Pacific Northwest
NATIONAL LABORATORY

PNNL-38003

Connecting Minds: AI Use Cases to Bridge Power Systems and Large Language Models for Practical Applications

May 2025

Yousu Chen Alex Anderson

U.S. DEPARTMENT of ENERGY

Prepared for the U.S. Department of Energy
Under contract DE-AC05-76RL01830

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from
the Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062
www.osti.gov
ph: (865) 576-8401
fox: (865) 576-5728
email: reports@osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
or (703) 605-6000
email: info@ntis.gov
Online ordering: http://www.ntis.gov

PNNL-38003

Connecting Minds: AI Use Cases to Bridge Power Systems and Large Language Models for Practical Applications

May 2025

Yousu Chen Alex Anderson

Prepared for
the U.S. Department of Energy
Under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99352

Executive Summary

Recent advances in artificial intelligence (AI) and development of large language models (LLMs) present the opportunity to develop a new generation of power systems applications. In contrast with early power system AI applications based on structured numerical data, LLMs offer unique capabilities to perform logical reasoning using text documents, unstructured data, and application programming interface (API) calls to computational software. This report seeks to bridge the knowledge gap between power systems engineers and LLM developers through a crosscutting explanation of use cases, characteristics, requirements, practical considerations from the perspectives of both LLM capabilities and industry needs.

In translating the capabilities of LLMs to enterprise problems in the power systems domain, this report aims to provide a pathway for development of the next generation of LLM-based applications with a direct path to implementation in utility control room environments by:

- Providing an overview of LLM concepts for power system engineers.
- Explaining key power system domain characteristics for LLM specialists.
- Identifying high-impact use cases where LLMs can bridge the gap between AI capabilities and power system expertise.
- Discussing practical considerations, challenges, and future research for LLM applications in power systems.

By connecting these two domains, this report serves as a practical guide for users seeking to harness LLMs to enhance grid reliability, improve affordability, and accelerate decision-making across power system operations, maintenance, planning, and analysis. It should be noted that this work is not intended to serve as a review paper of existing work and advancements in LLM, but rather serve as forward-looking overview of novel, unsolved problems that have not been examined but are well-suited for research, development, and demonstration of new LLM-based applications.

After introducing the architecture of LLMs and unique challenges of the power systems domain, this paper proposes twenty representative LLM applications grouped into categories of 1) power system operations, 2) asset management, 3) system planning and analytics, and 4) energy management and protection systems. Specific focus is given to applications that can be realistically deployed by electric utilities. Five use cases are presented within each category with descriptions of the motivation, objectives, approaches, example inputs / outputs, and benefits of each use case.

Acknowledgments

The authors wish to thank several industry stakeholders who participated in discovery interviews with the authors that helped inform the use cases presented in this work, including Dan Kopin of Vermont Electric Power Company (VELCO), Yang Feng of Siemens, Eric Meier and John Mosely of the Electric Reliability Council of Texas (ERCOT), Todd Viegut of AspenTech, Faiz Sulaiman of Microsoft, and Eamonn McCormick of Utilicast.

Contents

1.0 Introduction

The past decade has witnessed unprecedented advancements in artificial intelligence and machine learning (AI/ML), revolutionizing industries across the globe. Among these breakthroughs, Large Language Models (LLMs) have emerged as particularly transformative, with systems like OpenAI's GPT-4 [1], Anthropic's Claude, Google's Gemini, and Meta's Llama [2] demonstrating remarkable capabilities in understanding, generating, and reasoning with human language. These models have enabled widely-adopted applications such as ChatGPT, which reached 100 million users within two months of its launch, representing one of the fastest technology adoptions in history [3]. Contemporary LLMs now demonstrate sophisticated reasoning abilities, code generation capabilities, and multimodal understanding that extends beyond text to images and structured data. Trained on vast corpora of text data, they have evolved from simple statistical approaches to sophisticated neural architectures capable of complex tasks including text generation, translation, summarization, and knowledge extraction [4].

The ongoing evolution of these technologies has expanded their application domains from general consumer services to specialized technical fields including medicine [5], law [6], finance [7], and quantum chemistry [8]. This expansion is particularly significant for industries with extensive technical documentation and specialized knowledge requirements, where the ability to process unstructured information at scale offers substantial operational advantages.

The electric power industry represents one such domain, generating massive volumes of unstructured textual data, including operational logs, alarm messages, field reports, equipment manuals, industry standards, planning documents, and regulatory filings. This information contains valuable insights but has traditionally been difficult to analyze systematically due to its unstructured nature, specialized terminology, and sheer volume.

Over the past four decades, numerous attempts have been made to build AI/ML analysis and recommender tools, starting with Wollenberg & Sakaguchi's foundational 1987 paper [9] and hundreds of papers since, which are surveyed by [10]–[16]. However, previous generations of AI tools have struggled with unstructured nature of data, daily changes in grid topology, and procedure-based nature of operations. As a result, these tools have seen limited deployment by electric utilities beyond anomaly detection and solar/load forecasting. As power systems become increasingly complex with advanced monitoring systems, grid-edge resources, and evolving regulatory frameworks, the challenge of efficiently processing this unstructured information continues to grow.

LLMs provide powerful new ways to extract insights from unstructured power systems data. Unlike traditional rule-based or statistical methods, LLMs can interpret context, recognize patterns, and generate insights from text with unprecedented accuracy, as demonstrated by AI/ML applications in power systems summarized by [17]. As a result, there exists significant potential to develop LLM-based tools that assist with a range of planning, operations, asset management tasks, which are examined from the perspective of twenty novel use cases summarized in Table 1.

Specific focus is given to applications that can be realistically deployed in utility control room environments given the restrictions imposed by North American Electric Reliability Corporation

(NERC) Critical Infrastructure Protection (CIP) standards regarding 15-minute impacts on the bulk electric system (BES) [18]. Although it may be possible to develop LLMs for real-time energy management system (EMS) applications as suggested by [19], the latest guidance issued by NERC [20] does not permit deployment of such applications due to cybersecurity and grid reliability concerns.

The remainder of this paper is organized as follows: Section II presents the fundamental concepts of LLMs and their relevance to power system applications, addressing both power engineers unfamiliar with LLMs and LLM specialists new to power systems. Section III explores applications in control center operations and event analysis. Section IV examines asset management applications.

Table 1: Candidate LLM Use Cases in Power Systems

Application Categories	**Use Cases**
Control Center Operations and Event Analysis	• Intelligent Alarm Processing and Root Cause Analysis • Operations Log Analysis and Knowledge Management • Event Report Generation and Compliance Documentation • Switching Order Creation and Validation • Shift Turnover Summarization
Asset Management	• Intelligent Analysis of Maintenance Records • Equipment Manual and Technical Document Intelligence • Automated Processing of Field Work Orders • Registration and Interconnection of Customer-owned Resources • Validation of Asset Data Records
System Planning and Analytics	• Accelerated Generation Interconnection Studies • Outage Report Mining and Reliability Analysis • Enhanced Integrated Resource Planning • Technical Standards and Compliance Navigation • Assisted System Planning Decisions
Energy Management and Protection Systems	• SCADA and Control System Log Analysis • Energy Management System (EMS) Event Analytics • Protection Setting Verification and Mis-coordination Prevention • Day-ahead Outage Coordination • DER-informed Load Shedding Scheme Design

Section V covers system planning and analytics. Section VI describes energy management and protection systems applications. Section VII details practical considerations for LLM deployment. Finally, Section VIII concludes the paper with a discussion of challenges and research directions.

2.0 Fundamentals of Large Language Models for Power System Applications

This section bridges the knowledge gap between power system engineering and machine learning domains. We provide key concepts from both fields to establish a common foundation for understanding the applications discussed later.

2.1 LLM Architecture and Capabilities: A Power Engineer's Guide

2.1.1 What Makes LLMs Different from Traditional Software

Traditional power system software follows deterministic rules and equations. Whether calculating load flow or simulating protection operations, these programs produce consistent, predictable results given the same inputs. LLMs work differently. They are neural networks that learn patterns from data rather than following explicit rules. They are trained by processing massive text collections and learning statistical relationships between words and concepts [1]. They recognize patterns such as "this combination of words often follows that combination" based on the training data. This fundamentally different approach allows them to handle ambiguity and variation in ways traditional software cannot. This probabilistic approach explains both their flexibility with unstructured data and their occasional unpredictability.

