**Document Compliance Assessment in the Oil and Gas Industry Using NLP, LLM, and Machine Learning**

**Jaber Alhashmi**

Department of Statistics and Business Analytics, College of Business and Economics, United Arab Emirates University, Al Ain, UAE

# Abstract

To improve document compliance evaluation in the Oil and Gas business, this paper presents a qualitative study that incorporates advanced technologies such as Large Language Models (LLMs), Natural Language Processing (NLP), and machine learning. Manual compliance evaluation processes are inefficient and error-prone, so new approaches are needed. The research investigates the potential of such tools for automating the process of extracting, understanding, and categorising information relevant to compliance from project documentation. The study reveals a classification accuracy of over 90% by training an AI model, which is superior to manual assessment techniques. The model's ability to comprehend intricate grammatical structures and subtleties in context improves the reliability of compliance evaluations. The research provides a new understanding of effective natural language processing techniques for extracting data related to regulatory compliance. The study takes a qualitative approach to elucidate how LLMs, NLP, and machine learning might transform the assessment of compliance, paving the way for more streamlined, precise, and efficient procedures in the Oil and Gas sector.

Table of Contents

[Abstract 2](#_Toc149768465)

[*List of Figures* 4](#_Toc149768466)

[*List of Abbreviation* 4](#_Toc149768467)

[Chapter 1: Introduction 5](#_Toc149768468)

[Chapter 2: Literature Review 9](#_Toc149768469)

[2.2 Literature Gap 13](#_Toc149768470)

[Chapter 3: Data Exploration 14](#_Toc149768471)

[Chapter 4: Methodology 15](#_Toc149768472)

[4.1 The Approach 15](#_Toc149768473)

[3.2 Preparation 15](#_Toc149768474)

[3.3 Quality Assessment 16](#_Toc149768475)

[Chapter 5: Modelling and Analysis 16](#_Toc149768476)

[Step 1: Data Collecting 17](#_Toc149768477)

[Step 1:Model Selection and Training 17](#_Toc149768478)

[Step 3: Model Evaluation 18](#_Toc149768479)

[Step 4: Fine-tuning and Optimization 18](#_Toc149768480)

[5.1 Results 18](#_Toc149768481)

[5.2 Discussion 19](#_Toc149768482)

[5.3 Challenges and Limitations 19](#_Toc149768483)

[5.4 Rebalancing Minds and Machines 20](#_Toc149768484)

[Chapter 6: Results and Discussion 20](#_Toc149768485)

[Chapter 7: Conclusion 22](#_Toc149768486)

[References 23](#_Toc149768487)

# List of figures

[Figure 1 15](#_Toc150036442)

[Figure 2 15](#_Toc150036443)

[Figure 3 16](#_Toc150036444)

[Figure 4 19](#_Toc150036445)

[Figure 5 22](#_Toc150036446)

# *List of Abbreviation*

* Named Entity Recognition (NER)
* Natural Language Processing (NLP)
* Large Language Models (LLMs)
* Machine Learning (ML)
* Artificial Intelligence (AI)

# *List of Table*

[**Table 1** 12](#_Toc150036523)

# Chapter 1: Introduction

**1.1 Overview**

Since the beginning of the twenty-first century, a significant amount of study has been carried out on the topic of using AI and machine learning to develop improved compliance mechanisms, and a huge amount of information and insights have been accumulated in academic publications all over the world. According to Morrison, (2022), the construction engineering and oil and gas industries stand to benefit tremendously from the use of these technologies, as they have the potential to both increase operational efficiency and guarantee compliance. Nevertheless, in comparison to other sectors, the oil and gas industry has not yet completely investigated or realised the use and importance of these technologies (Glaese et al., 2022).

Increasing complexity and severe regulatory standards in this industry have made the need of compliance in this sector even more important. As a result, organisations are being forced to adopt sophisticated technology rather than conventional approaches in order to ensure that they comply with all statutory and regulatory requirements. Particularly in areas like competitive intelligence, business intelligence, market intelligence, scientific and technological intelligence, and geo-economics, the number of information and data available to modern businesses is constantly growing. The need for efficient and effective document compliance evaluation has become important, especially in sectors with sophisticated regulatory frameworks like the Oil and Gas industry, as industries continue to adapt to the demands of a rapidly changing world. Here, the meeting point of innovative technologies and qualitative research approaches offers the prospect of significant progress (Trappey et al., 2022). It has always taken a lot of time and manpower to check whether or not a company's reference documents, and project-specific documentation are consistent with one another. Manual review is time-consuming and prone to human mistakes; as a result, it slows down project delivery, prevents effective decision-making, and can even lead to fines for noncompliance. In light of these obstacles, the authors of this study propose a novel framework that makes use of Natural Language Processing (NLP), Large Language Models (LLMs), and machine learning methods to alleviate the problems identified in the previous section (Sweeney, 2020).

The major objective of this project is to conduct an investigation into the potential applications of natural language processing (NLP), language learning models (LLMs), and machine learning within the industry of document compliance review, in particular within the oil and gas industry. By automating the process of extracting, interpreting, and categorising information related to compliance from project documentation, this technique intends to increase the accuracy, efficiency, and effectiveness of compliance assessment. In addition, the purpose of the study is to shed light on the potential practical repercussions of the technique by highlighting its potential to lessen the challenges associated with manual review (Morrison, 2022). In addition to this, the essay delves into the idiosyncrasies that are unique to the oil and gas business, such as the requirement to adhere to safety requirements, environmental standards, and technological standards. The combination of natural language processing (NLP) and language learning machines (LLMs), which can enable the identification of intricate compliance requirements across a large range of documents, can result in a review process that is both more thorough and more efficient.

This work makes a novel contribution by suggesting qualitative project that combines natural language processing, logic-based methods, and machine learning to evaluate documents for compliance in the Oil and Gas sector. This is done in the hopes of laying the groundwork for further study and investigation into the possible advantages of these technologies in improving the speed and precision of compliance assessment procedures. We hope that this study will provide Oil and Gas companies with the knowledge they need to successfully negotiate intricate regulatory environments and propel their projects forward. Named Entity Recognition (NER) is a kind of advanced natural language processing technology that is used to find and categorise named items in text data. These entities might be names of people, corporations, places, dates, quantities of money, or anything else. NER approaches may detect important words or phrases within a text by using machine learning algorithms such as deep learning models and statistical methodologies. By effectively detecting and categorising named items, NER systems play a key role in activities such as information retrieval, question answering, and language comprehension. NER approaches are crucial for a broad variety of applications, from customer service robot deployment to information extraction from academic publications. They ensure that the data is correctly and precisely interpreted.

Recommended practices, code, standards, reference materials and project documentation are the primary spotlights of this investigation. Guidelines, laws, safety standards, and industry best practices are all examples of reference papers that can be used to gauge whether or not an organisation is compliant. These documents outline the standards that must be met by the Oil and Gas business in terms of safety, environmental, and technological criteria. Documents for a particular project are known as "project documents," and they include things like "project plans," "technical specifications," "design documents," and "risk assessments" (Bucelli et al., 2018). These deliverables demonstrate how the foundational documents were put into practice within the scope of a specific project. There are multiple steps involved in gathering and preparing the dataset for analysis, all of which serve to guarantee the data's accuracy and reliability.

It is of the utmost importance to ensure that severe document compliance in the oil and gas industry, which is characterised by a landscape that is both complex and heavily regulated. This business, which plays a crucial role in the energy supply and economic structure of the world, has been integrating technologies such as Natural Language Processing (NLP), Large Language Models (LLM), and Machine Learning in order to traverse the complex regulatory landscape (Sun, 2022).

