Article

Not peer-reviewed version

An Industrial Application of a Large Language Model to Enhancing Asset Integrity and Process Safety Management

Benard Ongwae *

Posted Date: 17 March 2025

doi: 10.20944/preprints202503.1172.v1

Keywords: asset integrity; process safety; large language model; generative pre-trained transformers; artificial intelligence; asset integrity; process safety; key performance indicators; predictive analytics

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2025 doi:10.20944/preprints202503.1172.v1

Article

An Industrial Application of a Large Language Model to Enhancing Asset Integrity and Process Safety Management

Benard Ongwae

Department of Industrial and Systems Engineering, University of Pretoria, Pretoria 0002, South Africa; u25477618@tuks.co.za

Abstract: This research incorporates artificial intelligence (AI) into asset integrity and process safety (AIPS) management, aiming to revolutionize conventional methods. It facilitates the automation of risk assessments, enhances predictive analytics, and supports the development of proactive measures to mitigate potential incidents. It explores the application of a generative pre-trained transformer (GPT) based large language models (LLM) to analyse and classify AIPS indicators from vast datasets to generate actionable recommendations to prevent future incidents. A comparative study between two onshore liquefied natural gas (LNG) plants; one utilizing AI-driven AIPS management and the other relying on manual data analysis is presented. The results indicate that AI-driven approaches significantly enhance the accuracy and speed of incident classifications, reducing data processing times. The test model effectively predicts potential future failures by analysing past incident patterns, enabling informed decision-making to prevent and mitigate future failures. The findings highlight the importance of adopting AI-driven AIPS management as a standard practice. It also emphasises the need for stronger collaboration between academia and industry in AI solutions to drive technological advancements for sustainability.

Keywords: asset integrity; process safety; large language model; generative pre-trained transformers; artificial intelligence; asset integrity; process safety; key performance indicators; predictive analytics

---

1. Introduction

Asset Integrity and Process Safety (AIPS) Management is a structured framework for ensuring that the integrity of hazardous processes are maintained by implementing best engineering, operational and maintenance strategies. The primary objective is to prevent and control incidents that could result in the release of hazardous materials or energy with the potential of leading to catastrophic consequences, such as fatalities and irreversible property and environmental damage. This research investigates the potential of customizing a large language model (LLM) for a specific industrial case of AIPS to advance sustainability of onshore refineries and petrochemical installations. The case involves analysing vast databases to identify AIPS incidents, classifying them into their respective categories, and to generate matching mitigating recommendation for each. The objective is to leverage AI-driven insights to make timely decisions to avoid future incidents for sustainability.

The process industry is experiencing a swift technological evolution, driven by the accelerated adoption of artificial intelligence (AI) to promote sustainability efforts. This trend was prominently highlighted at the Abu Dhabi International Petroleum Exhibition and Conference (ADIPEC) held between 4th and 7th November 2024. ADIPEC is the largest and most comprehensive annual global event that brings together the world's key players in the energy sector under one roof to share knowledge and showcase technological advancements. The 2024 conference focused on AI-driven technologies as the key enabler of innovation to power the transition towards a more sustainable energy sector [1].

 doi:10.20944/preprints202503.1172.v1

Conversely, there is a reluctance in embracing AI, particularly large language models (LLMs), as learning tools to support research and reduce students' workload. LLMs can be trained to understand, predict, and generate human-like text that are both coherent and contextually relevant in response to inquiries. There are genuine concerns on their impact on students critical thinking, originality, and on academic integrity, leading to the cautious approach to their adoption [2,3]. However, it is imperative to fully embrace AI as a compliment to human intellect rather than supplant. Ref. [4], emphasised that integration of AI can only foster critical thinking among students allowing them to focus on higher-order cognitive skills essential for navigating complex problems. From this perspective, Ref. [5] claims that embracing AI could lead to innovation that actively involves critical, analytical thought. This positions the academia and the manufacturing industry to collaborate in harnessing the promises of AI while at the same time reinforcing the foundational principles of critical inquiry that drives academic research and innovation. One potential area of such collaboration is the integration of AI into AIPS management strategies.

One of the LLM tools is generative pre-trained transformers (GPT), that can analyse large data sets to identify trends and predict anomalies to facilitate quick decision-making [6,7]. To illustrate, by taking advantage of their natural language pro-cessing capacities, safety critical information such as loss of primary containment (LOPC) and instrumented functions trips can be extracted from large unstructured data sources to improve the quality of reporting [8]. However, to fully harness the potential of this technology, it is important to first address employee competencies, particularly in accurately capturing quality data using the various tools employed in the industry. Some of these tools are summarised below.

1.1. Industrial Data Management Tools

The optimisation of operations in the petrochemical sector depends much on data management and decision-making tools. These technologies include enterprise resource planning (ERP), laboratory information management systems (LIMS), asset management and manufacturing control software. ERP integrates workstreams for data flow across functions facilitating simplified operations by managing inventories, tracking production schedules and optimising maintenance and supply chain logistics [9]. Systems, Applications and Products in Data Processing (SAP) is the reference ERP, with its main modules being the main sources of data. For instance, SAP PM (Plant Maintenance) is used to plan maintenance tasks, schedules, workload distribution, and track asset performance metrics. These metrics include mean time between failures (MTBF), maintenance backlogs, and mean time to repair (MTTR). SAP MM (Materials Management) handles procurement and inventory management, ensuring that optimal levels of materials and spare parts are available when needed reducing costs associated with overstocking or stockouts; SAP PP (Production Planning) optimises manufacturing processes by ensuring that production activities are aligned with demand forecasts and resource availability; SAP FI (Finance) for financial transactions that include cost tracking, accounting, and financial reporting. Of the most importance to this work is SAP EHS that supports risk management through structured incident database management and root cause analysis to arrive at actionable corrective and preventative actions (CAPA). Integrating SAP EHS with AI technologies can maximise its capacity to providing predictive insights for better operational effectiveness and safety. [10].

LIMS are for laboratories' data management for product quality control. By automating data collection and information accessibility, LIMS helps compliance with set industrial standards such as ISO 9001. They provide valuable information on the quality and consistency of products, minimizing errors and time of operational inactivity [11].

Health, safety, and environmental management systems (HSEMS) drive compliance to established statutory and regulatory protocols while facilitating risk mitigation strategies. The incorporation of data-driven insights into process control and HSEMS can enhance the identification of potential hazards to implement appropriate mitigation measures before deviations or incidents can occur [12–14].

 doi:10.20944/preprints202503.1172.v1

Table 1 presents a selection of key data management tools along with their industrial applications. A subset of these tools serves as data sources utilized in this work for demonstration purposes.

Table 1. Examples of industrial data management tools.

Category	System/Tool	Description
Enterprise Resource Planning (ERP)	SAP	Integrated management of core business processes such as procurement, production, inventory, and finance.
	Oracle ERP	Modules for asset management, supply chain, and operational efficiency.
	IFS Applications	Asset management, project management, and real-time operational visibility.
	Microsoft Dynamics 365	Cloud-based solutions for operational data, financials, and asset management.
Process Control and Automation	Honeywell Process Solutions (HPS)	Process automation, real-time monitoring, and optimization.
	Emerson DeltaV	Distributed Control System (DCS) for production and operational control.
	Schneider Electric's EcoStruxure	Manages operational data with energy and asset optimization.
Laboratory Information Management Systems (LIMS)	LabWare LIMS	Manages lab samples, test results, and compliance.
	STARLIMS	Quality control and data tracking for testing petrochemical products.
	PerkinElmer Informatics	Solutions for research, development, and production analytics.
Asset and Maintenance Management	Maximo (IBM)	Asset management software for predictive maintenance and reliability.
	Infor EAM	Lifecycle and performance management of physical assets.
Production and Operation Management	AspenONE (AspenTech)	Process optimization; production planning.