LLMs also differ from both traditional software and prior AI/ML in their ability to perform *logical reasoning*, which refers to the ability to generate coherent sequences of text to solve problems requiring structured thought. Unlike conventional deep learning and reinforcement learning methods based on implicit pattern recognition and input-output mappings, LLMs are able to use explicit step-by-step decomposition of problems through chain-of-thought (CoT) [21] and tree-of-thought (ToT) [22] prompting. This enables LLMs to exhibit human-like problem solving approaches applying deductive, inductive, and abductive logic with task-specific retraining. LLMs can also be integrated with symbolic solvers and external application programming interfaces (APIs) for deterministic inference through reasoning+acting (ReAct) implementations [23].

These two differences enable LLMs to tackle procedures-based tasks (e.g. day-ahead planning) that involve a combination of document text (e.g. a planned outage) and user requests to develop a sequential set of subtasks (e.g. processing the document, editing a network model file, making an API call to power flow simulation tool, and summarizing the results for rapid decision-making).

2.1.2 Transformer Architecture: The Engine Behind Modern LLMs

The breakthrough enabling modern LLMs came in 2017 with the introduction of the *Transformer* architecture [24]. This design processes text differently than earlier approaches:

- **Parallel Processing:** While traditional AI/ML data processing of power systems often process events in sequence, Transformers analyze all words in a text block simultaneously.
- **Attention Mechanism:** The core of the Transformer is its attention mechanism, which weighs the relationships between different words in context, allowing the model to focus on the most relevant keywords for better understanding.
- **Contextual Understanding:** Earlier models treated words as isolated units with fixed meanings. Transformers understand that meaning depends on context, which is critical for technical domains such as power engineering where terms like "trip," "fault," or "bus" have specific meanings different from everyday usage.
- **Multimodal Integration:** The textual understanding capabilities of LLMs can be integrated with a large vision model (such as GPT-4 [1]) to form multimodal versions. These models are able to interpret images from documents, flowcharts, substation one-line diagrams, and photographs of physical equipment, thereby enabling comprehensive interpretation and reasoning across various domains.

2.1.3 How LLMs Learn: From Raw Text to Specialized Applications

LLM development typically follows several stages that help explain their capabilities and limitations:

- **Pre-training:** The initial training phase exposes the model to diverse text spanning books, articles, reports, websites, and software codes. This develops general language capabilities but with limited specialized knowledge.
- **Domain Adaptation:** Models can be further trained on technical literature relevant to specific fields. For power applications, this might include IEEE papers, engineering textbooks, and technical standards, helping the model understand domain terminology.
- **In-Context Learning:** LLMs are able to learn and adapt to new information provided in the context window without additional training through few-shot prompting techniques [4] where the LLM is given several input/output examples from which it is able to infer patterns with high accuracy.
- **Retrieval Augmented Generation (RAG):** The model can be supplemented with a vector database of documents, API calls, and task-specific data that is used for retrieval of accurate information during text generation. Graph-RAG approaches [25] use an additional graph database to handle complex domain ontology and network topology concepts.
- **Fine-tuning:** The model may receive focused training on specific task types (i.e. summarization, classification, or question-answering) with human-labeled examples of desired outputs. Several thousand question-answer pairs are typically required for effective fine-tuning.

- **Deployment Customization:** After deployment, models can be enhanced with external knowledge (e.g. RAG or API interfaces to numerical solvers) or configured with specialized prompts that guide their behavior for particular applications.

2.1.4 Key Practical Considerations

Several practical aspects of LLMs deserve particular attention from power engineers:

- **Context Window:** This refers to how much text the model can consider at once. Longer contexts allow processing entire standards documents, or event logs in a single analysis but require more computational resources.
- **Accuracy:** LLMs can occasionally generate plausible-sounding but incorrect information, particularly for specialized technical details outside their training data. This makes validation crucial for critical applications.
- **Computational Requirements:** Running advanced LLMs requires significant computing resources, affecting deployment options (cloud vs. on-premises) and response times for interactive applications.
- **Integration Approaches:** Most power systems LLM applications combine language models with other technologies such as document libraries, structured databases, and traditional computational software.

2.2 Power System Data Characteristics: Insights for LLM Practitioners

LLM specialists approaching the power domain should understand several unique characteristics that distinguish it from more common LLM application areas like customer service or content creation.

2.2.1 Safety-Critical Infrastructure with Unique Terminology

Electric power systems form critical infrastructure with immediate physical and societal consequences from software and hardware failures. The unique challenges of this domain are reflected with documentation structures and formats:

- **Precision Requirements:** Power documentation prioritizes precise technical accuracy over readability or style. A single incorrect parameter in a protection setting or switching procedure could cause a blackout.

- **Specialized Vocabulary:** The industry uses highly technical terminology that even educated non-specialists would struggle to interpret. Terms such as "sympathetic tripping" or "coordination time interval" have precise meanings within the domain.
- **Abbreviations and Codes:** Documentation contains dense abbreviations that may be inconsistent across organizations or regions (e.g., DS/89/ISOL all referring to disconnect switches, but with subtle distinctions in usage).
- **Mixed Numerical and Textual Information:** Critical information often combines text with measurements, equations, and parameters. A transmission fault record might state, "A phase-to-phase fault occurred at 73% of the line length, with a fault current of 6.8 kA, clearing in 85 milliseconds".

2.2.2 Document Types and Their Characteristics

Power utilities manage diverse documents with different structures and purposes:

- **Event Logs and Alarm Records:** Terse, timestamp-oriented records of system changes, often in cryptic formats with minimal context. Example of a breaker trip event: "12:15:08.672 SUB_14 CB_52 OPEN PHASE_A I=5321A V=121kV"
- **Standards and Procedures:** Prescriptive documents with numbered requirements, explicit definitions, and precise obligations. Often structured hierarchically with cross-references to other documents.
- **Equipment Documentation:** Technical specifications, test results, and maintenance logs for long-lasting equipment, sometimes decades old. These records may include data from manufacturers that no longer exist.
- **Operational Directives:** Clear, step-by-step instructions for field crews or operators, including mandatory safety checks and confirmations. These directives follow a structured format with numbered steps and required verbal acknowledgments.
- **Engineering Analyses:** Technical assessments combining narrative explanations with calculations, simulation results, and recommendations. May include complex diagrams, or other visual data.

2.2.3 Multi-Timeframe, Multi-System Information

Power system operations span dramatically different timeframes, with information dispersed across multiple systems:

- **Protection Operations:** Millisecond-level events recorded in relay logs
- **Control Actions:** Second-to-minute timeframe operations documented in field measurement and operator logs
- **Operational Planning:** Hour-to-day decisions in operating procedures and scheduling systems
- **Asset Management:** Month-to-year cycles in maintenance records and equipment databases
- **Long-term System Planning:** Multi-year horizons in planning studies and regulatory filings

Understanding system events often requires correlating information across these different timeframes and systems a natural application for LLMs' ability to identify relationships in unstructured text.

2.3 Key LLM Capabilities for Power System Applications

Several LLM capabilities show particular promise for addressing power system information challenges:

2.3.1 Unlocking Insights from Operational Documents

LLMs can transform how utilities gain insights from operational documentation:

- **Contextual Information Retrieval:** Beyond basic keyword search, LLMs understand query intent, such as *"What caused the voltage collapse last August?"*, and extract root causes from reports, logs, and event records.
- **Event Reconstruction:** By analyzing alarm sequences, operator logs, and protection records, LLMs can automatically piece together event timelines that would otherwise take engineers hours or days to compile.
- **Pattern Detection:** LLMs can uncover recurring trends in equipment failures, false alarms, or near-miss events across thousands of reports, helping utilities improve reliability and prevent future incidents.

2.3.2 Automating Standards Interpretation and Compliance

Managing regulatory compliance requires analyzing complex regulatory documents to ensure adherence to industry standards:

- **Requirement Extraction:** LLMs can quickly identify specific obligations in standards documents, distinguishing between mandatory requirements and recommendations.
- **Gap Analysis:** By automatically comparing internal procedures against regulatory requirements, LLMs can flag potential compliance gaps or conflicts between documents.
- **Interpretation Assistance:** When engineers have questions about applying standards to specific situations, LLMs can retrieve specific clauses within standards, provide interpretations, and reference relevant precedents.