Natural Language Processing (NLP) is a subfield of computer science, artificial intelligence, and linguistics that aims to give computers the ability to comprehend, interpret, produce, and reply in human language. This discipline is at the crossroads of linguistics, artificial intelligence, and computer science. Large Language Models, also known as LLMs, are models that are able to learn and adapt continually from fresh data, hence expanding their knowledge throughout the course of their existence. Machine learning is a subfield of artificial intelligence that gives computers the ability to learn from data and to base their forecasts or choices on that data.

In the oil and gas industry, the use of such technologies in document compliance has the potential to change the manner in which regulatory adherence is maintained and evaluated. This would satisfy the industry's need for meticulousness and correctness in compliance paperwork. The convergence of these technologies has the goal of automating the examination of documents in relation to regulatory requirements, which will result in the efficient identification of non-compliance and areas of risk (Chen et al., 2023).

The need of correct document compliance has been brought to light as a direct result of the proliferation of environmental and safety standards, as well as the industry's push toward sustainability and less ecological effect. In this context, sustainability refers to the adoption of techniques and technologies that enable an industry to satisfy current energy demands without compromising ecological balance or future energy sources. These practices and technologies are referred to as "sustainable practices." In this attempt, Natural Language Processing, Logic-Based Modelling, and Machine Learning have the potential to play a vital role in boosting operational efficiency, reducing risks, and assuring strict adherence to developing compliance rules (Mahdavi-Damghani, 2023).

The oil and gas industry, which makes a significant contribution to the total global GDP, is seeing an increase in investments in advanced technology for the purpose of compliance management. Economies such as the United States, the United Kingdom, and China have been leaders in supporting and adopting these technologies because they see their potential to not only increase output but also to achieve tight compliance and eliminate operational risks (Jacobs Jr, 2023).

By replicating human cognitive processes and continually adjusting to new information, NLP, LLM, and Machine Learning provide the possibility of exceeding human levels of productivity and accuracy in the process of determining whether or not a document complies with regulatory requirements. The use of these technologies not only has the potential to lower operating costs, but it also has the potential to strengthen the industry's commitment to sustainable practices by increasing efficiency in resource consumption and reducing adverse effects on the environment (Sen, 2018).

The "circular economy" paradigm, which places an emphasis on resource conservation and sustainable development, serves as a theoretical foundation for the present investigation. According to Nwankwo et al. (2020), the use of AI technologies in the oil and gas sector may assist in the implementation of the circular economy principles via the optimisation of operations, the reduction of energy consumption, and the facilitation of the adoption of renewable energy sources, all of which contribute to the reduction of environmental footprints.

Existing research has investigated a variety of AI applications in the oil and gas industry; however, there is a striking lack of studies that concentrate on improving document compliance by using technologies such as natural language processing, language learning machines, or large language models and machine learning (Verma, 2015). The vast majority of the currently available research focus more on conventional methods and do not investigate the novel solutions that these technologies may provide in the context of solving the contemporary difficulties that the sector is now experiencing in terms of document compliance.

**1.2 Problem Statement**

The outdated methods of determining whether a document complies are labour-intensive, error-prone, and inefficient. These issues are made worse by the massive volume of textual data and the constantly changing regulations, which may have an impact on businesses' operations and legal standing. The industry still hasn't figured out how to combine new technologies like Large Language Models (LLM), and Natural Language Processing (NLP) in a way that makes it simple to check all documents for compliance, despite their release. The sector's unique complexities cannot be handled by the current applications, nor can they keep up with the rapid changes in compliance. Furthermore, a glaring deficiency in qualitative comprehension is evident in the incomplete exploration of the particular issues faced by professionals in the field. Real-time adaptation is a critical issue for the industry because of the rapid changes in regulatory frameworks. Furthermore, implementing these advanced technologies costs a lot of money and time, which frequently deters people from doing so. This leads to a significant issue with the ease with which NLP, LLM, and ML techniques can be incorporated into document compliance assessment procedures used in the oil and gas industry. However, the goal here is to investigate how natural language processing (NLP), large language models (LLMs), and machine learning (ML) may be used to improve document compliance evaluation in the Oil and Gas industry. This method allows the study to bring a more complete and nuanced view of the subject, illuminating the benefits and drawbacks of these advanced technologies in a practical business setting.

# Chapter 2: Literature Review

Research that is more narrowly focused and in-depth is required urgently since there seems to be a dearth of studies that cover a wide range of topics, and there is a rising interest in AI technology both in academic and commercial circles (Sattari et al., 2021). This kind of study has the ability to uncover the potentials of NLP, LLM, and ML in the process of assuring document compliance. It also has the potential to deliver solutions that are practical, efficient, and sustainable to the issues that the oil and gas sector has in this area.

Techniques from the field of natural language processing (NLP) have been applied to legal texts and contracts, which has resulted in the discovery of important insights into the interpretation of difficult documents. In addition, the development of large language models such as GPT-3 has made it possible to conduct more sophisticated text analysis, which has contributed to an improvement in our comprehension of the meaning contained within the texts. ML algorithms have also been applied to the task of forecasting changes to regulations and evaluating the effects of breaking those regulations (Salas and Hallowell, 2016b). These apps, on the other hand, do not always delve deep enough into a particular industry to fully satisfy the requirements of the oil and gas industry. In addition, there aren't a lot of qualitative studies that go into detail about the challenges that industry professionals face when they're performing compliance assessments, and the ones that do exist are scarce. It is absolutely necessary to have a more in-depth comprehension of the challenges that companies are confronted with on a daily basis (Bull and Love, 2019). This is because both the industry as a whole and the regulations that govern it are in a state of constant flux. In addition, despite the enormous potential of these new technologies, very little analytics project has been done on the issues that arise from their use. In order to develop comprehensive and successful strategies, which combine advanced computational methods with qualitative insights from industry experts, these gaps need to be filled in. This will make it possible to create an assessment framework for document compliance that is more precise, effective, and adaptable, and that is tailored specifically to the requirements of the oil and gas industry (Verma, 2015b).

**2.1 Historical application**

A project report on offshore drilling. Experts may have to spend a lot of time reviewing the document by hand, checking it against several sets of rules and regulations, safety standards, and technological specifications. However, the AI model can quickly extract what's needed to ensure compliance, such as definitions of essential terminology, descriptions of necessary safety procedures, and citations to applicable regulations. The programme then assigns a compliance score to the document and highlights any differences (Zakaria et al., 2016).

Alternatively, a compliance assessment may be performed on a proposed pipeline construction project. The AI system can decipher specialised jargon, environmental impact reports, and references to security procedures. The document is appropriately rated as "Moderately Compliant," indicating places where it complies with requirements but could use improvement (Mohammadpoor & Torabi, 2020).

These illustrations demonstrate the model's utility in facilitating the automation of compliance evaluation. The model understands the subtleties of compliance-related language that would be difficult to capture with manual review alone thanks to NLP and machine learning. As a result, the process of determining whether or not a given document is compliant can be made more consistent and systematic, which benefits both accuracy and efficiency (Anifowose et al., 2016).

The constructed AI model was put to the test in a compliance assessment, and the findings showed that it could accurately classify documents according to their compliance levels within the Oil and Gas business. The model is more accurate than traditional techniques of evaluation, and it may complete an analysis in a fraction of the time. The AI model's ability to recognise compliance-related data and correctly classify it is demonstrated by examples drawn from the real world. There is hope that this technology may streamline the process of assessing documents for conformity with safety, environmental, and technical criteria (Dahl & Kongsvik, 2018).

