Article

Not peer-reviewed version

The Identification of a Promising Research Topic in Applying Generative Artificial Intelligence to Petroleum Engineering Problems

Boris Chigarev *

Posted Date: 17 January 2025

doi: 10.20944/preprints202501.1294.v1

Keywords: promising research topic; generative artificial intelligence; adversarial attacks; Scopus bibliometric records; GSDMM algorithm; Scimago Graphica program

Article

The Identification of a Promising Research Topic in Applying Generative Artificial Intelligence to Petroleum Engineering Problems

Boris Chigarev

Oil and Gas Research Institute, Russian Academy of Sciences, Moscow, Russia; bchigarev@ipng.ru

Abstract: This study aims to identify a promising research topic using generative artificial intelligence in the petroleum sector. It involves collecting information on generative artificial intelligence in the fields of engineering and computer science, systematizing publications using the GSDMM algorithm, and searching for OnePetro platform publications related to the identified research objectives. A total of 12424 Scopus bibliometric records were analyzed. Title and annotation fields were used to cluster the records. What gave 21 clusters. For each cluster, stacked histograms of the difference in term occurrence for that cluster and other clusters were constructed using the Scimago Graphica program. A promising research topic could be adversarial attacks by compromising generative models by manipulating input data. This topic is underrepresented in petroleum publications, but has significant research potential as it has been extensively written about in publications from other subject areas. The results of this study may help petroleum industry professionals to broaden their search for publications on generative models and thereby improve their knowledge of this research area.

Keywords: promising research topic; generative artificial intelligence; adversarial attacks; Scopus bibliometric records; GSDMM algorithm; Scimago Graphica program

---

Introduction

The Objective and Tasks of Research

The aim of this study is to identify a promising research topic in using of generative artificial intelligence to solve engineering tasks in the oil and gas sector.

In this paper, the following tasks are accomplished to achieve the objective:

- Collection of information on generative artificial intelligence related to the fields of engineering and computer science indexed in Scopus in 2024.
- Systematization of publications by clustering using the GSDMM algorithm applied to the title and abstract texts of bibliometric records exported from Scopus.
- Selecting a visual representation of the clustering results to determine which terms best describe the subject matter of each cluster to potentially expand the literature search.
- Selection of publications based on their cluster membership score assigned by the GSDMM algorithm and terms reflected in visual analysis diagrams.
- Find examples of OnePetro platform publications that can be attributed to this cluster's topics.
- Identification of promising research topic that are poorly or not represented in OnePetro platform publications but have potential based on publications in more general databases.

A Brief Literature Review

The importance of the topic under study is well reflected in the McKinsey reports. According to the latest McKinsey Global AI Survey 2024, 65% of respondents said their organizations regularly use

 doi:10.20944/preprints202501.1294.v1

generative AI, nearly double the number in the previous survey conducted just 10 months ago. The survey also provides insight into the risks associated with generative AI, particularly its inaccuracy. For the past six years, the adoption rate of AI in respondents' organizations has hovered around 50%. This year, according to the study, that figure jumped to 72%. Respondents say investments in generative and analytic AI are starting to pay off. [1]. A recent McKinsey Global Survey found that employees are far ahead of their organizations in their use of generative AI. Companies need a holistic approach to transforming the way the entire organization works with generative AI; technology alone will not create value. Generative AI must be applied by rethinking employee skills and making management and infrastructure changes [2].

GAI engineering solutions are widely used in medicine. An example is the article [3] which reviews advances in medical LLMs, their integration with Retrieval-Augmented Generation (RAG) and operational engineering, and their application to improve diagnostic accuracy and educational utility, and concludes that their safe integration requires continuous evaluation.

The feasibility of analyzing bibliometric data using generative artificial intelligence to identify trends in the subject matter of European Resuscitation Council (ERC) congress recommendations over the last decade is presented in [4].

Let us give some examples to show how much attention is given to generative artificial intelligence in scientific publications.

Thus, according to a simple query "generative AI" to the abstract database The Lens, the number of articles in which the term appears in the titles and abstracts amounted to 251 in 2022, 2210 in 2023, and 5664 scientific papers in 2024. In the dimensions.ai database on a similar request, in 2022 — 677, in 2023 — 4300, and in 2024 — 12324 publications (not only articles). Actual as of 27.12.2024. Thus, the actuality of the topic is reflected not only in the data of Scopus database.

The main areas of research, according to the Scopus database query used in this article, are: Computer Science (10167); Engineering (5959); Mathematics (3008); Social Sciences (1208); Physics and Astronomy (1101); Decision Sciences (958); Materials Science (927); Medicine (604); Earth and Planetary Sciences (476); Energy (472); Arts and Humanities (399); Business, Management and Accounting (318).

Research in Computer Science, Engineering and Mathematics is applicable to various branches of knowledge and may be of most interest to professionals in the oil and gas sector. But knowledge from other fields, such as Decision Sciences and Materials Science, can also be adapted to their interests.