2.3.3 Knowledge Management and Transfer

With experienced personnel retiring, LLMs can help preserve institutional knowledge:

- **Technical Question Answering:** LLMs can answer questions about equipment, procedures, and historical events by drawing on enterprise document collections.
- **Procedure Generation:** By learning from existing procedures and standards, LLMs can help draft new procedures or update existing ones for equipment changes or new regulations.
- **Training and Documentation Support:** LLMs can generate training materials, explanations, and documentation that captures expert knowledge in accessible formats.

These foundational capabilities enable practical, high-value applications across the power system. The following sections explore how LLMs tackle specific challenges in operations, maintenance, and planning. Through real-world examples, this discussion aims to provide both power engineers and LLM specialists with a clearer understanding of how these technologies contribute to power system reliability, resiliency, security, and affordability.

3.0 Control Center Operations and Event Analysis

3.1 Intelligent Alarm Processing and Root Cause Analysis

Motivation: Control center operators face overwhelming alarm floods during system disturbances, with modern SCADA systems generating 200+ alarms within seconds of significant events [26]. Traditional rule-based processors struggle with novel scenarios, cascading effects, and changing system topologies. Operators may spend up to 85% of their cognitive resources processing alarms during major events, limiting their capacity for actual problem-solving.

Objective: Develop LLM-based systems to process alarm sequences, identify root causes, and provide actionable recommendations in natural language. These systems must adapt to novel situations and present findings clearly to reduce operator cognitive load.

Approach: The implementation architecture integrates several key components:

- Domain-adapted LLMs that understand power system alarm syntax and timing relationships
- Retrieval-augmented generation to incorporate system topology and protection schemes
- Causality detection capabilities using power system dependency models
- Confidence-scored outputs with multiple interpretation options for complex events
- Natural language interfaces enabling operators to query event details

Implementation challenges include strict latency requirements (analysis needed within seconds), integration with existing systems, and building operator trust.

Example Input: Alarm sequence from a line trip event (46 alarms in 5 seconds):

```
08:47:12.158 LINE_18_CB_A TRIP
08:47:12.160 LINE_18_CB_B TRIP
08:47:12.175 ZONE2_FAULT OPER
08:47:12.210 GEN_5_UNDERVOLT ALARM
08:47:12.245 BUS3_VOLT_LOW ALARM
08:47:12.350 LINE_22_FLOW HIGH
...
08:47:14.732 XFRMR2_TEMP HIGH
08:47:18.915 GEN_5_TRIP ALARM
```

Example LLM Analysis: "Root cause: at 08:47AM, a Phase-to-phase fault on Line 115-18 (57% of line length), causing protection to operate as expected, isolating the fault. Voltage dip triggered generator undervoltage and subsequent trip. Recommended actions: 1) Inspect poles #28-33, 2) Verify generator ride-through settings, 3) Assess bus voltage regulation for future mitigation."

Benefits: This approach reduces operator cognitive load, accelerates root cause identification, ensures consistent analysis regardless of experience level, and enhances decision-making during critical system events.

3.2 Operations Log Analysis and Knowledge Management

Motivation: As experienced operators retire, utilities lose irreplaceable operational knowledge. While operator logs record this expertise, they are scattered across text files, emails, and various databases in inconsistent formats [27]. Finding relevant past decisions during an emerging situation is nearly impossible with keyword searches, and new operators cannot easily learn from historical events.

Objective: Create practical tools to extract insights from operator logs, making past operational knowledge searchable and useful for current decisions and training.

Approach: Effective implementation requires several key components:

- Log standardization tools to clean, format, and unify records from diverse sources, ensuring consistency and accuracy.
- Power system-specific language models capable of understanding industry jargon and abbreviations.
- Advanced search functionality that retrieves relevant past events regardless of terminology variations.
- Pattern detection to recognize recurring operational scenarios and their resolutions.
- Intuitive interfaces that allow operators to query past events using natural language.
- Validation of extracted procedures by human subject matter experts

Example User Query: "What's worked best for handling western corridor contingencies during summer peaks?"

Example LLM Response: "Based on operator logs from 2020-2023, the most effective approaches for western corridor contingencies during summer peaks have been: 1) Quick generation redispatch from southern units has prevented most cascading issues 2) Network reconfiguration at junction A when generation response is delayed 3) Temporary load reduction at western substations as a last resort"

Benefits: This approach helps utilities preserve operational knowledge, train new operators more quickly, and allow operators rapidly access relevant historical experiences.

3.3 Event Report Generation and Compliance Documentation

Motivation: After major outages, utilities face tight deadlines to file detailed reports (such as DOE Form 417 [28]) with regulators. Engineers must collect evidence from dozens of systems while simultaneously managing restoration efforts. A typical NERC-reportable event requires assembling data from SCADA, PMUs, relay logs, operator notes, and voice recordingsoften demanding many engineering hours during the critical 72-hour window following an event required by NERC standard EOP-004-4 [29].

Objective: Develop LLM systems to automate initial disturbance report generation by extracting and organizing relevant information from operational logs, alarm records, measurement summaries, and event files that meet regulatory requirements.

Approach: Successful implementation requires:

- Pre-configured templates matching specific regulatory formats
- Data connectors that pull relevant records
- Timeline construction that accurately sequences events when timestamp conventions differ across systems
- Automatic identification of reportable thresholds
- Draft generation with clear indication of confidence levels and data gaps for engineer review

Example Input: SCADA historian, digital fault recorder captures, PMU data, operator logs, and voice recordings.

Example LLM Analysis: "Loss of 380MW load affecting three substations. Event initiated at 15:32:04.233 by B-phase fault on AA-BB 230kV line. Protection operated correctly with successful three-phase tripping (Zone 1, 87% of line). Full restoration achieved at 06:07:02. No generators tripped; frequency remained within acceptable limits."

Benefits: Engineers can produce accurate disturbance reports quickly, maintain focus on critical restoration activities, ensure consistent reporting formats across events, and reduce the risk of missing compliance deadlines.

3.4 Switching Order Creation and Validation

Motivation: Transmission system operators spend significant amounts of time and effort developing and executing switching orders for substation reconfiguration, coordinating equipment outages, and ensuring the safety of field crews [30]. Distribution operators perform similar tasks for load rollover, system restoration, etc. through switching of poletop and pad-mounted sectionalizers and reclosers. Writing switching orders is a highly complicated task, with dozens of steps for opening and closing multiple circuit breakers and airgap switches in a precise order.

Objective: Develop LLM applications to develop switching order documents based on clearance requests from operators and validate them against existing clearances, hot-line work orders, and utility procedures.

Approach: The implementation requires several key components:

- System one-line diagrams, normal switching configurations, and transmission node-breaker / distribution topology models
- Named entity recognition for switching components based on classification as protection-based switches, isolation switches, and grounding switches
- Graph traversal and tree-of-thought reasoning steps mimicking operator cognitive models of "make-before-break" logic for ensuring that airgap switches are not opened while under load
- Retrieval of current network topology and generate switching sequences for equipment isolation and restoration
- Retrieval of current clearance documents and nodal voltages to ensure switching does not create any hazardous switching conditions

Example User Query: "Develop a switching order to isolate 69-12kV transformer #2 in south-bell substation"

Example LLM Response: "Analysis of South-Bell Substation topology and clearances reveals that transformer bank 2 can be isolated as below:

- Shut off LTC taps for Transformer SB-HVMV-69-11-2 (bank #2)
- Balance Transformer SB-HVMV-69-11-1 (bank #1) and shut off tap changer
- Close breaker SB-11-3B (make loop)
- Open, lock breaker SB-11-5 (break loop from low-side)
- Open, lock breaker SB-69-7 (isolate from high-side)

- Open, lock disconnector switches SB-11-5A2 and SB-69-7A1 (create physical airgap)
- Return Transformer SB-HVMV-69-11-1 taps to auto
- Issue clearance between switches SB-69A-7A1 and SB-11-5A2."

Benefits: Assistance in document creation can reduce operator fatigue, improve restoration times, and train new operators more quickly.

3.5 Shift Turnover Summarization

Motivation: Control room operators typically work rotating 12-hour shifts. The standard shift handover process is for the outgoing operator to sit down with the incoming operator to review key activity during the previous shift, highlight current activity, and review future tasks during the incoming operators shift [31]. Verbal communication gaps during the handover process can result in poor situational awareness and operating mistakes during the next shift.