1.2. Open-Source LL-GPT Models

This section examines the capabilities, strengths and limitations of open-source LL-GPT models to evaluate their suitability to manage AIPS performance.

GPT-Neo can provide simple summaries by processing both structured and semi-structured inputs. However, the need for domain-specific pretraining limits its efficacy in handling large data. For accurate results, these constraints can be overcome with dataset fine-tuning. Despite being flexible, it is not as strong as other LLM models since it relies on external predictive tools for risk analysis and computing efficiency [15].

GPT-J is an improved version of GPT-Neo with additional natural language processing capabilities. It has the potential to process large datasets compared to GPT-Neo in addition to reports generation. It can interpret specific industrial terminologies relevant to the oil and gas applications with additional fine tuning. However, for advanced predictive modelling, it still requires integration with external risk assessment tools [16].

Bloom can be well for AIPS management practices in global operations and cross-border compliance because of its ability to generate reports in several languages. This makes it ideal for multinational corporations managing asset integrity across different geographical locations. However, without additional analytical tools, it suffers limitations in providing predictive insights.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2025 doi:10.20944/preprints202503.1172.v1

Regardless of these constraints, the tool has linguistic versatility makes it advantageous over other applications in multilingual industrial environments [17].

LLaMA excels in the real-time processing of structured and semi-structured data with lower computational and resource demands. Like most, it also required fine tuning to improve its precision to identify specific terminologies. Moreover, it can be integrated with external systems for predictive risk analysis. This can facilitate the identification of potential risks by evaluation of historical data. Because of its computational efficiency and lightweight design, it is appealing for real-time reporting and risk assessment with fewer computational demands [18].

ChatGPT-DataAnalyst performs better than the others. It offers precise information, is quicker, and is easy to use. It produces accurate reports with little fine-tuning, handling both structured and unstructured data. It is a top option for AIPS management due to its scalability and stated benefits [19].

Table 2. Summary of considered open-source LL-GPT Models.

LLM GPT Model	AIPS KPI Reporting	Predictive Risk Insights	Strengths	Limitations
GPT-Neo	Basic KPI reporting requires extensive fine-tuning for domain-specific applications.	Limited capabilities require external tools for predictive modelling.	Accessible, flexible.	Limited contextual understanding and scalability.
GPT-J	Generates detailed, context-aware reports; handles terminology with fine-tuning.	Better pattern recognition; still relies on external predictive frameworks.	Improved contextual capabilities.	Lacks built-in advanced predictive features.
Bloom	Multilingual reporting for diverse geographies; effective for large-scale operations.	Limited predictive capabilities; relies on supplementary tools.	Multilingual, contextually rich.	High computational requirements for real-time use.
LLaMA	Efficient real-time structured data reporting; highly adaptable to fine-tuning.	Supports integration with external tools for predictive insights.	Lightweight, efficient for real-time tasks.	Predictive capabilities require supplementary tools.
ChatGPT-DataAnalyst	High-speed, precise reporting for structured and unstructured data.	Strong pattern recognition for proactive risk assessment.	Scalable, minimal fine tuning needed.	Less established for specific industrial terminologies.

1.3. Literature Review

The reviewed articles identified five common overlapping themes on AI-driven applications AIPS management as outlined below.

Proactive Risk Management: Includes real-time system monitoring and predictive maintenance to identify irregularities early.

Digitalization of Safety Data: Analysis of extensive data on safety incidents to aid decision making.

Optimization of Alarm Management: This is where alarms are optimised by lowering the number of false positives and nuisances so that operators may concentrate on the import ones.

Integration with Human Factors: Using AI-powered simulations in training programs to improve human experiences.

Resilience Modelling: Advancing process resilience by foreseeing possible interruptions and implementing preventative measures in advance.

 doi:10.20944/preprints202503.1172.v1

Ref. [20] relied on bow tie analysis to visualise failure scenarios, which enabled optimised risk-based inspections. This made it easier to switch from reactive to predictive maintenance strategies and provided a structure to identify possible weaknesses in dangerous operations. In contrast, Ref. [21] questioned the use of operators to control the functionality of critical equipment and systems (CES), emphasising the need to set realistic expectations for human capabilities. Despite Ref. [20] focusing on technical systems and Ref. [21] on human factors, both underscore the importance of proactive strategies to address system vulnerabilities.

Table 3. Main components of a Bow-Tie model [20].

Bow-Tie Component	Description
Hazard and Top Event	A hazard is the potential source of risk (e.g., loss of containment of hydrocarbons). The top event is the incident that results from loss of control over the hazard (e.g., an oil spill or gas explosion).
Threats and Preventive Barriers	Threats are causes or initiating events that lead to the top event. These include equipment failures and human errors. Preventive barriers are safeguards to reduce the likelihood of these threats. These include engineering controls such as pressure relief systems and operational procedures.
Consequences and Mitigative Barriers	Consequences are potential negative outcomes, such as environmental damage, financial losses and human casualties. Mitigative barriers are measures that reduce the severity of these consequences. These include emergency response plans and fire suppression systems.

The works of [22] and Ref. [23] investigated the reliability and sustainability frameworks in industrial operations. Ref. [22] proposed a hierarchical model to integrate both leading and lagging indicators to prioritize risks to make risk-based decisions. In contrast, [23] adopted a broader perspective that aligned risks to their effects, regardless of the magnitude of their impacts. While [22] emphasize the quantification of risks through performance indicators, [23] extend their focus to encompass the broader impacts of asset management on sustainability. Both [22,23] provide alternative methods for managing AIPS and provide insightful information that may be incorporated with AI to create a well-rounded strategy.

The work of Ref. [24] investigated the function of preventive maintenance (PM) in a wastewater treatment facility, focusing on the ways that proactive maintenance lowers environmental hazards and minimises interruptions. Conversely, Ref. [25] proposed an integrated maintenance system that incorporates a several techniques such as reliability-centred maintenance (RCM) and risk-based inspections (RBI). Ref. [24] focused on environmental challenges in resource-intensive systems, while Ref. [25] addressed complex industrial maintenance needs. The significance of PM techniques in lowering operational risks for sustainability and tailored AI-driven maintenance strategies are emphasised by both studies.

Refs. [26,27] concur that using creative approaches may improve adherence to safety regulations. A simpler Markov-based method that lowers the complexity of safety modelling was introduced in Ref. [26], which concentrated on Safety Integrity Level (SIL) verification. Deficits in AIPS information handling that may impact research like SIL were discussed in Ref. [27]. To effectively manage risks, the study emphasised the need for customised safety data that corresponds with real-world process situations. LL-GPT capabilities can leverage their shared focus to simplify complex procedures and tailor strategies to real-world scenarios.

Studies by [28,29] have the common objectives of minimising industrial incidents through proactive measures emphasising the significance of integrating safety measures early in a project

 doi:10.20944/preprints202503.1172.v1

lifecycle. Ref. [28] proposed an evidence-based data driven incident prevention framework to bridge the gap between research and industrial practice, while Ref. [29] proposed Inherent System Safety Index (ISSI) to evaluate safety at the design stages. Both [28,29] had creative ideas, but they failed to emphasise how AI could be incorporated into their work.

