of the hallucinations we observe? What else could contribute?

Fine tuning and free lunch

- Fine-tuning is a great way to get improved performance for your applications! However, it comes at a price.
  - Both Hao & Neil fine-tuned in what they showed today.

- What are your thoughts on the pros and cons of fine tuning?

- When does it make sense to fine tune and when should we try to engineer alternate solutions?

Pushing the computational limit

In the tutorial, Neil made a traffic theory analogy: "If you build more highways, people will fill them. Transformers are kind of like that too – you build a more efficient algorithms, people will throw more at it".

Do you have any thoughts on computational best practices or future trends?
What about computational requirements to consider?

TLP COI Meeting – Key themes
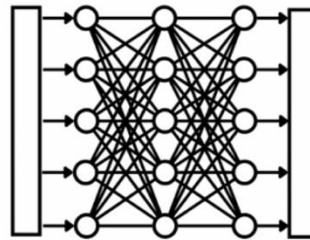
Want of standardization and commonality of tools

1. Traditional standards models don't work with new technologies
2. Guides to HOW to share the data

AI model Framing for business leaders

Understanding "where does it fail?" "Where does it not work?"
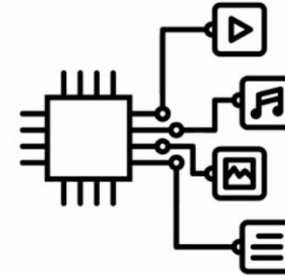
Challenges with LLMs and how to overcome

A simple framing for AI models

**Narrow AI** performs **specific tasks** with high proficiency.

Good for repetitive processes and patterns with narrow scope for which data is available
- Classification
- Translation
- Digital assistants
- Email filtering

Generative AI large language models perform **wide range** of tasks.

Good for generating responses by considering hundreds of potential completions for a given **prompt** and choosing the most likely one based on its training.
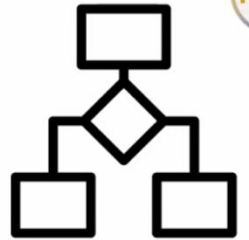
Logic-based AI is for **reasoning**.

Good when rigor is needed to draw conclusions from facts and rules, apply logical inference and for semantic interoperability.

ICONS from Noun Project: 6348230, 6418841, 4814710

11

From Melinda Hodkiewicz from UWA

Whova App

- What fallback mechanisms or human-in-the-loop strategies are in place when LLM recommendations are incorrect or incomplete?

- What is the anticipated impact on maintenance personnel training and skill development when providing a co-pilot that may create a crunch for new personnel?
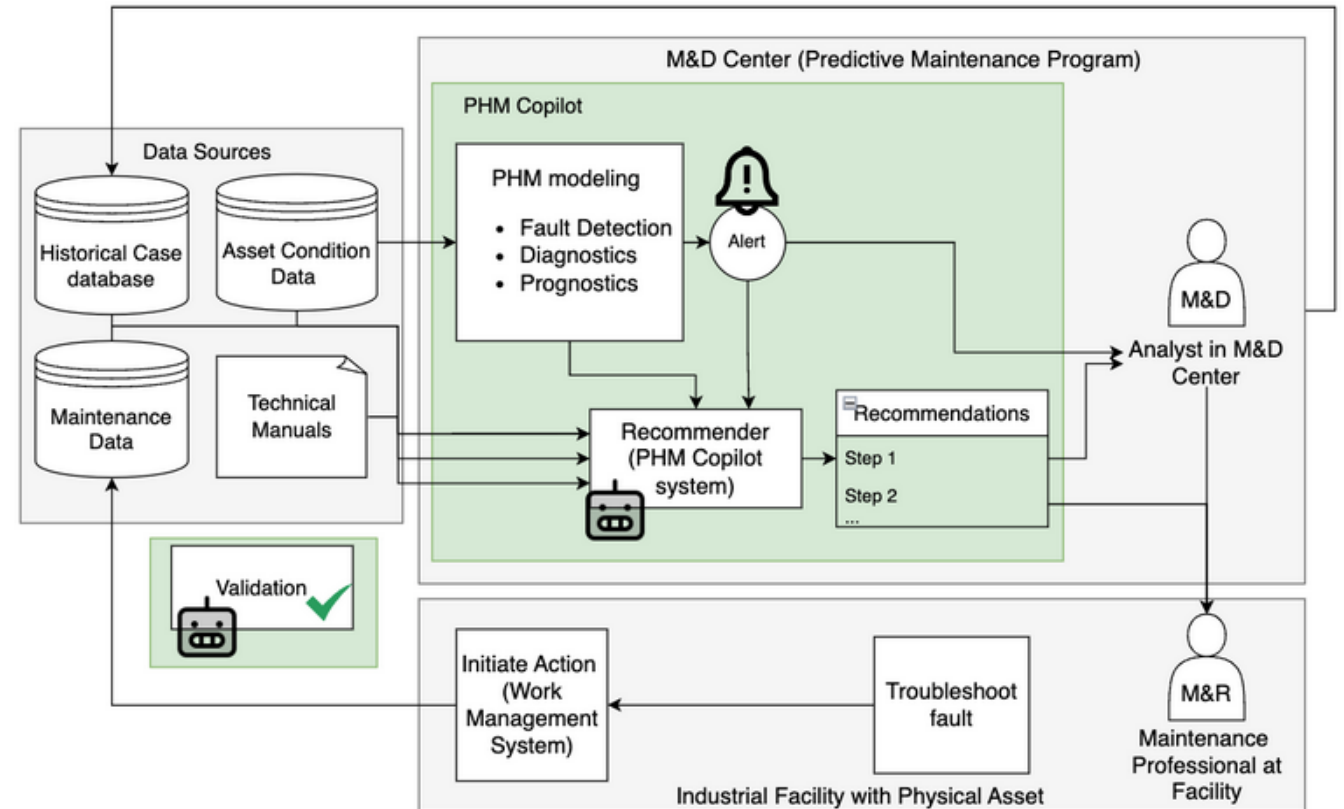
What are the useful areas/applications of LLMs or Multi-modal language models for PHM?

Thoughts on Typos

- In the tutorial this morning, Neil miss-typed "re entry" when the question was "recently calibrated". It didn't matter. The LLM system "knew" he meant recently.

- This behavior is a strength of the in-context information captured in the embeddings.

- Thoughts?