

Light Water Reactor Sustainability Program

Business Case Analysis for Artificial Intelligence-Large Language Model Technology Integration

Casey Kovesdi, Ahmad Al Rashdan
Idaho National Laboratory

Alex Tylecote, Morgan Schadeegg, Luke Martin
ScottMadden



June 2025

U.S. Department of Energy
Office of Nuclear Energy

DISCLAIMER

This information was prepared as an account of work sponsored by an agency of the U.S. Government. Neither the U.S. Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness, of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. References herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the U.S. Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the U.S. Government or any agency thereof.

Business Case Analysis for Artificial Intelligence- Large Language Model Technology Integration

**Casey Kovesdi, Ahmad Al Rashdan
Idaho National Laboratory**

**Alex Tylecote, Morgan Schadeegg, Luke Martin
ScottMadden**

June 2025

**Idaho National Laboratory
Light Water Reactor Sustainability
Idaho Falls, Idaho 83415**

<http://lwrs.inl.gov>

**Prepared for the
U.S. Department of Energy
Office of Nuclear Energy
Under DOE Idaho Operations Office
Contract DE-AC07-05ID14517**

Page intentionally left blank

ABSTRACT

Artificial intelligence (AI) assisted processes are expected to enhance operational efficiency and improve decision-making, supporting the long-term economic viability of nuclear power plants. However, detailed business analyses of AI-generated cost savings are rarely performed. Given the recent industry interest in large language models (LLMs), the U.S. Department of Energy Light Water Reactor Sustainability Program has conducted a comprehensive business case analysis of LLM AI implementation in nuclear plant engineering workflows.

The research employed three complementary business case approaches to evaluate the impact of an LLM, using three representative engineering processes as use cases: boric acid corrosion evaluations, maintenance rule evaluations, and 10 CFR 50.59 screenings. Through detailed workload analyses and structured interviews, the study quantified significant efficiency improvements of 11% to 59% across these processes.

The research further considers how these efficiency gains could translate into tangible reliability improvements through enhanced engineering capacity. An analysis of historical plant trip data indicates that enabling engineers to focus on proactive reliability activities could provide substantial financial benefits through avoided outages, potentially generating greater value than the direct efficiency improvements alone.

By documenting successful applications, implementation challenges, and strategic opportunities, this research provides nuclear utilities with a practical framework for evaluating the value of AI technology in supporting long-term operations through advanced digital technologies.

Page intentionally left blank

EXECUTIVE SUMMARY

The nuclear industry currently faces several converging challenges and opportunities that make artificial intelligence (AI) adoption particularly timely. On the one hand, the sector faces mounting pressure; a rapidly aging workforce is accelerating knowledge attrition, while decades of accumulated, unstructured information remain locked in outdated document systems, complicating knowledge transfer and operational efficiency. On the other hand, surging electricity demand, fueled by AI-driven technologies and widespread electrification, has sparked renewed interest in nuclear power as a reliable energy source with strong energy security attributes. Finally, AI technologies have reached a critical tipping point in capability and accessibility, making them viable solutions for nuclear applications.

In response to these challenges, the Light Water Reactor Sustainability (LWRS) Program has undertaken a comprehensive assessment of large language model (LLM) implementation in nuclear plant engineering workflows. This research, conducted in partnership with Vistra at their Comanche Peak Nuclear Power Plant, provides a rigorous, data-driven evaluation of AI technologies' potential to enhance operational efficiency, improve engineering productivity, and contribute to plant reliability.

This research used a systematic methodology to assess AI implementation at Comanche Peak, specifically targeting LLMs. Researchers began by identifying and mapping potential LLM use cases across engineering functions at the plant. Three representative engineering processes were selected for detailed analysis as indicative use cases: boric acid corrosion (BAC) evaluations, maintenance rule evaluations, and 10 CFR 50.59 screenings. The team conducted structured interviews with engineers at multiple experience levels and performed a comprehensive workload analysis to quantify time savings and efficiency gains for the three use cases. This was validated through in-person working sessions.

The analysis revealed significant potential for time savings across the use cases studied. BAC evaluations showed an 11%–22% time reduction, maintenance rule evaluations demonstrated 19%–39% time savings, and 50.59 screenings exhibited the most substantial improvement with 38%–59% time reduction. Projected annual time savings across these three use cases range from 235 to 405 hours, representing an approximate financial value of \$23,000–\$39,000 per year. These estimates are likely conservative for several reasons: they reflect the current maturity of the LLM and broader AI models, are based on initial user proficiency levels, were derived primarily from engineer estimates where many participants had limited

direct experience with the tool, and importantly, Comanche Peak was still in the pilot phase of implementation with full deployment in Q3 of 2025.

The study employed three complementary business case approaches to assess AI implementation value. The first approach analyzed three representative engineering processes in detail to establish a benchmark for efficiency improvements. This analysis at this project maturity found that each engineering use case gained approximately 110 annual hours in efficiency (ranging from 78 to 135 hours), valued at roughly \$10,500 annually per use case. The analysis shows that, with approximately 15–25 applications, the LLM implementation creates a positive financial return. At 20 use cases, the net present value (NPV) ranges from -\$1.3M to +\$0.6M depending on efficiency assumptions, while at 50 use cases, the value grows to between \$0.4M and \$3.4M.

The second business case approach explores how these efficiency improvements translate to overall engineering productivity. The analysis estimates that documentation and research activities amenable to LLM assistance constitute approximately 30%–50% of an engineer’s workload. With LLM-enabled efficiency improvements of 25%–40% on these tasks, this could translate to overall productivity gains of 7.5%–20% across the engineering function. For a typical plant with 60 engineers, this represents between 8,640 and 23,040 annual hours in enhanced capacity, valued at \$833k to \$2.2M annually. When modeled over a 10 year period, this approach yields NPV values ranging from \$3.8M in the low case to \$14.8M in the high case.

The third approach examines how this enhanced engineering capacity can deliver reliability improvements through more proactive technical work. An analysis of historical plant trip data from a specific turbine control system issue revealed multiple events causing substantial lost generation over a 4 year period, with an average event value of \$3.8M. By enabling engineers to focus on activities such as enhanced system health monitoring, technical backlog reduction, and strategic equipment upgrades, LLM implementation can help prevent equipment failures that lead to forced outages. For example, one senior engineer cited a long-standing issue with a turbine control system that had caused several trips over a period of nearly 20 years. If an enhanced engineering capacity from LLM implementation could prevent just 1–3 plant trips over a 10 year period, the resulting NPV would range from \$0.4M to \$5.5M.

The study also highlighted important implementation considerations. Accuracy and quality concerns emerged as a primary consideration, with potential for LLM-generated inaccuracies requiring careful verification. Researchers identified risks of over-reliance on LLM and broader AI tools, particularly regarding the preservation of engineering competency and development of basic analysis and problem-solving skills among younger engineers. Organizational change management presented another important challenge, with cultural adaptation and technology integration requiring thoughtful approaches with experienced and junior engineers alike.

The long-term value of these efficiency improvements ultimately lies in broader business transformation. These include creating a greater centralization of engineering expertise, enabling more efficient campaign maintenance strategies, and shifting toward a more agile workforce model. The immediate focus on freeing up time for higher value activities establishes the foundation for broader

organizational changes that can eventually evolve staffing models in line with natural attrition, without disrupting operations or losing critical expertise. These organizational changes represent the final stage of value realization, completing the journey from targeted efficiency improvements to comprehensive business transformation.

This research provides a foundation for future studies that could further validate benefits through post-implementation data collection, quantify impacts across additional use cases in operations, maintenance, and licensing functions, and develop methodologies to measure reliability improvements attributable to enhanced engineering capacity. The findings suggest that, while limited LLM implementations may not show positive financial returns when analyzed in isolation, comprehensive enterprise deployments that leverage the same LLM tool across numerous applications create substantial economic and strategic value for nuclear utilities.

Page intentionally left blank

ACKNOWLEDGMENTS

The authors wish to thank the Light Water Reactor Sustainability Program for funding this work. We especially thank Vistra and the Comanche Peak Nuclear Power Plant participants and engineers who provided interviews and invaluable insights essential to the development of this business case analysis.

Page intentionally left blank

CONTENTS

1.	INTRODUCTION.....	3
1.1	Introduction to the LWRS Program and ION.....	3
1.2	Project Background	3
1.3	Current Project.....	4
1.3.1	Specific Objectives	4
1.4	Approach.....	5
1.5	Scope: Focus on LLMs	6
2.	USE CASE SELECTION	6
2.1	Overview of Key LLM Capabilities	6
2.1.1	Engineering Report Autodrafting	6
2.1.2	General Engineering Chatbot	7
2.1.3	Historical Failure Analysis	7
2.2	Key Criteria for Use-Case Selection.....	8
2.3	Process of Identifying and Evaluating Use Cases for Inclusion.....	9
2.4	Use Cases Considered for Inclusion	9
2.4.1	Use Cases Related to Other Plant Functions	10
2.5	Selected Use Cases and Rationale for Selection.....	11
2.5.1	BAC Evaluations	11
2.5.2	Maintenance Rule Evaluations	11
2.5.3	10 CFR 50.59 Screening.....	12
3.	WORKLOAD ANALYSIS.....	12
3.1	Introduction.....	12
3.1.1	Workload Analysis Methodology.....	12
3.1.2	Interview Participants	13
3.1.3	AI Impact Rating	13
3.2	BAC Evaluation.....	14
3.2.1	Manual Process Time Assessment.....	14
3.2.2	Artificial-Intelligence-Assisted Impact Assessment.....	15
3.2.3	Artificial-Intelligence-Assisted Time Impacts	17
3.3	Maintenance Rule Evaluations	17
3.3.1	Manual Process Time Assessment.....	17
3.3.2	Artificial-Intelligence-Assisted Process	19
3.3.3	AI-Assisted Time Impacts	23
3.4	10 CFR 50.59 Screenings	24
3.4.1	Manual Process Time Assessment.....	24
3.4.2	AI-Assisted Process	25
3.4.3	AI-Assisted Time Impacts	27
3.5	Patterns of High-Impact LLM Applications in Engineering Tasks.....	28
4.	BUSINESS CASE RESULTS, BENEFITS, AND CHALLENGES.....	28
4.1	Summary of Workload Efficiency Benefits.....	28
4.1.1	AI Efficiency Gains: Estimated Percentage and Time Improvements	29

4.1.2	Annual Engineering Evaluations	30
4.1.3	Total Annual Savings from Three Example Use Cases.....	31
4.1.4	Inherent Conservatism of Results	31
4.2	Other Use Cases	32
4.2.1	Engineering Use Cases	32
4.2.2	Regulatory Compliance	32
4.2.3	Operations Use Cases	33
4.2.4	Maintenance Use Cases	33
4.2.5	Other Plant Functions	33
4.3	Indirect Benefits and Long-Term Strategic Value.....	34
4.3.1	Overview.....	34
4.3.2	Shifting from Reactive to Proactive Engineering.....	34
4.3.3	Enhanced System Walkdowns and Field Presence	34
4.3.4	Long-Term Plant Reliability Improvements.....	35
4.3.5	Improved Knowledge Transfer and Mentoring	35
4.4	Risks and Challenges	36
4.4.1	LLM Tool Accuracy and Reliability Concerns	36
4.4.2	Tool Over-Reliance and Impact on Engineer Competency	36
4.4.3	Cultural Adaptation Challenges and Managing Organizational Change.....	36
4.5	Business Case Approaches and Economic Analysis	37
4.5.1	Overview.....	37
4.5.2	Use Case Efficiency Benchmark	37
4.5.3	Implementation Costs	38
4.5.4	Business Case Approach 1: Expanding Use Cases Across the Plant	39
4.5.5	Business Case Approach 2: Engineer Value Framework.....	40
4.5.6	Business Case Approach 3: Avoided Cost Reliability Benefit Framework	41
4.5.7	Future Value Realization: Plant- and Fleet-Level Business Transformation	43
5.	RECOMMENDATIONS AND NEXT STEPS	44
5.1	Use-Case Expansion and Opportunities for Further Research	44
5.1.1	Post-Implementation Business Case Validation.....	44
5.1.2	Validating the Engineer Value Framework	44
5.1.3	Quantifying Benefits for Additional Engineering Use Cases.....	44
5.1.4	Quantifying Benefits in Other Plant Functions	44
5.1.5	Quantifying Reliability and Avoided Cost Benefits.....	44
5.1.6	Measuring the Impact of AI Through Alternative Success Metrics	45
5.2	Fleet-Level Business Transformation Strategies	46
5.2.1	Breaking Down Organizational Silos	46
5.2.2	Organizational Enablement and Governance	46
5.2.3	Knowledge Sharing and Best Practices	46
5.2.4	Workforce Evolution and Champion Development	46
5.2.5	Scaling Strategies and Resource Planning.....	47
5.2.6	Balancing Standardization with Customization.....	47
6.	REFERENCES.....	47
	Appendix A DATA COLLECTION AND STATISTICAL MODELING METHODOLOGY	49
	A-1. DATA COLLECTION AND ANALYSIS PROCESS	51
	A-2. MONTE CARLO SIMULATION DEVELOPMENT	51

A-3. MANUAL TIME DISTRIBUTION	51
A-4. SPECIAL CONSIDERATIONS FOR MAINTENANCE RULE EVALUATIONS	52
A-5. AI IMPACT FACTOR CORRELATION	52
A-6. ANNUAL IMPACT CALCULATION USING NESTED MONTE CARLO SIMULATION	53

FIGURES

Figure 1. Maintenance rule evaluation annual count by task.	19
Figure 2. Breakdown of worker hours by task for maintenance rule evaluations for manual process.	19
Figure 3. Worker hours per use case without LLM.	28
Figure 4. AI efficiency gains: Estimated percentage improvements.	29
Figure 5. LLM efficiency gains: Estimated time improvements.	30
Figure 6. Annual engineering evaluations for two-unit plant.	30
Figure 7. Total annual worker hours savings.	31

TABLES

Table 1. AI impact rating scale to percent time savings.	13
Table 2. BAC evaluation manual process time.	14
Table 3. Boric acid evaluation: Summary table of AI-assisted time impacts.	17
Table 4. Maintenance rule evaluation manual process time.	18
Table 5. Maintenance rule evaluation: Summary table of AI-assisted time impacts.	23
Table 6. 50.59 screenings manual process time.	25
Table 7. 50.59 screenings: Summary table of AI-assisted time impacts.	27
Table 8. NPV from worker hour savings from three use cases.	31
Table 9. Annual Efficiency Savings	39
Table 10. NPVs and IRRs with enterprise implementation.	40
Table 11. Engineer value framework: Efficiency gains and financial benefits.	41
Table 12. Reliability event historical data from turbine control system issues.	42
Table 13. Reliability impact scenario modeling results.	42

Page intentionally left blank

ACRONYMS

AD	Applicability Determination
AI	Artificial Intelligence
BAC	Boric Acid Corrosion
CFR	Code of Federal Regulations
DBD	Design Basis Document
FF	Functional Failure
FSAR	Final Safety Analysis Report
ICES	INPO Consolidated Events Database
INPO	Institute of Nuclear Power Operations
ION	Integrated Operations for Nuclear
LLM	Large Language Model
LWRS	Light Water Reactor Sustainability
MBSE	Model-Based Systems Engineering
MPFF	Maintenance Preventable Functional Failure
MRRP	Maintenance Rule Review Panel
NPV	Net Present Value
NRC	Nuclear Regulatory Commission
OT	Operational Technology
PCN	Procedure Change Notice
RMPFF	Repetitive Maintenance Preventable Functional Failure
SSCs	Structures, Systems, and Components
UFSAR	Updated Final Safety Analysis Report

Page intentionally left blank

DEFINITIONS

- 10 CFR 50.59: A regulatory process established by the Nuclear Regulatory Commission that defines criteria for determining whether changes, tests, or experiments at a nuclear facility require prior Nuclear Regulatory Commission approval. This process allows licensees to make certain changes without formal license amendments if specific safety criteria are met.
- Applicability Determination (AD): The process for determining which regulatory reviews and evaluations are required for a proposed activity, design change, or procedure modification at a nuclear facility.
- Artificial Intelligence (AI): Refers to computer systems and technologies designed to perform tasks that typically require human intelligence, including learning, reasoning, problem-solving, and decision-making.
- Artificial Narrow Intelligence: Denotes AI systems that are designed to perform a narrow, specific task rather than functioning as general-purpose intelligence.
- Boric Acid Corrosion (BAC): A technical evaluation process used to assess the effects of boric acid accumulation on plant components, ensuring that material degradation does not compromise structural integrity.
- Condition Based Maintenance: A proactive maintenance strategy that relies on continuous monitoring of equipment conditions to determine the optimal timing for maintenance, thereby reducing unnecessary maintenance actions.
- Final Safety Analysis Report (FSAR): A comprehensive technical document that outlines the design, operational procedures, and safety features of a nuclear power plant, serving as a key reference for ensuring regulatory compliance.
- Functional Failure (FF): In the context of maintenance rule compliance, a failure of a structure, system, or component (SSC) that renders it incapable of performing its intended function as defined in the maintenance rule scoping documents. This serves as the initial determination in the maintenance rule evaluation process.
- INPO Consolidated Events Database (ICES): A database maintained by the Institute of Nuclear Power Operations that collects, analyzes, and shares information about equipment failures, human performance issues, and other events at nuclear power plants to facilitate industrywide learning and improvement.
- Integrated Operations for Nuclear (ION): A framework for leveraging centralized digital infrastructure, advanced analytics, and improved operational processes to optimize the performance and safety of nuclear power plants.
- Large Language Model (LLM): A type of AI that has been trained on massive datasets to generate human-like text responses. They are particularly effective in automating document creation and natural language processing tasks.
- Light Water Reactor Sustainability (LWRS): A program dedicated to improving economics and extending the operational lifetime of existing nuclear power plants through technological and operational enhancements.
- Maintenance Preventable Functional Failure (MPFF): An FF that could have been prevented through appropriate maintenance activities. These failures are particularly significant for maintenance rule compliance as they indicate potential inadequacies in maintenance strategies or implementation.