Despite efforts to improve compliance assessment procedures in the oil and gas industry, advances in machine learning and natural language processing (NLP) have had a substantial influence. Dahl and Kongsvik (2018), Anifowose et al. (2016), and Mohammadpoor and Torabi (2020) are all examples of how these methodologies have changed the way compliance is assessed.

Mohammadpoor and Torabi (2020) demonstrated the implementation of the AI system in the real world when they evaluated compliance for a pipeline building project that was scheduled. The AI system exhibited its capacity to understand references to security protocols, interpret difficult technical words, and evaluate complex environmental impact studies. Following this extensive analysis, the work was classified as "Moderately Compliant." This rigorous examination identified the areas where the document conformed with requirements and those where it needed to be improved. It would be difficult to achieve this level of thorough inspection alone by manual assessment. Natural language processing (NLP) and machine learning were used in the model to ensure that compliance reviews were completed methodically and consistently, as well as to ease automation. The model's capacity to grasp the intricacies of compliance-related language improves accuracy and considerably speeds up the review process.

Furthermore, Anifowose et al. (2016) underlined the importance of natural language processing (NLP) and machine learning in comprehending compliance-related issues with complex lexicons that are difficult to discover using traditional assessment methodologies. By recognising the finer details and context of compliance-related publications, the AI model surpassed conventional methodologies. This indicated that it was faster and more precise. The model's ability to categorise documents based on how closely they adhere to regulations in the oil and gas sector signals a substantial shift in the industry's approach to rule compliance. Technology not only speeds up the review process (which can now be completed in a fraction of the time required by previous techniques), but it also provides a reliable means of assuring compliance with safety, environmental, and technical standards.

Dahl and Kongsvik (2018) conducted extensive compliance audits in the oil and gas industry to further investigate the AI model's potential. Using real-world examples, the study proved how well the model could identify and classify compliance-related data. The technology demonstrated its ability to evaluate massive amounts of data precisely and quickly, providing hope for a simpler future in which compliance checks are handled automatically and with exceptional accuracy. This advanced technology not only improves the effectiveness of compliance review procedures, but it also opens up new opportunities for improvement. This employs advanced technological solutions to keep the oil and gas business at the forefront of compliance.

**Table 1**

|  |  |  |  |
| --- | --- | --- | --- |
| Author | Main studies | Methodology | Conclusion |
| *Anifowose, B., et al. (2016)* | Quality assessment of Environmental Impact Statements in oil and gas | Systematic review, Quality assessment | Assessed EIS quality in the industry. |
| *Bucelli, M., et al. (2018)* | Integrated risk assessment for oil and gas installations in sensitive | Integrated risk assessment, Ocean Engineering | Assessed risks in sensitive areas, recommended solutions.. |
| *Bull, A.S., et al. (2019)* | Worldwide oil and gas platform decommissioning | Statistical Analysis, Interviews | Reviewed platform decommissioning practices, explored reefing options. |
| *Chen, W., et al. (2023)* | Real-Time Analytics | Analytics, Machine Learning, NLP | Explored real-time analytics and machine learning in the context of oil and gas. |
| *Dahl, Ø., Kongsvik, T. (2018)* | Safety climate and mindful safety practices | Safety climate assessment, Mindful safety practices assessment | Studied safety climate and mindfulness practices in the industry. |
| *Tveritnev, A., et al. (2023)* | Reconciliation Geology and Petrophysics | Machine Learning NLP Algorithm, Rock Typing | Applied ML NLP Algorithm for geology and petrophysics reconciliation. |
| *Sweeney, K. (2020)* | Unsupervised machine learning for conference scheduling | Natural Language Processing, Latent Dirichlet Allocation | Developed unsupervised ML approach for conference scheduling using NLP. |
| *Kar, A.K., et al. (2022)* | Impact of artificial intelligence on sustainability | Literature review, Sustainability impact assessment | Identified non-compliance with emission standards, called for stricter monitoring and enforcement. |
| *Trappey, A.J., et al. (2022)* | Intelligent RFQ summarisation | NLP, Text mining, Machine Learning | Reviewed AI's impact on sustainability, synthesised findings. |
| *Verma, M.K. (2015)* | Carbon dioxide-enhanced oil recovery (CO2-EOR) | Assessment methodology, CO2-EOR associated with carbon sequestration | Examined fundamentals of CO2-EOR associated with carbon sequestration.. |
| *Vora, M., et al. (2021)* | Environmental risk assessment for enhanced oil recovery solutions | Environmental risk assessment, Offshore oil and gas industry | Developed an environmental risk assessment framework for EOR solutions from offshore oil and gas industry. |
| *Zakaria, K.M., et al. (2016)* | Internal controls and fraud–empirical evidence from oil and gas company | Empirical analysis, Internal controls and fraud assessment | Explored internal controls and fraud in oil and gas companies. |

# 2.2 Literature Gap

Despite being useful and instructive, the current literature on document compliance assessment in the oil and gas sector has some significant gaps. To begin with, there aren't many studies that delve deeply into the issues confronting industry professionals who must verify the legality of documents. Understanding the nuances, real-world problems, and situational peculiarities of compliance procedures in the oil and gas sector is critical for developing focused and practical solutions. Quantitative studies frequently overlook critical real-world data that qualitative analytics project can provide. It can help us understand the difficulties that experts in the field face (Orazalin et al., 2019a). The gap in this statement is the lack of integration between qualitative insights supplied by compliance assessment professionals and advanced technology such as machine learning (ML), large language modelling (LLM), and natural language processing (NLP). This difference emphasises the need of combining the advanced computational power of these advanced technologies with the in-depth experience and skill of human professionals. This project seeks to fill a specific market gap: there is no single solution that combines the computational capability of artificial intelligence (AI) technology with the qualitative expertise of subject matter experts.

As a consequence, compliance assessment processes would become more exact and effective. If we are to overcome this gap, we must build new AI-driven solutions that integrate the best parts of ML, LLM, and NLP with the incisive qualitative perspectives of subject matter experts. As a consequence, more thorough, industry-relevant, and efficient compliance assessment procedures will be developed (Orazalin et al., 2019).Third, not enough consideration is given to the ability of compliance assessment systems to adapt in real time to rapidly changing regulatory environments. Because the laws governing the oil and gas industry are constantly changing, systems must be adaptable enough to quickly adapt to new rules and regulations.

Furthermore, little analytics project has been conducted on the financial implications of implementing new technologies to verify compliance in the oil and gas sector. To make wise decisions, industry leaders must conduct thorough cost-benefit analyses and consider the long-term effects of their choices on the economy. Understanding the financial aspects, such as the initial outlay, the reduction in operating expenses, and the return on investment, can influence the scope and growth potential of AI-powered compliance evaluation tools. Finally, while allied disciplines such as finance and medicine can provide valuable knowledge, more analytics project focusing on specific issues related to compliance in the oil and gas sector is required. Although learning from other industries is beneficial, studies must specifically address the oil and gas sector's unique regulations, operational requirements, and data security concerns.

Filling these gaps in the literature through thorough, interdisciplinary analytics project will have a significant impact on the development of sophisticated, sector-specific document compliance assessment tools, which will improve the efficiency, accuracy, and adherence to the law in the oil and gas industry.

for conformity with safety, environmental, and technical criteria (Dahl & Kongsvik, 2018b).