According to Scopus bibliometric data, the main publication sources for 2024 are: Lecture Notes In Computer Science Including Subseries Lecture Notes In Artificial Intelligence And Lecture Notes In Bioinformatics (499); IEEE Access (309); ACM International Conference Proceeding Series (259); Proceedings Of Machine Learning Research (249); Communications In Computer And Information Science (176); Multimedia Tools And Applications (162); Lecture Notes In Networks And Systems (161). Interestingly, the term IEEE appears 47 times in the full list of publishers, indicating that it is possible to extend the literature search on the subject under consideration also in the IEEE Xplore system, which provides open access to its bibliometric data.

The topic is most represented by authors from China (4649); United States (2455); India (1300); United Kingdom (760); South Korea (546); Germany (533). By contrast, the number of publications attributed to Russia is only 66.

There is nothing surprising in the data presented above, for example, according to FUNDING SPONSOR data, 4 Chinese resources alone are funding the following number of studies related to the topic under study in 2024: National Natural Science Foundation of China (2190); Ministry of Science and Technology of the People's Republic of China (1772); National Key Research and Development Program of China (571); Ministry of Education of the People's Republic of China (277).

Equally interesting is some of the data on funding in the U.S. — U.S. Department of Defense (182); U.S. Department of Energy (68); U.S. Navy (60); U.S. Army (55); U.S. Air Force (50).

 doi:10.20944/preprints202501.1294.v1

The relevance of the topic under consideration for the oil and gas sector can be reflected by the fact that the query "generative ai" to the OnePetro database yielded 147 Search Results, of which Journal Articles (24) and Proceedings Papers (116). According to the personal observations of the author of this article, a significant excess of Proceedings Papers over Journal Articles is more characteristic of new topics.

The significant difference in the number of publications on the query "generative ai" to the OnePetro database and general abstract databases, indicates the possibility of knowledge transfer from general engineering research to the oil and gas sector.

It is not necessary to give a brief overview of the topic "generative ai" in the introduction, because even in one ScienceDirect, 21 review articles in the field of Computer Science and Engineering were published in 2024 for the simple query "generative ai". And for the query "Title, abstract, keywords: generative adversarial network" in 2024 there were 1394 papers published, of which review articles —143. However, adding "AND OnePetro" or "AND GSDMM" used in our work or even "AND VOSviewer" to the query will give us zero results. Note: VOSviewer [5] — the most widely used program for bibliometric analysis in recent times. Current as on December 28, 2024.

Regarding the use of the GSDMM algorithm, it is most appropriate to refer to the work of the authors of the algorithm [6] and its implementation posted by Ryan Walker on GitHub [https://github.com/rwalk/gsdmm].

Due to the strict specificity of the tasks of this paper, the identification of their direct analogs in publications was not found.

Materials and Methods

The main data used in this paper were bibliometric records exported from Scopus on query: "(TITLE-ABS-KEY ("generative artificial intelligence" OR "generative AI" OR "AI generated" OR "Generative Adversarial Networks" OR "Generative Model" OR "GenAI" OR "Simulative Generation" OR "Creative Generation" OR "Interactive Generation") AND (LIMIT-TO (DOCTYPE,"ar") OR LIMIT-TO (DOCTYPE,"cp") OR LIMIT-TO (DOCTYPE,"re")) AND (LIMIT-TO (SUBJAREA,"COMP") OR LIMIT-TO (SUBJAREA,"ENGI") OR LIMIT-TO (SUBJAREA,"MATH") OR LIMIT-TO (SUBJAREA,"DECI") OR LIMIT-TO (SUBJAREA,"ENER") OR LIMIT-TO (SUBJAREA,"EART")) AND (LIMIT-TO (PUBYEAR,2024)) AND (LIMIT-TO (LANGUAGE,"English")))". There are 12471 results for the query. Data is current as of November 7, 2024.

Since the title and annotation fields were used to cluster the records, entries without the annotation field were excluded from consideration, leaving a total of 12424 results.

The GSDMM algorithm implemented by Ryan Walker on Rust was used to cluster the records. [https://github.com/rwalk/gsdmm-rust].

The GSDMM algorithm was used with the following parameters: alpha=0.07, beta=0.37, K=100, maxit=500, vocab size=1227. What gave 21 clusters.

The parameters were selected by simple testing. Attention was paid to the number of obtained clusters and sufficiently uniform distribution of records (documents) across the clusters. To ensure convergence of the obtained results, overestimated values of K=100, maxit=500 were used.

The dictionary for the GSDMM algorithm was formed from Author keywords. Index keywords were not used because they are often redundant and may occur less frequently or not at all in the texts of titles and abstracts on which clustering was performed. Dictionary lemmatization was used to normalize the spelling of keywords, which, unlike stemming, preserves the readability of terms. The latter is very important for the graphical presentation of the analysis results.

Title and annotation texts were subjected to the same lemmatization as keywords.

For complex keywords, the space character was replaced by an underscore so that the keyword is treated as a single term by the algorithm.

The Scimago Graphica program was used to create the diagrams [7].

 doi:10.20944/preprints202501.1294.v1

For each cluster, stacked histograms of the difference in term occurrence for that cluster and other clusters were constructed.