Model-Based Systems Engineering (MBSE): An engineering methodology that utilizes digital models as the primary medium for information exchange among teams, enhancing design, analysis, and communication.

Net Present Value (NPV): A financial calculation that aggregates future cost savings or benefits, discounted to their present value, to assess the overall economic viability of an investment.

Operational Technology: Encompasses the hardware and software systems employed to monitor and control industrial processes, including those vital to nuclear plant operations.

Owner: The nuclear utility or operating company that holds the operating license for a nuclear power plant and is responsible for its safe operation, regulatory compliance, and business decisions. In this report, “Owner” refers to Vistra, the entity implementing the LLM technology at Comanche Peak Nuclear Power Plant and making strategic decisions about its deployment across their nuclear fleet.

Repetitive Maintenance Preventable Functional Failure (RMPFF): A failure that occurs on the same component or component type, performing the same function, for the same maintenance-related cause within a 24-month period. They often trigger additional regulatory oversight and may require systems to be placed in (a)(1) status under the maintenance rule.

Work Reduction Opportunity: Identifies processes within nuclear operations where automation or process improvements can significantly reduce manual labor, leading to efficiency gains.

Business Case Analysis for Artificial Intelligence- Large Language Model Technology Integration

1. INTRODUCTION

1.1 Introduction to the LWRS Program and ION

The U.S. Department of Energy Light Water Reactor Sustainability (LWRS) Program conducts research and develops technologies and demonstrations that support the long-term operation of the nation's nuclear power plants. The program addresses technical, regulatory, and economic challenges to ensure the continued safe and economically viable operation of these vital energy assets. A key initiative within the LWRS Program is the Plant Modernization Pathway, which aims to develop technologies, methods, and guidance to help nuclear plants implement digital instrumentation and control systems, advanced automation, and data analytics solutions while maintaining safety and reliability requirements.

The Integrated Operations for Nuclear (ION) effort was established by the LWRS Program to help nuclear operators evaluate technological advancements to transform business models and work processes [1]. ION focuses on the integration of digital infrastructure, advanced analytics, and improved human performance to create sustainable cost reductions through work reduction opportunities. These work reduction opportunities represent areas where technological innovations can eliminate unnecessary work, automate routine tasks, and enhance decision-making through a more effective utilization of plant data and operational information. The ION framework provides a structured approach for nuclear utilities to identify, evaluate, and implement digital technologies that can reduce operational costs while maintaining or enhancing safety and reliability performance [2].

More recently, potential impacts of automation and artificial intelligence (AI) in the nuclear industry across several domains have come under particular focus for research. Recent research has examined how AI technologies can enhance operations, maintenance, and engineering functions while ensuring alignment with the industry's strict safety and regulatory requirements [3]. Nuclear energy AI applications of interest for the LWRS Program include predictive maintenance [4], knowledge preservation, operational decision support, and enhanced data analytics for system monitoring [5]. These technologies show particular promise for addressing industry challenges, such as workforce demographics, operational costs, and the need for more efficient knowledge transfer and decision-making processes.

In this context, the LWRS Program has recognized the need for practical, data-driven assessments of emerging AI technologies in actual nuclear plant environments. By understanding how these technologies perform in real-world settings, researchers can develop more effective implementation strategies and provide nuclear operators with credible business cases to support AI adoption decisions. This foundational research serves as the basis for exploring targeted applications of AI within the broader ION framework, connecting theoretical capabilities to practical operational improvements that deliver measurable value to nuclear utilities.

1.2 Project Background

Beginning in 2023, Vistra (the plant owner) began exploring AI applications for engineering functions at Comanche Peak Nuclear Power Plant (referred to as the "Reference Plant" in the remainder of this document), motivated by the significant time engineers spent on documenting findings rather than technical analysis. A critical early decision was to focus implementation on engineering documentation tasks where AI could deliver immediate value while maintaining existing quality control processes and regulatory compliance.

This implementation approach involved preprocessing over 20 years of human-verified historical data, which included removing irrelevant information such as signatures, extraneous formatting, and nonessential content from the training data to leave only essential technical information. The team aggregated approximately 1.4 million documents containing an estimated 20 million pages of technical content for training the LLM model. The use of human-verified training data was essential to prevent the AI from learning and incorporating erroneous information in the text of reports generated by the tool.

The implementation focused on a seamless integration with existing engineering workflows, requiring minimal procedural changes while providing significant efficiency benefits. This structured approach to AI implementation emphasized practical utility and regulatory compliance while addressing the unique challenges of applying AI technology in the nuclear industry context, establishing the groundwork for the current business case analysis.

An important design principle of this LLM implementation has been its positioning as a human-assisted tool rather than a replacement for engineering expertise. The ultimate deployment goal is to enhance engineer efficiency by automating routine documentation tasks, enabling engineers to focus their specialized knowledge on higher value tasks. It is not intended to remove the essential role of qualified engineers in verification, technical judgment, and final approval.

1.3 Current Project

Building on the initial implementation of LLM technology at the Reference Plant, LWRS researchers from Idaho National Laboratory (INL) identified an opportunity to conduct a comprehensive assessment of the business case for LLM integration in engineering functions. An initial assessment in 2024 identified promising applications for LLM in reducing engineering workload and achieving significant time savings in routine engineering tasks, such as evaluation documentation and technical research [6].

The LWRS Program expanded this research initiative to develop a comprehensive business case for integrating LLM technologies into engineering functions. This effort focused on measuring both quantitative efficiency gains and qualitative benefits achieved through LLM-assisted workflows. The project team established a methodical approach to analyze selected use cases, compare traditional and LLM-assisted processes, and develop a comprehensive business case that could inform future technology investment decisions.

This report documents the findings of this collaborative effort, providing insights into the practical application of LLM technology in nuclear engineering functions. The analysis presented here contributes to the broader LWRS Program objective of enhancing the economic competitiveness of existing nuclear plants while maintaining safety and reliability. By documenting both successful applications and implementation challenges, this research provides a foundation for nuclear operators considering similar technology deployments across their organizations.

It should be noted that, at the time of this assessment, the Reference Plant was still in the pilot phase of LLM implementation, with full deployment planned but not yet implemented. This limited scope of adoption may have influenced the potential conservatism observed in user feedback.

1.3.1 Specific Objectives

This project was designed with several specific objectives to ensure a comprehensive evaluation of the LLM implementation at the Reference Plant:

- **Identify and Prioritize Use Cases:** Select representative engineering use cases that demonstrate the potential for LLMs to create efficiency gains and quality improvements in nuclear operations, focusing on activities with high documentation burden and standardized formats.

- **Analyze Workforce Impacts:** Establish baseline task durations and quality metrics for both novice and experienced engineers, comparing traditional workflows with LLM-assisted approaches to quantify the actual time savings and performance improvements.
- **Develop Comprehensive Business Case:** Quantify both direct efficiency gains (time savings) and indirect benefits (improved documentation quality, knowledge preservation, standardization) to establish the total value proposition for LLM implementation.
- **Identify Implementation Considerations:** Document the technical, organizational, and cultural factors that influence successful LLM adoption, including training requirements, change management approaches, and quality control processes.
- **Create Scalable Assessment Framework:** Develop a methodology that can be applied to evaluate similar LLM implementations across other nuclear plants and functional areas, providing a foundation for broader industry adoption.
- **Establish Future Research Directions:** Identify opportunities for expanded application of LLM technology beyond the initial engineering functions, supporting the long-term vision of integrated operations for nuclear power plants.

By achieving these objectives, the project aims to provide practical, actionable insights that nuclear operators can use to make informed decisions about LLMs and broader AI investments while establishing a foundation for future research into advanced applications of these technologies in nuclear plant operations.

It is important to emphasize that, while this study focuses on three specific engineering use cases, the true transformative potential of LLM technology lies in its broad implementation across dozens of engineering activities. When deployed comprehensively, the cumulative time savings can create a generalized efficiency improvement for engineering staff, enabling them to redirect significant portions of their workday from administrative documentation to higher value technical activities.

1.4 Approach

This study employed a systematic approach to assess the business case for implementing LLM technology in nuclear engineering workflows at the Reference Plant. The approach consisted of the following steps:

1. The research began by examining the LLM implementation at the Reference Plant, documenting the available functionality, technical architecture, and potential applications across engineering workflows. This initial assessment provided a foundational understanding of how the technology currently functions and its capabilities.
2. Through interviews with plant personnel and a review of current implementations, the complete range of potential use cases were identified for LLM technology within the engineering function. This mapping exercise yielded nine distinct engineering activity categories where LLM assistance could provide value.
3. From the identified totality of use cases, three representative activities were selected for detailed workload analysis. These selections—boric acid corrosion (BAC) evaluations, maintenance rule evaluations, and 50.59 screenings—were chosen based on implementation maturity, frequency of occurrence, documentation structure, and potential for measurement.
4. For each selected use case, structured workload analyses were conducted to quantify the specific percentage efficiency gains and time savings achieved through LLM assistance. This analysis disaggregated each process into component steps to identify precisely where and how an LLM creates value.

5. A statistical model was developed to analyze the efficiency gains across different evaluation types and complexity levels. This approach accounted for the natural variability in engineering tasks and captured how LLM assistance provides different levels of benefit depending on task complexity. See Appendix A for more information.
6. Using the quantified time savings, the net present value (NPV) of the LLM implementation was modeled for the three use cases. This model incorporated both direct efficiency gains and implementation costs to determine the overall financial value proposition.
7. The research extrapolated findings from the three analyzed use cases to develop projections for the broader universe of potential applications. This provided a more comprehensive view of the potential organization-wide impact of LLM technology.
8. Beyond direct time savings, several indirect benefits were identified that, while more difficult to quantify, represent significant potential value. These indirect benefits point to opportunities for future research to develop more comprehensive valuation methodologies.

1.5 Scope: Focus on LLMs

This study specifically evaluates LLM capabilities and their application in nuclear plant engineering workflows. LLMs represent a category of AI that excels at understanding and generating human-like text based on patterns learned from vast amounts of training data, making them particularly well-suited for documentation-heavy tasks. It is important to note that this study does not assess other AI technologies that could also benefit nuclear operations, such as traditional rule-based natural language processing for categorizing work orders or matching conditions to historical records, computer vision for equipment inspections, predictive analytics for maintenance optimization, or expert systems for safety-critical decision support.

2. USE CASE SELECTION

2.1 Overview of Key LLM Capabilities

An LLM at the Reference Plant offers a suite of capabilities specifically designed to address common challenges in engineering workflows. Unlike general-purpose commercial AI systems, this implementation has been tailored to the requirements of the nuclear industry. The system leverages an extensive knowledge base of plant-specific documentation combined with nuclear industry standards and regulatory requirements to provide targeted assistance across multiple engineering functions.

2.1.1 Engineering Report Autodrafting

The report autodrafting capability represents a core function of the LLM implementation at the Reference Plant. This functionality enables the automatic generation of structured draft reports for various engineering evaluations based on the site's historical documentation, technical specifications, and industry regulatory guidance.

Unlike generic commercially available LLMs, the LLM has been trained specifically on plant documents from the Reference Plant's document repository. This plant-specific training allows the generation of documents that align with existing protocols, procedures, and component-specific information.

The LLM implementation at the Reference Plant utilizes an architecture that enables different approaches to report generation.

- A "one-shot" generation, where a complete document is created from a single prompt, and more sophisticated "agentic workflows" that break complex tasks into manageable steps.

- For complex tasks, like 50.59 screenings, a “manager model” that orchestrates different expert personas, including final safety analysis report (FSAR) researchers, technical specification researchers, and domain-specific drafting models. This approach mimics the natural workflow of experienced engineers by systematically gathering the required information before drafting documentation.

This serves as a starting point for engineers, who can then focus on technical verification and analysis rather than administrative documentation. As described by the plant staff, the LLM functions like “a decent intern”—providing valuable initial drafts while still requiring qualified review. This saves significant time, particularly for routine evaluations that follow standard formats and protocols.

A key feature of the autodrafting capability is attributing sources, which provides engineers with references to the original documents from which information was pulled. This allows engineers to verify the accuracy of generated content and trace information back to approved plant documentation.

2.1.2 General Engineering Chatbot

The general engineering chatbot capability provides a conversational interface for engineers to query plant-specific information and receive contextually relevant responses based on the plant’s document repository. While commercially available LLMs provide general knowledge, an engineering chatbot has been trained specifically on plant documentation, giving it focus on site-specific inquiries.

A chatbot provides several key functions that support engineering work:

- **Information retrieval:** Engineers can ask direct questions about plant systems, components, procedures, or historical events without needing to navigate through multiple document repositories. For example, an engineer can ask “What room is CCP-101 in?” or “What are the technical specification functions of this component?” and receive specific answers with references to source documentation.
- **Document search and summarization:** The system can quickly locate relevant information across thousands of documents, summarizing findings and providing direct references to source materials. This significantly reduces time spent searching through documentation, which can be a large portion of engineering evaluations.
- **Historical analysis:** Engineers can request information about previous failures, maintenance history, or operational events related to specific components or systems, helping to inform current decision-making with historical context.
- **Technical guidance:** The chatbot can provide procedural guidance, regulatory requirements, and technical specifications relevant to specific engineering tasks based on its knowledge of plant documentation and nuclear industry standards.

According to feedback from engineers at the Reference Plant, a chatbot is particularly valuable for rapidly accessing and synthesizing information across multiple document repositories that would typically require extensive manual searches or consultations with experienced staff. While engineers value the ability to rapidly access information, they also noted the importance of verification, as accuracy issues could occur with chatbots. These concerns, which are further detailed in Section 4.4.1, emphasize the continued need for engineering judgment when using LLM-generated responses.

2.1.3 Historical Failure Analysis

The historical failure analysis capability leverages the LLM’s ability to analyze and synthesize information from past operational experience, maintenance records, and condition reports to provide insights into equipment failure modes, root causes, and maintenance strategies. This extends beyond simple document retrieval to provide analytical support for engineering decisions:

- **Component failure history retrieval:** Engineers can quickly access comprehensive failure histories for specific components or systems across the plant’s operational history. For example, an engineer could ask, “Give me the failure history of bearing leaks on the main feed pump,” and receive a chronological summary of relevant events.
- **Pattern recognition:** The LLM can identify recurring issues or common failure modes across similar components or systems by analyzing patterns in historical maintenance records and condition reports.
- **Root cause analysis support:** By synthesizing information from previous similar failures, the system can provide engineers with potential contributing factors and failure mechanisms to consider during root cause analyses.
- **Maintenance strategy recommendations:** Based on historical performance data, the system can suggest maintenance approaches that have proven effective for similar components or conditions.

This capability is particularly valuable for supporting engineering decisions by providing rapid access to operational experience that might otherwise be difficult to compile manually. It functions as an institutional memory tool, making the collective experience of the organization readily available to engineers regardless of their personal experience level.

This capability helps address one of the significant challenges in the nuclear industry—knowledge retention and transfer as experienced personnel retire. While the LLM does not replace engineering judgment, it provides a valuable tool for less experienced engineers to quickly access relevant historical context.

Unlike the more structured report autodrafting capability, the historical failure analysis function operates more effectively through conversational queries, allowing engineers to explore failure histories iteratively and follow different analytical paths based on initial findings.

2.2 Key Criteria for Use-Case Selection

Several critical criteria guide the selection of use cases for detailed analysis:

- **LLM maturity:** The primary criterion was the maturity of the LLM functionality for specific applications. Selected use cases needed to align with capabilities that were well-developed.
- **Engineering application:** The second criterion is the use case’s direct application to engineering workflows and processes. Selected cases needed to demonstrate clear value in supporting core engineering functions.
- **Measurable impact:** The third criterion focuses on clear efficiency potential. Selected use cases must demonstrate obvious opportunities for time savings while maintaining or improving quality. Early results showing 15–20 minute draft generation times versus hours of traditional work illustrate the type of efficiency gains sought. Selected use cases needed to demonstrate measurable outcomes that could be documented and analyzed.
- **Broad applicability:** The fourth criterion considered the potential for broader application across the nuclear industry. Selected use cases needed to represent common activities that other facilities might benefit from implementing.
- **Documentation availability:** The final criterion was the availability of comprehensive documentation about the implementation and results. Selected use cases needed sufficient documentation to support a thorough analysis and validation of reported benefits.

2.3 Process of Identifying and Evaluating Use Cases for Inclusion

The identification and evaluation process described in this section focused specifically on selecting a small, representative set of LLM use cases at the Reference Plant for a detailed workload analysis. The project team needed to identify focused use cases that would meet the criteria mentioned in Section 2.2.

The identification process began with gathering input through initial discussions with relevant plant leadership on the status of the LLM implementation at the Reference Plant. These discussions provided critical insights into which implemented use cases were the most promising and would be suitable candidates for quantitative assessment. A review of reporting from a previous LWRs Program initiative with the Reference Plant was also conducted, providing an overview of the initial LLM model deployment at the plant [6].