Refs. [30,31] share a focus on enhancing hazard identification and risk prioritisation through structured and systematic methodologies. Ref. [30] combined Fuzzy Multi-Criteria Decision-Making (FTOPSIS) with HAZOP for risk evaluation and supporting preventive strategies. On the other hand, [31] used Layers of Protection Analysis (LOPA) for SIL determination for cumulative risk assessment. Both [30,31] support proactive risk management and address the limitations of traditional risk assessment methods. While [30] focused on a specific case of a biogas process unit, [31] gave a global approach that harness the capabilities of AI across diverse industries.

The researches Refs. [32,33] looked into safety improvements through advanced monitoring. Both studies tackle vulnerabilities that could otherwise lead to catastrophic failures if not managed well. Ref. [32] addressed alarm flooding that hindered operator response and proposed a two-level Intelligent Alarm Management Framework (IAMF) to filter redundant alarms and diagnoses their causes. On the other hand, [33] investigated cyber security risks on process control systems, proposing the identification of Basic Process Control Systems (BPCS) and Safety Instrumented Systems (SIS) vulnerabilities. Both studies accentuate on the importance of root-cause identification for preventing cascading failures and maintaining operational safety. Together, these works highlight the need for robust strategies to address evolving risks in industrial systems, from alarm management to cyber security.

Ref. [34] underlined the transformative potential of AI and Industry 4.0 technologies, such as digital twins to advance AIPS management. They showcased how these innovations can manage safety critical equipment (SCE) and support the development of their performance standards. However, the authors do not provide a concrete case study to validate their recommendations implementation in the oil and gas industry.

Traditionally, risk evaluation has relied on historical data and expert judgment, an approach that falls short in today's fast-paced environments. AI offers a solution by rapidly analysing large datasets, enabling accurate and timely risk assessments. As [35] demonstrated, AI technologies can identify potential risks by analysing data that human analysts might overlook. For instance, machine learning algorithms can process sensor data in real-time, calculate the likelihood of incidents, and support proactive preventive measures. While the study offers valuable theoretical perspectives, it lacks practical case studies to illustrate the real-world benefits of the proposed approaches, especially in the oil and gas industry.

Refs. [36,37] discussed leveraging AI to predict equipment failures based on usage patterns, maintenance history, and environmental to identify weaknesses in existing safety practices and propose effective solutions. Ref. [38,39] outlined leveraging AI to enrich employee learning experiences, for example, utilizing virtual reality platforms for realistic on the job scenarios that employees might encounter. Although the authors present valuable theoretical insights, they do not support their proposals with empirical case studies to demonstrate the mentioned capabilities with specific relevance to the industry.

The foreseen challenges for implementing AI in AIPS management include organizational resistance, as some employees may fear job displacement. To counter employees' fears, AI tools should be designed as complementary tools to enhance human performance rather than to replace them. Ref. [40] suggests that integrating AI while enhancing employees' professional skills leads to more efficient work practices leading to workforce acceptance. Privacy and data security also pose significant risks, as large data and information are analysed, some of which may be sensitive and confidential. As emphasized by Ref. [41], establishing robust data governance frameworks protect sensitive information, without which organizations become vulnerable to data breaches.

By analysing extensive datasets from various operational aspects, companies can uncover actionable insights that foster innovation and efficiency. This enables the optimization of resources aligning with the broader sustainability goals of the sector. Emerging technologies such as Blockchain and distributed detection systems are beginning to play a role in improving the integrity and traceability of data. These technologies promise to further improve operational efficiencies by guaranteeing data reliability and facilitating real -time operations monitoring. As these systems continue to advance, their roles in shaping and improving decision-making processes becomes increasingly vital [42–44].

2. Materials and Methods

Two similar sites were analysed comparatively over a period of 1 year between January 2024 to January 2025, to assess their performance using a set of the set of AIPS KPIs in Table 4. These indicators provided valuable insights on operational risks, AIPS, and reliability. The KPIs for Management of Change (MoC) and the maintenance of SCE were consolidated into a single averaged KPI, as both showed similar trends.

Table 4. AIPS Key performance indicators.

KPI Code	KPI Description	Notes
#LOPC (KPI1)	Number of Loss of Primary Containments	An increasing number indicates poor plant safety and integrity performance
#PRDActiv (KPI2)	Number of Pressure Relieve Device (PRD) Activations	An increasing number indicates operational challenges and system challenges requiring immediate attention. PRDs, including safety valves, are designed to release excess pressure when they exceed safe operating limits, acting as a safeguard to prevent equipment failure.
# SISAct(KPI3)	Number of Safety Instrumented Systems Activations	SIS activations reflect the effectiveness of safety mechanisms in responding to unsafe conditions. Frequent activations signal operational risks and inefficiencies, while low or decreasing activations indicate stable and well-managed processes.
# Ttrip (KPI4)	Number of Trips leading to Shutdown of Equipment, Unit or Whole plant	Frequent equipment trips indicate process inefficiencies, or inadequate maintenance, while a reduction in trips reflects improved system stability and operational control. This KPI provides critical insights into the health of the system, helping to minimize downtime, prevent damage, and maintain consistent production levels.
# PCI (KPI5)	Number of Primary Containment Inspections and Testing Results Outside Acceptable Limits	This applies specifically to vessels, reactors and heat exchangers, where an increasing number reflects declining equipment integrity, potentially leading to catastrophic failures.
#SOLExc (KPI6)	Number of Safe Operating Limits (SOL) excursions	A higher frequency of SOL excursions suggests that the process is consistently operating at or beyond its established safety boundaries, often indicating underlying issues such as inadequate process control, subpar equipment performance, or deficient maintenance practices.
#FALSECAlam (KPI7)	Number of False Critical Alarms Activations	A rising frequency of false critical alarm activations suggests potential shortcomings in alarm management and process control systems. This may be due to alarm thresholds that are set too loosely or too tightly, inaccurate sensor calibrations, or malfunctioning control loops, all of which contribute to misleading alarm signals that do not accurately reflect the system's true status.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2025 doi:10.20944/preprints202503.1172.v1

%OverdueSCEMain (KPI8)	Percentage of Safety Critical Equipment Overdue for Maintenance	A growing backlog of overdue maintenance on safety-critical equipment clearly signals systemic deficiencies in asset management and maintenance scheduling. This trend not only points to potential resource limitations and misaligned organizational priorities but also undermines the overall reliability and integrity of essential safety systems.
%MoCComp (KPI9)	Percentage of Management of Change Non-Compliance	A higher frequency of Change non-compliance events signals a failure in the organization's change control process modifications to processes, equipment, or procedures are not being adequately reviewed, documented, or approved.

A structured analysis to assess the capabilities of several open-source LLM GPT models was conducted to identify the most suitable for automating AIPS KPIs reporting and to generate predictive insights. ChatGPT-DataAnalyst was selected over GPT-Neo, GPT-J, Bloom, and LLaMA because it offers better fine-tuning options and has a higher scalability in generating precise insights [45].

2.1. KPI Data Sourcing

The data used in this study was obtained from four primary databases outlined in Table 1.