For a better understanding of cluster topics, the diagrams are accompanied by examples of publications that reveal cluster topics. Articles with a high cluster membership score obtained using the GSDMM algorithm are given. To find publications related to the petroleum sector, we used keyword queries of this cluster to the OnePetro service. It should be noted that there is a certain subjectivity in the selection of publications for the example. That is, these publications are close to the topic of the cluster, but cannot match a large number of terms related to the cluster. The queries to OnePetro were not time-limited, but the publications found were new, which only emphasizes the novelty of the topic. In the oil and gas industry, the use of generative models has not yet gained much momentum, the reliability of new approaches is not yet comparable to older methods, and this is a natural limitation for their implementation. Therefore, an important aspect is the demonstration of realizations of solutions to problems that could not be realized by older, well-tested methods. The identification of such works can be of great value for bibliometric research. In addition, after examples of publications from the OnePetro platform, the Subject to which the platform assigns the publication is indicated. Note: the most specific Subjects were selected, e.g., if an article is categorized by the system as 'Subjects: Artificial intelligence, Artificial Lift Systems, Beam and related pumping techniques, Information Management and Systems, Neural networks, Well & Reservoir Surveillance and Monitoring', then the terms 'Artificial Lift Systems', 'Beam and related pumping techniques' and 'Well & Reservoir Surveillance and Monitoring' were selected to reflect the scope of 'Artificial intelligence'.

Results and Discussions

General Description of the Occurrence of Terms in Clusters

The list below only contains terms with 200 or more occurrences. Cluster #97 does not contain any specific high-frequency terms and is not shown in the list. The use of pivot tables was attempted to represent the occurrence of terms in clusters, but due to their sparsity it was not reasonable to place them in the text of the article, so the data is presented as a list.

cluster # → Terms (Occurrences)

17 → accuracy (210); model (204)

18 → image (287); model (235); dataset (211); generative_adversarial_network (202)

20 → model (300); generative_model (258); diffusion_model (220); image (217); train (215)

25 → model (328); dataset (247); generative_adversarial_network (222); train (216); synthetic_datum (207)

28 → design (308); ai (306); generative_ai (267)

31 → model (512); generative_model (416); learn (313); train (244); challenge (237); framework (215); dataset (213)

34 → image (1033); generative_adversarial_network (913); model (750); dataset (640); network (594); train (516); gan (493); quality (470); generator (434); challenge (426); design (355); accuracy (348); discriminator (330); application (314); algorithm (312); art (309); deep_learn (308); learn (280); resolution (249); noise (245); loss_function (232); attention (229); framework (227); architecture (218); reconstruction (209); super-resolution (201)

46 → model (520); large_language_model (478); llm (363); dataset (300); challenge (283); ai (269); generative_ai (250); application (217)

48 → model (352); generative_adversarial_network (310); dataset (265); train (250); network (244); accuracy (234); security (201)

49 → image (656); model (567); dataset (519); train (489); generative_adversarial_network (424); gan (309); challenge (296); accuracy (284); quality (262); application (225); deep_learn (219); synthetic (211); network (205); medical_image (201); segmentation (201)

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 January 2025 doi:10.20944/preprints202501.1294.v1

50 → model (433); generative_adversarial_network (367); prediction (236); accuracy (227); challenge (217); dataset (214); train (206)

60 → model (406); dataset (368); learn (282); train (248); challenge (206)

66 → model (298); image (286); generative_model (211)

82 → generative_ai (435); ai (387); technology (366); challenge (328); application (286); model (271); artificial_intelligence (224); generative_artificial_intelligence (219)

84 → model (359); design (294); train (267); generative_adversarial_network (249)

87 → student (595); ai (563); generative_ai (559); learn (513); chatgpt (423); education (387); technology (382); generative_artificial_intelligence (358); design (343); challenge (297); teach (285); artificial_intelligence (237); application (235); model (235)

89 → model (501); generative_adversarial_network (444); accuracy (423); dataset (394); train (371); challenge (248); classification (232); network (231); gan (224); application (200)

92 → model (242)

94 → image (608); model (546); generative_adversarial_network (414); dataset (402); quality (400); train (349); gan (336); art (314); challenge (312); generation (300); learn (288); network (274); generative_model (250); design (244); generator (229); framework (209); semantic (201)

Many terms, such as: image; model; dataset, are present in many clusters, so when plotting the occurrence of terms in a cluster, we focused on the difference between the occurrence of a term in that cluster and other clusters.

Cluster Term Occurrence Diagrams and Publication Examples

The cluster numbering is kept in the order that they were obtained using the GSDMM algorithm.

Figure 1 shows the term occurrence diagram for cluster #17 containing 273 documents and 461 terms obtained by GSDMM algorithm.