To build a comprehensive understanding of candidate use cases for assessment, a thorough review was conducted of relevant Reference Plant documentation, including procedures, forms, and examples of actual engineering evaluations. This documentation provided essential context about existing workflows and identified which activities had sufficient documentation to support before and after comparisons.

The selection process for workload analysis candidates was coordinated with senior leaders in engineering and other leaders supporting LLM engineering initiatives across the utility's fleet. Their input ensured that the selected assessment use cases would provide insights applicable to high-priority activities throughout the organization, maximizing the value of the analysis findings.

To incorporate the broader industry perspective, the project team leveraged insights from multiple utilities who worked to establish consensus on major use cases. This forum helped identify which use cases would have the greatest applicability across the industry, ensuring that findings from the Reference Plant workload analysis would provide value to the broader nuclear community. Additionally, regular roundtable discussions with nuclear industry leaders on LLM implementation, facilitated by ScottMadden, provided ongoing insights into which use cases were showing the most promise across different organizations, helping to validate the selection of assessment candidates.

2.4 Use Cases Considered for Inclusion

Based on the identification process outlined in Section 2.3, the project team considered a range of engineering use cases and use cases from other plant functions that could benefit from LLM assistance and would be suitable for quantitative workload analysis. The following use cases were considered as candidates for detailed assessment:

- **BAC Evaluations:** Formal assessments that document boric acid deposits on plant components, analyze potential corrosion impacts, determine structural integrity implications, and specify required corrective actions. These evaluations follow standardized templates and require an integration of component-specific information with technical guidance on corrosion mechanisms.
- **Maintenance Rule Reviews:** Documentation processes required under 10 CFR 50.65 that evaluate whether equipment failures constitute functional failures, if failures were preventable, and if established performance criteria have been exceeded. These reviews involve significant research into component history and technical specifications.
- **(10 CFR) 50.59 Screening:** An initial evaluation process designed to determine if proposed changes to a facility's procedures, tests, or modifications require prior Nuclear Regulatory Commission (NRC) approval by systematically reviewing relevant technical documentation. These screenings involve complex regulatory analyses and detailed documentation of engineering judgments.
- **License Renewal Documentation:** Comprehensive technical evaluations, aging management programs, and analyses prepared to support extending a nuclear plant's operating license beyond the

original term. This documentation requires extensive research across design basis documents (DBDs) and industry standards.

- **Operating Experience Analyses:** Systematic reviews of plant-specific and industrywide events to identify lessons learned, emerging trends, and potential vulnerabilities affecting plant operations. These analyses involve synthesizing information from diverse sources to identify applicable insights.
- **Simple Failure Evaluations:** Targeted assessments of individual component failures that document the circumstances, impacts, immediate corrective actions, and recommendations for preventing recurrence. These evaluations require an integration of observed conditions with component design information.
- **Work Order Development:** The formal process of creating and authorizing documents that direct the maintenance, testing, and modification activities with all necessary specifications. This process requires an accurate incorporation of component-specific information and procedural requirements.
- **Operational Decision-Making Meeting Documentation:** Records of meetings that capture evaluations of emerging plant conditions, including risk assessments and action plans to address off-normal events. These documents require a structured presentation of technical information to support operational decisions.
- **Cause Analysis Documentation:** Structured processes for documenting failure analyses to identify root causes and contributing factors of equipment failures or operational events. These analyses involve systematic evaluation methodologies and require the careful documentation of evidence and conclusions.
- **System Health Reports:** Periodic assessments of system performance, reliability, and maintenance effectiveness requiring the compilation of system data, trend analysis, and documentation of improvement plans.
- **Equipment Qualification Documentation:** Technical documents validating that equipment can function under specified environmental conditions through researching design specifications and environmental compatibility validation.
- **Operability Evaluations:** Assessments determining if degraded conditions affect a component's ability to perform its function, requiring research of technical specifications and development of compensatory measures.

2.4.1 Use Cases Related to Other Plant Functions

The Reference Plant staff also identified several high-potential applications beyond core engineering functions. These use cases were considered for inclusion but were deferred to subsequent implementation phases to allow for proper focus on engineering documentation processes. Engineers and operations staff indicated these applications would likely be prioritized in future expansion of the LLM implementation:

- **Operational Decision-Making Meeting Documentation:** Records capturing evaluations of emerging plant conditions, risk assessments, and action plans. Key activities include documenting technical discussions, capturing risk assessments, and recording action assignments.
- **Cause Analysis Documentation:** Structured processes for documenting failure analyses to identify root causes and contributing factors. This involves applying systematic evaluation methods, documenting evidence, and developing conclusions.
- **Maintenance Strategy Reviews:** Periodic assessments of maintenance approaches for specific components or systems. These reviews require researching maintenance history, evaluating performance trends, and documenting strategy changes.

- **Regulatory Compliance Reporting:** Documentation demonstrating adherence to regulatory requirements and commitments. This includes researching regulatory standards, collecting compliance evidence, and documenting conclusions.
- **Fire Impairment Documentation:** Records documenting temporary fire protection system impairments and compensatory measures. This involves identifying affected systems, determining required compensatory measures, and documenting compliance.
- **Procedure Revisions:** Updates to plant procedures based on operational experience, regulatory changes, or design modifications. Key activities include researching technical bases, identifying affected procedure sections, and documenting change justification.
- **Plant Data Trending Analysis:** Assessment of operational data to identify patterns and potential degradation. This requires compiling historical data, identifying trends, and documenting recommendations.

2.5 Selected Use Cases and Rationale for Selection

The evaluation process identified three primary engineering use cases that not only promise significant efficiency gains but also align closely with the Reference Plant's engineering priorities. These use cases were chosen for their standardized documentation processes, high potential for automation (i.e., currently time-intensive, tedious, burdensome, or error prone), and broad applicability across plant systems.

2.5.1 BAC Evaluations

BAC evaluations represent an ideal candidate for LLM implementation due to several factors. These evaluations follow a highly standardized format that must be completed regularly for various plant components. The evaluation process involves gathering component information, determining material composition, assessing corrosion potential, and documenting findings in a consistent format.

The standardized nature of BAC evaluations makes them particularly suitable for automation. The process follows clear procedural requirements, utilizes established technical references, and produces documentation with consistent structure. Additionally, there exists a substantial historical record of completed evaluations that can inform the development and validation of LLM-generated content.

BAC evaluations support equipment reliability monitoring under the Maintenance Effectiveness Monitoring Program and are applicable across numerous systems and components. The evaluation process relies heavily on industry-standard technical guidance, such as the Electric Power Research Institute (EPRI) Boric Acid Corrosion Guidebook, making it a common and consistent activity across all nuclear plants [7]. This combination of standardized industry guidance, and universal applicability across the nuclear fleet makes efficiency improvements in this area particularly valuable as a demonstration case with broadly applicable findings.

2.5.2 Maintenance Rule Evaluations

Maintenance rule evaluations were selected as the second use case due to their regulatory importance and structured assessment requirements. These reviews are essential for ensuring compliance with 10 CFR 50.65 by systematically evaluating equipment performance against established criteria. The review process includes several standardized steps, including:

- Determining whether equipment failures affect maintenance rule functions
- Evaluating functional failure (FF) classification
- Assessing preventability
- Determining performance criteria impacts.

Like BAC evaluations, maintenance rule evaluations follow a structured process with consistent documentation requirements. The technical complexity of these reviews makes them an excellent test case for more sophisticated LLM applications that must integrate multiple types of plant information and regulatory requirements. The regular frequency of these evaluations means that efficiency improvements can deliver significant cumulative benefits over time.

2.5.3 10 CFR 50.59 Screening

This use case involves conducting initial evaluations to determine if proposed changes to plant procedures, tests, or modifications require prior NRC approval. The process relies on structured screening questions and systematic justifications, making it an excellent candidate for LLM integration. The high volume of research, documentation, and the need for consistent evaluation criteria underscore the potential for efficiency improvements through automation.

Together, these use cases were selected because they offer clear, measurable benefits and serve as a scalable template for further LLM applications across the nuclear industry. Their implementation is expected to drive significant time savings, reduce human error, and enhance overall operational efficiency by leveraging standardized processes and rich historical data.

3. WORKLOAD ANALYSIS

3.1 Introduction

The workload analysis component of this business case assessment follows a structured approach designed to quantify efficiency gains achieved through AI-assisted engineering processes. This assessment aims to establish baseline task durations for engineers of varying experience levels and compare performance improvements between traditional and AI-assisted workflows.

3.1.1 Workload Analysis Methodology

The workload analysis methodology for conducting the workload analysis incorporated the following steps:

1. **Document Manual and AI-Assisted Processes:** The analysis began by mapping traditional workflows and those assisted by an LLM for various engineering activities, identifying specific process steps where AI could potentially create efficiencies. During this process, activities with no potential for LLM assistance were deliberately excluded to focus measurement efforts on relevant workflows. This documentation was completed prior to engineer interviews to establish a framework for subsequent data collection.
2. **Define Metrics and Data Collection Approach:** Time taken to perform the task was established as the primary quantitative metric, while also designing the approach to capture the underlying reasons for any efficiency gains. This included examining whether time savings came from reduced errors, fewer revisions, or other workflow improvements. Methods were developed to collect data on broader positive and negative consequences beyond simple time metrics, such as quality impacts and user experience factors.
3. **Data Collection:** Interviews were conducted with engineers to gather quantitative data on both manual and AI-assisted processes. These sessions collected time estimates for traditional approaches, projected time savings with AI assistance, and qualitative feedback on the impact of the technology on workflow and output quality. These interviews followed a structured format to ensure consistent data collection across participants.
4. **Analysis:** After completing interviews, collected data was consolidated and analyzed, bringing together quantitative time estimates and qualitative feedback. This analysis process involved calculating average time estimates for manual processes, projecting time savings based on engineer

feedback, and categorizing qualitative insights into key themes. Initial findings documentation was prepared for subsequent validation.

5. **Validation:** In-person follow-up meetings were conducted to validate findings with the engineers who provided the initial data. During these sessions, the accuracy of time estimates and efficiency projections was verified while exploring qualitative aspects in greater depth. Potential organizational benefits were examined from the projected time savings, such as increased capacity for higher value work and potential quality improvements.

3.1.2 Interview Participants

Interviews were conducted following a structured approach based on the workload analysis tool. Engineers with hands-on experience provided time estimates and qualitative feedback for traditional processes, while also assessing potential LLM impacts. Management interviews focused on quality assessment, review processes, and broader implementation considerations. This dual perspective approach allowed the analysis to capture both operational efficiencies and quality implications across the BAC evaluations, maintenance rule evaluations, and 10 CFR 50.59 screenings.

For each selected use case, multiple engineers with varying experience levels were interviewed to provide insight into how the tool impacts users with different expertise. Additionally, management interviews provided an organizational perspective on quality control considerations and broader implementation implications.

The structured interview approach ensured consistent data collection across participants while allowing for qualitative insights that helped contextualize the quantitative findings. This methodology provided a robust basis for estimating potential efficiency gains while identifying important quality and implementation considerations.

3.1.3 AI Impact Rating

To standardize the assessment of AI's impact across different engineering activities, a five-point rating scale was used based on interviews with engineers and an analysis of pilot implementation data. This scale translates qualitative assessments of AI assistance potential into quantified percentage time savings ranges, enabling consistent evaluation across different tasks and processes.

The rating system considers factors including task complexity, repetitiveness, research intensity, and documentation requirements. Lower ratings (1–2) represent activities where an LLM provides minimal to modest assistance, typically involving tasks requiring significant engineering judgment or field observations. Higher ratings (4–5) represent activities where an LLM creates substantial time savings, particularly in research-intensive or documentation-heavy tasks where the LLM's information retrieval and drafting capabilities deliver maximum value.

Table 1. AI impact rating scale to percent time savings.

AI Impact Rating	Percent Time Savings
1	0%–5%
2	6%–15%
3	16%–35%
4	36%–60%
5	61%–80%

3.2 BAC Evaluation

BAC evaluations represent a standardized engineering process with well-defined workflow steps, making them an ideal candidate for AI assistance. The findings below detail the current manual process, AI impact assessment, and projected time savings.

3.2.1 Manual Process Time Assessment

The manual process for BAC evaluations demonstrates significant variability in time requirements. The total process time ranges from approximately 30 minutes for simple cases to more than 4 hours for complex evaluations. The process begins with condition report initiation, where engineers review existing work orders to check if the leak has been previously identified. This can take from 5 minutes for a single location to 2–3 hours for walkdowns that identify 20–30 locations, which includes labeling pictures and reviewing work management system records. The actual workflow is well-structured, with engineers now using a standardized template that streamlines documentation.

The most time-intensive components of the evaluation involve research activities. Engineers reported spending between 2 and 20 minutes per evaluation identifying affected materials and components, often needing to reference drawings to determine whether materials are stainless steel or carbon steel. Researching component material composition and susceptibility requires additional lookup in the EPRI guidebook or general searches, adding another 4–15 minutes. Calculating corrosion rates and time to critical degradation demands examining temperature, BAC, and material thickness, which takes between 5 and 30 minutes depending on complexity. The final documentation phase is relatively efficient due to templating, with engineers reporting that drafting the evaluation takes between 10 and 15 minutes for most cases but can extend to 30–60 minutes for complex scenarios. Senior engineers performing reviews typically spend 5–20 minutes verifying calculations and component identification.

Table 2. BAC evaluation manual process time.

Task	Activity	Manual Time Estimate (Minutes)		
		Low	High	Median
Identification & Inspection	Initiate condition report	4	75	18.5
BAC Evaluation	Identify affected materials & components	2	17.5	7.5
	Research component material composition & susceptibility to BAC	3.5	15	8.5
	Compare visual observations with reference criteria	3.5	20	11
	Perform calculations (estimating time to critical degradation)	5	22.5	12.5
	Draft evaluation and update condition report	10	67.5	26
Review and Corrective Actions	Review completed condition report	5	25	13.5
	Ensure appropriate corrective actions are assigned	2	17.5	9.5
	Create and submit work orders	2	5	3
TOTAL		37	265	110

3.2.2 Artificial-Intelligence-Assisted Impact Assessment

3.2.2.1 Overview

The integration of an LLM into the BAC evaluation process presents varied opportunities for efficiency improvements and quality enhancements. Based on structured interviews with engineers of different experience levels and engineering management at the Reference Plant, the impact potential of LLM assistance varies significantly across different tasks within the evaluation workflow. The most significant impact opportunities exist in the report drafting, while activities requiring visual inspection, engineering judgment, and plant-specific knowledge show more limited potential for LLM augmentation.

3.2.2.2 Identification and Inspection

Initiate Condition Report

The potential for LLM assistance in initiating condition reports appears limited. Engineers reported this process is already streamlined, typically requiring only 5–15 minutes per report, with complex cases involving multiple locations requiring up to 2–3 hours mainly due to field activities. The condition report format does not involve complex drafting that would significantly benefit an LLM tool. The process primarily involves looking up information in existing systems such as the plant work management system followed by basic documentation of observations.

LLM Impact Rating: 1 (0%–5% time savings).

3.2.2.3 Task 2: BAC Evaluation

Identify Affected Materials and Components

An LLM could provide moderate assistance through the automated lookup of component information from drawings and reference materials. One opportunity identified was helping match components observed in the field to reference documentation. However, visual inspection and engineering judgment remain essential, particularly for complex valve configurations where multiple materials may be involved.

LLM Impact Rating: 2 (6%–15% time savings).

Research Component Material Composition and Susceptibility

Material composition research shows valuable automation potential. The current process involves engineers consulting material specifications and the EPRI guidebook to determine susceptibility to corrosion. This typically takes 5–10 minutes for straightforward cases but can extend to 20 minutes for complex components. An LLM tool could streamline document searches and material lookups, though engineers emphasized that verification would remain necessary. The process is typically straightforward for experienced engineers but could provide more substantial benefits for less experienced personnel.

LLM Impact Rating: 2 (6%–15% time savings).

Compare Visual Observations with EPRI Reference Criteria

This activity presents moderate potential for LLM assistance in helping engineers interpret and categorize boric acid conditions based on visual observations. Engineers must determine whether leaking boric acid is wet or dry, document its color, and correlate these observations with reference data. Visual assessment remains an area where engineering expertise is critical, with an LLM serving as a potential assistant rather than a replacement for human judgment. If an LLM were capable of advanced image analysis, the LLM impact could be greater.

LLM Impact Rating: 2 (6%–15% time savings).

Perform Calculations (Estimating Time to Critical Degradation)

Calculations related to corrosion rates show moderate potential for LLM automation. These calculations typically take engineers between 5 and 20 minutes, occasionally extending to 30–40 minutes for complex scenarios involving multiple components or varying operational conditions. While not extremely complex, they require accurate inputs and an appropriate selection of reference values. Engineers indicated an LLM could automate standard calculations while emphasizing the need for human verification of results.

LLM Impact Rating: 2 (6%–15% time savings).

Draft Evaluation and Update Condition Report

Report drafting represents the most significant opportunity for LLM impact. Engineers reported this activity typically takes 10–40 minutes for routine evaluations, with complex cases requiring up to 2 hours. Senior engineering management identified grammar issues and procedure adherence as the primary quality concerns in manual evaluations. The current process uses a templated approach. An LLM could significantly reduce drafting time while improving quality by ensuring comprehensive coverage of all required procedural elements.

LLM Impact Rating: 4 (36%–60% time savings).