Objective: Develop LLM applications to develop written summaries of the previous shift, including current system conditions (weather, outages, operating reserves, etc.), maintenance jobs that are in progress, and possible contingencies about which the incoming shift needs to be concerned.

Approach: The LLM can develop structured summaries to improve situational awareness through:

- Compilation of system operating conditions, such as generation setpoints and interchange schedules
- Summarization of log entries and alarms during the previous shift, including system disturbances, network congestion, and weather impacts
- Retrieval of ongoing / planned outages, field work activities, abnormal configurations, and tagged equipment
- Summarization of operating directives received and changes to generation / voltage schedules
- Retrieval of EMS application issues, including unsolved contingencies, state estimation failures, and communications network issues

Example Input: SCADA historian, operator logs, voice recordings, switching orders, operating directives

Example LLM Analysis: "Analysis of events and activities from the previous shift reveals:

- TWINBR 21A0 breaker was outaged by unknown cause and is currently isolated by field crews under clearance 25-234768
- Temporary interchange restriction with the east-west interface limited to 750MW
- Maintenance is planned at NWCARL with planned outage of MAIN BUS 1.
- Protection systems are to be reconfigured for OLIVE-NWCARL-1 and NWCARL-SOUTHB-1 as a long-line through the NWCARL transfer bus
- Ongoing lack of comms and unresolved alarms to GOSHEN 23-052 breaker due to bad RTU
- All EMS applications are converging"

Benefits: Assistance in document creation can reduce operator workload and increase grid reliability through improved situational awareness.

4.0 Asset Management

4.1 Intelligent Analysis of Maintenance Records

Motivation: Utilities maintain extensive records of equipment maintenance activities, often spanning decades and containing valuable information about failure modes, effective repairs, and maintenance intervals [32]. However, these records frequently exist as unstructured text that is difficult to analyze systematically.

Objective: Develop LLM applications to extract structured information from maintenance narratives, identify patterns, and provide insights to optimize maintenance practices and asset management decisions.

Approach:

- Processing of historical maintenance records
- Finetuning of LLMs to understand domain-specific maintenance terminology
- Named entity recognition for components, failure modes, maintenance actions, and results
- Development of classification models for maintenance types, root causes, and outcomes
- Recommendation systems for optimal maintenance intervals and practices

Example Input: 10 years of transformer maintenance records

Example LLM Analysis: "Analysis of 1000 transformer maintenance records reveals:

- 30% of issues stem from cooling system failures.
- Brand X pumps show a 50% failure rate after 5 years.
- Oil degrades faster after overloads exceeding 110% for over 10 hours.
- **Action:** Adjust preventive maintenance for 7 high-risk units."

Expected Benefits: Early identification of systematic issues, optimized maintenance intervals, and improved equipment reliability.

4.2 Equipment Manual and Technical Document Intelligence

Motivation: Field technicians waste critical hours searching for equipment information during outages. A typical substation contains devices from dozens of manufacturers, spanning installation dates over decades [27]. For example, when a 30-year-old breaker fails at 2AM, finding the right manual among thousands of poorly-indexed documents can extend outage time by hoursespecially when the only staff member familiar with the equipment retired years ago.

Objective: Build LLM-based search tools that quickly deliver exact maintenance procedures, specifications, and troubleshooting steps from technical documentation directly to field personnel when needed.

Approach: Key system elements include:

- Document processing pipeline that handles various document formats
- Technical recognition system that understands circuit diagrams (devices, parts), procedures, and tables
- Query interface, especially for mobile use in field
- Equipment relationship modeling to connect related devices, parts, and procedures
- Version control of procedures

Example User Query: "How do I manually close GE breaker in substation 21 without power?"

Example LLM Response: "Manual closing procedure for GE breaker (ID #S8-CB-21) without power:

1. Confirm breaker OPEN 2. Open right side panel 3. Remove charging handle 4. Insert handle into charging port and pump clockwise until CHARGED indicator appears 5. Press yellow manual close button 6. Verify GREEN indicator shows breaker closed

See attached diagram for component locations. Last maintenance: 04/2023."

Benefits: Field crews restore equipment faster, inexperienced staff can confidently maintain unfamiliar devices, critical institutional knowledge remains accessible, and human errors are reduced.

4.3 Automated Processing of Field Work Orders

Motivation: Unlike transmission system models which are typically infrequently, distribution feeder models require daily updates due to ongoing field work, such as pole replacement and temporary cuts and jumpers. Distribution utilities typically maintain a geographic information system (GIS) database as the system of record from which all planning and operations models are developed [33]. Updates to GIS records are performed manually through paper field work orders, resulting in delayed updates and discrepancies between models and the physical system.

Objective: Develop LLM applications to automate processing of field work orders to update GIS records and power flows models used by .

Approach: Effective implementation requires:

- Pre-configured templates match specific maintenance routines and field work documents
- Data connectors to GIS, planning, and operations tools
- Finetuning on distribution asset configurations, datasheets, and prior maintenance documents
- Multi-agent pipelines to ingest field work orders, identify correct equipment, and update GIS and engineering models based on work performed

Example Input: Field work order document for resizing a fuse

Example LLM Analysis: "Analysis of work order indicates the following changes:

- Field crews replaced fuse 9478, which was previously rated at 50A with a new 75A fuse
- Updated GIS record 9478 in FUSES table with new RATING attribute of 75 A
- Updated CYME model for protection element between node 34798 and 34799 with new current rating of 75.0
- Updated FLISR application database with new fuse record"

Benefits: Engineering teams can benefit from reduction in effort required to update records manually and from significant increase in model accuracy due to timely updates of as-built models.

4.4 Registration and Integration of Customer-owned Resources

Motivation: Distribution utilities face significant challenges in mapping behind-the-meter (BTM) resources to the correct feeder models for power studies, control, and aggregation to the bulk transmission level. DER data is typically spread across a variety of customer billing, GIS, and engineering databases. As a result, transmission entities are unable to gain information about the quantity of DERs connected to a given substation bus, while residential customers face long interconnection approval queues due to the burden of manual data entry and study creation.

Objective: Develop LLM applications to automate processing of connection requests by inserting DER data into appropriate engineering models, running planning studies, and performing aggregation for data exchange need to support implementation of Federal Energy Regulatory Commission (FERC) Order 8888 and FERC Order 901 [34].

Approach: Effective implementation requires:

- Pre-configured templates matching specific interconnection forms, design documents, and approval letters issued to customers
- Data connectors that pull relevant records of existing and proposed DERs, network models, and system constraints
- Attribute-based access control of DER data to ensure personally identifiable information of customers is not shared with the LLM
- Finetuning and retrieval of software-specific modeling formats used by planning and operations tools
- Multi-agent workflows to update models, develop studies, run numerical simulations, and generate reports upon ingest of DER interconnection request batches
- Agentic retrieval and aggregation of DER capacities and controls for market aggregation and exchange with transmission entities

Example Input: Interconnection request, GIS model, metering database, unbalanced planning model, load and generation profiles

Example LLM Analysis: "Analysis of interconnection feasibility shows

- Customer wishes to install a 150kW PV and 50kWh battery on feeder 19-3 bus M2378 as part of an aggregation program
- During springtime minimum load conditions, the highest bus voltage will be 1.07 pu. This can be lowered to 1.04 pu by requiring an IEEE 1547-2018 category B volt-var curve
- During summer peak loading conditions, a 110% overload is expected on line L23785, which will require reconductoring to 397 ACSR
- The DER will be connected to transmission planning bus 379842 and raise total installed DER capacity to 1.39 MW."

Benefits: Engineering teams can benefit from significant reductions in the effort required for updating planning models, running studies, and creating reports. Customers can benefit from faster interconnection times and reduced planning fees, and lower energy costs from automated processing of requests.

4.5 Intelligent Validation of Asset Management Records

Motivation: Distribution utilities typically own tens of thousands of individual assets (including poles, transformers, capacitor banks, and switchgear), which are typically tracked in a GIS and a separate asset management database. Asset data and definitions from disparate sources typically contain errors ranging from inconsistent naming to typographic errors in data entered manually. Cleanup and validation of asset data needed to construct accurate power flow models is labor-intensive and very time-consuming, even with formal data-validation rules and automation scripts [35].

Objective: Develop LLM applications to construct validation rules and analyze asset datasets to identify inconsistencies and errors in GIS and asset database records.