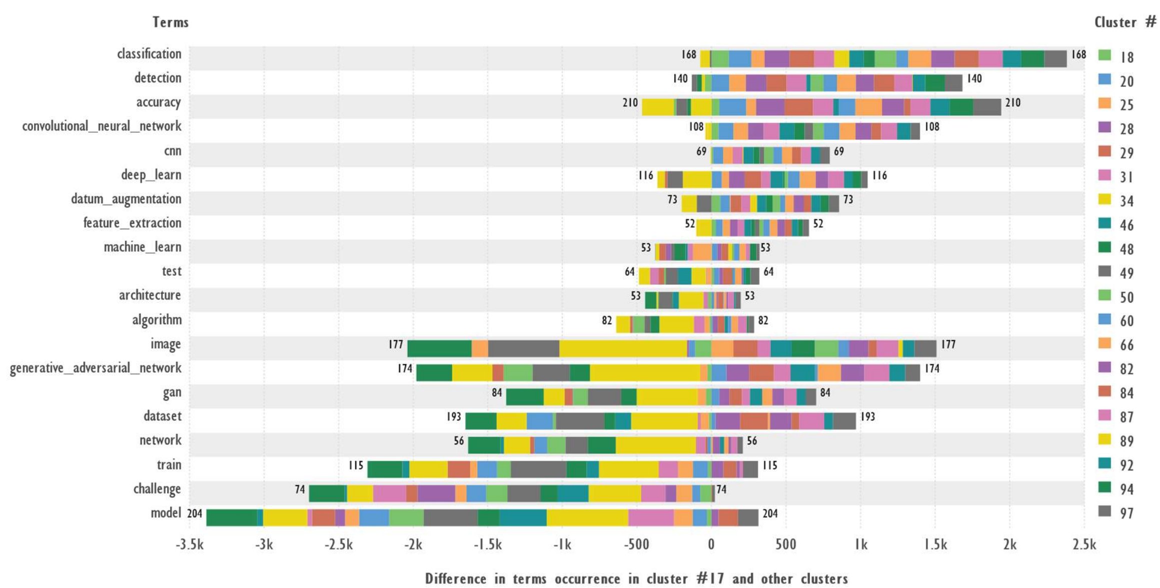


Figure 1. Stacked histogram of difference in term occurrence for cluster #17.

Examples of publications that contain the top terms of this chart and have a high score of belonging to this cluster.

"Detection of Diseases in Tomato Leaves Using Deep Learning Models: A Survey" [8], "War Snake Optimisation Algorithm with deep Q-Net for COVID-19 classification" [9].

Note. Publications with the maximum cluster membership score do not necessarily contain the top terms from the figure; cluster membership is evaluated by a set of term similarities.

 doi:10.20944/preprints202501.1294.v1

According to the query to the OnePetro platform: '"generative adversarial network'' ''detection'' ''classification'' ''accuracy'' ''feature extraction'' ''data augmentation''' 4 results were obtained, of which two publications are the most relevant to the topics of this cluster — "Imbalanced Working States Recognition of Sucker Rod Well Dynamometer Cards Based on Data Generation and Diversity Augmentation" [10] (Subjects: Beam and related pumping, Well & Reservoir Surveillance and Monitoring.) and "Automated Hyperparameter Optimization of Convolutional Neural Network (CNN) for First-Break (FB) Arrival Picking" [11] (Subjects not specified).

The chart shows that it is advisable to create an abbreviation dictionary with 'CNN'→'Convolutional Neural Network' and 'GAN'→'Generative Adversarial Network'.

Figure 2 shows the term occurrence diagram for cluster #18 containing 428 documents and 532 terms obtained by GSDMM algorithm.

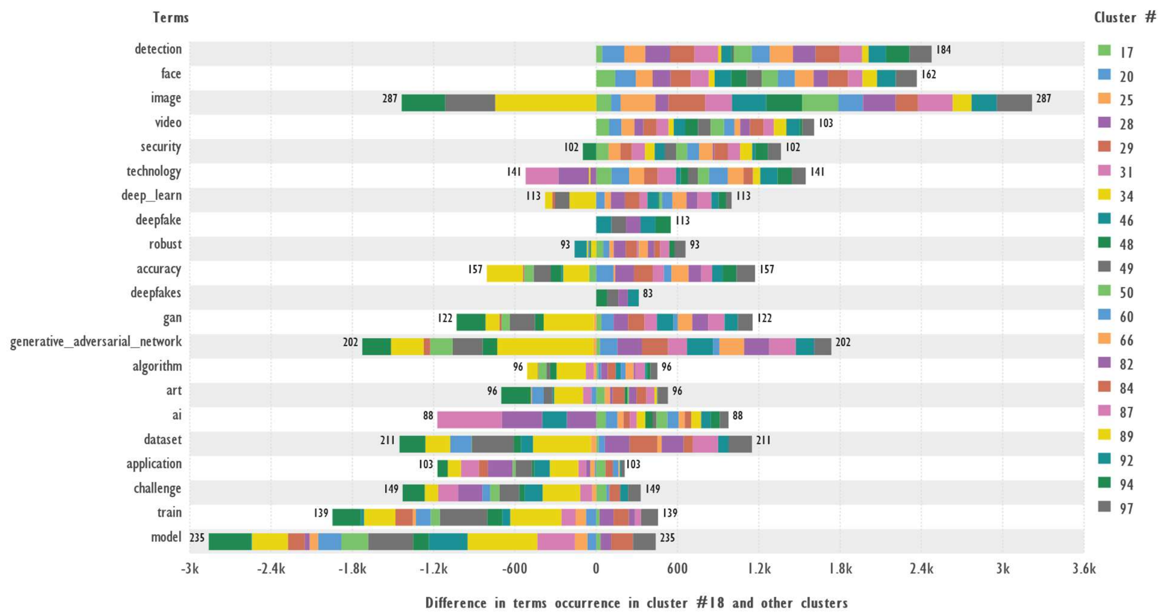


Figure 2. Stacked histogram of difference in term occurrence for cluster #18.