3.2.2.4 Review and Corrective Actions

Review Completed Condition Report

Review activities show moderate potential for LLM impact through improved readability and procedural adherence. Senior engineers reported spending approximately 5–20 minutes reviewing each evaluation, focusing on verifying component identification, corrosion rate calculations, and procedural compliance. An LLM could potentially reduce review cycles by improving the initial quality of submitted evaluations, though engineers emphasized that thorough human review would remain essential.

LLM Impact Rating: 2 (6%–15% time savings).

Ensure Appropriate Corrective Actions are Assigned

This activity shows minimal potential for LLM impact as it relies heavily on plant-specific maintenance knowledge and engineering judgment. Engineers reported this typically takes 1–5 minutes for straightforward cases but can extend to 20–30 minutes when researching historical maintenance records to determine the appropriate corrective approach. While an LLM could assist with retrieving historical information, the judgment regarding appropriate corrective measures remains an area requiring human expertise.

LLM Impact Rating: 1 (0%–5% time savings).

Create and Submit Work Orders

The creation and submission of work orders presents minimal opportunity for LLM impact for this specific tool. Engineers reported this activity typically takes only 3–5 minutes and follows a straightforward process that is already efficient. Engineers did not identify significant inefficiencies that would warrant automation, suggesting resources would be better directed toward higher impact areas within the evaluation process.

LLM Impact Rating: 1 (0%–5% time savings).

3.2.3 Artificial-Intelligence-Assisted Time Impacts

Table 3. Boric acid evaluation: Summary table of AI-assisted time impacts.

Task	Activity	Manual Time Estimate (Minutes)			LLM Impact				
		Low	High	Median	Rating	Low (%)	High (%)	Low (Minutes)	High (Minutes)
Identification & Inspection	Initiate condition report	4	75	19	1	0	5	0.0	3.8
BAC Evaluation	Identify affected materials and components	2	18	8	2	6	15	0.1	2.6
	Research component material composition and susceptibility to BAC	4	20	9	2	6	15	0.2	3
	Compare visual observations with EPRI reference criteria	4	20	11	2	6	15	0.2	3.0
	Perform calculations (time to critical degradation)	5	23	13	2	6	15	0.3	3.4
	Draft evaluation and update condition report	10	68	26	4	36	60	3.6	40.5
Review and Corrective Actions	Review completed condition report	5	25	14	2	6	15	0.3	3.8
	Ensure appropriate corrective actions are assigned	2	18	10	1	0	5	0.0	0.9
	Create and submit work orders	2	5	3	1	0	5	0.0	0.3
TOTAL		37	265	110	2	11	23	5	61

3.3 Maintenance Rule Evaluations

3.3.1 Manual Process Time Assessment

Maintenance rule evaluations exhibit significant time variability, ranging from approximately 155 minutes (2.6 hours) to 1,935 minutes (32 hours) depending on complexity. The process begins with INPO Consolidated Events Database (ICES) report determination (from 10 minutes to 8 hours) followed by FF evaluation, which constitutes the most time-intensive portion. Engineers must document event descriptions, identify affected components and applicable maintenance rule functions, and provide

detailed failure analysis documentation. For complex cases, this core evaluation can require up to two full days as engineers research regulatory guidance, technical calculations, vendor information, and work order history, often consulting with other personnel.

Table 4. Maintenance rule evaluation manual process time.

Task	Task Activity	Manual Time Estimate (Minutes)		
		Low	High	Median
ICES Report Determination	Determine if ICES report is required	10	255	60
	Provide coordinator with required input	10	75	40
Event Information and FF Evaluation	Document event description	5	37	20
	Identify associated systems, structures, and components (SSCs)	5	37	18
	Identify applicable maintenance rule functions	3	13	7
	Document basis for FF determination	10	370	55
Maintenance Preventable FF (MPFF) Determination	Identify cause of FF	8	650	140
	Document MPFF determination with justification	5	30	15
	Document corrective actions	5	30	15
Repetitive MPFF (RMPFF) Evaluation	Research previous MPFFs	42	275	110
	Document RMPFF determination	7	20	10
Performance Criteria Evaluation	Evaluate plant-level performance monitoring	20	82	30
	Review function reliability performance criteria	12	29	20
	Create condition report for MRRP review	14	33	20
TOTAL		155	1935	560

Plant data show that, out of 481 maintenance rule evaluations over the 5-year period analyzed, only 77 (16%) were identified as a functional failure (FF) and continued past the FF evaluation stage. Of these, 37 were determined to be maintenance preventable FFs (MPFFs), and only four (0.8% of total) were classified as repetitive MPFFs (RMPFFs). Figure 1 illustrates the annual average of these evaluation types.

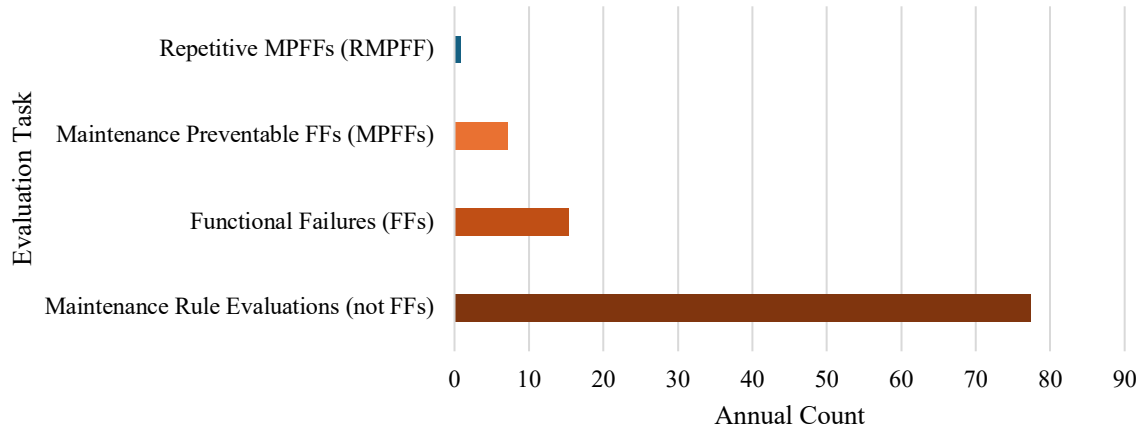


Figure 1. Maintenance rule evaluation annual count by task.

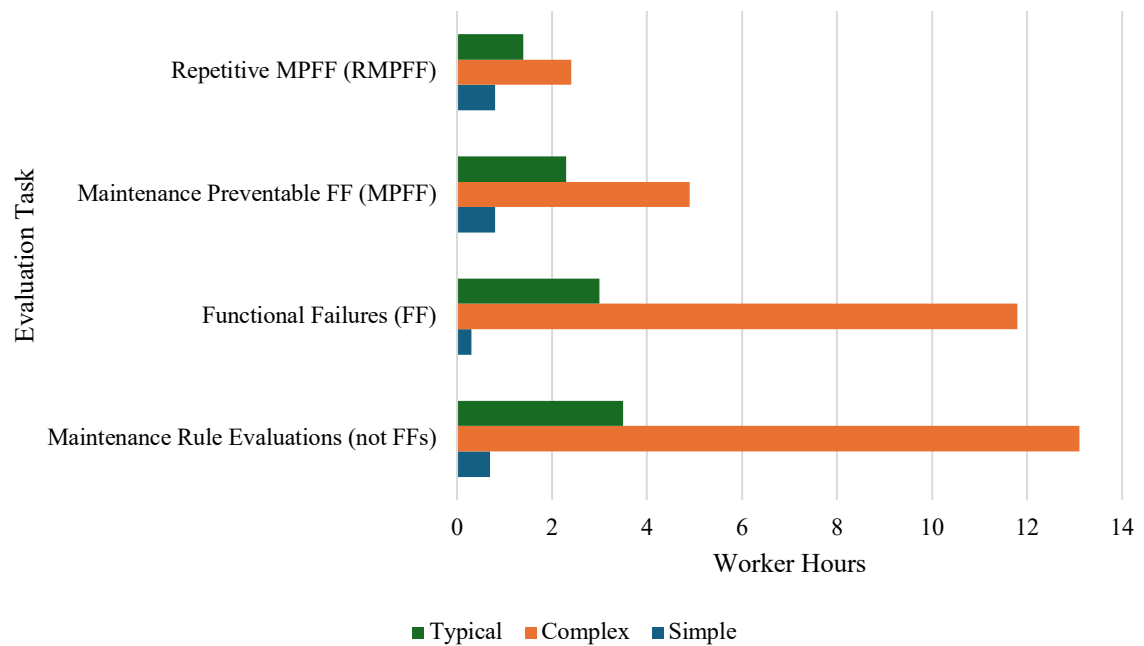


Figure 2. Breakdown of worker hours by task for maintenance rule evaluations for manual process.

3.3.2 Artificial-Intelligence-Assisted Process

3.3.2.1 Overview

The application of an LLM in maintenance rule evaluations represents a significant opportunity to streamline workflows, enhance consistency, and reduce the time engineers spend on routine research and documentation tasks. Through structured interviews with system engineers and management at the Reference Plant across varying experience levels, clear patterns emerged regarding where an LLM could most effectively augment the existing process while preserving essential engineering judgment and oversight.

The maintenance rule evaluation process involves multiple steps where engineers must retrieve information from diverse sources, analyze complex data sets, make nuanced regulatory determinations,

and document findings in standardized formats. The LLM-assisted approach automates many of the time-consuming research activities while providing consistent formatting and ensuring thorough procedural compliance, particularly for the most labor-intensive aspects of the evaluation.

Engineers identified several key areas where an LLM provides substantial value, including automated document retrieval, pattern recognition across historical failures, standardization of documentation, and generation of properly formatted reports. The following sections detail how an LLM capabilities align with specific maintenance rule evaluation tasks and the potential impact on both efficiency and quality.

3.3.2.2 Task 1: ICES Database Report Determination

Determine if ICES Database Report is Required

The LLM provides moderate assistance by automatically checking against limiting conditions for operation (LCO) criteria and regulatory requirements. For straightforward cases where an LCO is less than 72 hours, a determination can be made quickly with or without an LLM. For complex scenarios, the tool can accelerate the process by checking multiple criteria simultaneously. The LLM impact is most substantial for complex cases that might otherwise take up to 8 hours to research. An LLM can significantly reduce this time by automatically cross-referencing applicable regulations and station-specific reporting requirements.

LLM Impact Rating: 3 (16%–35% time savings).

Provide Coordinator with Required Input

An LLM can standardize format and ensure consistency. This activity typically takes 10 minutes to 2 hours in the manual process, with longer durations occurring when additional information gathering is needed. While an LLM can significantly reduce documentation time, it cannot eliminate the need for human interface when gathering information about recent events that are not yet fully documented.

LLM Impact Rating: 3 (16%–35% time savings).

3.3.2.3 Task 2: Event Information and FF Evaluation

Document Event Description

An LLM demonstrates meaningful impact in standardizing documentation format and ensuring consistency. In the manual process, engineers typically spend 5–30 minutes on this activity, with variation largely depending on the quality of the initial condition report. The tool can extract key details from condition reports and other sources to create standardized, objective descriptions that address all required elements. This standardization helps reduce the subjectivity that less experienced engineers might introduce in reports.

LLM Impact Rating: 3 (16%–35% time savings).

Identify Associated Systems, Structures, and Components

Engineers noted moderate potential for an LLM in identifying systems, structures, and components (SSCs) associated with an event. The manual process typically takes 5–20 minutes, with longer durations needed when more complex failures involve multiple subcomponents. An LLM can help correlate event data with plant systems and map reported issues to the correct SSCs. By automatically identifying these relationships, particularly for complex systems with multiple interfaces, an LLM can reduce the time required while ensuring more comprehensive identification of all affected components.

LLM Impact Rating: 2 (6%–15% time savings).

Identify Applicable Maintenance Rule Functions

With the recent implementation of new scoping criteria,^a identifying applicable maintenance rule functions has become more complex. Engineers reported this typically takes 5–30 minutes manually, with variation based on their familiarity with the systems involved. An LLM provides valuable assistance by mapping components to maintenance rule functions and retrieving supporting documentation. This is particularly valuable as engineers adapt to the new multilevel functional hierarchy introduced in the updated scoping criteria.

LLM Impact Rating: 3 (16%–35% time savings).

Document Basis for FF Determination

This activity represents the most substantial opportunity for an LLM. Engineers consistently identified this as the most time-consuming aspect of maintenance rule evaluations, requiring anywhere from 20 minutes for simple failures to 2–3 days for complex issues.

An LLM can dramatically reduce research time by automating lookups across multiple sources and databases, including:

- FSAR tables and equipment specifications
- DBDs
- Equipment history and similar failures
- Industry operating experience
- Regulatory guidance.

By automating these searches and compiling relevant information into a structured format, an LLM tool can potentially reduce the time for this activity by 40%–60%, particularly for complex cases.

LLM Impact Rating: 4 (36%–60% time savings).

3.3.2.4 Task 3: MPFF Determination

Identify Cause of FF and Determine if Failure Should Have Been Prevented by Maintenance

For determining whether a failure should have been prevented by maintenance, an LLM offers significant value by accessing electronic work packages and leveraging historical maintenance data. Engineers reported this activity typically takes between 10 minutes and 5 hours manually, with complex cases requiring extensive research. An LLM can automatically retrieve and analyze information from work packages, maintenance history, and similar past failures. Engineers particularly valued an LLM ability to automatically check maintenance history records for patterns, helping engineers determine whether failures were preventable.

LLM Impact Rating: 3 (16%–35% time savings)

Document MPFF Determination with Justification

An LLM offers moderate assistance in standardizing evaluation writing with consistent format and requirements. The manual process typically takes 10–60 minutes, depending on the complexity of the failure and the amount of supporting documentation required. The tool can suggest appropriate language based on similar cases, ensuring that all regulatory requirements are addressed. An LLM can draft the determination while ensuring proper documentation of the justification, which helps maintain consistency across different engineers and reduces review cycles.

LLM Impact Rating: 2 (6%–15% time savings).

^a According to an expert at the Reference Plant, the site implemented new scoping criteria that established a multilevel functional hierarchy, increasing the complexity of component-to-function mapping during evaluations.

3.3.2.5 Task 4: RMPFF Evaluation

Research Previous MPFFs

Researching previous MPFFs represents another area with significant LLM impact potential. In the manual process, engineers spend between 10 minutes and 4 hours reviewing maintenance rule databases for similar failures. An LLM can automatically search for and correlate historical failure data from multiple databases, presenting potential repetitive failures for engineering review. By automating this research process, an LLM can dramatically reduce search time by efficiently mining historical records, identifying patterns, and presenting relevant information in a structured format.

LLM Impact Rating: 4 (36%–60% time savings).

Document RMPFF Determination

For documenting RMPFF determinations, an LLM helps ensure a uniform evaluation approach and generates standardized documentation. Engineers reported that this activity typically takes 5–45 minutes manually, depending on the complexity of the determination. The tool's impact is primarily in standardizing documentation and ensuring that all required elements are addressed.

LLM Impact Rating: 3 (16%–35% time savings).

3.3.2.6 Task 5: Performance Criteria Evaluation

Evaluate Plant-Level Performance Monitoring

An LLM provides moderate assistance by automatically analyzing event data against plant performance criteria and flagging potential impacts. Engineers reported that this activity typically takes 20–82 minutes manually, with longer durations for events with potential plant-level impacts. An LLM can automatically check whether the event resulted in an unplanned transient, unplanned reactor trip, or unplanned increase in risk color ranking. By automating this assessment, an LLM helps ensure a comprehensive evaluation against all required criteria while saving time.

LLM Impact Rating: 3 (16%–35% time savings).

Review Function Reliability Performance Criteria and Determine If Current Failure Exceeds Performance Criteria

An LLM provides moderate assistance in evaluating whether a current failure exceeds established performance criteria. Engineers typically spend 12–29 minutes reviewing reliability criteria and comparing them against established thresholds. An LLM can help automate the lookup of performance criteria from maintenance rule databases and compare current failures against historical patterns. While the final determination requires engineering judgment, an LLM streamlines the data collection and initial assessment steps.

LLM Impact Rating: 2 (6%–15% time savings).

Create Condition Report for MRRP Review if Performance Criteria Exceeded

An LLM offers valuable assistance by generating standardized reports with consistent formatting and required elements. This activity typically takes 1–30 minutes manually, with variation based on the complexity of the performance criteria exceedance. The LLM ensures procedure compliance and improved language and clarity in these reports, which helps maintain consistency across different engineers and reduces the likelihood of review comments.

LLM Impact Rating: 3 (16%–35% time savings).

3.3.3 AI-Assisted Time Impacts

Table 5. Maintenance rule evaluation: Summary table of AI-assisted time impacts.