(1) *General Incident Logs:* This includes data from the Distributed Control System (DCS), Management of Change (MOC) logs, Root Cause Analysis (RCA) logs, overrides of process safeguarding systems, and downgraded situations, all of which are recorded in a standardized Excel-based spreadsheet for efficient tracking and analysis. The incidents are categorized into specific types, including near-misses, emergency incidents, equipment failures and process deviations.

(2) *SAP Asset and Maintenance Management*: This is a tool for optimizing the lifecycle performance and reliability of physical assets across industries. It integrates tools for planning, monitoring, and executing maintenance tasks. The solution has powerful features like managing work orders, keeping track of assets, and planning preventative maintenance. It also helps with regulatory compliance by keeping detailed records of asset activities and speeding up safety processes.

(3) *SAP Incident Management Module:* This is part of the SAP S/4HANA suite for managing safety, health, environmental, and operational performance incidents. The module provides a functionality for incident categorization and an actions tracker.

(4) *Meridiam Incident Database:* Meridian software offers solutions designed to optimize engineering document control, incident management, regulatory compliance, and asset lifecycle management. It centralizes and secures critical equipment performance data while ensuring seamless access, making it an essential tool for industries that rely on complex engineering workflows.

2.2. Data Preparation

Insights from [46] underscore that meticulous attention to data preparation is beneficial for robust analysis and for generating reliable, actionable insights. Given the complexity and volumes of the data, the steps below were taken to clean, standardize, and transform it data from the various sources into a format suitable for analysis. This process is presented by Figure 1.

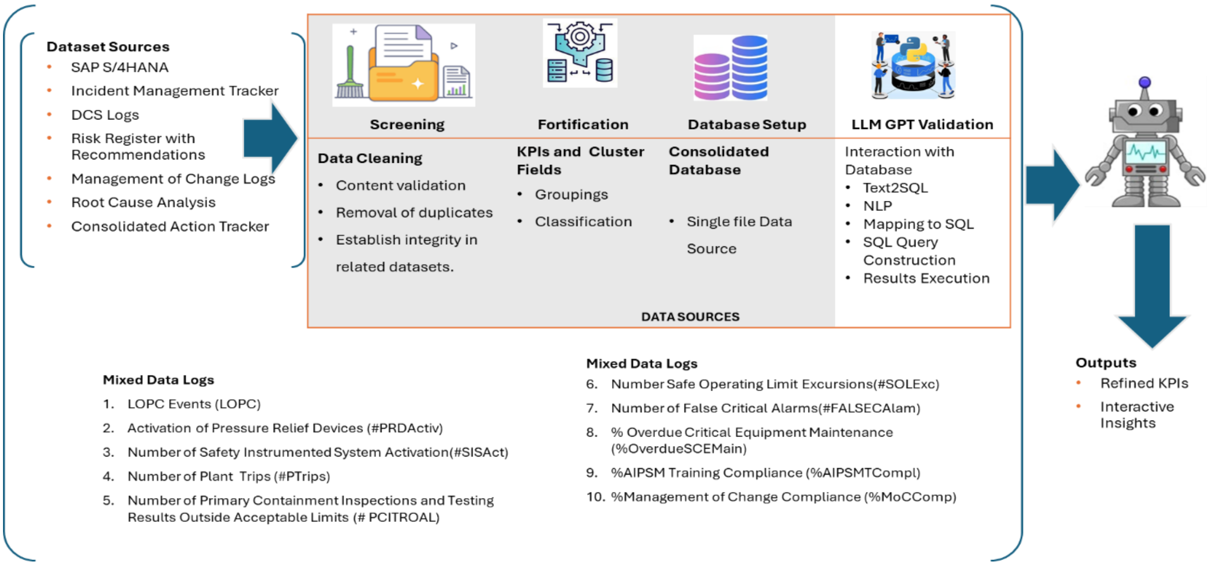

Figure 1. Data preparation workflow for LLM analyses.

2.3. Data Refinement:

This entailed removing undesirable or inappropriate indicators, correcting spacing problems, and making inputs uniform. Records with missing or ambiguous values were synchronised. Additionally, any anomalies such as duplicate entries or incorrect classifications, are identified and removed to maintain the integrity of the dataset.

Once the data was cleaned, it was transformed into a structured format suitable for LLM model processing. This required reformatting the content of within incident descriptions, performance notes, and safety records so that they could be incorporated into the natural language processing features. For example, incident descriptions that were tokenized, and key terms related to AIPS (such as "barrier failure", "hazard", "loss of primary containment" etc.) were tagged for further analysis. This step ruled out the collection of irrelevant information and arranged it to make it easier to categorize and analyse [47].

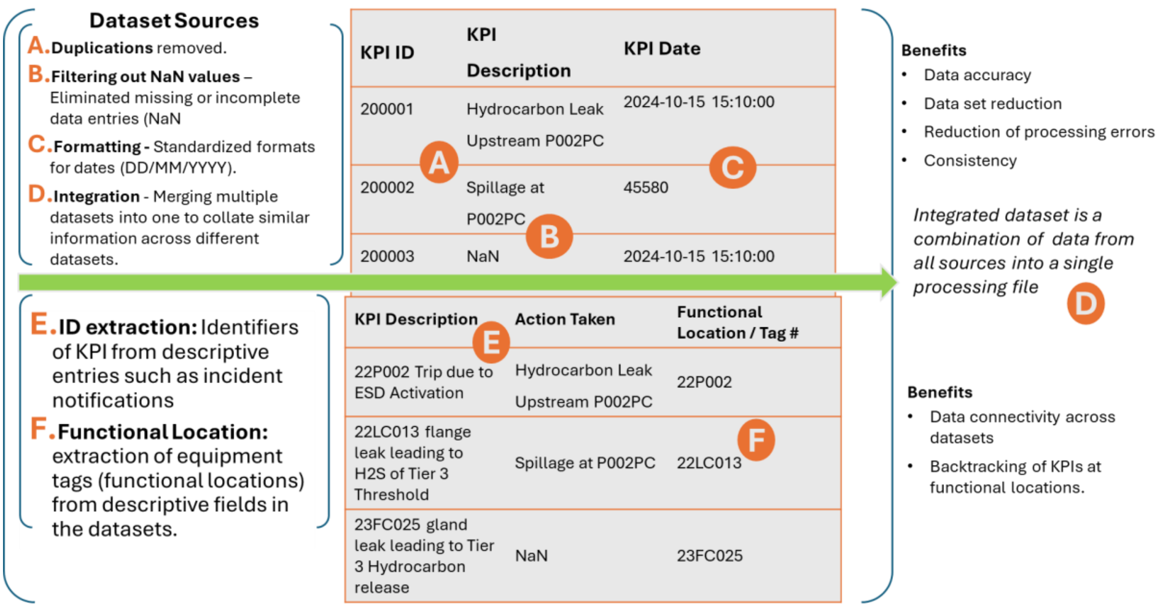

Figure 2. Data refinement workflow for accurate LLM outputs.

 doi:10.20944/preprints202503.1172.v1

2.4. LLM GPT Application

The model was utilized to analysed and categorize AIPS data and KPIs, as outlined in Ref. [48–52]. This was aimed at enhancing the accuracy and clarity of the outputs, which were subsequently employed to generate predictive insights.