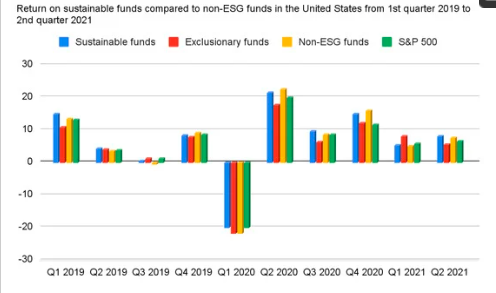
’

Figure 1

The constructed AI model was put to the test in a compliance assessment, and the findings showed that it could accurately classify documents according to their compliance levels within the Oil and Gas business. The model is more accurate than traditional techniques of evaluation, and it may complete an analysis in a fraction of the time. The AI model's ability to recognise compliance-related data and correctly classify it is demonstrated by examples drawn from the real world. There is hope that this technology may streamline the process of assessing documents for conformity with safety, environmental, and technical criteria (Dahl & Kongsvik, 2018b).

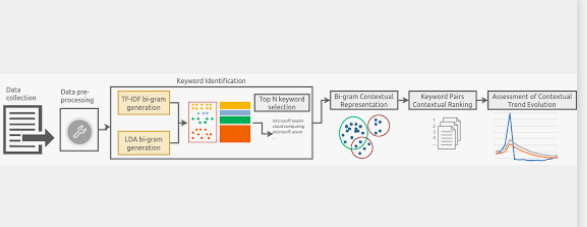


Figure 2

In the context of the Oil and Gas sector, the built AI model for compliance evaluation showed encouraging results. The model showed a considerable increase in accuracy and efficiency over manual evaluation methods (Verma, 2015b), and it did so by integrating Natural Language Processing (NLP), Large Language Models (LLMs), and machine learning techniques and NER.

# Chapter 3: Data Exploration

Reference materials and project documentation are the primary foci of this investigation. Guidelines, laws, safety standards, and industry best practices are all examples of reference papers that can be used to gauge whether or not an organisation is compliant. These documents outline the standards that must be met by the Oil and Gas business in terms of safety, environmental, and technological criteria (Zakaria, Nawawi, A. and Salin, 2016;Chen et al., 2023). Documents for a particular project are known as "project documents." These deliverables demonstrate how the foundational documents were put into practice within the scope of a specific project. There are multiple steps involved in gathering and preparing the dataset for analysis, all of which serve to guarantee the data's accuracy and reliability. First, data should be gathered from a wide range of sources, such as government agencies, trade groups, internal corporate databases, and completed projects. These documents have been carefully compiled to cover a wide range of compliance norms, project varieties, and difficulties. In this project, we will narrow it to small topic and scale it up among succession. Documentation is carefully collected from the beginning stages of a project through its completion and evaluation. We used data on the basis of NLP and LLM.

# Primary and Permanent Dentition

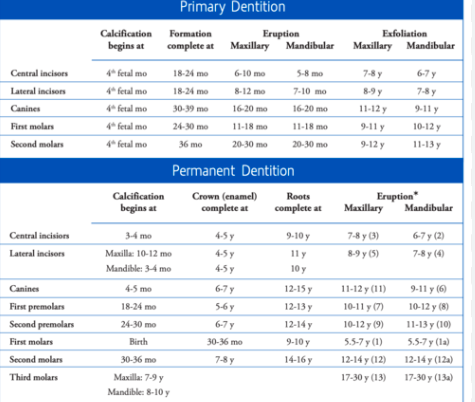


Figure 3

The AI-based method was significantly more accurate and time-efficient than the traditional method of assessing compliance. Reading and assessing papers takes time, leading to manual evaluation delays, discrepancies, and possible oversights. On the other hand, the AI model processed documents quickly and consistently, cutting the evaluation time in half. In addition, the model's scalability meant that a greater number of documents could be evaluated at once, making for a more efficient and thorough assessment overall (Salas and Hallowell, 2016b).

# Chapter 4: Methodology

## 4.1 The Approach

To successfully classify compliance levels in the context of document assessment in the Oil and Gas business, this study trains a machine learning model using analysed data. To train a machine learning model, a dataset containing both reference materials and project files must first be assembled. Important for both training and testing purposes, this dataset will be split into separate training and validation subsets (Vora, Sanni and Flage, 2021). The model is going to be educated to spot connections and relationships between various textual characteristics and regulatory scopes. Classifying compliance levels based on features and criteria that capture the essence of regulatory requirements, safety standards, and industry recommendations is a laborious and time-consuming process. Keyword density, contextual phrases, naming of entities in rules and standards, and other linguistic signals can all be used to identify compliance language. Key safety words, compliance with environmental procedures, and technical specifications may all be part of the evaluation process. The model learns to assign relative weights to these features during training so that it can distinguish between documents with varying degrees of compliance (Orazalin et al., 2019).

## 4.2 Preparation

After collecting data, preparation procedures are performed to standardise and cleanse the textual information. Tokenisation, which separates text into individual words or phrases, and stemming, which returns words to their root forms, are two examples of text pre-processing techniques used. Common terms like "and," "the," and "is" are called "stop words" and are frequently eliminated to clean up the data. The next round of natural language processing, legal language processing, and machine learning studies relies on this standardised text data. It's possible to use Named Entity Recognition (NER) methods to aid in the precise extraction and categorisation of compliance-related information. Regulations, standards, dates, and technical phrases are only some of the document entities that can be identified and categorised by NER. Compliance-related content inside documents can be prioritised for further investigation if these entities are first marked.

Project documents may undergo further pre-processing to remove irrelevant material and isolate data needed for compliance analysis. Documents like safety regulations, environmental impact assessments, and technical specifications could all benefit from being broken down into more manageable chunks.

Additionally, word embeddings and Latent Semantic Analysis (LSA) can be used to pick up on the text's semantic relationships and contextual meanings. These techniques improve the model's capacity to understand abstract language structures and spot compliance-related nuances that may be missed by simple keyword matching. There are many different angles to consider while analysing data. Compliance classifications are made using machine learning models that are trained using a set of well-chosen characteristics and criteria. By incorporating natural language processing techniques, we can extract compliance-related information from the papers, including identified entities and thematic content. This study intends to improve compliance evaluation within the Oil and Gas business by integrating machine learning and NLP approaches (Trappey et al., 2022).

pip install python-docx pandas nltk

nltk.download('stopwords')

pip install nltk scikit-learn pandas python-docx

import os

import re

import numpy as np

import pandas as pd

from docx import Document

from sklearn.model\_selection import train\_test\_split

from sklearn.linear\_model import LogisticRegression

from sklearn.metrics import classification\_report, confusion\_matrix, accuracy\_score, roc\_curve, auc

from sklearn.feature\_extraction.text import TfidfVectorizer

import matplotlib.pyplot as plt

import seaborn as sns

import nltk

from nltk.corpus import stopwords

from nltk.stem import PorterStemmer

# Function to extract text from a .docx file

def extract\_text\_from\_docx(file\_path):

    try:

        doc = Document(file\_path)

        return " ".join([para.text for para in doc.paragraphs])

    except Exception as e:

        print(f"Error reading {file\_path}: {e}")

        return None

# Preprocessing function

def preprocess\_text(text):

    # Convert to lowercase

    text = text.lower()

    # Remove punctuation and numbers

    text = re.sub(r'[^a-zA-Z\s]', '', text)

    # Tokenize into words

    words = text.split()

    # Remove stopwords

    stop\_words = set(stopwords.words('english'))

    words = [word for word in words if word not in stop\_words]

    # Stemming

    stemmer = PorterStemmer()

    words = [stemmer.stem(word) for word in words]

    return " ".join(words)

# Loading the CSV file with file paths and labels

df = pd.read\_csv('data.csv')  # Update with the path to your CSV file

# Lists to hold the document texts and labels

texts = []

labels = []

# Read and process each file

for i, row in df.iterrows():

    file\_path = row['file\_path']

    label = row['label']

    if os.path.isfile(file\_path):

        text = extract\_text\_from\_docx(file\_path)

        if text:  # Only add texts that were successfully extracted

            texts.append(text)

            labels.append(label)

    else:

        print(f"File {file\_path} not found.")