Approach: Effective implementation requires:

- Document processing and retrieval pipelines that handle asset records from GIS, asset management databases, and relevant documents
- Technical recognition system that interprets asset datasheets, nameplates, and database records to standardize naming conventions and asset identifiers
- Contextual understanding of asset usage and configurations to create data validation rules
- Couple with an anomaly detection AI/ML agent to identify inconsistent and invalid values across multiple systems of record
- Multimodal capabilities to process imagery from vehicle-mounted cameras to identify assets that are missing from existing records
- Highlighting of modeling errors and recommendation of corrected values from retrieved documents

Example Input: GIS database, asset datasheets, equipment nameplates, asset management database, field work orders and test records

Example LLM Analysis: "Analysis of transformer GIS records shows the following errors for Voltage Regulator R9458:

- The apparent power rating is listed as 165kVA in the asset database, which may be incorrect. Standard ratings of GE-VR-FC regulators are 144kVA and 167kVA
- The short circuit leakage reactance is -0.12%, which should be a positive value as this is not an equivalanced three-winding transformer
- The phasing information in GIS shows that it is connected to phase A, while a field work order indicates seasonal rephasing was completed in October to place it on phase C."

Benefits: Engineering teams can take advantage of more accurate asset models for use in planning and operations models and reduce time spent achieving convergence of power flow solutions.

5.0 System Planning and Analytics

5.1 Accelerated Generator Interconnection Studies

Motivation: Interconnection of new generation resources require a significant number of studies to be performed by system planning engineers to check for system violations and changes in dynamic behavior caused by the proposed generator. Typically, a power flow study, short-circuit study, transient stability study, and relay coordination study must be performed before the interconnection request can be approved.

Objective: Develop LLM applications to automate model editing, numerical simulations, and preparation of study results. The LLMs would insert the proposed generator in the utility models for each separate study tool in the correct data format and with the correct parameters. The LLM would then call the API interfaces of each numerical solver to run the studies, extract results, and prepare reports for decision-makers.

Approach: Effective implementation requires:

- Technical language processing that recognizes power system domain terms and generation parameters from interconnection requests
- Finetuning and retrieval of solver-specific data formats (e.g. .raw and .dyn) and API interfaces used by each solver
- Ingest of interconnection applications and data extraction / mapping LLM agents to identify modeling parameters
- Multi-agent pipelines with chain-of-thought reasoning to add the proposed generator to the network model files used by each study tool
- Automation of simulation studies through API calls generated and executed by solver agents
- Identification of violations and remediation strategies to enable interconnection

Example Input: Interconnection application, transmission planning model, transient stability model

Example LLM Analysis: "The interconnection agreement seeks to add a new 300MW gas turbine plant at BEAVER substation

- Added 300MW unit to bus 85. No overloads or violations occur during the peak-load case.
- Loss of the BEAVER-CLINCH line during peak load results in a 130% overload on the FREMON-CLINCH line.

- Added GGOV1 model for proposed unit and completed 23 transient stability simulations summarized in beaver_sims.pdf
- Added unit and synchronous machine to short circuit model. Bus fault current levels are increased by 157% and upgrades will be required for BEAVER 138kV breakers

Benefits: Engineering teams can benefit from significant reductions in time and effort needed to assemble models and perform simulation studies required to process interconnection requests. Generation developers and customers benefed from reduced queues, resulting in more reliable and affordable energy delivery.

5.2 Outage Report Mining and Reliability Analysis

Motivation: Utilities produce detailed reports for significant system outages and equipment failures, containing valuable information about failure modes, system vulnerabilities, and effective recovery strategies. While engineers might remember recent events, patterns spanning years remain hiddenleading to repeated failures that could have been prevented by connecting historical dots.

Objective: Develop LLM applications to extract actionable patterns from historical outage reports to identify systemic vulnerabilities, prioritize investments, and enhance reliability planning.

Approach: Effective implementation requires:

- Unified repository of historical reports
- Technical language processing that recognizes power system domain terms
- Pattern detection across timeframes and geographical regions to identify recurring issues (spatio-temporal relationship)
- Root cause extraction that differentiates initiating events from consequential outcomes
- Risk quantification tools that estimate likelihood and consequence of vulnerability patterns

Example Input: 500 transmission outage investigations (2010-2022)

Example LLM Analysis: "Analysis of ten years of outage data reveals critical vulnerability patterns:

1. Thermal overload protection operated conservatively on 11 occasions due to ambient temperature sensor failures, causing unnecessary line trips during summer peaks. 2. Breaker failure protection unexpected operations under high fault conditions led to unnecessary substation outages 9 times. 3. Risk model indicates addressing protection coordination issues would prevent 37% of customer outage-minutes with estimated $4.2M annual savings."

Benefits: Engineering teams can target investments toward highest-impact vulnerabilities, improve reliability metrics, and reduce capital spending by focusing on root causes rather than symptoms.

5.3 Enhanced Integrated Resource Planning

Motivation: Resource planners face widening uncertainty bands in both load and generation forecasts. A utility's 20-year plan might contain hundreds of narrative assumptions buried in thousands of pages with no systematic way to trace forecast misses to specific assumptions [36]. When actual conditions inevitably diverge from projections, planners struggle to identify which narrative assumptions need adjustment and how they affect both load and resource adequacy.

Objective: Create practical LLM tools to track forecast assumptions, quantify their impacts, generate alternative futures, and enable rapid scenario updates as conditions evolve.

Approach: Effective implementation includes:

- Centralized repository of planning assumptions from load and generation forecasts with corresponding assumptions
- Connection between narrative elements to specific numerical impacts
- Scenario generation capabilities allowing planners to quickly create consistent "what-if" cases, including atypical weather events
- Directional impact analysis showing how shifts in key variables affect resource needs

Example Inputs: 2019 IRP documents and actual data through 2024

Example LLM Analysis: "Your 2019 resource plan missed the following items:

- Industrial growth: Overestimated by 3.3% annually (projected 4.2%, actual 0.9%)
- Gas unit availability: 7% lower than modeled during peak seasons
- Solar capacity factor: 3.2% higher than projected

Refresh your base case with 3 alternative scenarios addressing: 1) reduced industrial load. 2) accelerated electrification with distributed storage growth. 3) extreme weather impacts on peak demand and disturbances."

Benefits: Planners can create more realistic scenarios and better align generation resources with emerging trends, reducing both reliability risks and unnecessary spending.

5.4 Technical Standards and Compliance Navigation

Motivation: Power utilities must comply with a list of technical standards (IEEE, IEC, NERC, etc.) and internal procedures, each with their own testing intervals, documentation needs, and technical specifications. Non-compliance with NERC standards can result in monetary penalties and fines of up to $1 million USD per day [37]. As a result, engineers spend substantial time researching interpretations and ensuring consistency across multiple documents, which is quite challenging.

Objective: Provide efficient access to relevant compliance requirements and verify internal procedures satisfy all applicable standards and regulations.

Approach: Practical implementation requires:

- Processing of all technical documents with different formats and resources
- Standards relationship mapping between requirements
- Automatic version tracking to evolving standards
- Gap analysis between internal procedures and external requirements

Example User Query: "Are our 230-kV transfer limits compliant with NERC standards?"

Example LLM response: "Compliance risk identified: Your transmission capacity limits may not meet NERC TPL-001-5 reliability criteria.

Your transfer limits are based on N-1 contingencies, but NERC TPL-001-5 requires that steady-state thermal limits must consider N-1-1 contingencies.

RECOMMENDATION: To include N-1-1 contingency analysis and re-run planning studies to avoid potential compliance violations."

Benefits: Minimize compliance violation risk, reduce unnecessary over-compliance work, and improve work efficiency.

5.5 Assisted System Planning Decisions

Motivation: Under the requirements of NERC Standard TPL-001-5 [38], transmission planners and planning coordinators are required to perform near-term and long-term planning studies examining upgrades to transmission facilities needed to accommodate load growth, plant retirements, new interconnections, etc. under a variety of system conditions. System planning engineers spend significant amounts of time developing study and performing repetitive model edits for what-if analysis of alternative configurations. Development of the most recent MISO long-range transmission plan required over 40,000 staff hours to complete the study [39]. Automation of scenario creation and comparison of planning alternatives could significantly reduce the analysis burden placed on utility staff.