Task	Activity	Manual Time Estimate (Minutes)			LLM Impact				
		Low	High	Median	Rating	Low (%)	High (%)	Low (Minutes)	High (Minutes)
ICES Report Determination	Determine if ICES report is required	10	255	60	2	6	15	0.6	38.3
	Provide coordinator with required input	10	75	40	2	6	15	0.6	11.3
Event Information and FF Evaluation	Document event description	5	37	20	3	16	35	0.7	12.8
	Identify associated SSCs	5	37	18	2	6	15	0.3	5.5
	Identify applicable maintenance rule functions	3	13	7	2	6	15	0.2	2.0
	Document basis for FF determination	10	370	55	4	36	60	3.6	222.0
MPFF Determination	Identify cause of FF	8	650	140	3	16	35	1.3	171.5
	Document MPFF determination with justification	5	30	15	2	6	15	0.3	60.5
	Document corrective actions	5	30	15	3	16	35	0.8	10.5
Repetitive MPFF Evaluation	Research previous MPFFs	42	275	110	4	36	60	15.0	165.0
	Document RMPFF determination	7	20	10	3	16	35	1.1	7.0
Performance Criteria Evaluation	Evaluate plant-level performance monitoring	20	82	30	3	16	35	3.3	28.6

	Review function reliability performance criteria	12	29	20	2	6	15	0.7	3.5
	Create condition report for MRRP review	14	33	20	3	16	35	2.2	12.5
TOTAL		155	1935	560	3	19	39	31	751

3.4 10 CFR 50.59 Screenings

3.4.1 Manual Process Time Assessment

The 10 CFR 50.59 screening process evaluates whether proposed changes to nuclear plant systems, structures, or components require prior NRC approval. The process demonstrates substantial time variation, ranging from 218 minutes (3.6 hours) to 1,345 minutes (22.4 hours) based on complexity. The most time-consuming component of the screening process is Part 1, which involves searching the FSAR and DBDs for potentially affected design functions and performance requirements. This research phase ranges from 2–3 hours for simple component replacements to 1–2 full days for system-level modifications, with engineers noting that critical information may be scattered across multiple sections of extensive documents.

Part 2 of the screening, addressing the four screening questions, generally requires 1–2 days for complex cases or as little as 30 minutes for simpler evaluations, with almost all effort concentrated on Question 1. Engineers reported that approximately 90% of the total process time is devoted to research activities, with only 10% spent on documenting findings. This research-heavy process presents a significant opportunity for LLM-assisted efficiency gains, particularly in searching large technical documents and compiling relevant information across multiple sources. Engineers consistently identified document searching as the area with the greatest potential for time savings through an LLM implementation.

Table 6. 50.59 screenings manual process time.

Task	Activity	Manual Time Estimate (Minutes)		
		Low	High	Median
Activity Identification & Documentation	Identify proposed activity by facility design authorization, procedure change notice (PCN), or procedure number	5	30	10
	Provide summary description of the proposed activity	15	75	40
	Document applicable processes from the applicability determination (AD) review	5	30	15
Part 1: 50.59 Screen Review	Search the FSAR and updated FSAR (UFSAR) for design function and performance requirements	120	840	330
	Document FSAR locations and method of evaluation for affected SSCs	18	50	34
Part 2: 50.59 Screening Questions	Answer Part 2 Question 1: Impact on design function and performance requirements	45	270	90
	Answer Part 2 Questions 2–4: Evaluation methodology, controlling procedure, and evaluation results impacts	10	50	15
TOTAL		218	1345	534

3.4.2 AI-Assisted Process

3.4.2.1 Overview

Based on interviews with design engineers and reviewers at the Reference Plant, the integration of an LLM has significant potential to improve efficiency across the 10 CFR 50.59 screening workflow. This section examines how LLM technology could assist with each activity in the process, identifying where the greatest impacts are likely to occur.

3.4.2.2 Task 1: Activity Identification and Documentation

Identify Proposed Activity by PCN or Procedure Number

An LLM provides limited assistance for this primarily administrative task. Design engineers indicated that identifying the proposed activity by procedure change notice (PCN) or procedure number is straightforward and requires minimal time. For design engineers, the identification step is particularly efficient as the activity is already clearly defined, limiting the value an LLM can add.

LLM Impact Rating: 2 (6%–15% time savings).

Provide Summary Description of the Proposed Activity

An LLM demonstrates moderate potential for assisting with drafting comprehensive summaries of proposed activities. Engineers reported this activity typically takes between 10 and 30 minutes depending on the complexity of the change. An LLM can help ensure all relevant aspects of the modification are described and potential impacts captured. It can also standardize description formats, improving clarity and consistency, which was noted as an issue by engineering management.

LLM Impact Rating: 2 (6%–15% time savings).

Document Applicable Processes from AD Review

Engineers identified limited potential for an LLM assistance in documenting applicable processes from the applicability determination (AD) review. The task is primarily a straightforward documentation of already-identified processes. This administrative step primarily involves transferring information that has already been processed during the AD phase, making it a low-value target for LLM assistance.

LLM Impact Rating: 1 (0%–5% time savings).

3.4.2.3 Task 2: Part 1—FSAR Screen Review

Search the FSAR and UFSAR for Design Function, Performance Requirements

Engineers consistently identified FSAR research as having the highest potential for LLM assistance in the entire 50.59 process. In the manual process, this activity takes anywhere from 2–3 hours for simple component replacements to 1–2 days for complex system-level changes. An LLM can dramatically reduce search time by extracting relevant information across multiple large technical documents, including not just the FSAR but also DBDs. Engineers estimated that an LLM could potentially reduce research time by 65%–75% for complex modifications.

LLM Impact Rating: 4 (36%–60% time savings).

Document FSAR Locations and Method of Evaluation for Affected SSCs

For documenting FSAR locations and evaluation methods, an LLM provides moderate assistance with organization and standardized documentation of findings. Engineers indicated this activity itself requires minimal time as it follows directly from the previous research step. While the documentation itself is straightforward, an LLM can help ensure consistency in formatting and assist with structured documentation, including completeness checks and proper referencing of source material, potentially reducing documentation errors.

LLM Impact Rating: 4 (36%–60% time savings).

3.4.2.4 Task 3: Part 2—50.59 Screening Questions

Answer Part 2 Question 1: Impact on Design Function and Performance Requirements

Engineers reported the highest potential for LLM assistance when answering Question 1 regarding impacts on design functions. This activity typically takes 30 minutes to 1 hour for simple replacements but can extend to 3–4 hours for complex modifications. An LLM can substantially accelerate this analysis by comparing proposed changes against documented design functions and highlighting potential conflicts. Engineers estimated that, for complex evaluations, an LLM could reduce the time from 4 hours to approximately 1 hour, representing 75% time savings.

LLM Impact Rating: 5 (61%–80% time savings).

Answer Part 2 Questions 2–4: Evaluation Methodology, Controlling Procedure, and Impacts on Evaluation Results

For the remaining screening questions, engineers indicated limited time is typically spent as these questions are more straightforward and result in negative responses in most cases. An LLM can assist with drafting standardized responses and ensuring consistent formatting across the fleet. While an LLM can efficiently create an initial response templates with proper language structure and regulatory framework, engineers emphasized they will still need to scrutinize and review LLM-generated content for technical accuracy.

LLM Impact Rating: 4 (36%–60% time savings).

3.4.3 AI-Assisted Time Impacts

Table 7. 50.59 screenings: Summary table of AI-assisted time impacts.

Task	Activity	Manual Time Estimate (Minutes)			LLM Impact				
		Low	High	Median	Rating	Low (%)	High (%)	Low	High
Activity Identification & Documentation	Identify proposed activity by Facility Design Authorization, PCN, and procedure number	5	30	10	2	6	15	0	4.5
	Provide summary description of the proposed activity	15	75	40	2	6	15	1	11.25
	Document applicable processes from the AD review	5	30	15	1	0	5	0	1.5
Part 1: 50.59 Screen Review	Search the FSAR and UFSAR for design function and performance requirements	120	840	330	4	36	60	43	504
	Document FSAR locations and method of evaluation for affected SSCs	18	50	34	4	36	60	6	30
Part 2: 50.59 Screening Questions	Answer Part 2 Question 1: Impact on design function and performance requirements	45	270	90	5	61	80	27	216
	Answer Part 2 Questions 2–4: Evaluation methodology, controlling procedure, and impacts on evaluation results	10	50	15	4	36	60	4	30
TOTAL		218	1345	534	4	38	59	82	797

3.5 Patterns of High-Impact LLM Applications in Engineering Tasks

Across the three use cases analyzed in detail, certain patterns emerge regarding which tasks are most conducive to LLM assistance. Tasks receiving the highest LLM impact ratings (4–5) consistently involve substantial information retrieval, synthesis of data from multiple sources, and standardized documentation activities. Specifically, research-intensive tasks that require engineers to search through large technical documents (like FSARs or design basis records), identify relevant sections, and compile this information showed the most dramatic efficiency improvements.

Document drafting activities also demonstrated high potential for LLM assistance, particularly when following standardized formats with consistent structure and terminology. These patterns suggest that the most significant efficiency gains from LLM implementation occur when LLM tools can leverage their strengths in information retrieval, pattern recognition, and natural language generation to assist with the most time-consuming research and documentation activities. The efficiency improvements observed in these high-impact tasks provide a basis for identifying additional use cases that share similar characteristics across other engineering functions and plant operations.

4. BUSINESS CASE RESULTS, BENEFITS, AND CHALLENGES

4.1 Summary of Workload Efficiency Benefits

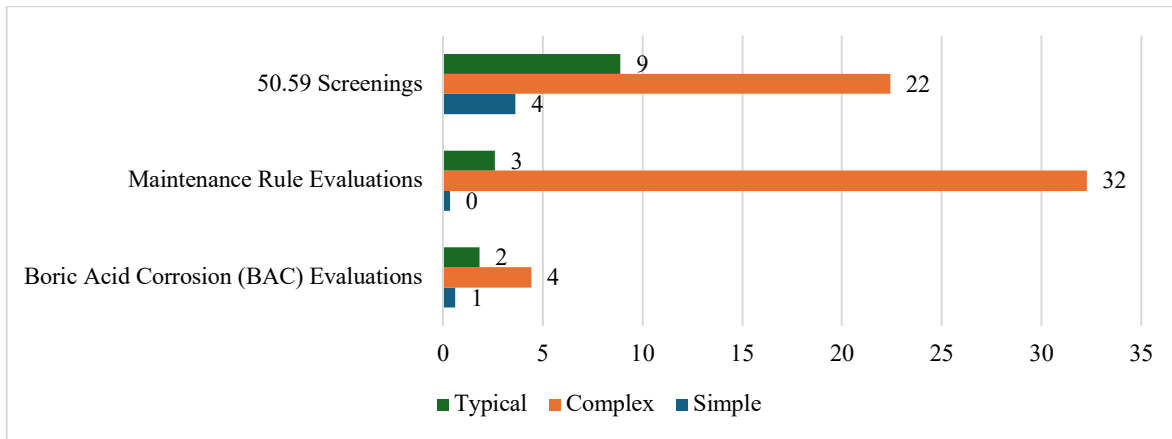


Figure 3. Worker hours per use case without LLM.

Figure 3 presents the worker hours required per use case without LLM implementation. The detailed workload analysis conducted for each evaluation type revealed that time requirements can vary substantially between simple and complex cases:

- **BAC Evaluations:** Range from 37 minutes for simple cases to 265 minutes (4.4 hours) for complex evaluations, with an average of 110 minutes (1.8 hours)
- **Maintenance Rule Evaluations:** Range from 12 minutes (0.2 hours) to 1,935 minutes (32.3 hours), with a typical evaluation taking 84 minutes (1.4 hours)^b
- **50.59 Screenings:** Range from 218 minutes (3.6 hours) to 1,345 minutes (22.4 hours), with an average of 534 minutes (8.9 hours).

^b As explained in Section 3.4, this very large range is related to the fact that most maintenance rule evaluations do not comprise of a complete evaluation, but end after FF evaluations.

4.1.1 AI Efficiency Gains: Estimated Percentage and Time Improvements

The LLM introduction can substantially reduce the time required for several of the tasks identified in the manual process analysis.

The analysis revealed that LLM implementation delivers substantial efficiency improvements across all three evaluation types, with the impact varying by task and evaluation type. As illustrated in Figure 4, the percentage time savings demonstrates clear patterns:

- BAC Evaluations: 11%–23% overall time reduction
- Maintenance Rule Evaluations: 19%–39% overall time reduction
- 50.59 Screenings: 38%–59% overall time reduction.

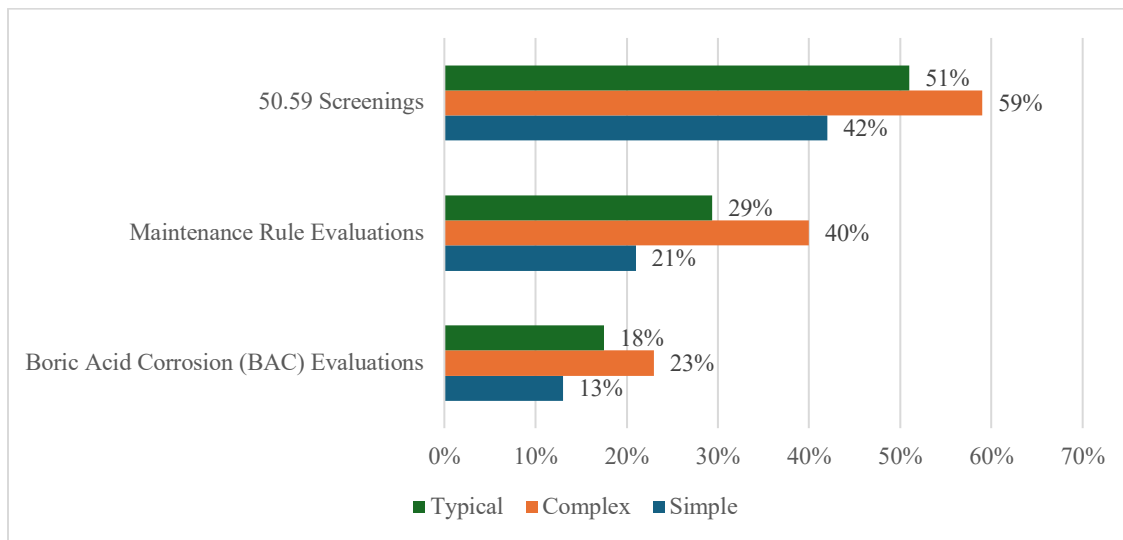


Figure 4. AI efficiency gains: Estimated percentage improvements.^c

These percentage improvements translate to significant absolute time savings, as shown in Figure 5:

- BAC Evaluations: 5–61 minutes saved per evaluation (typical 20 minutes)
- Maintenance Rule Evaluations: 31–751 minutes saved per evaluation (typical case of 391 minutes or 1.4 hours)
- 50.59 Screenings: 82–797 minutes saved per evaluation (typical case of 309 minutes or 5.2 hours).

^c These efficiency improvements align with documented results across technical industries. Recent studies validate similar productivity gains from AI implementation: McKinsey Global Institute (2023) found 30%–45% improvements for technical professionals using AI assistants, MIT Technology Review (2023) demonstrated 25%–40% efficiency gains in technical documentation tasks, Stanford and Microsoft Research (2023) documented 20%–35% efficiency gains in technical work, and IDC’s survey of 1,500+ engineering organizations found an average 28% efficiency increase. The nuclear industry’s results of 11%–59% fall within these established ranges, confirming that nuclear applications follow broader industry patterns rather than being outliers. Most organizations report achieving a return on investment within 6–12 months, with continued improvement as users develop better prompting skills and workflows.

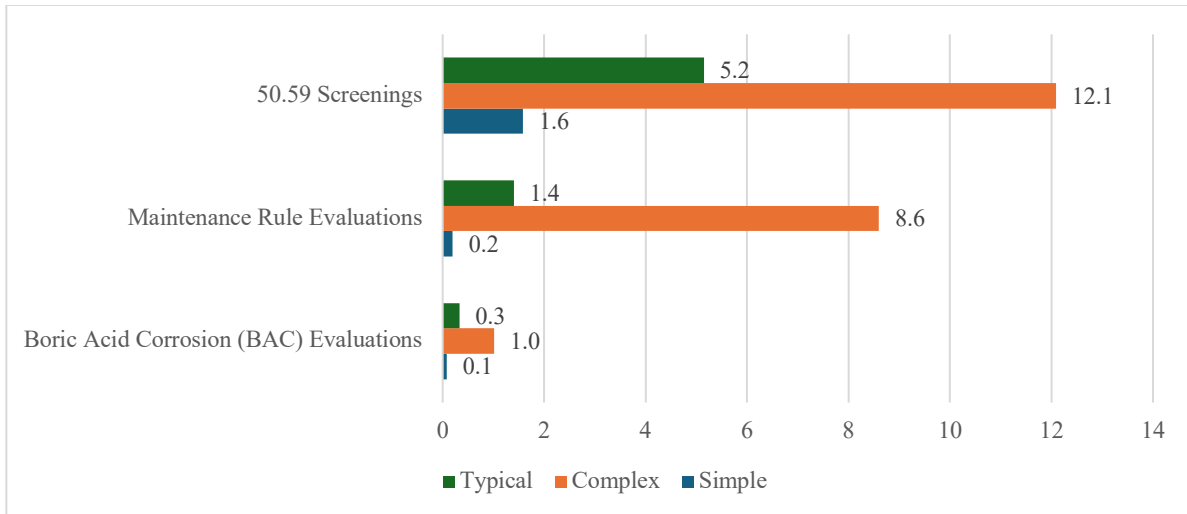


Figure 5. LLM efficiency gains: Estimated time improvements.

4.1.2 Annual Engineering Evaluations

To understand the aggregate impact of these efficiency improvements, historical data was analyzed on evaluation volumes at the Reference Plant over a 5-year period. Figure 6 illustrates the annual frequency of each evaluation type:

- BAC Evaluations: Approximately 50 evaluations performed annually
- Maintenance Rule Evaluations: Approximately 93 evaluations performed annually
- 50.59 Screenings: Approximately 35 evaluations performed annually.