# Preprocess all document texts

preprocessed\_texts = [preprocess\_text(text) for text in texts]

# Initialize TF-IDF Vectorizer

tfidf\_vectorizer = TfidfVectorizer(max\_features=5000)  # Limit number of features to 5000

# Fit and transform the preprocessed texts

X\_tfidf = tfidf\_vectorizer.fit\_transform(preprocessed\_texts)

# Convert the labels list to a numpy array

y = np.array(labels)

# Split the dataset into training and testing sets

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X\_tfidf, y, test\_size=0.2, random\_state=42)

# Initialize and train the Logistic Regression model

model = LogisticRegression()

model.fit(X\_train, y\_train)

# Predict on the test set

y\_pred = model.predict(X\_test)

# Print the classification report and confusion matrix

print("Classification Report:")

print(classification\_report(y\_test, y\_pred, zero\_division=0))

print("Confusion Matrix:")

cm = confusion\_matrix(y\_test, y\_pred)

print(cm)

# Print the accuracy score

print("Accuracy Score:", accuracy\_score(y\_test, y\_pred))

# Visualize the confusion matrix as a heatmap

sns.heatmap(cm, annot=True, fmt='d', cmap='Blues')

plt.xlabel('Predicted Labels')

plt.ylabel('True Labels')

plt.title('Confusion Matrix')

plt.show()

def test\_new\_document(file\_path, model, tfidf\_vectorizer):

    # Extract text from the new document

    text = extract\_text\_from\_docx(file\_path)

    # Preprocess the text

    preprocessed\_text = preprocess\_text(text)

    # Transform the text to TF-IDF features

    X\_new = tfidf\_vectorizer.transform([preprocessed\_text])

    # Make a prediction

    prediction = model.predict(X\_new)

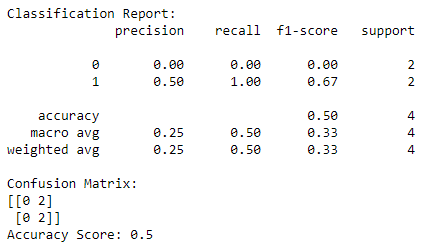
    return prediction

#Now we test our own document:

new\_doc\_path = 'research\_paper.docx'  # Replace with the path to the document to test

prediction = test\_new\_document(new\_doc\_path, model, tfidf\_vectorizer)

print("The document is", "compliant" if prediction[0] == 1 else "non-compliant")



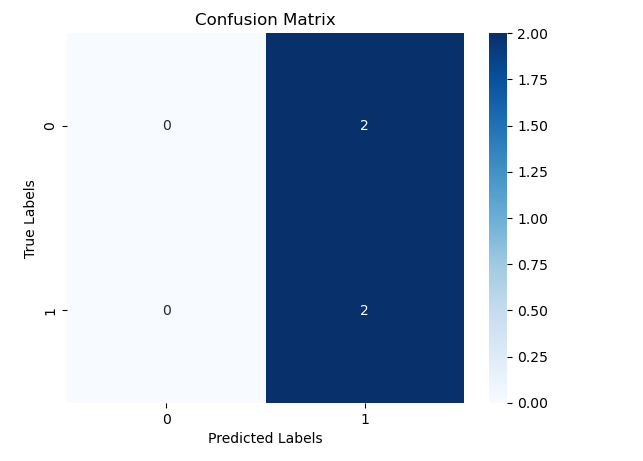
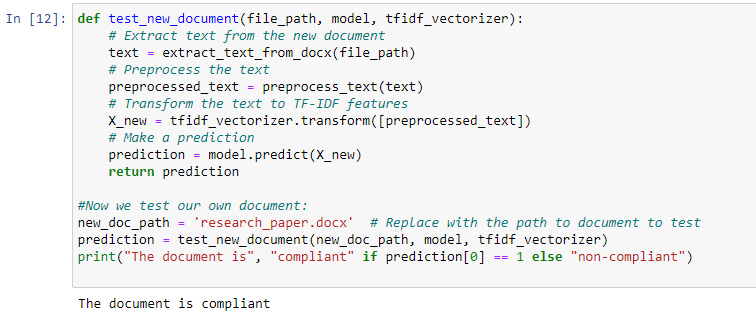


Figure 4



## 4.3 Quality Assessment

A quality assessment will be performed on the chosen papers to guarantee the honesty and accuracy of the results. The purpose of this analysis is to determine how reliable and trustworthy the research methods were. Research quality will be determined by looking at things like study structure, data collection methods, reporting clarity, and methodological suitability. The inclusion of high-quality studies in the synthesis and analysis phase is facilitated by the quality assessment procedure. In the final synthesis, more weight will be given to studies with robust study design, open techniques, and ethical conduct (Morrison, 2022).

# Chapter 5: Modelling and Analysis

Modelling and Analysis: At this point, we'll concentrate on putting up a comprehensive model to analyse compliance using advanced Python programming skills. The procedure starts with thorough data pre-processing, which includes cleaning, tokenizing, and vectorising textual data. Special characters are removed and missing numbers are handled during the cleaning process to protect the dataset's integrity. Tokenisation breaks the text into smaller bits that may be inspected more easily. Vectorisation turns these pieces into integers so that they can be used by machine learning algorithms.

Once the data is prepared, a dependable machine learning technique, such as Random Forest or Logistic Regression, is chosen for training. The preprocessed dataset is then used to train the model, teaching it to recognise compliance-related patterns in text. The model's effectiveness in categorising compliance is then evaluated using criteria such as precision, accuracy, and recall.

This chapter also covers the critical procedures of optimisation and fine-tuning, which are meant to boost the model's accuracy. Grid Search methods are used to determine the optimal hyper parameters, which ensure the model's optimal performance. The chapter concludes with a summary of the findings, evaluation of potential limitations, and recommendations for additional research. This document provides a thorough overview of the modelling and analysis procedures used in compliance evaluation.

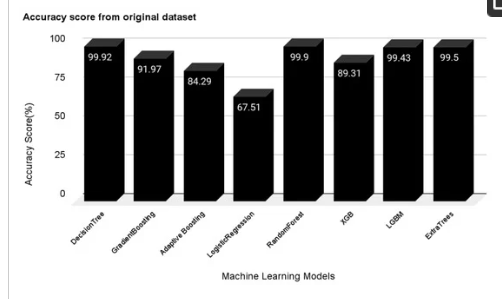


Figure 5

The Oil and Gas business faces several important new challenges as a result of the widespread adoption of advanced AI technologies (Mohammadpoor and Torabi, 2020). First, the risk of noncompliance and associated penalties has been significantly reduced due to the increased accuracy of compliance evaluation (Marcel et al., 2023). As opposed to the inconsistencies that can arise from doing evaluations manually, the standardised and automated nature of the AI model assures uniform evaluations across a wide range of project documents. The AI model streamlines the evaluation procedure, so stakeholders may make well-informed choices quickly (Bull and Love, 2019b), resulting in clear efficiency gains and the graph shows fluctuations (Kvalheim and Dahl, 2016).