Objective: Develop LLM applications to assist with preparation of planning models, development of API calls to numerical simulation tools, identification of system upgrades to mitigate system violations, and creation of alternatives-analysis reports

Approach: Effective implementation requires:

- Technical language processing that converts user requests into model adjustments
- ReAct training on recognition of causes of voltage and thermal violations using self-training via simulator APIs and planning procedures documents
- Scenario generation capabilities allowing planners to create consistent "what-if" cases for peak-load, minimum-load, etc.
- Multi-agent workflows applying tree-of-thought reasoning to evaluate multiple alternative planning options by inserting new elements, executing API calls to run simulation studies, and evaluating effectiveness of alternatives
- Cost allocation and decision-making support capabilities providing summarization and comparison of simulation results

Example User Query: "Under summer peak-load conditions for 2035, provide a set of alternatives to mitigate overloading of EASTLI 345-138kV transformer."

Example LLM Response: "Analysis of EASTLI substation under summer peak-load indicates overloads on the transformer between buses 37 and 38 can be caused by two 345kV line contingencies and four 138kV line contingencies.

- Upgrade the substation with a new 1000 MVA transformer, which will resolve all N-1 and N-1-1 contingencies at a cost of $17M.
- Install a second 500 MVA bank in parallel with the existing transformer, which will cost $9M. Installation of another bank will require additional upgrades to substation buses, new breakers, and protection systems which are not included in the cost estimate.
- Install a 100 MW battery storage at WESTLI at a cost of $55M that will also alleviate DER export constraints during minimum-load conditions. Additional upgrades may be necessary."

Benefits: Engineering teams can use an LLM interface to analyze planning alternatives more efficiently and quickly, reducing fatigue and burnout being faced by current system planners currently overwhelmed by the number of studies required to meet current load growth projections.

6.0 Energy Management and Protection Systems

6.1 SCADA and Control System Log Analysis

Motivation: Modern control systems generate vast amounts of log data documenting system operations, user actions, and state changes. These logs contain valuable information for troubleshooting, security monitoring, and performance optimization, but their volume and complexity make manual analysis impractical.

Objective: Develop LLM applications to analyze control system logs, identify abnormal patterns, extract operational insights, and support root cause analysis for system issues.

Approach:

- Collection and normalization of logs from diverse control systems (SCADA, EMS, DMS, etc.)
- Training of models to understand control system terminology, error codes, and operational states
- Anomaly detection to identify unusual event sequences or system behaviors
- Temporal pattern recognition to correlate events across different timeframes
- Creation of natural language interfaces for operators and engineers to query system behavior
- Design of visualization tools to represent complex event sequences in comprehensible formats

Example Input: 3 days of EMS logs surrounding a system failure event

Example LLM Analysis: "Root cause analysis indicates the EMS failure at 14:23 on March 17 was preceded by a sequence of relevant events:

1. Database synchronization errors began 37 hours prior to failure (March 16, 01:42) with increasing frequency (log pattern shown)
2. Memory utilization on redundant servers diverged significantly 8 hours before failure, with primary server showing 4% higher utilization per hour
3. ICCP link failures to 3 neighboring systems occurred sporadically in the 4 hours before system failure, with connection retry patterns indicating network instability

4. Final trigger was application deadlock between state estimator and contingency analysis processes at 14:22:07"

Expected Benefits: Faster troubleshooting of complex issues, proactive detection of developing problems, better understanding of system behavior, and reduced system downtime.

6.2 Energy Management System (EMS) Event Analytics

Motivation: Power system control center generates extensive log data documenting system operations, user actions, state changes, and events. These logs contain valuable information for troubleshooting and performance optimization, but their volume and complexity make manual analysis challenging. During system restoration, engineers often searching through logs from multiple sources under time pressure, increasing the risk of missing critical details until post-event analysis days later.

Objective: Turn overwhelming system logs into actionable intelligence, uncovering developing issues before they cause outages and pinpointing root causes within minutes rather than days.

Approach: Effective implementation includes:

- Log aggregation hub incorporating various data sources, including handling proprietary formats from different vendors
- Pattern detection at multiple timescales
- Automated correlation between seemingly unrelated events across different subsystems
- Early warning system for developing issues based on historical failure patterns

Example User Query: "Identify why the state estimator diverged"

Example LLM Response: "Critical EMS issue detected:

- State estimator convergence time increasing
- Bad data flagging on 345kV tie-line measurements at Northern corridor
- Power flow solutions showing unrealistic line flows in Northern corridor

Recommended actions: 1) Switch to backup state estimator instance using previous configuration; 2) temporarily exclude suspicious measurements from solution."

Benefits: Control room teams identify developing problems before outages occur, accelerate root cause analysis, and build institutional knowledge.

6.3 Protection Setting Verification and Mis-coordination Prevention

Motivation: Protection miscoordination causes cascading outages that can leave thousands without power. A single incorrect relay setting can turn a minor fault into a regional blackout, yet most utilities still verify thousands of complex settings manually. Engineers check calculations by hand, cross-reference values against multiple standards, and attempt to visualize interdependencies between dozens of devicesa process so time-consuming that many settings receive only cursory review before implementation.

Objective: Create LLM tools that automatically verify protection settings against standards, coordination studies, and default factorial specifications, identifying potential reliability risks before they reach the field.

Approach: Successful implementation requires:

- Processing of protection coordination studies and relay setting sheets into structured formats
- Training of models to understand protection engineering concepts and setting relationships
- Requirement extraction from standards documents (e.g., PRC-023, PRC-024)
- Verification algorithms to compare actual settings against required values
- Natural language explanations of complex coordination relationships for engineers

Example Input: Protection settings study for a 115-kv line and actual relay settings

Example LLM Analysis: "Settings verification identified the following discrepancies between coordination study recommendations and actual settings: Zone 2 reach (125% of line) reduces coordination margin with Zone 3 to 0.3s Ground pickup threshold (0.75A) may miss high-resistance faults Reclosing time (45 cycles) creates potential conflict with downstream distribution reclosers (30 cycles)

Recommended setting adjustments: 1) adjust Zone 2 reach to 120%. 2) lower ground pickup to 0.5A. 3) increase reclose time to 60 cycles."

Benefits: Identify setting conflicts before they reach the field, prevent cascading outages from mis-coordination, and reduce engineering review time.

6.4 Planned BES Outage Coordination

Motivation: Reliability coordinators, transmission operators, and balancing authorities are required by NERC standard IRO-017-1 [40] to coordinate planned outages of BES facilities and equipment to ensure the reliability of grid. The outage coordination process is generally performed manually based on a set of standalone documents and schedules exchanged between operators and operations engineers at each entity via electronic notices and email. Depending on the duration of outage and criticality of the facility, outages may need to be studied between a few months to a few days before the outage is approved. The coordination study examines whether the outage will result in any contingency analysis violations, as well as impacts (if any) on voltage control, generation schedules, operation of protection systems,

Objective: Develop LLM applications to assist with outage scheduling by ingesting disparate schedules, assembling study models, executing API calls to simulation tools, and identifying system violations.

Approach: Effective implementation requires:

- Pre-configured templates matching specific outage coordination documents, schedules, and electronic notices
- Finetuning and retrieval of software-specific modeling formats and API calls used by planning and operations tools
- Processing of system one-line diagrams, normal switching configurations, and node-breaker models
- Ingest of generation schedules, planned outages, and forecasts and translation to network model updates
- Graph reasoning to develop updated contingency analysis definitions based on planned equipment outage
- Multi-agent workflows to update models, develop studies, write API calls to run contingency analysis studies, and generate reports with outcomes

Example Input: Node-breaker transmission model, generation dispatch schedule, load forecast, DER forecast, planned outage of breaker

Example LLM Analysis: "Analyzing a planned outage of breaker OLIVE-032-4A from 04/17 08:00 to 04/18 16:00:

- No other outages are planned at OLIVE or neighboring substations
- OLIVE-032-4A is a 138kV in a ring-bus configuration. Outage of the breaker itself will not impact steady-state power but does affect contingency definitions
- Loss of OLIVE-NWCARL line will also result in loss of OLIVE-SOUTHB line.
- New double contingency will result in overload of OLIVE-KANKAK line. Identifying mitigation options

- NWCARL generator is not scheduled to run. Generation re-dispatch to run NWCARL unit may resolve this violation.
- NWCARL generator is scheduled to run on 04/19. Rescheduling the outage to a later date may reduce generator dispatch costs."

Benefits: Utilities can coordinate outages more effectively resulting in improved grid reliability and reduced workloads. Energy affordability gains may be made through more effective outage scheduling by avoiding out-of-merit generation dispatch costs.