Text processing and categorization: The first task involved text preprocessing, whereby the model was prompted to read through the incident reports, safety observations, and operational data. From the textual records, ChatGPT-DataAnalyst interpreted these and recognizes features, including the type of incident, its severity, and anticipated causes. The model categorised the incidents into predefined tiers based given by API 754 standard, which divides incidents into four distinct categories: major incidents, serious incidents, near misses, and performance indicators.

Data enrichment and classification: After processing the data, LLM model was used to enrich the dataset by adding contextual insights. To visualize safety hardware barriers, events containing "loss of primary containment" were related to specific control actions depending on the failed equipment. For example, corrosion monitoring and risk-based inspections for vessels and mechanical seals replacements for large pumps. These outputs are presented in detail as part of the results in the next chapter.

3. Results

The results of this study are three-fold (1) generated predictive insights from the AI model; (2) a demonstration of AI integration in AIPS KPI analysis and (3) a comparative assessment of AIPS management between Site 1 and Site 2.

3.1. Predictive Insights

The generated predictive insights categorised AIPS events, identifying recurring failure patterns for proactive risk mitigation. By analysing leak sources, historical incident trends, LOPC data, the system pinpoints high-risk areas and gave maintenance efforts to prioritise. Ultimately, these insights shift from reactive incident management to proactive prevention to significantly reduce safety risks and operational disruptions.

Table 5. LLM generated mitigation insights with respect to an incident.

Event iD	Metric	Root Causes	LLM Generated Incident Mitigation Insights
200001	Hydrocarbon Leak Upstream P002PC	Pipeline material degradation, corrosion, lack of leak detection	**Process & Risk Assessment** Regular HAZOP/LOPA studies and scenario simulations **Mechanical Integrity** Use corrosion-resistant materials; implement NDT and stress testing **Instrumentation** Install smart leak detection sensors with automatic shutdown capabilities **Operational Practices** Routine inspections and operator training on emergency shutdown procedures
200004	Sulphuric Acid Leak at Tank V321 Main Outlet Valve	Material incompatibility, seal degradation, valve failure	**Process and Risk Assessment** Update risk analysis for corrosive substances **Mechanical Integrity**

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2025 doi:10.20944/preprints202503.1172.v1

			Upgrade valve materials and seals; perform frequent integrity tests **Instrumentation** Implement remote monitoring with early leak detection alarms **Operational Practices** Regular maintenance and pre-shift checklists for critical valves
200005	Large Steam Leak from Boiler 322 Main Outlet Manifold	High pressure/temperature stresses, joint failures	**Process and Risk Assessment** Include thermal stress scenarios in digital twins **Mechanical Integrity** Conduct regular pressure and temperature tests; ensure high-standard welds and bolted joints **Instrumentation** Use thermal imaging and smart sensors to detect anomalies **Operational Practices** Scheduled maintenance and operator drills for high-pressure systems
200006	H2S-Rich Lean Amine Leak from Pump 2-G-0303C Double Isolation Manifold	H2S-Rich Lean Amine Leak from Pump 2-G-0303C Double Isolation Manifold	**Process and Risk Assessment** Reassess double isolation design with risk ranking for H2S releases **Mechanical Integrity** Frequent testing of seals/gaskets and enhanced manifold design **Instrumentation** Install additional H2S sensors and alarms near critical points **Operational Practices** Strict maintenance schedules and emergency response training for H2S incidents
200007	Loss of High-Level Control at 554V101 Leading to Liquid Overflow	Control system failure, sensor malfunction, absence of redundancy	**Process and Risk Assessment** Incorporate control system failure modes in hazard analysis **Mechanical and Instrumentation Controls** Retrofit with dual redundant level sensors; implement automatic shutdown/diversion systems **Operational Practices** Regular calibration checks and pre-emptive inspections; install secondary overflow containment systems

3.2. AI assisted AIPS KPI Analysis

The categorization of AIPS KPIs from about 15,000 entries took five working days to complete manually. In contrast, the same task with validation took less than 50 - 70 minutes using LLM (GPTDataAnalys). The substantial time savings present a compelling justification for the adoption of

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2025 doi:10.20944/preprints202503.1172.v1

AI-driven data analysis to reduce human effort, accelerates decision-making, and to optimizes resource allocation.

The data validation presented in Table 6 involved a comparison to evaluate the accuracy of LLM-generated insights against manually verified data. DataAnalyst correctly categorized 19458 out of 21046 AIPS KPI, incorrectly categorising 1588 entries, yielding a confidence of 92%: the ratio between the two. This demonstrates a high level of agreement between AI-generated insights and human evaluations, with minor discrepancies.

Table 6. Data validation results.

Insights	#LOPC(KPI1)	#PRDActiv (KPI2)	# SISAct(KPI3)	# Ttrip (KPI4)	(#PCIROAL(KPI5)	#SOLExc (KPI6)	#FALSECAlam(KPI 7)	Total Count	Analysis Time
GPTDataAnalyst	3576	3368	4869	1846	2983	1217	1600	19458	50-70 minutes
Validated Results	3934	3714	5204	1982	3179	1289	1744	21046	4-5 days
Wrongly Classified	358	346	335	136	197	72	144	1588	-
% Error	9%	9%	6%	7%	6%	6%	8%	8%	-

The validated data generated by the LLM achieved 92% accuracy. Despite minor discrepancies, the significant time savings and consistency demonstrates the value of adopting AI for AIPS management.

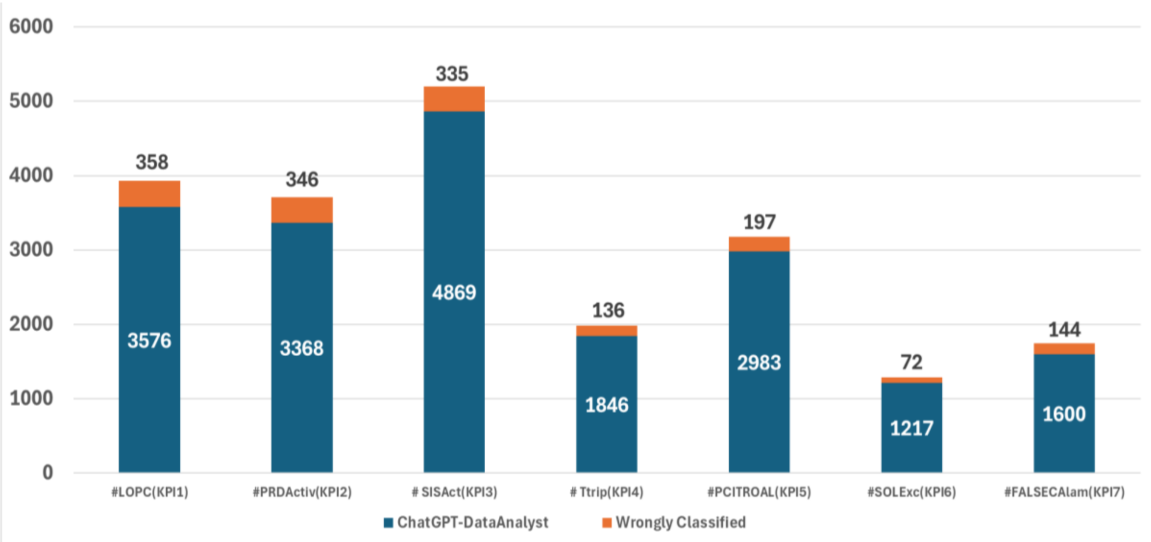

Figure 3. Distribution of GPTDataAnalyst classified data.