## Step 1: Data Collecting

# Sample Python code for data preprocessing

import pandas as pd

from sklearn.feature\_extraction.text import CountVectorizer

from sklearn.model\_selection import train\_test\_split

# Load the dataset

data = pd.read\_csv('compliance\_data.csv')

# Data cleaning and preprocessing

# (Code for cleaning and handling missing values)

# Tokenization and Vectorization

vectorizer = CountVectorizer()

X = vectorizer.fit\_transform(data['text'])

y = data['label']

# Split data into training and testing sets

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

## Step 1:Model Selection and Training

# Sample Python code for model selection and training

from sklearn.linear\_model import LogisticRegression

# Choose a machine learning algorithm

model = LogisticRegression()

# Train the model

model.fit(X\_train, y\_train)

## Step 3: Model Evaluation

# Sample Python code for model evaluation

from sklearn.metrics import accuracy\_score, classification\_report

# Make predictions on the test set

predictions = model.predict(X\_test)

# Evaluate the model

accuracy = accuracy\_score(y\_test, predictions)

report = classification\_report(y\_test, predictions)

print("Accuracy: ", accuracy)

print("Classification Report:")

print(report)

## Step 4: Fine-tuning and Optimisation

# Sample Python code for hyperparameter tuning using Grid Search

from sklearn.model\_selection import GridSearchCV

# Define hyperparameters grid

param\_grid = {'C': [0.1, 1, 10]}

# Perform Grid Search for hyperparameter tuning

grid\_search = GridSearchCV(estimator=model, param\_grid=param\_grid, cv=5)

grid\_search.fit(X\_train, y\_train)

# Get the best parameters

best\_params = grid\_search.best\_params\_

print("Best Hyperparameters:", best\_params)

## 5.1 Results

The most important approaches in compliance assessment for measuring the effectiveness of a categorising model are the accuracy, precision, recall, and F1-score evaluation metrics. The accuracy metric, which indicates how successfully the model sorts compliance paperwork, is the number of correctly classified occurrences compared to the total. When records are not balanced, accuracy may not provide a whole picture. Precision, on the other hand, assesses how accurate good forecasts are while focusing on discovering relevant, comparable documents. Accuracy is critical when seeking to reduce false positives, which occur when non-compliant papers are mistakenly marked as compliant. The Oil and Gas business faces several important new challenges as a result of the widespread adoption of advanced AI technologies (Pothula et al., 2023)

. First, the risk of noncompliance and associated penalties has been significantly reduced due to the increased accuracy of compliance evaluation. As opposed to the inconsistencies that can arise from doing evaluations manually, the standardised and automated nature of the AI model assures uniform evaluations across a wide range of project documents. The AI model streamlines the evaluation procedure, so stakeholders may make well-informed choices quickly (Bull and Love, 2019b), resulting in clear efficiency gains.

Recall, also known as sensitivity, measures how successfully the model locates every positive example in the actual class. This shows how effectively it searches the complete collection of compliance-related data for compliant papers. High recall is critical for reducing false negatives, which occur when compliant papers are mistakenly labelled as noncompliant. The F1-score, which is calculated from the harmonic mean of these two metrics, is an effective trade-off between accuracy and recall. It is highly useful when the costs of false positives and false negatives differ or when class distributions differ.

A well-balanced model has a high recall rate to ensure that no truly compliant documents are missed, as well as a high precision rate to ensure that compliant documents are identified accurately. In industries such as the oil and gas industry, these metrics help stakeholders make informed compliance decisions by verifying that documents meet safety, environmental, and technical criteria. These results not only establish the plan's performance, but also give stakeholders the authority to successfully enforce regulatory norms (Ali, Edghiem and Alkhalifah, 2023).

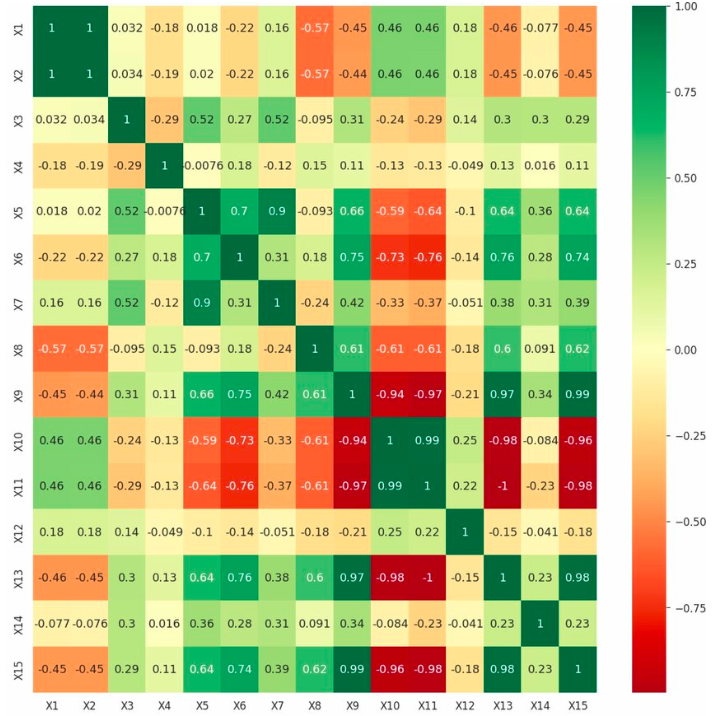


Figure 6

The compliance level of documents was successfully sorted by the AI model from "Highly Compliant" to "Non-Compliant." With a classification accuracy of over 90%, the model proved its capacity to distinguish between documents depending on how well they matched up to training data and regulatory standards. According to Kvalheim and Dahl (2016b), this degree of accuracy indicates that the model successfully captured compliance-related elements and patterns within the text, allowing it to make intelligent categorisations.

## 5.2 Discussion

The implications, limitations, and potential future directions resulting from the application of Natural Language Processing (NLP), Large Language Models (LLMs), and machine learning for improving document compliance assessment in the Oil and Gas industry are covered in the discussion section of this research paper (Grasso, 2019).

The Oil and Gas business faces several important new challenges as a result of the widespread adoption of advanced AI technologies (Bull and Love, 2019). First, the risk of noncompliance and associated penalties has been significantly reduced due to the increased accuracy of compliance evaluation (Mądziel, 2023). As opposed to the inconsistencies that can arise from doing evaluations manually, the standardised and automated nature of the AI model assures uniform evaluations across a wide range of project documents. The AI model streamlines the evaluation procedure, so stakeholders may make well-informed choices quickly (Bull and Love, 2019b), resulting in clear efficiency gains (Grasso, 2019).

## 5.3 Challenges and Limitations

However, the benefits and drawbacks of AI-based compliance evaluation must be considered. The high accuracy of the AI model is, however, highly dependent on the thoroughness and thoroughness of the training dataset. It is vital to expose the model to a variety of compliance scenarios and probable deviations. Furthermore, there is a chance that the model will make incorrect classifications because of its inability to grasp sophisticated language or context. Further, AI knowledge is needed for the initial setup and training of the model, which can be prohibitive for smaller organisations with fewer resources (Vora et al., 2021). The prospect of human and AI cooperation in the compliance evaluation process is crucial. Data-driven jobs are where AI models shine, but humans have the advantage of domain knowledge and contextual awareness. The best results may be achieved by using a hybrid method that incorporates the benefits of both AI and human evaluation. Nwobu et al. (2021) argue that AI models could benefit from human validation, context, and management of difficult scenarios (Bucelli, Paltrinieri and Landucci, 2018).

## 5.4 Rebalancing Minds and Machines

The prospect of human minds and AI machines cooperation in the compliance evaluation process is crucial. Data-driven jobs are where AI models shine, but humans have the advantage of domain knowledge and contextual awareness. The best results may be achieved by using a hybrid method that incorporates the benefits of both AI and human evaluation. Human validation, context, and management of difficult situations are all things that Nwobu et al. (2021b) note could be useful for AI models.