These volumes reflect typical annual requirements for the dual unit plant and may vary based on outage schedules, plant modifications, and other operational factors. The distribution highlights that, while 50.59 screenings demonstrate the highest percentage efficiency gains, their lower frequency means that the aggregate impact from maintenance rule evaluations is potentially more significant due to their much higher volume.

Importantly, the evaluation volumes used in this analysis represent only the three processes studied in detail. Additional engineering evaluation types identified in Section 4.2 would further increase the potential aggregate benefits from LLM implementation across engineering functions.

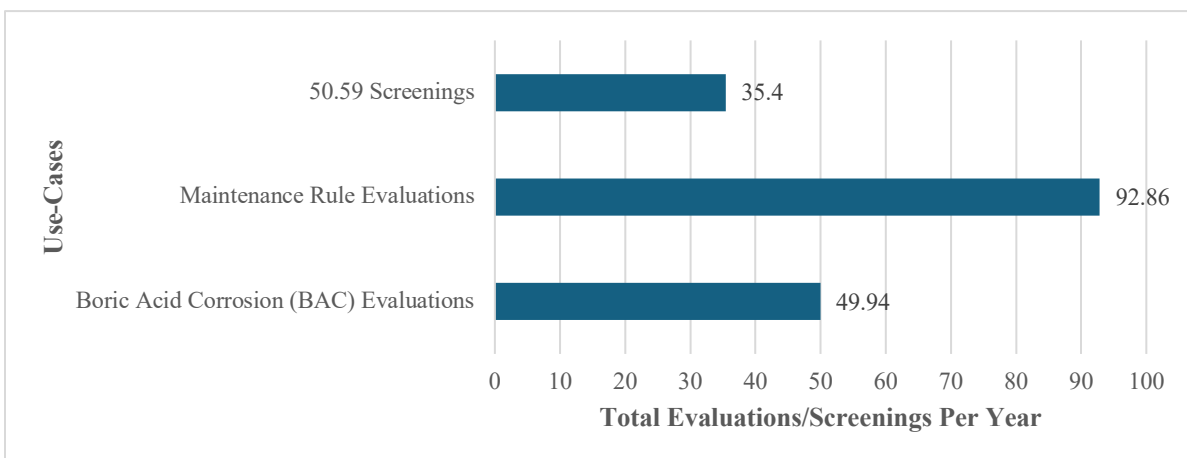


Figure 6. Annual engineering evaluations for two-unit plant.

4.1.3 Total Annual Savings from Three Example Use Cases

By extrapolating the task-based time savings across the volume of evaluations conducted on an annual basis, the analysis estimates that the implementation of LLM technology could save approximately 330 hours per year in an average-case scenario. In the best-case scenario, total annual savings might reach 405 hours, while in the worst-case scenario, around 235 hours per year could be recovered.

It is important to emphasize that these time savings represent efficiency gains rather than harvestable cost reductions, at least in the short and medium term. The primary value lies in freeing up engineers to perform more value-added technical work that directly impacts plant safety, reliability, and performance. This redirection of skilled resources toward higher value activities creates a strategic benefit beyond simple time savings.

When quantifying these efficiency gains using the fully loaded engineering cost of about \$96/hr, the annual value of recovered time is approximately \$30,788 in the average case, \$37,969 in the best-case scenario, and \$21,966 in the worst-case scenario. These figures help illustrate the magnitude of the opportunity when these efficiency improvements are scaled across the plant's operations over a year, while recognizing that the true value extends beyond financial metrics to encompass the broader strategic benefits detailed in Section 4.3.

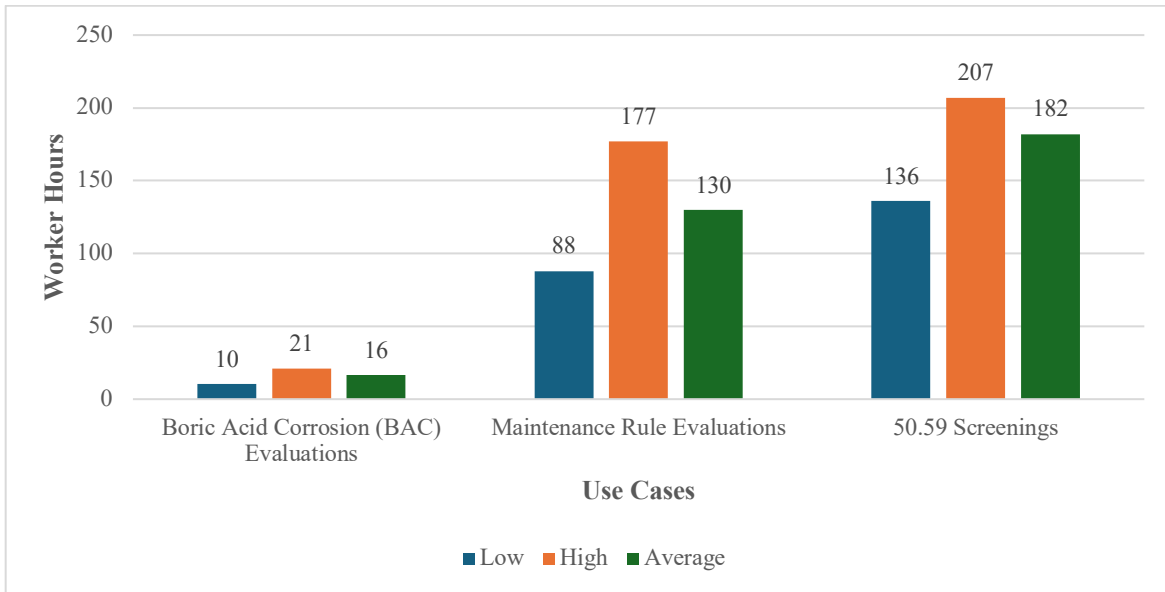


Figure 7. Total annual worker hours savings.

Table 8. NPV from worker hour savings from three use cases.

Scenario	Annual Hours	Annual Value
Expected	330 hours	\$32,000
High	405 hours	\$39,000
Low	235 hours	\$23,000

4.1.4 Inherent Conservatism of Results

These efficiency estimates are likely conservative for several reasons. The analysis was conducted during the early implementation phase when engineers were still becoming familiar with LLM

capabilities. As users gain experience with LLMs and develop more sophisticated prompting techniques, efficiency gains are expected to increase. Additionally, the estimates were primarily derived from interviews with engineers who had varying levels of experience with LLM tools, with more senior engineers typically providing more conservative assessments than junior staff. Finally, LLMs and wider AI continue to evolve rapidly, with improvements in model accuracy and functionality likely to enhance time savings in future iterations.

4.2 Other Use Cases

Additional use cases identified for LLM integration beyond the primary examples include both engineering and non-engineering applications—such as license renewal documentation, operating experience analyses, simple failure evaluations, and various administrative functions—that demonstrate similar opportunities for efficiency gains and process standardization. Each use case listed has been validated to offer comparable direct time savings and indirect strategic benefits, reinforcing the overall potential for expanding LLM-driven work reductions across the nuclear enterprise.

4.2.1 Engineering Use Cases

4.2.1.1 System Performance Reporting

System engineers regularly prepare reports documenting the overall health and performance of their assigned systems. These reports require compiling operational data, maintenance history, and performance trends to assess system reliability. An LLM could automate data collection, standardize report formats, and assist in identifying significant trends or anomalies that require engineering attention.

4.2.1.2 Equipment Qualification Documentation

These technical documents validate that equipment can function under specified environmental conditions, particularly during and after design basis accidents. The process involves researching design specifications, documenting testing parameters, and validating environmental compatibility. An LLM can assist by extracting specifications from design documents, standardizing documentation formats, and ensuring comprehensive coverage of regulatory requirements.

4.2.1.3 Operability Evaluations

Operability evaluations assess whether degraded conditions affect a component's ability to perform its safety function. Engineers must research technical specifications, apply regulatory guidance, and develop compensatory measures when needed. An LLM could streamline research activities, suggest precedent evaluations, and assist in documenting the technical basis for determinations.

4.2.1.4 Plant Data Trending Analysis

Assessment of operational data to identify patterns and potential degradation, whether through plant engineering or a monitoring and diagnostics center. An LLM could help compile historical data, identify trends, and document recommendations based on data analysis.

4.2.2 Regulatory Compliance

4.2.2.1 License Renewal Documentation

License renewal documentation involves comprehensive technical evaluations, aging management programs, and analyses to support extending a nuclear plant's operating license beyond its original term. This process requires extensive research across DBDs, regulatory requirements, and industry standards. An LLM could significantly accelerate the research phase by efficiently extracting relevant information from thousands of pages of documentation and by standardizing document formatting based on successful precedents.

4.2.2.2 *Regulatory Compliance Reporting*

Documentation demonstrating adherence to regulatory requirements and commitments. An LLM tool could help research regulatory standards, compile compliance evidence, and document conclusions with appropriate regulatory language.

4.2.3 *Operations Use Cases*

4.2.3.1 *Operational Decision-Making Meeting Documentation*

These meetings capture evaluations of emerging plant conditions, including risk assessments and action plans to address off-normal events. Documentation must be thorough, technically accurate, and traceable. An LLM could assist in drafting meeting summaries, documenting technical discussions, and ensuring action assignments are clearly defined and tracked.

4.2.4 *Maintenance Use Cases*

4.2.4.1 *Work Order Development*

Work order development involves creating and authorizing documents for maintenance, testing, or modification activities with all necessary specifications. An LLM could help develop technical specifications, identify procedural requirements, and document safety considerations based on similar work orders.

4.2.4.2 *Maintenance Strategy Reviews*

Periodic assessments of maintenance approaches for specific components or systems, typically conducted by engineering with maintenance department input. These reviews require researching maintenance history, evaluating performance trends, and documenting strategy changes. An LLM could assist in compiling historical data, identifying patterns, and drafting technical justifications for program adjustments.

4.2.5 *Other Plant Functions*

4.2.5.1 *Procedure Revisions*

Updates to plant procedures are required based on operational experience, regulatory changes, design modifications, or simply changes in plant processes. The process involves researching technical bases, identifying affected procedure sections, and documenting change justifications. An LLM could accelerate the research process for administrative, maintenance, and technical procedures, ensure consistent formatting, and help identify potential impacts on related procedures. For operations-specific procedures that involve critical safety functions, LLM assistance would likely focus more on formatting and reference identification rather than content generation.

4.2.5.2 *Operating Experience Analyses*

Systematic reviews of plant and industry events to identify lessons learned and potential vulnerabilities. These analyses involve research across diverse information sources, synthesis of applicable insights, and development of plant-specific recommendations. An LLM could accelerate the research process, identify relevant insights, and assist in documentation.

4.2.5.3 *Cause Analysis Documentation*

Structured processes for documenting failure analyses to identify root causes and contributing factors. An LLM could assist in applying systematic evaluation methods, documenting evidence, and developing conclusions based on established methodologies.

4.2.5.4 Fire Impairment Documentation

Records documenting temporary fire protection system impairments and compensatory measures. An LLM could assist in identifying affected systems, determining required compensatory measures, and documenting compliance with fire protection requirements.

4.3 Indirect Benefits and Long-Term Strategic Value

4.3.1 Overview

While the quantitative analysis provides a rubric of immediate economic benefits through time savings, the full value of LLM implementation extends beyond these direct measurements. Calculating value by multiplying worker hours saved by hourly wage rates captures only the first-order benefits of this technology.

Additional strategic value lies in how these time savings transform engineering work and enable more effective use of specialized engineering talent. By automating routine documentation tasks, LLMs free engineers to redirect their expertise toward higher value activities that better leverage their technical knowledge and experience. This reallocation of skilled resources can act as a potential value “multiplier” beyond the direct time savings calculated in our quantitative analysis. Engineers can refocus their time on activities that directly impact plant reliability, safety, and performance rather than administrative tasks.

The following sections explore these indirect benefits in greater detail. While the current study has focused primarily on quantifying direct time savings, these broader strategic benefits likely represent the greater long-term value to nuclear utilities.

The transformation from “time saved” to “value created” happens when engineers can apply their specialized knowledge to activities that more directly leverage their expertise. This shift can create cascading benefits throughout the organization. These indirect benefits become particularly significant as the nuclear fleet continues to age and face both economic and technical challenges that require more effective utilization of engineering resources.

4.3.2 Shifting from Reactive to Proactive Engineering

One of the most significant indirect benefits of implementing LLM technology is the potential shift from reactive to proactive engineering practices. The time savings identified in the analysis create opportunities for engineers to devote more attention to forward-looking activities rather than being consumed by administrative documentation of recent events.

With reduced time spent on routine documentation tasks, engineers can allocate more time to:

- Trend analysis and pattern identification across plant systems
- Predictive maintenance planning
- System health monitoring and improvement initiatives
- Design optimization and enhancement projects
- Proactive regulatory compliance planning.

4.3.3 Enhanced System Walkdowns and Field Presence

The time savings achieved through LLM implementation create opportunities for an increased engineering presence in the field. During interviews, engineers noted that administrative documentation requirements often compete with field time, forcing difficult prioritization decisions.

With reduced documentation time, engineers can:

- Conduct more frequent and thorough system walkdowns

- Increase face-to-face collaboration with maintenance and operations personnel
- Directly observe equipment performance in various operating conditions
- Identify potential issues through physical observation that might not be captured in data reports
- Provide more timely technical support for operations and maintenance activities.

This enhanced field presence offers several potential benefits, including improved equipment condition assessment, better cross-functional collaboration, and more comprehensive technical oversight of plant activities. Engineers with greater field presence gain deeper system knowledge through direct observation, which can inform more effective engineering decisions and recommendations.

The value of increased field presence is particularly significant in an aging fleet where subtle degradation mechanisms may not immediately be apparent through remote monitoring or periodic testing alone.

4.3.4 Long-Term Plant Reliability Improvements

The combination of time savings, more proactive engineering focus, and enhanced field presence creates conditions for long-term plant reliability improvements. By redirecting engineering resources toward activities that directly impact equipment performance and system health, utilities can potentially realize benefits including:

- Reduced equipment failures through earlier identification and resolution of degraded conditions
- More effective implementation of industry operating experience
- Enhanced component and system optimization
- Improved maintenance strategy development
- More comprehensive aging management
- Better integration of design and operational considerations.

Engineers highlighted how the reduced administrative burden could enable them to address emerging issues before they develop into significant problems. This shift toward proactive engineering creates potential for substantial cost avoidance through the early identification and mitigation of equipment reliability issues, reducing forced outage risks and optimizing maintenance scheduling.

The value of this shift extends beyond immediate operational improvements to potentially impact broader plant performance indicators, including improved net capacity factors, maintenance costs, and regulatory performance metrics.

4.3.5 Improved Knowledge Transfer and Mentoring

Another significant indirect benefit is the potential for improved knowledge transfer in an industry facing demographic challenges with an aging workforce. The LLM tool not only captures institutional knowledge from historical documentation but also creates a capacity for more effective mentoring and knowledge transfer among staff.

With reduced administrative burden, experienced engineers can:

- Dedicate more time to mentoring less experienced staff
- Participate in knowledge transfer activities like procedure validation and enhancement
- Document insights and lessons learned more comprehensively
- Develop improved training materials and reference documents
- Engage in more thorough technical reviews of work products from less experienced engineers.

The time savings also benefit new engineers by providing them with LLM-assisted drafting capabilities that incorporate institutional knowledge, potentially accelerating their learning curve and reducing the risk of knowledge loss as experienced personnel retire.

During interviews, engineering managers highlighted this benefit, noting that reduced administrative burden on senior engineers created more opportunities for them to work directly with junior staff, sharing insights that might otherwise remain undocumented. This “multiplier effect” of experienced engineer knowledge being more effectively transferred throughout the organization represents a strategic benefit with long-term implications for maintaining technical excellence and operational safety.

4.4 Risks and Challenges

4.4.1 LLM Tool Accuracy and Reliability Concerns

Engineers consistently expressed concerns about the accuracy and reliability of information generated by LLMs, which can have a negative impact on its trustworthiness. A particular concern was the common problem of AI “hallucinations” where an LLM tool generates plausible but fabricated information, including model numbers and technical specifications.

These LLM accuracy issues directly impact the time-saving benefits, as engineers must thoroughly verify all LLM-generated information.

The technical complexity of nuclear operations introduces additional reliability challenges. There are concerns about whether LLMs could properly interpret system conditions, such as distinguishing between active and inactive boric acid leaks with their different colorations. Similarly, questions were raised about LLMs’ ability to analyze the temperatures, concentrations, and operational states essential for accurate corrosion analysis.

4.4.2 Tool Over-Reliance and Impact on Engineer Competency

A consistent concern across interviews was the potential negative impact on engineer competency through over-reliance on LLMs. Multiple interviewees expressed the opinion that excessive LLM and wider AI dependence could diminish critical thinking skills.

The potential skill loss was compared to historical technological transitions, suggesting legitimate concerns about fundamental skill degradation. There was particular concern about newer staff missing valuable learning opportunities, as the traditional hands-on approach to difficult problems provides comprehensive learning that might be bypassed with AI assistance.

The interviews revealed specific concerns about critical thinking development. The research process for evaluations provides valuable learning opportunities that might be lost with LLM automation. Engineers worried that staff might not develop the depth of understanding necessary for complex safety evaluations if they rely too heavily on LLM-assisted research. This concern may also introduce possible risks of mistrust where engineers incorrectly trust the LLM tool in circumstances where it should not be trusted (i.e., it provides factually incorrect information). Engineers recognized the need for clear guidance on effective LLM usage, with some suggesting a structured implementation approach with appropriate guardrails for newer engineers.