3.3. Comparative Assessment Between Site 1 and Site 2

Figures 4 and 5 illustrate the results from Site 1 and Site 2, respectively. In June 2024, Site 1 implemented the AI-driven AIPS management outlined in this study, whereas Site 2 continued relying on traditional, manual spreadsheet-based methods. The adoption of AI for AIPS management at Site 1 marked the onset of significant improvements in safety outcomes and enhanced KPIs. Prior June 2024, Site 1demonstrated poor KPI trends like those of Site 2, which reflected safety and

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2025 doi:10.20944/preprints202503.1172.v1

operational inefficiencies. Leveraging AI-generated insights to proactively apply incident mitigation measures at Site 1 led to consistent, month-to-month improvements across all KPIs.

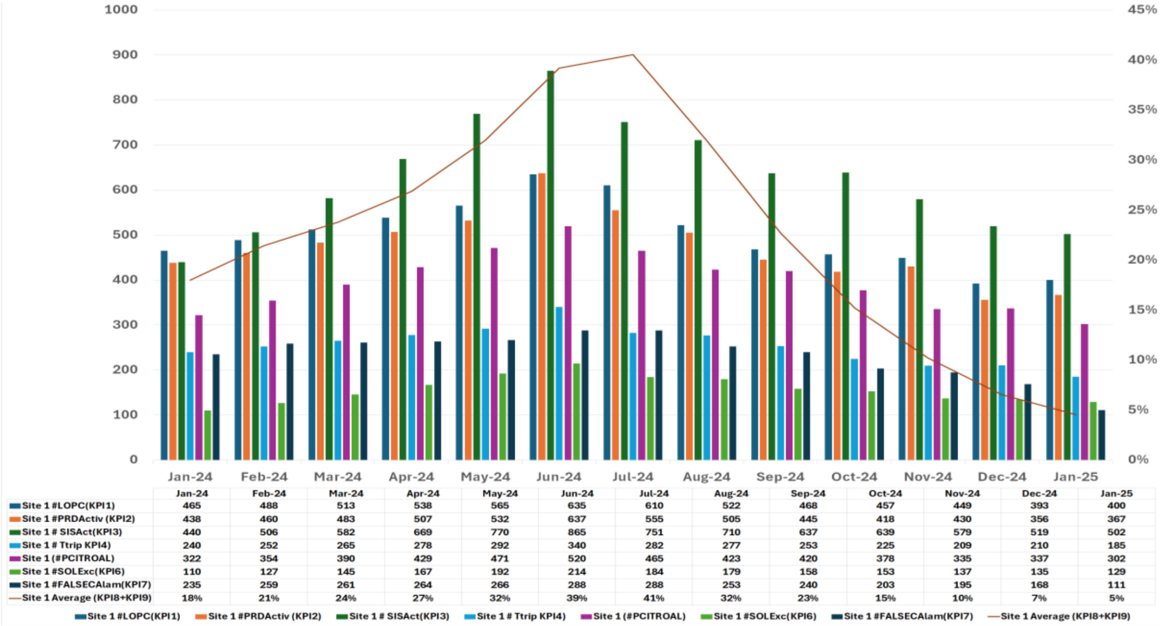

Figure 4. AIPS KPI data at Site 1 with notable improved performance post June 2024.

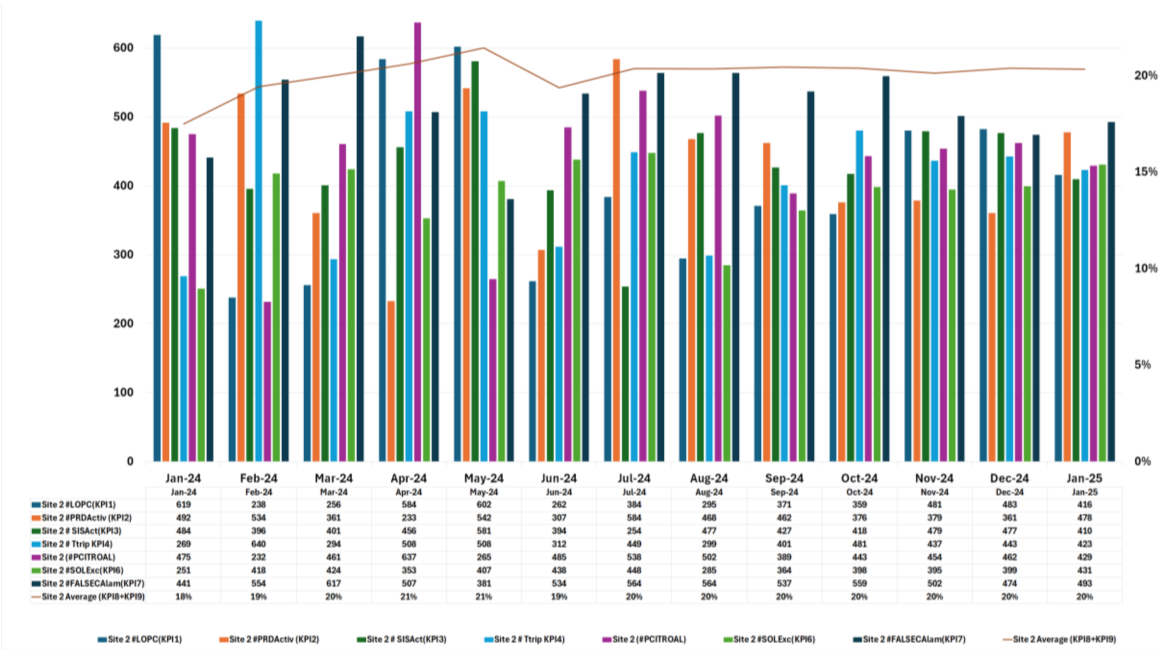

Figure 5. AIPS KPI data at Site 2. Note the stagnant performance.

4. Discussion

This work introduces a novel approach that customises a LLM model to improve AIPS performance for sustainable onshore petrochemical installations. Two liquefied natural gas plants have been used to comparatively access the efficacy of the model. The model can accurately and consistently classify AIPS incidents from several large databases reducing the reliance on subjective manual interpretations reducing the time it takes to manually process data. By scaling down on manual interventions, human errors are greatly minimised.

 doi:10.20944/preprints202503.1172.v1

This is an additional tool for text-based intelligence to compliment sensor based monitoring and predictive analytics to improve safety performance. This is a significant bridge that closes the gap between structured safety data from digital sensor logs and unstructured data from manual sources such as incident reports and shift logs.

Furthermore, the study demonstrates how to fine GPT to assist in decision making processes in AIPS management by providing safety recommendations for proactive safety management. Normally, safety assessments depend expert judgments, which are prone to human error in addition to them being intensive. In summary, these are the key pointers from this work:

i. *Automated incident categorization:* analysis of large volumes of shift logs, incident reports, and sensor alerts to classify AIPS incidents into predefined categories (e.g., equipment failure, human error, hazardous material release).

ii. *AI-assisted decision-making for AIPS: interpretation of historical safety reports, to generate contextual safety recommendations.*

iii. *Knowledge-driven AIPS framework: Synthesising lessons learned from previous incidents, providing* actionable insights for continuous AIPS improvement and risk reduction.

5. Conclusions

This research emphasises for the adoption of AI as best practice in AIPS management highlighting its transformative impact on operational risk mitigation, and sustainability. The findings strongly reinforce that if effectively implemented, AI technologies will drive efficiency in AIPS management and contribute to long-term sustainability through data-driven safety insights.