Examples of publications that contain the top terms of this chart and have a high score of belonging to this cluster.

"Enhancing Social Media Security: LSTM-Based Deep Fake Video Detection" [12], "A Novel Approach for Detecting Deepfake Face Using Machine Learning Algorithms" [13].

According to the query to the OnePetro platform: '"generative adversarial network" detection face video security deep fake' results most relevant to this cluster were as follows — "Personal Protective Equipment Detection Using Computer Vision Techniques" [14] and "The Role of Personalized Generative AI in Advancing Petroleum Engineering and Energy Industry: A Roadmap to Secure and Cost-Efficient Knowledge Integration: A Case Study" [15]. Subjects: Safety, Data security. Using the term deepfake instead of deep fake without quotation marks yielded no results. When referring to OnePetro, an approximate but related search is used rather than an exact search using the terms presented in the diagram.

Based on the chart, the following was entered into the vocabulary lemmatizer — deepfakes→deepfake.

Figure 3 shows the term occurrence diagram for cluster #20 containing 449 documents and 504 terms obtained by GSDMM algorithm.

 doi:10.20944/preprints202501.1294.v1

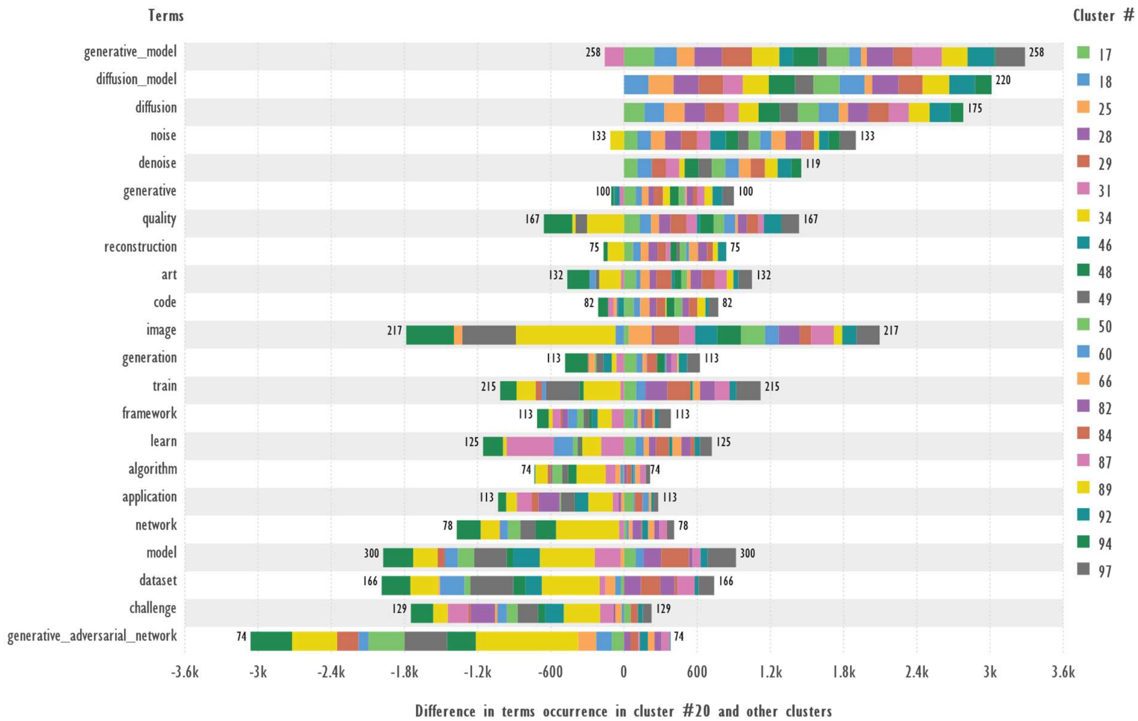


Figure 3. Stacked histogram of difference in term occurrence for cluster #20.

Examples of publications that contain the top terms of this chart and have a high score of belonging to this cluster.

"Mix-DDPM: Enhancing Diffusion Models through Fitting Mixture Noise with Global Stochastic Offset" [16], "Listening to the Noise: Blind Denoising with Gibbs Diffusion" [17].

According to the query to the OnePetro platform: '"generative model" "diffusion model" noise denoise reconstruction' 7 results were obtained, of which 3 publications are the most relevant to the topics of this cluster — "Integrating Multi-Source Experiment Data and Variational Diffusion Model for Intelligent Constructing Digital Core: A Case Study of Sandstone Reservoir in the Turgay Basin Central South Kazakhstanc" [18], and "Implementation of Denoising Diffusion Probability Model for Seismic Interpretation" [19], and "A Missing Well-Logs Imputation Method Based on Conditional Denoising Diffusion Probabilistic Models" [20]. Subjects: Reservoir Characterization, Reservoir Fluid Dynamics, Seismic processing and interpretation.

Figure 4 shows the term occurrence diagram for cluster #25 containing 413 documents and 507 terms obtained by GSDMM algorithm.