6.5 DER-informed Load Shedding Scheme Design

Motivation: Under-frequency load shedding (UFLS) and under-voltage load shedding (UVLS) schemes are protection schemes used to trip distribution feeder breakers automatically during large system disturbances. Planning coordinators are required by NERC Standards PRC-006-5 [41] and PRC-010-2 [42] to conduct design assessments, maintain a database of modeling parameters, and keep dated records (i.e. spreadsheets) of the settings implemented by distribution providers. Current design practices generally are based on single-hour peak load of distribution feeders and do not include modeling of the quantity of DERs or their ride-through characteristics. Lack of accurate modeling of DER can result in misoperation of UFLS and UVLS schemes due to insufficient or excessive load shedding [43].

Objective: Develop LLM applications to improve UFLS / UVLS design through creation and simulation of more accurate planning models assembled from feeder loading spreadsheets, transmission planning models, substation one-line diagrams, and disparate DER records from distribution providers.

Approach: Effective implementation requires:

- Processing of protection coordination studies and relay setting sheets into structured formats
- Rulesets for estimation of DER ride-through characteristics based on timeframes of revisions to IEEE Std. 1547 and disparate regulatory implementations
- Pre-configured templates for aggregation of DERs by substation bus and estimated ride-through characteristics for each distribution provider
- Finetuning and retrieval of solver-specific data formats and API interfaces for transient stability simulation solver
- Document processing and retrieval pipelines to handle database maintenance, basecase creation, and compliance records managements

- Multi-agent pipelines to assemble simulation models from disparate records and execute API calls to run simulation studies for various initiating events
- Natural language interfaces for planning engineers to query simulated system behavior

Example Input: Transient stability model, aggregated DER spreadsheets, system online diagrams, regulatory documents, feeder load spreadsheets

Example LLM Analysis: "Analysis of system configuration and DER data from Lakeside Distribution Company shows:

- State regulations stated enforcement of IEEE 1547-2018 in 2022, with estimated 238 MW of DERs wide correct ride-through settings
- A total of 145 MW of DERs were installed prior to 2014 and may have must-trip settings
- System one-line diagram shows a new 3MW solar farm on distribution feeder WESTLI-3, which is currently enable to trip at 59.3 Hz.
- Mapped peak load and DER data from 184 distribution feeders to 76 planning model buses
- Simulation of peak-load case with a 25% generation loss event and existing UFLS settings results in a frequency nadir of 58.96 Hz"

Benefits: Improving modeling of DER impacts on system behavior helps reduce engineering review time and prevent cascading outages from mis-coordination of protection settings.

7.0 Practical Considerations for LLM Deployment

Development of LLM-based applications to support power systems planning and operations use cases involve several practical considerations related to implementation and deployment decisions that must be made by application developers.

7.1 Technical Implementation Approaches

Development of full custom LLM models is very expensive and frequently not required. It is recommended that developers consider simpler implementations and consider the following design choices:

- **Retrieval-Augmented Generation:** Most power system LLM applications rely on RAG, which retrieves domain-specific knowledge, such as fault logs, operational reports to enhance LLM responses.
- **Fine-tuning vs. Prompt Engineering:** Determining when to fine-tune LLMs on power system data versus using sophisticated prompting with foundation models.
- **Domain-Specific Embeddings:** Creating specialized embeddings that capture power engineering concepts and terminology relationships.
- **Integration with Structured Data Systems:** Approaches for connecting LLM insights with traditional power system databases, SCADA systems, and analytical tools.
- **Handling Mixed Numeric and Textual Content:** Strategies for processing documents containing both narrative text and critical numerical values.
- **Evaluation Metrics:** Domain-appropriate evaluation approaches for assessing LLM performance on power system tasks.

7.2 Deployment Considerations for Critical Infrastructure

The power systems domain contains large amounts of critical energy infrastructure information (CEII) data that is highly sensitive and cannot be exposed to commercial LLM infrastructure.

- **Security and Access Control:** Ensuring LLM systems have appropriate access controls and data handling protections for sensitive operational data.
- **Explainability Requirements:** Approaches for providing transparency into LLM reasoning for critical operational or compliance applications.

- **Verification and Validation:** Methodologies for testing LLM systems against power system domain knowledge and operational requirements.
- **On-premises vs. Cloud:** Balancing cloud-based AI flexibility with on-premises deployment constraints in real-time grid operations and NERC Critical Infrastructure Protection compliant environments as given by NERC CIP-002 standard [18] and latest guidance on deployment of BES applications in the cloud [20].
- **Human-in-the-Loop Design:** Effective interfaces for expert validation and oversight of LLM outputs in critical applications.
- **Training and Change Management:** Techniques for resolving inconsistencies in protection relay settings, outage logs, and SCADA event records across different utilities and vendors.

7.3 Data Quality and Preparation

Significant attention must be paid to collection and preparation of data used for finetuning of LLMs

- **Document Preprocessing:** Techniques for converting legacy documents, scanned materials, and diverse formats into structured, LLM-compatible inputs.
- **Domain-Specific Cleaning:** Approaches for resolving inconsistencies in power system datasets, such as mismatched abbreviations, vendor-specific naming conventions, and timestamp mis-alignments between SCADA logs, PMU records, and relay events.
- **Contextual Preservation:** Methods for preserving relationships across documents, ensuring that interconnected data remains linked during processing.
- **Training Data Creation:** Strategies for developing high-quality, structured datasets to enable fine-tuning of LLMs on specific tasks.
- **Synthetic Data Generation:** Using existing models to generate synthetic training data for rare system events to improve LLM performance on underrepresented scenarios.

8.0 Challenges and Future Directions

8.1 Key Barriers to LLM Adoption in Power Systems

Beyond deployment concerns, there still exist fundamental limitations in LLM capabilities that must be considered, such as

- **Domain-Specific Terminology Gaps:** General-purpose LLMs struggle with highly specialized, often arcane power systems terminology, particularly for legacy equipment documentations, manufacturer-specific codes, and rare operational scenarios. Fine-tuning on industry-specific datasets is needed to bridge this gap.
- **Numerical Reasoning Limitations:** LLMs alone lack precision in complex power engineering calculations and can not replace numerical solvers. Hybrid approaches integrating LLMs with symbolic reasoning and domain-specific traditional solvers are essential for reliable, practical grid management.
- **Access to Historical Knowledge:** Many critical utility records span decades and exist in scanned PDFs, handwritten logs, and legacy formats. Need advanced techniques extract actionable insights.
- **Challenges in Operational Technology Integration:** Deploying LLMs within operational technology environments raises security concerns, real-time processing constraints, and interoperability challenges.
- **Validation for Mission-Critical Applications:** Ensuring LLM-generated recommendations align with engineering standards and regulatory requirements demands rigorous testing, human-in-the-loop verification, and explainability frameworks.

8.2 Research Directions for LLM-Powered Grid Applications

Usage of LLMs for power systems applications is an emerging research domain, with many unexplored topics including

- **Power System-Specific Foundation Models:** Development of foundation models pretrained on power system technical literature, standards, and operational documents.
- **Multimodal Understanding:** Integration of text, diagrams, schematics, and time-series data for comprehensive power system understanding.
- **Spatio-Temporal Reasoning:** Advanced capabilities for understanding cause-effect relationships across wide areas with different operational timeframes.

- **Expert-Guided Fine-Tuning:** Methodologies for efficiently incorporating domain expert knowledge into model training without extensive labeled datasets.
- **Combined Symbolic and Neural Approaches:** Hybrid systems that integrate traditional power system physics-based models with LLM capabilities.

References

[1] J. Achiam, S. Adler, S. Agarwal, *et al.*, "Gpt-4 technical report," *arXiv preprint arXiv:2303.08774*, 2023.

[2] H. Touvron, T. Lavril, G. Izacard, *et al.*, "Llama: Open and efficient foundation language models," *arXiv preprint arXiv:2302.13971*, 2023.

[3] K. Parveen, T. Q. B. Phuc, A. A. Alghamdi, *et al.*, "Unraveling the dynamics of chatgpt adoption and utilization through structural equation modeling," *Nature Scientific Reports*, vol. 14, no. 1, p. 23 469, 2024.

[4] T. Brown, B. Mann, N. Ryder, *et al.*, "Language models are few-shot learners," *Advances in neural information processing systems*, vol. 33, pp. 1877–1901, 2020.