The study emphasises more collaboration between the academia and the industry in AI research and application to drive innovation for sustainable development with the academia contributing to theoretical insights and the industry contributing practical expertise, resources, and real-world applications. Such collaboration will bridge the gap between theoretical advancements and their practical implementation. The work provides a foundation for such collaboration by exploring AI's role in sustainable safety management in the industry.

Author Contributions: Conceptualization, B.O.; methodology, B.O.; software, B.O.; validation, B.O. and B.O.; formal analysis, B.O.; investigation, B.O.; resources, B.O.; data curation, B.O.; writing—original draft preparation, B.O.; writing—review and editing, B.O. and M.A.; visualisation.; supervision, B.O. The author has read and agreed to the published version of the manuscript.

Funding: This research received no external funding from any source.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analysed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. ADIPEC. ADIPEC 2024: ADNOC EnergyAI (Nov, 2024). Available online: https://www.adipec.com/energy-ai/ (accessed on 28 November 2024).
2. Kulkov, I., Kulkova, J., Rohrbeck, R., Menvielle, L., Kaartemo, V. and Makkonen, H., 2024. Artificial intelligence-driven sustainable development: Examining organizational, technical, and processing approaches to achieving global goals. *Sustainable Development*, *32*(3), pp.2253-2267.
3. Oke, A. and Fernandes, F.A.P., 2020. Innovations in teaching and learning: Exploring the perceptions of the education sector on the 4th industrial revolution (4IR). *Journal of Open Innovation: Technology, Market, and*

Complexity, *6*(2), p.31.Darwin, Rusdin, D., Mukminatien, N., Suryati, N., Laksmi, E.D. and Marzuki, 2024. Critical thinking in the AI era: An exploration of EFL students' perceptions, benefits, and limitations. *Cogent Education*, *11*(1), p.2290342.

5. Rampersad, G., 2020. Robot will take your job: Innovation for an era of artificial intelligence. *Journal of Business Research*, *116*, pp.68-74.
6. Rane, N., 2023. ChatGPT and Similar Generative Artificial Intelligence (AI) for Smart Industry: role, challenges and opportunities for industry 4.0, industry 5.0 and society 5.0. *Challenges and Opportunities for Industry*, *4*.
7. Ahmadi, A., 2024. Unravelling the Mysteries of Hallucination in Large Language Models: Strategies for Precision in Artificial Intelligence Language Generation. *Asian Journal of Computer Science and Technology*, *13*(1), pp.1-10.
8. Sharma, R., 2023. The transformative power of AI as future GPTs in propelling society into a new era of advancement. *IEEE Engineering Management Review*.
9. Sternad, S. and Bobek, S. (2013) 'Enterprise Resource Planning Acceptance Model (ERPAM)', *Enterprise Resource Planning*, pp. 407–432. doi:10.4018/978-1-4666-4153-2.ch023.
10. Liu, Y. et al. (2018) 'Brief valuation of systems, Applications & Products (SAP) corporation', *Global journal of Economics and Business Administration* [Preprint]. doi:10.28933/gjeba-2018-03-0801.
11. Schuchardt, K. and Schymik, R. (2000a) 'Laboratory Information and Management Systems (LIMS)', *Ullmann's Encyclopedia of Industrial Chemistry*. doi:10.1002/14356007.b06_307.
12. Ikegwu, A.C., Nweke, H.F., Anikwe, C.V., Alo, U.R. and Okonkwo, O.R., 2022. Big data analytics for data-driven industry: a review of data sources, tools, challenges, solutions, and research directions. *Cluster Computing, 25(5),* pp.3343-3387.
13. Ahmad, R.W., Salah, K., Jayaraman, R., Yaqoob, I. and Omar, M., 2022. Blockchain in oil and gas industry: Applications, challenges, and future trends. Technology in society, 68, p.101941.
14. Aderamo, A.T., Olisakwe, H.C., Adebayo, Y.A. and Esiri, A.E., 2024. AI-driven HSE management systems for risk mitigation in the oil and gas industry. *Comprehensive Research and Reviews in Engineering and Technology*, *2*(1), pp.1-22.
15. Sufi, F.K., 2024. A Systematic Review on the Dimensions of Open-Source Disaster Intelligence using GPT. *Journal of Economy and Technology*.
16. Sufi, F., 2024. An innovative GPT-based open-source intelligence using historical cyber incident reports. *Natural Language Processing Journal*, *7*, p.100074.
17. Seger, E., Dreksler, N., Moulange, R., Dardaman, E., Schuett, J., Wei, K., Winter, C., Arnold, M., Ó hÉigeartaigh, S., Korinek, A. and Anderljung, M., 2023. Open-sourcing highly capable foundation models. *Research paper, Centre for the Governance of AI*.
18. Katlariwala, M. and Gupta, A. (2024) 'Product recommendation system using large language model: Llama-2', *2024 IEEE World AI IoT Congress (AIIoT)*, pp. 0491–0495. doi:10.1109/aiiot61789.2024.10579009.
19. *How to use CHATGPT's Advanced Data Analysis Feature* (2024) *MIT Sloan Teaching & Learning Technologies*. Available at: https://mitsloanedtech.mit.edu/ai/tools/data-analysis/how-to-use-chatgpts-advanced-data-analysis-feature/ (Accessed: 14 February 2025).
20. Schmitz, Peter, Swuste, Paul, Reniers, Genserik, & Nunen, Karolien van 2020, 'Mechanical integrity of process installations: Barrier alarm management based on bowties', *Process Safety and Environmental Protection*, vol. 138, pp. 139-147, ISSN 0957-5820, Elsevier BV, https://doi.org/10.1016/j.psep.2020.03.009
21. McLeod, Ronald W. 2017, 'Human factors in barrier management: Hard truths and challenges', *Process Safety and Environmental Protection*, vol. 110, pp. 31-42, ISSN 0957-5820, Elsevier BV, https://doi.org/10.1016/j.psep.2017.01.012
22. Hassan, Jakiul, & Khan, Faisal 2012, 'Risk-based asset integrity indicators', *Journal of Loss Prevention in the Process Industries*, vol. 25, no. 3, pp. 544-554, ISSN 0950-4230, Elsevier BV, https://doi.org/10.1016/j.jlp.2011.12.011
23. Ratnayake, R. M. Chandima, & Markeset, T. 2012, 'Asset integrity management for sustainable industrial operations: measuring the performance', *International Journal of Sustainable Engineering*, vol. 5, no. 2, pp. 145-158, ISSN 1939-7038, Informa UK Limited, https://doi.org/10.1080/19397038.2011.581391