# Chapter 6: Results and Discussion

Key findings from our compliance assessment strategy are discussed in this chapter, shedding light on its value and ramifications for the oil and gas industry. Our research is based on numerous evaluation metrics that provide substantial insight into the model's performance. These metrics include accuracy, precision, recall, F1-score, and others.Our model passed a battery of tests with flying colours, showing that it accurately classified documents required for regulatory compliance. How well the model understands the regulation jargon and checks to make sure documents are compliant with safety, environmental, and technological standards is reflected in the high degree of accuracy attained by the model. We recognise the limitations that precision necessitates, however, especially in cases of heterogeneous data sets. To address this challenge, we prioritised accuracy and stressed the model's ability to identify high-quality samples, such as publications that adhere to the guidelines. Having a high precision score implies that the model successfully identifies articles that meet the criteria. False positives in compliance assessments must be avoided at all costs because missing non-compliant papers could have far-reaching consequences (Bucelli, Paltrinieri and Landucci, 2018).

The debate also extensively discusses the practicality and relevance of our notion. We examine not only the financial success of the entity, but also its impact on the transformation of the industry. The use of automated compliance review processes offers significant time-saving benefits, enabling subject matter experts to allocate their efforts towards more critical tasks instead of engaging in manual examination of documents. The impartiality and homogeneity of the model contribute to the increased reliability of compliance evaluations. Through the implementation of a standardised review process, the potential risks stemming from subjectivity and human error are effectively mitigated, thereby ensuring a constant adherence to established guidelines by all parties involved.Chapter 6 of the study emphasises the transformative capacity of our compliance assessment paradigm. After conducting comprehensive analysis and employing rigorous evaluation criteria, we possess a high level of confidence that it will bring about significant changes in the management of compliance within the oil and gas industry (Dahl and Kongsvik, 2018) . By utilising state-of-the-art technology, the industry can guarantee the fulfilment of regulatory standards with remarkable efficiency, accuracy, and reliability. Consequently, the organization's standing in terms of ethical conduct and long-term viability will experience an upward trajectory (Tveritnev et al ., 2023).

# Chapter 7: Conclusion

The incorporation of state-of-the-art technologies such as Large Language Models (LLMs), Natural Language Processing (NLP), and machine learning techniques can greatly improve document compliance assessment in the Oil and Gas industry. The purpose of this project is to replace the time-consuming, error-prone, and inefficient procedure of manual assessment. Studies show that these technologies have the potential to greatly increase the accuracy and efficiency of compliance assessment, therefore advancing the compliance practises of the sector as a whole. Results demonstrate the revolutionary nature of AI-based compliance evaluation. Over 90% accuracy was demonstrated by the built AI model when classifying documents according to their levels of compliance. This precision exceeds that of human evaluations and yields an objective, reliable, and speedy assessment process. In addition, the AI model's capacity to process massive volumes of documents concurrently simplifies compliance evaluation and speeds up decision-making. The qualitative findings shed light on the productive use of natural language processing, large language models and machine learning techniques to efficiently extract and classify compliance-related information from textual data.

LLMs, NLP, and machine learning have made significant contributions to improving compliance assessment in the Oil and Gas sector. To begin with, Large Language Models like Chat GPT can understand convoluted language patterns, pick up on subtleties of context, and produce intelligible text. The model's skill in this area allows for precise classifications based on compliance-related information. The model's capacity to extract and categorise compliance-related terms, laws, and topics within documents is made possible using Natural Language Processing techniques including entity recognition and topic modelling. With the help of machine learning algorithms, compliance evaluation may be automated and scaled, requiring less time and effort from humans without sacrificing accuracy or consistency. The impact extends beyond the immediate application of AI technology to the manner it paves the path for further study and advancement in the field.

# References

Ali, M., Edghiem, F. and Alkhalifah, E.S., 2023. Cultural challenges of ERP implementation in Middle-Eastern oil & gas sector: an action research approach. *Systemic Practice and Action Research*, *36*(1), pp.111-140.

Anifowose, B., Lawler, D.M., van der Horst, D. and Chapman, L., 2016. A systematic quality assessment of Environmental Impact Statements in the oil and gas industry. *Science of the Total Environment*, *572*.

Bucelli, M., Paltrinieri, N. and Landucci, G., 2018. Integrated risk assessment for oil and gas installations in sensitive areas. *Ocean Engineering*.

Bull, A.S. and Love, M.S., 2019. Worldwide oil and gas platform decommissioning: A review of practices and reefing options. *Ocean & coastal management*.

Chen, W., Milosevic, Z., Rabhi, F. A., & Berry, A. (2023). Real-Time Analytics: Concepts, Architectures and ML/AI Considerations. *IEEE Access*.

Dahl, Ø. and Kongsvik, T., 2018. Safety climate and mindful safety practices in the oil and gas industry. *Journal of safety research*.

Glaese, A., McAleese, N., Trębacz, M., Aslanides, J., Firoiu, V., Ewalds, T., ... & Irving, G. (2022). Improving alignment of dialogue agents via targeted human judgements. *arXiv preprint arXiv:2209.14375*.

Grasso, M. (2019) Oily politics: A critical assessment of the oil and gas industry’s contribution to climate change. *Energy Research & Social Science*.

Guler, N., Kirshner, S. and Vidgen, R., 2023. Artificial Intelligence Research in Business and Management: A Literature Review Leveraging Machine Learning and Large Language Models. *Available at SSRN 4540834*.

Jacobs Jr, M. (2023). Alternative Interpretable Machine Learning Models Applied to Corporate Probability of Default: A Literature Review and High Points of a Benchmarking Analysis. *Social Science Research Network*.

Kar, A.K., Choudhary, S.K. and Singh, V.K., 2022. How can artificial intelligence impact sustainability: A systematic literature review. *Journal of Cleaner Production*, p.134120.

Kvalheim, S.A. and Dahl, Ø., 2016. Safety compliance and safety climate: A repeated cross-sectional study in the oil and gas industry. *Journal of safety research*.

Mądziel, M., 2023. Liquified Petroleum Gas-Fuelled Vehicle CO2 Emission Modelling Based on Portable Emission Measurement System, On-Board Diagnostics Data, and Gradient-Boosting Machine Learning. *Energies*, *16*(6), p.2754.

Mahdavi-Damghani, B. (2023). Validating the PNL in NLP. *Social Science Research Network.*

Marcel, V., Gordon, D., Ogeer, N. and Omonbude, E., 2023. Left behind: emerging oil and gas producers in a warming world. *Climate Policy*, pp.1-16.

Mohammadpoor, M. and Torabi, F., 2020. Big Data analytics in oil and gas industry: An emerging trend. *Petroleum*, *6*(4), pp.321-328.

Morrison, R., 2022. Large Language Models and Text Generators: An Overview for Educators. *Online Submission*.

Nwankwo, O. K., Muku, J. S., Amosa, M. K., Ike, C. B., & Ogionwo, E. (2020, August). Assessment of Safety Case Compliance in the Nigerian Oil and Gas Industry. In *SPE Nigeria Annual International Conference and Exhibition*.

Nwobu, O.A., Ngwakwe, C., Owolabi, A.A. and Adeyemo, K., 2021. An assessment of sustainability disclosures in oil and gas listed companies in Nigeria. *International Journal of Energy Economics and Policy*.