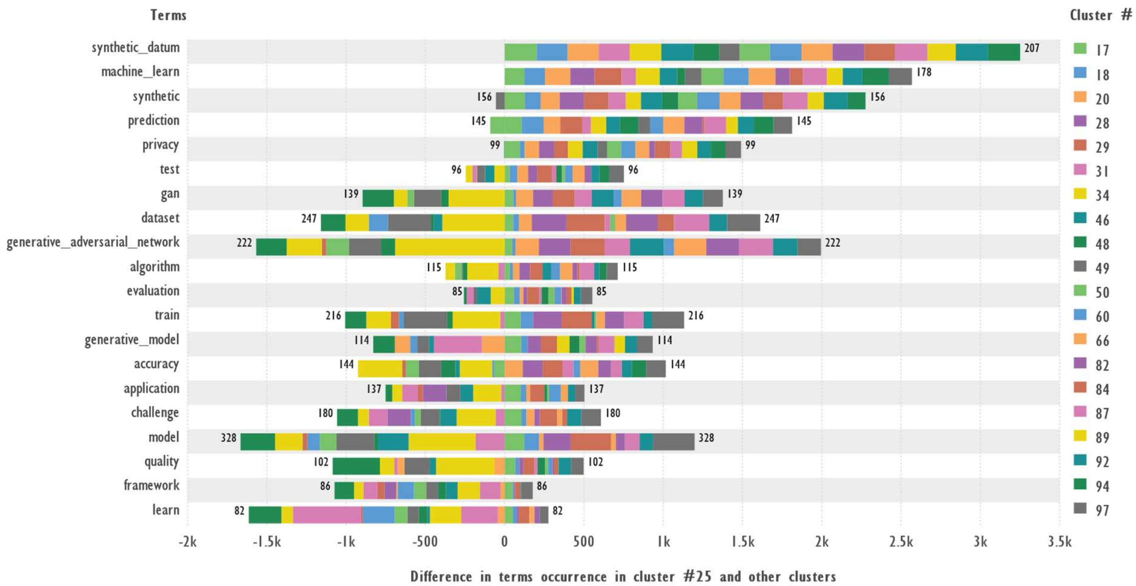


Figure 4. Stacked histogram of difference in term occurrence for cluster #25.

Examples of publications that contain the top terms of this chart and have a high score of belonging to this cluster.

“Generating Synthetic Time-Series Data on Edge Devices Using Generative Adversarial Networks” [21], “From shallows to depths: unveiling hybrid synthetic data modeling for enhanced learning with privacy considerations in naturally imbalanced datasets” [22].

According to the query to the OnePetro platform: ‘“generative adversarial network” “synthetic data” “machine learn” privacy’ results most relevant to this cluster were as follows — “Machine Learning-Based CO2 Saturation Tracking in Saline Aquifers Using Bottomhole Pressure for Carbon Capture and Storage CCS Projects” [23] (Subjects: Storage Reservoir Engineering, CO2 capture and sequestration) and previously mentioned publication [15].

Figure 5 shows the term occurrence diagram for cluster #28 containing 543 documents and 458 terms obtained by GSDMM algorithm.

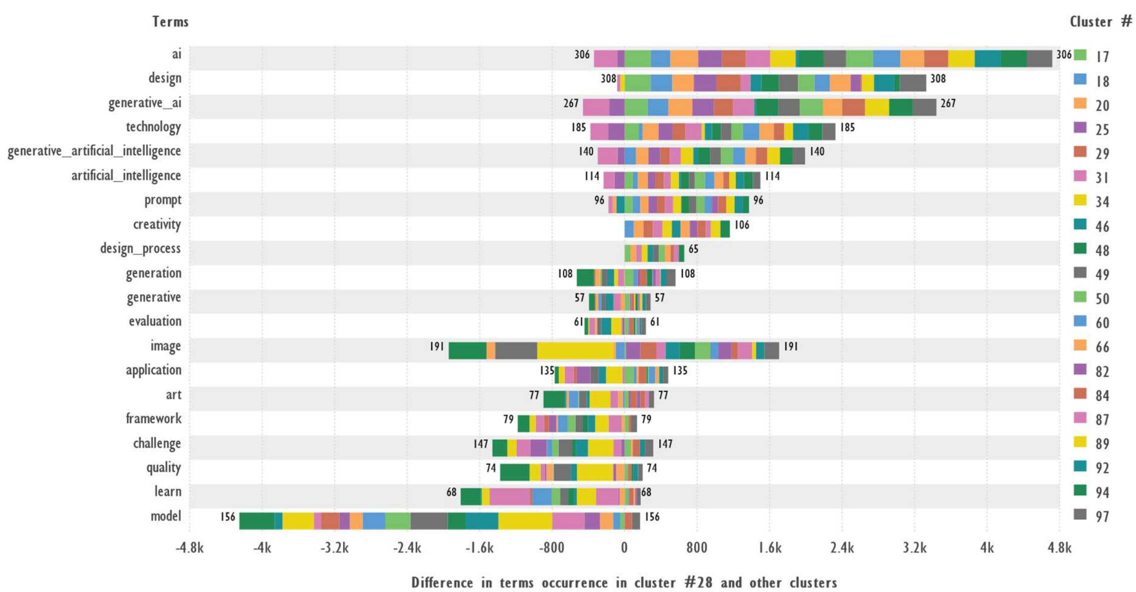


Figure 5. Stacked histogram of difference in term occurrence for cluster #28.