4.4.3 Cultural Adaptation Challenges and Managing Organizational Change

The interviews revealed significant cultural adaptation challenges within the organization. Engineers exhibited varying levels of acceptance toward LLM and broader AI technology, with a possible pattern emerging across different demographics. More experienced engineers tended to be more cautious about AI adoption, while there were concerns that younger engineers might be too quick to trust the outputs without proper verification.

4.5 Business Case Approaches and Economic Analysis

4.5.1 Overview

The NPV analysis for the LLM AI implementation project provides a financial framework for evaluating the long-term economic benefits of technology investment. This analysis examines both direct efficiency gains (time savings) and their potential to unlock broader strategic value through the redeployment of engineering resources to higher value activities.

The business case includes three complementary frameworks. Business Case Approaches 1 and 2 provide alternative frameworks for quantifying the direct efficiency gains from LLM implementation—one based on specific use cases, and one based on overall engineering workload. Business Case Approach 3 then examines how these efficiency improvements could be translated into specific indirect benefits through enhanced reliability. While the frameworks analyze these benefits separately for clarity, they represent connected aspects of the same value proposition rather than competing alternatives.

The approach to this financial assessment recognizes that the primary value of efficiency improvements lies not in immediate cost reductions but in the redeployment of engineering talent toward activities that enhance plant safety, reliability, and performance. This aligns with the industry's focus on maintaining critical engineering expertise while improving work processes. The direct value of time savings calculated in Business Case Approaches 1 and 2 creates the foundation for the indirect benefits examined in Business Case Approach 3.

For this analysis, the identified time savings were valued using a pro rata calculation based on a fully loaded annual engineering salary of \$185,000, which translates to approximately \$96/hr assuming 1,920 working hours per year. This approach provides a conservative monetary valuation of the efficiency benefits while recognizing that the strategic value of redirected engineering effort likely exceeds this baseline calculation.

While the business case approaches presented here demonstrate substantial value potential in the medium term, the complete realization of these efficiency benefits will likely be unlocked through broader business transformation, as discussed in Section 4.5.7.

4.5.2 Use Case Efficiency Benchmark

Our detailed workload analysis of three specific engineering processes, BAC evaluations, maintenance rule evaluations, and 50.59 screenings, provides a foundation for understanding an LLM technology's efficiency impact. These use cases were selected for in-depth analysis due to their documentation requirements, standardized processes, and frequency of occurrence.

The analysis revealed substantial efficiency gains:

- BAC Evaluations: 11%–23% time reduction (5–61 minutes saved per evaluation)
- Maintenance Rule Evaluations: 19%–39% time reduction (31–751 minutes saved per evaluation)
- 50.59 Screenings: 38%–59% time reduction (82–797 minutes saved per evaluation).

In aggregate, these three engineering processes alone yield annual efficiency gains of approximately 330 hours in the expected case (ranging from 235 to 405 hours depending on assumptions). When valued using the engineering rate mentioned above, this represents an annual value of approximately \$32,000 in the expected case (ranging from \$23,000 to \$39,000).

On average, each engineering use case yields approximately 110 hours in annual efficiency gains (ranging from 78 to 135 hours per use case), representing an average efficiency improvement of 32% (ranging from 11% to 59% depending on the specific process). In monetary terms, this translates to approximately \$10,500 in annual value per use case (ranging from \$7,500 to \$13,000). This average

provides a useful benchmark for estimating potential benefits as LLM implementation expands to additional use cases throughout the organization.

4.5.3 Implementation Costs

When evaluating an LLM technology implementation, utilities must consider both the direct financial costs and the significant organizational effort required. The Reference Plant's experience highlights the importance of planning for both aspects.

The NPV analysis uses a discount rate of 6.3%, which represents the opportunity cost of capital for Vistra at the time of writing.

4.5.3.1 Initial Implementation Costs

These one-time expenses establish the foundation for LLM technology deployment:

- **Initial LLM Licensing:** Initial setup fee for the AI platform and environment, which includes the first year of access to the service
- **Data Preparation and Model Training:** Costs associated with collecting, cleaning, and processing plant documentation to train the LLM model
- **System Integration:** Technical work to connect the LLM platform with existing plant systems and document repositories.

4.5.3.2 Ongoing Operational Costs

Annual recurring expenses to maintain and enhance the LLM system include:

- **Technical Support and Maintenance:** Support for system operation and troubleshooting
- **Annual Licensing Fees:** Recurring costs for maintaining software licenses
- **System Updates and Model Refinement:** Expenses for periodic model retraining with new documentation
- **Continued Training:** Costs to onboard new staff and provide advanced training for existing users.

4.5.3.3 Staff Time for Implementation

The Reference Plant's experience revealed that internal resource requirements represent a significant and often underestimated aspect of implementation. The complete implementation timeline spans approximately four years before full benefits can be realized. Industry experience has shown that implementations can move significantly faster, with some utilities achieving full deployment in 2–3 years depending on resource allocation and organizational commitment. Vistra's approach at Comanche Peak with dedicated resources, including a full-time AI engineering lead and supporting data scientists, positions them to potentially accelerate this timeline:

Implementation Year 1: Initial Model Training and Documentation Preparation

Engineers must identify, organize, and review documentation for inclusion in the model training dataset. This typically takes a full year and involves significant effort to determine which documents contain relevant information and to ensure they are in appropriate formats for training. This phase also includes initial output validation where engineers must dedicate time to reviewing LLM-generated outputs and providing feedback for model refinement, often requiring multiple iterations.

Implementation Year 2: Validation and Refinement

Subject matter experts must spend time working with the LLM solution developers to explain nuclear-specific terminology, processes, and regulatory frameworks to ensure the model is properly

calibrated to the nuclear context. Engineers continue to validate outputs, test capabilities, and provide feedback for improvements.

Implementation Years 3–4: Process Integration and Change Management

Workflow adjustments are necessary to incorporate the LLM tool into existing engineering processes, requiring development and testing of new procedures. This phase typically takes 1.5–2 years. Staff time is required to develop training materials, conduct training sessions, and provide ongoing support to users as they adapt to the new technology. This phase includes organizational change management to address cultural adaptation challenges. By the end of Year 4, the system is expected to be fully implemented with benefits beginning to monetize.

Based on the Reference Plant experience, total staff implementation hours range from:

- Low Estimate: 200 hours (\$18,800)
- Moderate Estimate: 500 hours (\$47,000)
- High Estimate: 1,000 hours (\$94,000).

These staff implementation hours are monetized at the same \$94/hr rate and included in the first year costs.

4.5.4 Business Case Approach 1: Expanding Use Cases Across the Plant

This approach examines how the economic value of the LLM implementation changes substantially when deployment is expanded beyond the initial three engineering processes to address the broader range of applications discussed in Section 4.2. By distributing implementation costs across multiple functions and applications, the overall value proposition becomes significantly more favorable.

Based on the analysis of the three initial use cases, efficiency gains can be projected for additional use cases across three potential outcomes:

- **Low Case:** Approximately 78 hours saved per use case annually
- **Expected Case:** Approximately 110 hours saved per use case annually
- **High Case:** Approximately 135 hours saved per use case annually.

When these efficiency gains are valued at \$96/hr, the annual value per application ranges from \$7,500 (low) to \$10,500 (expected) to \$13,000 (high). Table 9 illustrates how the NPV transforms as implementation scales across the enterprise under each scenario:

Table 9. Annual Efficiency Savings

Number of Use Cases Deployed	Annual Efficiency Savings Per Use Case		
	Low Case	Expected Case	High Case
10	\$75k	\$105k	\$130k
20	\$151k	\$211k	\$260k
25	\$188k	\$264k	\$325k
30	\$226k	\$316k	\$390k
50	\$376k	\$527k	\$650k
100	\$753k	\$1,055	\$1,301

Table 10 presents the resulting NPV and internal rate of return (IRR) calculations over a 10 year period for various implementation scales. As shown in the table, smaller implementations (10–20 use

cases) show negative NPVs across most scenarios, while implementations at 25 use cases or more yield positive NPVs in the expected and high cases. The IRR analysis similarly demonstrates that larger scale implementations provide increasingly attractive returns, with 50 use cases implementations generating IRRs ranging from 94% in the expected case to 287% in the high case.

Table 10. NPVs and IRRs with enterprise implementation.

Number of Use Cases Deployed	NPV (10 Year Period)			IRR (10 Year Period)		
	Low Case	Expected Case	High Case	Low Case	Expected Case	High Case
10	-\$1.8M	-\$0.9M	-\$0.4M	N/A	N/A	N/A
20	-\$1.3M	-\$0.2M	\$0.6M	N/A	-7%	50%
25	-\$1.0M	\$0.2M	\$1.1M	N/A	18%	80%
30	-\$0.7M	\$0.6M	\$1.5M	N/A	35%	110%
50	\$0.4M	\$2.1M	\$3.4M	16%	98%	230%
100	\$3.1M	\$5.9M	\$8.1M	71%	250%	528%

4.5.5 Business Case Approach 2: Engineer Value Framework

The engineer value approach examines how LLM technology improves overall engineering productivity across the organization. Rather than focusing on specific use cases, this approach analyzes how the LLM implementation enhances efficiency across an engineer's entire workload, enabling a more effective utilization of specialized technical talent.

The detailed analysis of the three representative engineering processes discussed above demonstrated efficiency improvements ranging from 11% to 59% when using LLM technology, with an average of 32% across these specific activities. Translating these task-specific improvements into organization-wide impacts involves considering:

1. **Workload portion that can be assisted by LLM technology:** Based on interviews with engineers and management, it is estimated that documentation and research activities amenable to LLM assistance constitute approximately:
Low case: 30% of an engineer's total workload
Moderate case: 40% of an engineer's total workload
High case: 50% of an engineer's total workload
2. **Varying degrees of efficiency improvement:** While our analysis showed an average 32% improvement, we account for implementation variability:
Low case: 25% efficiency improvement on LLM-assisted tasks
Moderate case: 32% efficiency improvement on LLM-assisted tasks
High case: 40% efficiency improvement on LLM-assisted tasks
3. **Resulting organization-wide productivity improvement:**
Low case: 25% efficiency $\times$ 30% of workload = 6% overall efficiency improvement
Moderate case: 32% efficiency $\times$ 40% of workload = 11% overall efficiency improvement
High case: 40% efficiency $\times$ 50% of workload = 18% overall efficiency improvement.

For the Reference Plant, which has **60** engineers, this translates to the annual efficiency gains shown in Table 11.

Table 11. Engineer value framework: Efficiency gains and financial benefits.

Scenario	Calculation	Annual Hours	Annual Benefits	10 Year NPV	IRR (%)
Low	60 engineers × 1,920 hours × 7.5%	8,640 hours	\$833k	\$3.8M	93%
Expected	60 engineers × 1,920 hours × 13%	14,976 hours	\$1.4M	\$8.8M	362%
High	60 engineers × 1,920 hours × 20%	23,040 hours	\$2.2M	\$14.8M	950%

This approach provides a more holistic view of efficiency improvements by considering the overall engineering workload rather than individual use cases. It demonstrates how even relatively modest improvements in efficiency (7.5%–20%) across the entire engineering function can create substantial value when implemented across all engineers at a site or fleet.

4.5.6 Business Case Approach 3: Avoided Cost Reliability Benefit Framework

This approach builds upon the direct efficiency gains quantified in Approaches 1 and 2 to examine how these improvements could translate into substantial indirect benefits through plant reliability enhancements. While the direct efficiency gains documented in the previous frameworks provide measurable value, this approach explores how redirecting engineering capacity toward proactive activities can deliver even greater financial benefits through avoided plant trips and equipment failures.

This scenario considers how a comprehensive implementation of an LLM across multiple engineering functions could create efficiency savings that allow engineers to spend more time on high-value activities that enhance plant reliability and lead to tangible reductions in equipment failures that cause forced outages. As engineers gain more capacity through AI-assisted workflows across numerous use cases, they can redirect this capacity toward more proactive and strategic activities, such as monitoring equipment, addressing technical backlogs, and implementing strategic upgrades that directly impact plant reliability.

The relationship between increased engineering capacity and improved plant reliability is supported by feedback from plant personnel, though further analysis will be necessary to determine exactly when and how these reliability benefits would be realized. This scenario provides a framework for considering these potential benefits based on real plant experiences.

During interviews with engineering staff at the Reference Plant, one senior engineer provided a detailed example of a long-standing issue with a turbine control system that had caused multiple plant trips over a period of years. The engineer explained that, in the past, engineers had sufficient time to address such reliability issues proactively but increasing administrative and documentation burdens have reduced their capacity for such high-value work. According to this engineer, the technology to resolve this issue had been available since 2005, but human resource constraints delayed implementation for nearly two decades: *“Had we implemented that back when we started having these problems it probably would have eliminated ten to twelve plant trips because of that one issue.”*

To quantify the potential reliability impact, plant trip data from the Reference Plant related to this specific issue was obtained. The data revealed substantial lost generation over just a recent 4 year period, see Table 12.

Table 12. Reliability event historical data from turbine control system issues.

Year	Unit	Lost Generation	Value
2021	Unit 1	29,646 MWh	\$0.9M
2023	Unit 1	153,003 MWh	\$4.6M
2024	Unit 2	17,166 MWh	\$0.5M
2024	Unit 2	245,102 MWh	\$7.4M
2024	Unit 1	186,067 MWh	\$5.6M

The values above are calculated using an electricity price of \$30/MWh, resulting in an average event value of \$3.8M. These figures represent only direct lost generation revenue and do not include additional costs associated with recovery, restart, inspections, regulatory compliance activities, and system wear and tear.^d Using these data points, we modeled three scenarios for potential reliability improvements, see Table 13.

Table 13. Reliability impact scenario modeling results.

Scenario	Annual Reliability Value	10 Year NPV	IRR (%)
Low—1 trip avoided every 10 years	\$380k	\$0.4M	16
Moderate—2 trips avoided every 10 years	\$760k	\$3.8M	164
High—3 trips avoided every 10 years	\$1.14m	\$5.5M	360

The interviews revealed multiple mechanisms through which enhanced engineering capacity could improve reliability, including:

1. **Earlier Identification of Equipment Issues:** Engineers reported that, with more time for system walkdowns and field presence, they could identify degrading equipment conditions before they cause failures.
2. **Addressing Technical Backlogs:** One engineer noted that “every engineer carries a pretty heavy backlog of old CAP items they’re trying to close out,” and that time constraints often prevented them from resolving known issues until they became emergent problems.
3. **Strategic Technology Implementation:** The turbine control system example demonstrates how time for strategic planning and implementation of technology upgrades can prevent recurring issues.
4. **Enhanced Cross-Functional Collaboration:** Engineers emphasized that having more time for “working with the technicians that are actually working on the equipment” provides invaluable insights that cannot be gathered from documentation alone.
5. **Industry Engagement:** Engineers reported that time constraints sometimes prevented them from participating in user groups and industry collaboration, limiting their access to solutions developed at other plants.

^d The turbine control system trip values shown represent actual historical losses at Comanche Peak and may be lower than industry-average trip costs. Plant engineers estimate that average trip costs range from \$4.5M–\$5M when including typical repair activities. The conservative nature of these assumptions strengthens the business case, as more complex trips requiring extended outages or equipment replacement would yield significantly higher avoided cost benefits.

This reliability impact framework demonstrates that the indirect benefits realized through redirected engineering capacity may substantially exceed the direct efficiency gains quantified in Frameworks 1 and 2. While the connection between specific time savings and reliability improvements requires further validation, the plant data and engineer testimonials provide compelling evidence that even modest improvements in engineering capacity can translate to significant avoided costs over time. These reliability benefits represent a critical component of the medium-term value proposition for LLM implementation, bridging between the immediate efficiency gains and the longer term business transformation discussed in Section 4.5.7. This approach illustrates how the efficiency improvements quantified in earlier approaches can be translated into tangible business outcomes beyond simple time savings.

4.5.7 Future Value Realization: Plant- and Fleet-Level Business Transformation

As noted in Section 3.5 and further detailed in Section 4.2, the potential applications of LLM technology extend far beyond plant engineering functions. When viewed from a fleet-level perspective, these broader applications represent a substantial efficiency opportunity across multiple functional areas. Analysis by plant leadership indicates that approximately 35% of the nuclear workforce occupies roles that could benefit significantly from LLM and other AI technologies. In a typical nuclear fleet, this represents hundreds of staff members across diverse functions, including training, projects, oversight and audits, emergency planning, and maintenance planning.

These technical and administrative positions share similar documentation, research, and procedural requirements that make them suitable for AI assistance, though the efficiency gains would vary based on their specific task profiles. The non-engineering use cases identified in Section 4.2, such as work management, operational decision-making, and compliance reporting, illustrate how the same core AI capabilities can be applied across multiple functional areas to create value throughout the organization.

As discussed, to fully realize these identified efficiency gains, a structured and comprehensive approach to organizational modernization is required. Individual efficiency improvements, while valuable, achieve their maximum impact when integrated into a comprehensive business transformation that strategically redeploys capacity and reimagines work processes.

This transformation aligns closely with the ION framework discussed at the beginning of this report, which provides a strategic approach for modernizing nuclear operations through process optimization, technological innovation, workforce transformation, and governance adaptation. Such an approach would move beyond individual use case efficiency gains to reimagine how engineering resources are deployed, potentially reorganizing workflows to eliminate redundant activities and better utilize specialized skills.