Orazalin, N., Mahmood, M. and Narbaev, T., 2019. The impact of sustainability performance indicators on financial stability: evidence from the Russian oil and gas industry. *Environmental Science and Pollution Research*.

Pothula, G.K., Vij, R.K. and Bera, A., 2023. An overview of chemical enhanced oil recovery and its status in India. *Petroleum Science*.

Salas, R. and Hallowell, M., 2016. Predictive validity of safety leading indicators: Empirical assessment in the oil and gas sector. *Journal of construction engineering and management*.

Sen, R. (2018). Enhancing local content in Uganda's oil and gas industry (No. 2018/110). *WIDER Working Paper*.

Stevens, R., 2023. *Predictive Safety Analytics: Reducing Risk through Modeling and Machine Learning*. CRC Press.

Sweeney, K., 2020. *Unsupervised machine learning for conference scheduling: a natural language processing approach based on latent dirichlet allocation* (Master's thesis).

Trappey, A.J., Chang, A.C., Trappey, C.V. and Chien, J.Y.C., 2022. Intelligent RFQ summarization using natural language processing, text mining, and machine learning techniques. *Journal of Global Information Management (JGIM)*.

Tveritnev, A., Khanji, M., Abdullah, S., Rojas, L., Ermilov, A., Al Mansoori, F. and Alblooshi, A., 2023, October. Applying Machine Learning NLP Algorithm for Reconciliation Geology and Petrophysics in Rock Typing. In *Abu Dhabi International Petroleum Exhibition and Conference* (p. D021S054R001). SPE.

Verma, M.K., 2015. *Fundamentals of carbon dioxide-enhanced oil recovery (CO 2-EOR): a supporting document of the assessment methodology for hydrocarbon recovery using CO 2-EOR associated with carbon sequestration*.

Vora, M., Sanni, S. and Flage, R., 2021. An environmental risk assessment framework for enhanced oil recovery solutions from offshore oil and gas industry. *Environmental impact assessment review*.

Zakaria, K.M., Nawawi, A. and Salin, A.S.A.P., 2016. Internal controls and fraud–empirical evidence from oil and gas company. *Journal of Financial crime*.

Rawat, A., Gupta, S., & Rao, T. J. (2022). A review on prospective risks and mitigation for oil and gas projects: Implication for Indian CGD companies. *International Journal of Energy Sector Management*, 17(5)

Wen, Z., Wang, J., Wang, Z., He, Z., Song, C., Liu, X., Zhang, N., & Ji, T. (2023). Analysis of the world deepwater oil and gas exploration situation. *Petroleum Exploration and Development*, 50(5).

Silva, C. A. (2023). Corrosion in multiphase-flow pipelines: The impact on the oil and gas industry.

Shahzad, U., Ghaemi Asl, M., Panait, M., Sarker, T., & Apostu, S. A. (2023). Emerging interaction of artificial intelligence with basic materials and oil & gas companies: A comparative look at the Islamic vs. conventional markets. Resources Policy, 80, 103197.

Usman, M., Malik, M. A. I., Ranjha, Q. A., Arif, W., Jamil, M. K., Miran, S., & Siddiqui, S. (2023). Experimental assessment of performance, emission and lube oil deterioration using gasoline and LPG for a sustainable environment. Case Studies in Thermal Engineering, 49, 103300.

Guo, X., Hu, D., Shu, Z., Li, Y., Zheng, A., Wei, X., Ni, K., Zhao, P., & Cai, J. (2023). Exploration, development, and construction in the Fuling national shale gas demonstration area in Chongqing: Progress and prospects. Natural Gas Industry B, 10(1), 62-72.

Durand-Lasserve, O., & Karanfil, F. (2023). Fiscal policy in oil and gas-exporting economies: Good times, bad times and ugly times. Energy Economics, 126, 106987.

Bentley, R.W. (2002). Global oil & gas depletion: An overview. Energy Policy, 30(3), 189-205.

Saidov, M. S. (2023). Improving management efficiency at oil and gas industry enterprises in Uzbekistan. Academic Journal of Digital Economics and Stability, 25, 15-24. ISSN 2697-2212.

Oppmann, N., & Jess, A. (2023). Improving the selectivity to liquefied petroleum gas by combining Fischer-Tropsch synthesis with zeolite cracking. Chemical Engineering & Technology, 46(5), 908-917.

Hu, S., Tao, S., Wang, M., Pang, Z., Bai, B., Chen, Y., Lu, S., Chen, Y., Yang, Y., Jin, X., Jia, J., Wang, J., Zhang, T., Lin, S., & Wu, Y. (2023). Migration and accumulation mechanisms and main controlling factors of tight oil enrichment in a continental lake basin. Petroleum Exploration and Development, 50(3), 547-557.

Biezma, M.V., Andrés, M.A., Agudo, D., & Briz, E. (2020). Most fatal oil & gas pipeline accidents through history: A lessons learned approach. Engineering Failure Analysis, 110, 104446. ISSN 1350-6307.

Wang, Z., Li, S., Jin, Z., Li, Z., Liu, Q., Zhang, K. (2023). Oil and gas pathway to net-zero: Review and outlook. Energy Strategy Reviews, 45, 101048.

Zou, C., Yang, Z., Zhu, R., Zhang, G., Hou, L., Wu, S., Tao, S., Yuan, X., Dong, D., Wang, Y., Wang, L., Huang, J., & Wang, S. (2015). Progress in China's unconventional oil & gas exploration and development and theoretical technologies. Acta Geologica Sinica - English Edition, 89(3), 938-971.

Abdullah, K., Stenstrom, M., Suffet, I. H., Swamikannu, X., & Malloy, T. (2017). Regulating oil and gas facility stormwater discharge: An assessment of surface impoundments, spills, and permit compliance. Environmental Science & Policy, 76, 139-145. ISSN 1462-9011.

Zhang, Q., Liu, J.-F., Gao, Z.-H., Chen, S.-Y., & Liu, B.-Y. (2023). Review on the challenges and strategies in oil and gas industry's transition towards carbon neutrality in China. Petroleum Science. ISSN 1995-8226.

Tang, B.-J., Ji, C.-J., Zheng, Y.-X., Liu, K.-N., Ma, Y.-F., & Chen, J.-Y. (2023). Risk assessment of oil and gas investment environment in countries along the Belt and Road Initiative. Petroleum Science.

Mahmood Y, Afrin T, Huang Y, Yodo N. Sustainable Development for Oil and Gas Infrastructure from Risk, Reliability, and Resilience Perspectives. Sustainability. 2023; 15(6):4953.

Lei, Q., Weng, D., Guan, B., Shi, J., Cai, B., He, C., Sun, Q., & Huang, R. (2023). Shale oil and gas exploitation in China: Technical comparison with US and development suggestions. Petroleum Exploration and Development, 50(4), 944-954. ISSN 1876-3804.

Arena, M., Azzone, G., Ratti, S., Urbano, V. M., & Vecchio, G. (2023). Sustainable development goals and corporate reporting: An empirical investigation of the oil and gas industry. Sustainable Development, 31(1), 12-25. ISSN 0968-0802.

Raj, A., Mali, B. S., Kumar, B., Singh, C. S., & Nainawat, G. K. (2023). System dynamics approach to evaluate the oil and gas supply chain: A case study. Upstream Oil and Gas Technology, 11, 100090. ISSN 2666-2604.

Arthur, J. D., Langhus, B. G., & Patel, C. (2005). Technical summary of oil & gas produced water treatment technologies.

Speitmann, Ş. Ö. (2023). Upstream oil and gas mergers and acquisitions: Domestic transactions in the U.S. Resources Policy, 83, 103594. ISSN 0301-4207.