Examples of publications that contain the top terms of this chart and have a high score of belonging to this cluster.

"PlantoGraphy: Incorporating Iterative Design Process into Generative Artificial Intelligence for Landscape Rendering" [24], "Exploring the application path of AIGC technology in the styling design of traditional artifacts-a case study of Song Dynasty Lacquerware" [25].

According to the query to the OnePetro platform: '"generative artificial intelligence" design technology prompt creativity' one publication was found — "Comparative Analysis of Single and Multiagent Large Language Model Architectures for Domain-Specific Tasks in Well Construction" [26]. Subject: Information Management and Systems.

Figure 6 shows the term occurrence diagram for cluster #29 containing 118 documents and 303 terms obtained by GSDMM algorithm.

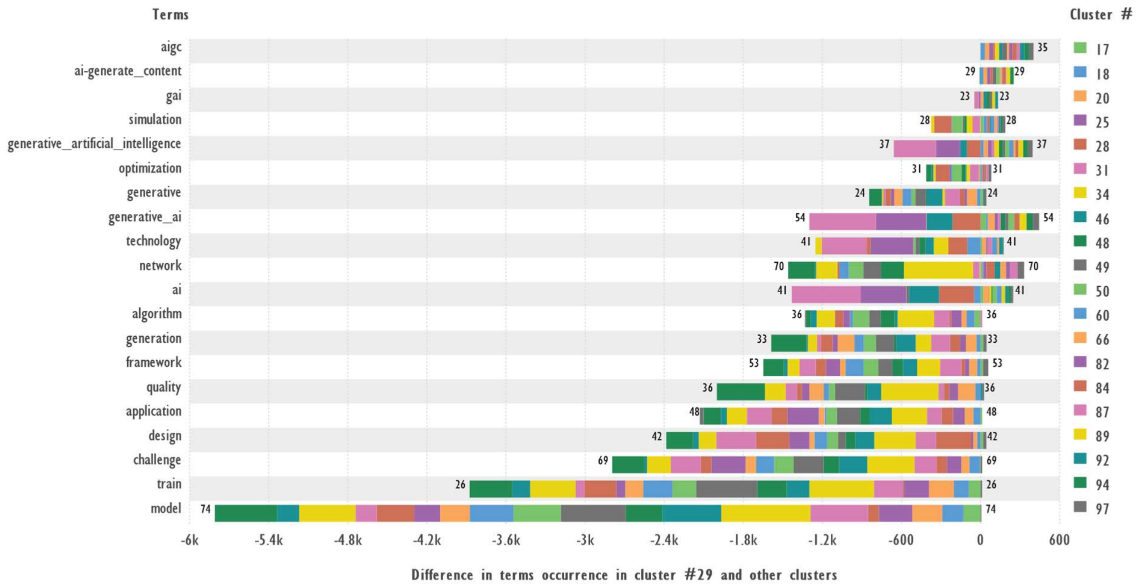


Figure 6. Stacked histogram of difference in term occurrence for cluster #29.

Examples of publications that contain the top terms of this chart and have a high score of belonging to this cluster.

"AI-Generated Content-as-a-Service in IoMT-Based Smart Homes: Personalizing Patient Care With Human Digital Twins" [27], "Semantic Communications for Artificial Intelligence Generated Content (AIGC) Toward Effective Content Creation" [28].

According to the query to the OnePetro platform: '"ai generate content" "generative artificial intelligence"' one publication was found — "Generative AI Enabled Conversational Chatbot for Drilling and Production Analytics" [29]. Subject: Information Management and Systems.

It should be noted that 'aigc' and 'gai' are abbreviations for the terms used in the query.

The small number of terms distinguishes this cluster from the others.

Figure 7 shows the term occurrence diagram for cluster #31 containing 730 documents and 641 terms obtained by GSDMM algorithm.