 doi:10.20944/preprints202503.1172.v1

24. Hernández-Chover, Vicent, Castellet-Viciano, Lledó, & Hernández-Sancho, Francesc 2020, 'Preventive maintenance versus cost of repairs in asset management: An efficiency analysis in wastewater treatment plants', *Process Safety and Environmental Protection*, vol. 141, pp. 215-221, ISSN 0957-5820, Elsevier BV, https://doi.org/10.1016/j.psep.2020.04.035
25. Wang, Qingfeng, Wenbin, Liu, Xin, Zhong, Jianfeng, Yang, & Qingbin, Yuan 2011, 'Development and application of equipment maintenance and safety integrity management system', *Journal of Loss Prevention in the Process Industries*, vol. 24, no. 4, pp. 321-332, ISSN 0950-4230, Elsevier BV, https://doi.org/10.1016/j.jlp.2011.01.008
26. Shu, Yidan, & Zhao, Jinsong 2014, 'A simplified Markov-based approach for safety integrity level verification', *Journal of Loss Prevention in the Process Industries*, vol. 29, pp. 262-266, ISSN 0950-4230, Elsevier BV, https://doi.org/10.1016/j.jlp.2014.03.013
27. Liaw, Horng-Jang 2019, 'Deficiencies frequently encountered in the management of process safety information', *Process Safety and Environmental Protection*, vol. 132, pp. 226-230, ISSN 0957-5820, Elsevier BV, https://doi.org/10.1016/j.psep.2019.10.015
28. Wang, Sun, Hao, Yang, Ming, & Haiqing 2022, 'Resilience-based approach to maintenance asset and operational cost planning', *Process Safety and Environmental Protection*, vol. 162, pp. 987-997, ISSN 0957-5820, Elsevier BV, https://doi.org/10.1016/j.psep.2022.05.002
29. Sultana, S. and Haugen, S. (2022) 'Development of an inherent system safety index (ISSI) for ranking of chemical processes at the concept development stage', *Journal of Hazardous Materials*, 421, p. 126590. doi:10.1016/j.jhazmat.2021.126590.
30. Severi, C.A. et al. (2022) 'Identification of critical operational hazards in a biogas upgrading pilot plant through a multi-criteria decision-making and Ftopsis-Hazop approach', *Chemosphere*, 307, p. 135845. doi:10.1016/j.chemosphere.2022.135845.
31. Baybutt, Paul 2012, 'Using risk tolerance criteria to determine safety integrity levels for safety instrumented functions', *Journal of Loss Prevention in the Process Industries*, vol. 25, no. 6, pp. 1000-1009, ISSN 0950-4230, Elsevier BV, https://doi.org/10.1016/j.jlp.2012.05.016
32. Hu, Jinqiu, & Yi, Yan 2016, 'A two-level intelligent alarm management framework for process safety', *Safety Science*, vol. 82, pp. 432-444, ISSN 0925-7535, Elsevier BV, https://doi.org/10.1016/j.ssci.2015.10.005
33. Iaiani, Matteo, Tugnoli, Alessandro, Macini, Paolo, & Cozzani, Valerio 2021, 'Outage and asset damage triggered by malicious manipulation of the control system in process plants', *Reliability Engineering & System Safety*, vol. 213, p. 107685, ISSN 0951-8320, Elsevier BV, https://doi.org/10.1016/j.ress.2021.107685
34. Ayomoh, M. and Ongwae, B., 2025. A Systematic Review of Asset Integrity and Process Safety Management Sustainability for Onshore Petrochemical Installations. *Sustainability*, *17*(1), p.286.
35. Moktadir, M.A., Ali, S.M., Kusi-Sarpong, S. and Shaikh, M.A.A., 2018. Assessing challenges for implementing Industry 4.0: Implications for process safety and environmental protection. Process safety and environmental protection, 117, pp.730-741.
36. Khan, Faisal, Rathnayaka, Samith, & Ahmed, Salim 2015, 'Methods and models in process safety and risk management: Past, present and future', *Process Safety and Environmental Protection*, vol. 98, pp. 116-147, ISSN 0957-5820, Elsevier BV, https://doi.org/10.1016/j.psep.2015.07.005
37. Pan, Y. and Zhang, L., 2021. Roles of artificial intelligence in construction engineering and management: A critical review and future trends. Automation in Construction, 122, p.103517.
38. Singh, S., Sharma, P.K., Yoon, B., Shojafar, M., Cho, G.H. and Ra, I.H., 2020. Convergence of blockchain and artificial intelligence in IoT network for the sustainable smart city. Sustainable cities and society, 63, p.102364.
39. Wong, L.W., Tan, G.W.H., Ooi, K.B., Lin, B. and Dwivedi, Y.K., 2024. Artificial intelligence-driven risk management for enhancing supply chain agility: A deep-learning-based dual-stage PLS-SEM-ANN analysis. International Journal of Production Research, 62(15), pp.5535-5555.
40. Manchana, R., 2022. Optimizing Real Estate Project Management through Machine Learning, Deep Learning, and AI. Journal of Scientific and Engineering Research, 9(4), pp.192-208.

41. Wamba, S.F. and Queiroz, M.M., 2020. Blockchain in the operations and supply chain management: Benefits, challenges and future research opportunities. International Journal of Information Management, 52, p.102064.
42. Ogbu, A.D., Ozowe, W. and Ikevuje, A.H., 2024. Data science's pivotal role in enhancing oil recovery methods while minimizing environmental footprints: An insightful review.
43. Ashry, I., Mao, Y., Wang, B., Hveding, F., Bukhamsin, A.Y., Ng, T.K. and Ooi, B.S., 2022. A review of distributed fiber–optic sensing in the oil and gas industry. Journal of Lightwave Technology, 40(5), pp.1407-1431.
44. Tabibian, S.S. and Sharifzadeh, M., 2023. Statistical and analytical investigation of methanol applications, production technologies, value-chain and economy with a special focus on renewable methanol. Renewable and Sustainable Energy Reviews, 179, p.113281.
45. Prandner, D., Wetzelhütter, D. and Hese, S. (2025) 'CHATGPT as a data analyst: An exploratory study on AI-supported Quantitative Data Analysis in empirical research', *Frontiers in Education*, 9. doi:10.3389/feduc.2024.1417900.
46. Luengo, J., García-Gil, D., Ramírez-Gallego, S., García, S., Herrera, F., Luengo, J., García-Gil, D., Ramírez-Gallego, S., García, S. and Herrera, F., 2020. Imbalanced data preprocessing for big data. *Big Data Preprocessing: Enabling Smart Data*, pp.147-160.
47. Fan, C., Chen, M., Wang, X., Wang, J. and Huang, B., 2021. A review on data preprocessing techniques toward efficient and reliable knowledge discovery from building operational data. *Frontiers in energy research*, *9*, p.652801.
48. Maharana, K., Mondal, S. and Nemade, B., 2022. A review: Data pre-processing and data augmentation techniques. *Global Transitions Proceedings*, *3*(1), pp.91-99..
49. Wang, S., Celebi, M.E., Zhang, Y.D., Yu, X., Lu, S., Yao, X., Zhou, Q., Miguel, M.G., Tian, Y., Gorriz, J.M. and Tyukin, I., 2021. Advances in data preprocessing for biomedical data fusion: An overview of the methods, challenges, and prospects. *Information Fusion*, *76*, pp.376-421.
50. Benhar, H., Idri, A. and Fernández-Alemán, J.L., 2020. Data preprocessing for heart disease classification: A systematic literature review. Computer Methods and Programs in Biomedicine, 195, p.105635.
51. Khan, F., Rathnayaka, S. and Ahmed, S., 2015. Methods and models in process safety and risk management: Past, present and future. Process safety and environmental protection, 98, pp.116-147.
52. Abioye, S.O., Oyedele, L.O., Akanbi, L., Ajayi, A., Delgado, J.M.D., Bilal, M., Akinade, O.O. and Ahmed, A., 2021. Artificial intelligence in the construction industry: A review of present status, opportunities and future challenges. Journal of Building Engineering, 44, p.103299.