Following the principles of the ION framework, enterprise wide implementation would leverage LLM technology alongside other digital tools to create an integrated ecosystem where information flows across traditionally siloed departments. Such an integration enables not just more efficient workflows but also new approaches to solving problems and making decisions that weren't previously possible. The LLM-enabled efficiency gains documented in this report directly support ION's core objective of transforming the workforce from task-based labor to knowledge-centric roles supported by digital systems.

The long-term transformation enabled by this comprehensive approach may ultimately influence organizational structure itself. While immediate workforce implications focus on freeing up time for higher value activities, this transformation may eventually create opportunities for staffing models to change in line with natural attrition, without disrupting operations or losing critical expertise. These organizational changes represent the final stage of value realization, completing the journey from targeted efficiency improvements to comprehensive business transformation.

5. RECOMMENDATIONS AND NEXT STEPS

5.1 Use-Case Expansion and Opportunities for Further Research

5.1.1 Post-Implementation Business Case Validation

Future research should focus on validating the efficiency gains identified in the workload analysis of the three engineering processes examined in this study. This validation would involve collecting detailed time measurement data from engineers using an LLM for BAC evaluations, maintenance rule evaluations, and 50.59 screenings. The validation should compare actual time savings against the projected improvements. Future validation should also track user adoption rates, quality improvement metrics, and impact on engineer field presence. In addition, structured feedback sessions with engineering teams and management would be valuable to assess qualitative improvements in documentation, communication, and review processes.

5.1.2 Validating the Engineer Value Framework

Future research should build upon the engineer value framework presented in Section 4.5.5 through systematic data collection and analysis across multiple sites and engineering roles. This research would involve developing metrics to quantify more precisely what percentage of engineering workload is applicable to LLM assistance across different engineering disciplines and how the efficiency gains vary by task type and complexity. Studies should track actual time allocation before and after LLM implementation.

5.1.3 Quantifying Benefits for Additional Engineering Use Cases

Although the current analysis focuses on three engineering use cases, the additional use cases identified in Section 4.1.4 suggest that many other engineering applications deliver similar levels of time savings. Future research should expand the quantitative analysis to include these additional engineering applications, such as license renewal documentation, system health reports, and operability evaluations.

This expanded analysis would allow for a broader economic assessment, capturing direct savings in worker hours across a wider spectrum of engineering functions. By applying the workload analysis methodology developed in this study to these additional use cases, researchers could confirm whether the observed percentage improvements (ranging from 11% to 59%) are consistent across different types of engineering work and provide a comprehensive picture of the aggregate value generated by LLM integration.

5.1.4 Quantifying Benefits in Other Plant Functions

LLM integration has implications that extend beyond engineering, with several non-engineering departments showing potential for similar efficiency improvements. Further studies could identify and document specific processes in administrative, regulatory, and operational support areas that could benefit from AI-enabled automation. These studies could estimate the time savings and enhancements in workflow efficiency within these departments. By consolidating the benefits realized in non-engineering functions with the savings in engineering, future research aims to provide an integrated view of the overall enterprise value generated by LLM integration.

5.1.5 Quantifying Reliability and Avoided Cost Benefits

Future research should develop methodologies to quantify the relationship between increased engineering capacity and plant reliability improvements. This could include a historical analysis of events where timely engineering intervention could have prevented forced outages or power reductions, establishing a framework for valuing these avoided costs.

5.1.6 Measuring the Impact of AI Through Alternative Success Metrics

While the primary quantitative analysis focused on direct labor savings, a more comprehensive evaluation of LLM integration should consider alternative success metrics to capture the “indirect” and broader impact of the technology (as discussed in Section 4.3). Together, these metrics provide a broader view of LLM value and effectiveness that reflects both direct and indirect benefits derived from LLM application.

5.1.6.1 User Adoption Rates and Tool Utilization

A critical measure of the LLM tool’s success is the rate at which engineering staff adopt and effectively use the technology. Tracking metrics, such as login frequency, active user counts, task completion rates, and user feedback, helps determine whether the tool is being integrated into daily workflows. High adoption rates and consistent utilization not only indicate user buy-in (i.e., trust in the AI) but also serve as leading indicators of the potential for broader operational impact. A further analysis can help determine whether high utilization correlates with improved performance across other operational metrics.

5.1.6.2 Quality Improvements in Documentation

Another key metric is the qualitative enhancement in the documentation produced by the LLM tool. Indicators here include reduced instances of typographical and grammatical errors, improved clarity, and enhanced technical accuracy. Engineers have noted that LLM-generated drafts may well require fewer revisions, thereby reducing the likelihood of misinterpretation. Future research should quantify the impact of improved document quality on overall process efficiency and decision-making effectiveness, providing a clear picture of benefits that extend beyond time saved.

5.1.6.3 Documentation Standardization and Consistency

Standardization is a direct outcome of using templated AI systems. An LLM enforces consistent formatting and language usage, which simplifies reviews and ensures that all technical reports adhere to predetermined standards. Tracking improvements in standardization can highlight how much the tool reduces variability across documents. Further research should seek to correlate standardization with improvements in overall operational performance.

5.1.6.4 Increased Engineer Productivity

A key indirect benefit of reducing administrative burdens is the increased capacity for engineers to spend more time on the plant floor. This metric will measure the reallocation of engineering hours from desk-based tasks to direct field engagement. Enhanced field presence can lead to more effective onsite inspections, quicker issue resolution, and improved proactive maintenance practices. Future research should employ observational studies and performance correlation analyses to quantify the impact of increased field presence on operational safety and overall plant performance.

5.1.6.5 Improvements in Plant Reliability

A central business case for LLM implementation, as detailed in Business Case Approach 3, is the potential for significant plant reliability improvements through redirected engineering capacity. This stands as a primary value driver that extends beyond efficiency metrics into direct financial impact through avoided forced outages and generation losses. Research should focus on establishing these causal relationships and creating frameworks that allow nuclear utilities to confidently attribute reliability improvements to the enhanced engineering capacity enabled by LLM implementation. This approach will provide a more complete picture of the technology’s full value proposition.

5.1.6.6 Reduction in Review Cycles and Rework

Streamlined review processes are a tangible benefit of LLM integration. By producing high-quality initial drafts that closely align with established templates and protocols, the tool reduces the need for multiple review iterations and extensive rework. Future studies should analyze whether these shorter cycles directly contribute to improved project turnaround times and enhanced decision-making, thereby adding layers of value beyond simple time savings.

5.1.6.7 Impact on Training and Onboarding Processes

The indirect benefits of LLM technology extend into training and onboarding. New and less experienced engineers gain access to well-crafted, standardized document templates that serve as both learning tools and benchmarks for quality. This metric will assess how quickly new staff reach proficiency and the extent to which the tool reduces the need for extensive one-on-one training. In doing so, it serves as an indicator of broader organizational benefits, as experienced engineers are freed up for mentorship and strategic tasks. Research in this area will quantify improvements in learning curves and the overall enhancement of knowledge transfer across teams.

5.2 Fleet-Level Business Transformation Strategies

To optimize LLM implementation as part of a comprehensive fleet-level transformation, utilities should move beyond site-specific deployments to capture the additional efficiencies and advantages of coordinated, enterprisewide implementation. Future INL research aligned with the ION framework discussed in Section 4.5.7 can support these utility efforts by examining optimal fleet-level deployment strategies, governance models, and organizational structures that enable the transition from isolated implementations to integrated, knowledge-centric operations across multiple sites.

5.2.1 Breaking Down Organizational Silos

LLM technology may enable the breakdown of traditional silos between departments and sites, facilitating more integrated operations and decision-making processes across the fleet. This transformation from task-based to knowledge-centric operations requires evolving organizational structures and governance models to align with the ION framework vision.

5.2.2 Organizational Enablement and Governance

Successful fleet-level transformation depends on establishing cross-functional AI leadership teams capable of managing deployment, updates, and data governance across multiple sites. Practical project management offices or delivery teams coordinate implementation while maintaining operational continuity. Implementing transparent processes for prioritizing use cases can ensure alignment with both site-specific needs and fleetwide strategic objectives.

5.2.3 Knowledge Sharing and Best Practices

Effective knowledge-sharing mechanisms help to capture and disseminate AI implementation lessons learned across fleets, accelerating adoption and avoiding duplicated efforts. This includes platforms and processes for sharing AI wins, challenges, and solutions across sites, creating a continuous improvement cycle.

5.2.4 Workforce Evolution and Champion Development

Early AI adopters serve as champions to drive broader organizational change by becoming AI advocates who lead adoption at other sites and within their departments. Aligning AI integration with natural workforce transitions through attrition enables role modernization while maintaining operational expertise.

5.2.5 Scaling Strategies and Resource Planning

Optimal staffing models balance initial investment with sustainable growth, starting with small and dedicated teams of between four and six 4-6 Full-Time Equivalents and scaling strategically as value is demonstrated. Connecting AI efforts to broader operational technology and digital transformation strategies maximizes organizational impact and prevents isolated technology implementations.

5.2.6 Balancing Standardization with Customization

Fleet-level implementation strategies balance standardization with site-specific customization, ensuring consistent implementation while accounting for the specific contexts of individual plants. Developing a roadmap that connects immediate efficiency gains at the plant level to long-term business transformation at the fleet scale can maximize the strategic value of LLMs.

6. REFERENCES

1. S. J. Remer, J. Hansen, C. Kovesdi, Z. Spielman, S. Lawrie, L. Martin, M. Dep, and B. Szews. 2022. “Integrated Operations for Nuclear Business Operation Model Analysis and Industry Validation.” INL/RPT-22-68671, Idaho National Laboratory.
<https://www.osti.gov/servlets/purl/1894935>.
2. S. J. Remer, J. Hansen, C. Kovesdi, Z. A. Spielman, S. Lawrie, L. Martin, A. Tylecote, and B. Szews. 2023. “Applying the ION Business Model to a Domestic Nuclear Plant: Assessment and Transformation Implementation Plan.” INL/RPT-23-73942, Idaho National Laboratory.
3. S. J. Remer, J. Hansen, C. Kovesdi, Z. Spielman, S. Lawrie, L. Martin, M. Dep, and B. Szews. 2022. “Integrated Operations for Nuclear Business Operation Model Analysis and Industry Validation.” INL/RPT-22-68671, Idaho National Laboratory.
<https://www.osti.gov/servlets/purl/1894935>.
4. V. Agarwal, K. A. Manjunatha, A. V. Gribok, T. J. Mortenson, and H. Palas. 2022. “Scalable Risk-Informed Predictive Maintenance Strategy for Operating Nuclear Power Plants.” Presented at PSAM 16, June 26–July 1, 2022, Honolulu, HI.
5. J. A. Farber, A. Y. Al Rashdan, and R. D. Reese. 2023. “Anomaly Detection and Identification Using a Leave-One-Variable-Out Method.” INL/CON-23-71470, Idaho National Laboratory.
<https://www.osti.gov/servlets/purl/2468627>.
6. Hunton, P. J., and S. Lawrie. 2024. “Integrated Operations for Nuclear: Work Reduction Opportunity Demonstration.” INL/RPT-24-80256, Idaho National Laboratory.
<https://www.osti.gov/servlets/purl/2475050>.
7. EPRI. 2012. “Boric Acid Corrosion Guidebook: Managing Boric Acid Corrosion Issues at PWR Power Stations – Revision 2.” Report No. 1000975, EPRI.

Page intentionally left blank

Appendix A
DATA COLLECTION AND STATISTICAL MODELING
METHODOLOGY

Page intentionally left blank

Appendix A

DATA COLLECTION AND STATISTICAL MODELING METHODOLOGY

A-1. DATA COLLECTION AND ANALYSIS PROCESS

The analysis began with structured interviews at the Representative Plant to establish baseline time requirements for completing three evaluation types: BAC evaluations, maintenance rule evaluations, and 50.59 screenings. Two to three engineers were interviewed for each use case to provide quantitative workload data, with additional qualitative insights gathered from engineering managers.

Engineers provided estimates for the minimum time (best case), maximum time (worst case), and typical time for completing each evaluation type manually. For analysis purposes, the minimum and maximum values were interpreted as approximating the 5th and 95th percentiles of the distribution, though this was an analytical assumption rather than an explicit statement from the engineers. The final dataset used the minimum of all low estimates and maximum of all high estimates across engineers, while the “typical” estimates were averaged to establish a central tendency.

After compiling the initial time estimates, the results were validated with the engineering team and minor adjustments were made to ensure the data accurately reflected real-world experience.

A-2. MONTE CARLO SIMULATION DEVELOPMENT

To account for the inherent variability in task durations and AI assistance benefits, a correlated Monte Carlo simulation with nested sampling was developed in Excel.

A-3. MANUAL TIME DISTRIBUTION

For each evaluation type, a right-skewed distribution was selected to reflect the reality of engineering tasks. Right-skewed distributions are characterized by a long tail extending toward higher values, resulting in a mean that is higher than the median. This pattern is particularly appropriate for task duration modeling because engineering evaluations typically follow this pattern: most can be completed in a predictable timeframe near the typical (median) value, but a small percentage require significantly more effort due to unforeseen complications or complexity, creating the extended tail toward longer durations.

The beta distribution was specifically chosen to model this pattern because it can be parameterized to create the desired right skew while still constraining values within a defined range. The parameters were calibrated to match the assumed 5th percentile (minimum), typical case (median), and 95th percentile (maximum) values reported by engineers for each evaluation type. For example, in BAC evaluations, the formula generated time values between 37 minutes (minimum) and 265 minutes (maximum), with most values clustered closer to the typical time of 110 minutes but occasionally producing longer durations to represent complex cases.

This approach realistically models the engineering work pattern where most evaluations take close to the typical time, but a small number of complex cases take significantly longer. The right-skewed distribution ensures the simulation captures these occasional complex cases that, though infrequent, represent a disproportionate amount of the total engineering effort and therefore present the greatest opportunity for efficiency gains.

A-4. SPECIAL CONSIDERATIONS FOR MAINTENANCE RULE EVALUATIONS

Maintenance rule evaluations required special modeling due to their sequential, multistage process and varying complexity levels. Plant data shows that, out of 481 maintenance rule evaluations over the 5-year time period analyzed, only 77 (16%) were identified as a functional failure (FF) and continued past the FF evaluation stage. Of these, 37 were determined to be maintenance preventable FFs (MPFFs), and only 4 (0.8% of total) were classified as repetitive MPFFs (RMPFFs).

To accurately model this distribution, the simulation first determined the evaluation type based on historical frequencies by generating a random number and comparing it against the cumulative probabilities for each type (84% non-FF, 8% FF, 7.2% MPFF, and 0.8% RMPFF). Then, based on the evaluation type, the appropriate sequence of activities was simulated by summing the time for relevant tasks. For example, a non-FF evaluation would only include initial screening tasks, while an RMPFF evaluation would include the complete sequence of tasks, including repetitive failure analysis and performance criteria evaluation.

This approach created moderate variability in task duration that accurately reflected the different evaluation paths. For instance, non-FF evaluations might only take 40–350 minutes, while the rare RMPFF evaluations could take 150–1,935 minutes, correctly modeling the full range of complexity present in real-world maintenance rule evaluations.

A-5. AI IMPACT FACTOR CORRELATION

Based on engineer feedback that AI assistance provides greater time savings for more complex evaluations, a correlation factor of approximately 0.8 was implemented between task duration and AI efficiency gains. This correlation reflects the observation that AI tools deliver their highest value in complex scenarios requiring extensive research and documentation.

The model calculated the AI impact factor by combining two components:

1. **A deterministic correlation component (weighted at 80%)** that increased the impact factor proportionally to the task duration. This component creates a direct relationship between task complexity (as measured by time) and AI impact, reflecting the insight that AI tools provide greater assistance when there is more information to process, more documentation to generate, or more research to conduct.
2. **A random variation component (weighted at 60%)** that added moderate variability to the AI impact. This component acknowledges that the relationship between task complexity and AI benefit is not perfectly deterministic—other factors such as task type, available documentation, and specific content also influence AI effectiveness.

It is important to note that this 0.8 correlation factor is an assumption of the model based on engineer observations and early implementation experience. This assumption could be debated and refined as more data becomes available from expanded AI implementation. If the actual correlation is lower, the model may overestimate the benefit of AI for complex cases; if the correlation is higher, the benefits for complex cases may be underestimated.

The impact of this approach is significant for the overall results: because the model assumes AI provides a greater percentage time savings on more complex tasks, and because these complex tasks represent a disproportionate amount of the total time, the calculated efficiency gains are notably higher than would be predicted by a model assuming uniform benefits across all task types. This assumption therefore has a substantial influence on the final estimated time savings, particularly for evaluation types with high variability in task duration.

A-6. ANNUAL IMPACT CALCULATION USING NESTED MONTE CARLO SIMULATION

To generate robust statistical results, Excel's Data Table function was utilized to implement a nested Monte Carlo simulation:

1. The first level generated 1,000 different evaluation scenarios capturing the range of possible evaluation complexities. Each scenario calculated both the manual time required and the time saved through AI assistance.
2. Excel's Data Table function then created 100 annual simulations based on this pool of 1,000 scenarios. Each annual simulation randomly selected the appropriate number of evaluations for each type (based on historical frequency data) and summed the time savings.
3. This nested approach effectively created 10,000 different simulation runs, providing high statistical confidence in the results.
4. The results were multiplied by the number of annual evaluations by type, based on work management data from the Representative Plant, to calculate the total annual time savings.
5. From these results, the mean time savings and percentage efficiency gains were calculated.