[5] K. Singhal, S. Azizi, T. Tu, *et al.*, "Large language models encode clinical knowledge," *Nature*, vol. 620, no. 7972, pp. 172–180, 2023.

[6] J. Lai, W. Gan, J. Wu, Z. Qi, and P. S. Yu, "Large language models in law: A survey," *AI Open*, 2024.

[7] J. Lee, N. Stevens, and S. C. Han, "Large language models in finance (finllms)," *Neural Computing and Applications*, pp. 1–15, 2025.

[8] H. W. Sprueill, C. Edwards, K. Agarwal, *et al.*, "CHEMREASONER: Heuristic search over a large language models knowledge space using quantum-chemical feedback," in *Proceedings of the 41st International Conference on Machine Learning*, 2024, pp. 46 351–46 374.

[9] B. F. Wollenberg and T. Sakaguchi, "Artificial intelligence in power system operations," *Proceedings of the IEEE*, vol. 75, no. 12, pp. 1678–1685, 1987.

[10] V. S. S. Vankayala and N. D. Rao, "Artificial neural networks and their applications to power systemsa bibliographical survey," *Electric power systems research*, vol. 28, no. 1, pp. 67–79, 1993.

[11] L. Cheng and T. Yu, "A new generation of ai: A review and perspective on machine learning technologies applied to smart energy and electric power systems," *International Journal of Energy Research*, vol. 43, no. 6, pp. 1928–1973, 2019.

[12] A. K. Ozcanli, F. Yaprakdal, and M. Baysal, "Deep learning methods and applications for electrical power systems: A comprehensive review," *International Journal of Energy Research*, vol. 44, no. 9, pp. 7136–7157, 2020.

[13] J. Xie, I. Alvarez-Fernandez, and W. Sun, "A review of machine learning applications in power system resilience," in *2020 IEEE power & energy society general meeting (PESGM)*, IEEE, 2020, pp. 1–5.

[14] U. Pandey, A. Pathak, A. Kumar, and S. Mondal, "Applications of artificial intelligence in power system operation, control and planning: A review," *Clean Energy*, vol. 7, no. 6, pp. 1199–1218, 2023.

[15] I. Alhamrouni, N. H. Abdul Kahar, M. Salem, *et al.*, "A comprehensive review on the role of artificial intelligence in power system stability, control, and protection: Insights and future directions," *Applied Sciences*, vol. 14, no. 14, p. 6214, 2024.

[16] G. L. Rajora, M. A. Sanz-Bobi, L. B. Tjernberg, and J. E. Urrea Cabus, "A review of asset management using artificial intelligence-based machine learning models: Applications for the electric power and energy system," *IET Generation, Transmission & Distribution*, vol. 18, no. 12, pp. 2155–2170, 2024.

[17] Y. Chen, X. Fan, R. Huang, Q. Huang, A. Li, and K. P. Guddanti, "Artificial intelligence/machine learning technology in power system applications (pnnl-35735)," Pacific Northwest National Laboratory (PNNL), Richland, WA (United States), Tech. Rep., 2024.

[18] *Cyber security: Bes cyber system categorization*. [Online]. Available: https://www.nerc.com/pa/stand/reliability%20standards/cip-002-5.1a.pdf.

[19] S. L. Choi, R. Jain, C. Feng, *et al.*, "Generative ai for power grid operations," National Renewable Energy Laboratory (NREL), Golden, CO (United States), Tech. Rep., 2024.

[20] "Bes operations in the cloud." (2023), [Online]. Available: https://www.nerc.com/comm/RSTC_Reliability_Guidelines/SITES_WhitePaper_BES_Ops_in_Cloud.pdf.

[21] J. Wei, X. Wang, D. Schuurmans, *et al.*, "Chain-of-thought prompting elicits reasoning in large language models," *Advances in neural information processing systems*, vol. 35, pp. 24 824–24 837, 2022.

[22] S. Yao, D. Yu, J. Zhao, *et al.*, "Tree of thoughts: Deliberate problem solving with large language models," *Advances in neural information processing systems*, vol. 36, pp. 11 809–11 822, 2023.

[23] S. Yao, J. Zhao, D. Yu, *et al.*, "React: Synergizing reasoning and acting in language models," in *International Conference on Learning Representations (ICLR)*, 2023.

[24] A. Vaswani, N. Shazeer, N. Parmar, *et al.*, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[25] S. Purohit, G. Chin, P. S. Mackey, and J. A. Cottam, "Graphaide: Advanced graph-assisted query and reasoning system," in *2024 IEEE International Conference on Big Data (BigData)*, IEEE, 2024, pp. 3485–3493.

[26] M. Kezunovic and A. Bose, "The future ems design requirements," in *2013 46th Hawaii International Conference on System Sciences*, IEEE, 2013, pp. 2354–2363.

[27] H. Zhang, S. Kincic, and S. Edwards, *Advanced power applications for system reliability monitoring*. Springer Nature, 2020.

[28] U. D. of Energy Form DOE-417, *Electric emergency incident and disturbance report*. [Online]. Available: https://doe417.pnnl.gov/files/DOE-417_Form.pdf (visited on 04/02/2025).

[29] *Event reporting*. [Online]. Available: https://www.nerc.com/pa/stand/reliability%20standards/eop-004-4.pdf.

[30] A. A. Anderson, S. Kincic, C. K. Fallon, and B. A. Jefferson, “Transmission operator workflows for real-time reliability studies: A review of control room practices and naturalistic decision making,” Pacific Northwest National Laboratory (PNNL), Richland, WA (United States), Tech. Rep. PNNL-36516, 2024.

[31] J. R. Clark, N. A. Stanton, and K. M. Revell, “Identified handover tools and techniques in high-risk domains: Using distributed situation awareness theory to inform current practices,” *Safety science*, vol. 118, pp. 915–924, 2019.

[32] A. Alquraidi and M. Awad, “Physical asset management for critical utilities-a systematic literature review,” *IEEE Access*, 2024.

[33] G. Cochenour, R. Ochoa, and V. Rajsekar, “Distribution network model readiness for advanced distribution management systems,” in *2014 IEEE PES T&D Conference and Exposition*, IEEE, 2014, pp. 1–5.

[34] *Reliability standards to address inverter-based resources*. [Online]. Available: https://www.ferc.gov/media/e-1-rm22-12-000.

[35] K. Oyoo and D. Berleant, “An automated data validation approach to enterprise asset management for power and utilities organizations,” in *2021 IEEE Electrical Power and Energy Conference (EPEC)*, IEEE, 2021, pp. 1–6.

[36] B. Biewald, D. Glick, S. Kwok, K. Takahashi, J. P. Carvallo, and L. C. Schwartz, “Best practices in integrated resource planning: A guide for planners developing the electricity resource mix of the future,” Lawrence Berkeley National Laboratory (LBNL), Berkeley, CA (United States), Tech. Rep., 2024.

[37] NERC, “Rules of procedure appendix 4b: Sanction guidelines,” North American Electric Reliability Corporation, Tech. Rep., 2021. [Online]. Available: https://www.nerc.com/AboutNERC/RulesOfProcedure/Appendix_4B_effective%2020210119.pdf.

[38] *Transmission system planning performance requirements*. [Online]. Available: https://www.nerc.com/pa/Stand/Reliability%20Standards/TPL-001-5.pdf.

[39] MISO, “Mtep24 long rang transmission plan,” Midland Independent System Operator, Tech. Rep., 2024. [Online]. Available: https://cdn.misoenergy.org/MTEP24%20Full%20Report658025.pdf.

[40] *Outage coordination*. [Online]. Available: https://www.nerc.com/pa/stand/reliability%20standards/iro-017-1.pdf.

[41] *Automatic underfrequency load shedding*. [Online]. Available: https://www.nerc.com/pa/stand/reliability%20standards/prc-006-5.pdf.

[42] *Undervoltage load shedding*. [Online]. Available: https://www.nerc.com/pa/Stand/Reliability%20Standards/PRC-010-2.pdf.

[43] A. Anderson, S. Datta, T. L. Vu, *et al.*, “Cloud-based testbed for adaptive under-frequency load shedding with high der penetration,” in *2024 IEEE Power & Energy Society General Meeting (PESGM)*, IEEE, 2024, pp. 1–5.

Pacific Northwest
National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99352
1-888-375-PNNL (7675)

www.pnnl.gov