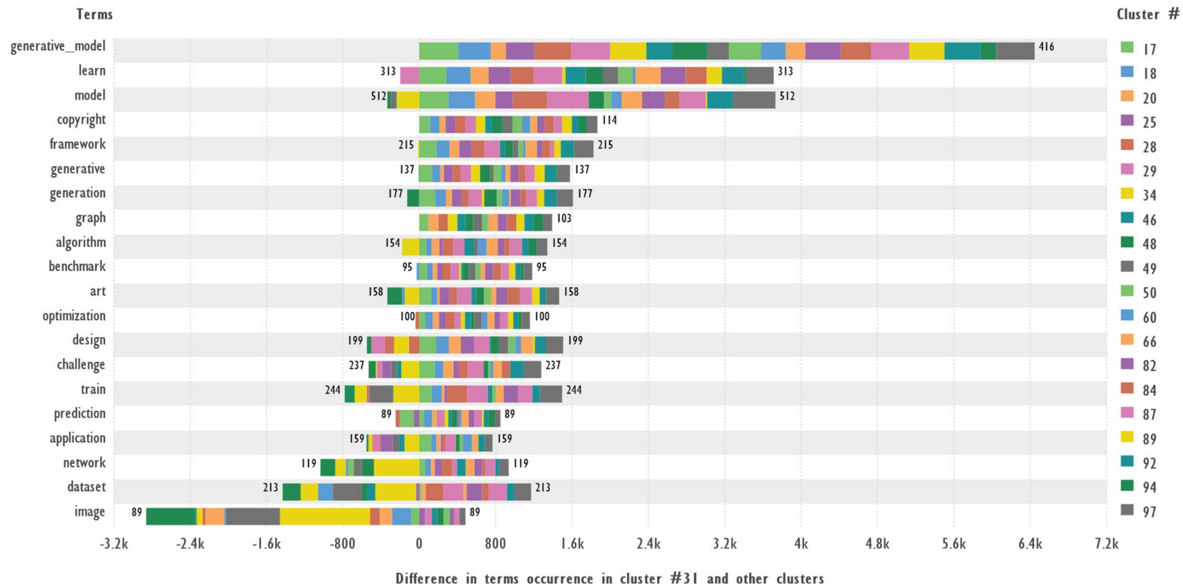


Figure 7. Stacked histogram of difference in term occurrence for cluster #31.

Examples of publications that contain the top terms of this chart and have a high score of belonging to this cluster.

“Structure-Based Drug Design with a Deep Hierarchical Generative Model” [30], “GRELinker: A Graph-Based Generative Model for Molecular Linker Design with Reinforcement and Curriculum Learning [31].

The problem of forming a query using the terms of this cluster was that the OnePetro search engine, when indicating the use of the exact meaning of the term “generative model” in the title of publications, also produced the term “generalized model”, which is probably due to the fact that the system uses the ’stemming generative→genera’ and ’generalized→genera’. Therefore, we manually selected two publications that somehow correspond to the terms describing this cluster.

“Generative Models for Production Forecasting in Unconventional Oil and Gas Plays” [32] and “Analyzing X-Ray CT Images from Unconventional Reservoirs Using Deep Generative Models” [33]. Subjects: Formation Evaluation & Management, Unconventional and Complex Reservoirs.

Figure 8 shows the term occurrence diagram for cluster #34 containing 1283 documents and 640 terms obtained by GSDMM algorithm.

 doi:10.20944/preprints202501.1294.v1

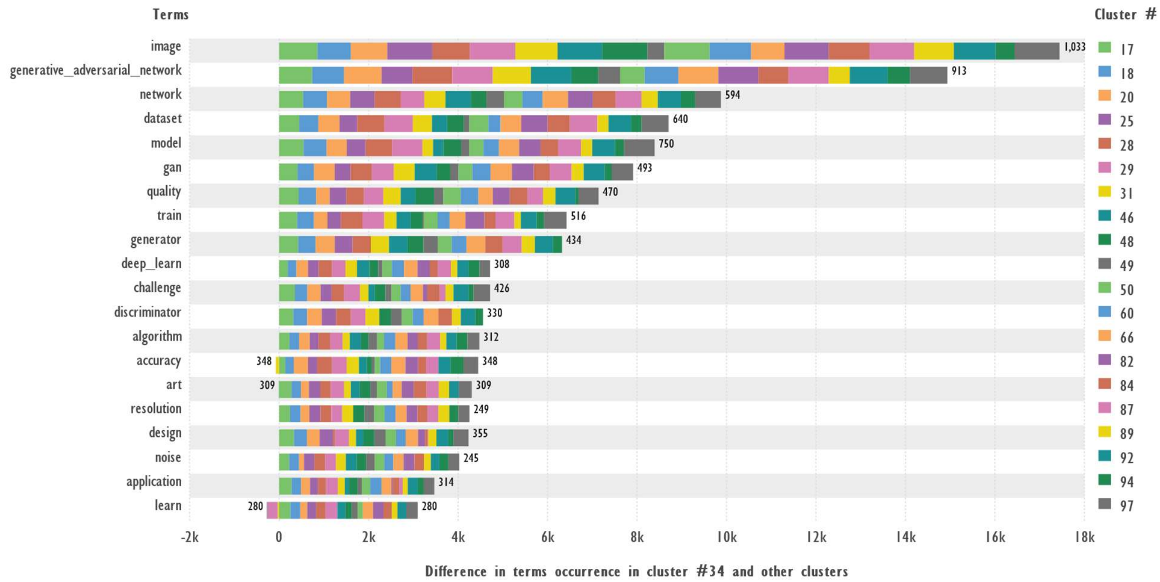


Figure 8. Stacked histogram of difference in term occurrence for cluster #34.

Examples of publications that contain the top terms of this chart and have a high score of belonging to this cluster.

“DAE-GAN: Underwater Image Super-Resolution Based on Symmetric Degradation Attention Enhanced Generative Adversarial Network” [34], “Trans-CNN GAN: Self-Attention Generative Adversarial Networkd for Remote Sensing Image Super-Resolution” [35].

According to the query to the OnePetro platform: ’image “generative adversarial network” dataset model quality’ — 135 Search Results obtained. Given the general nature of the subject matter of this cluster and the large output result, the following are examples of publications subjectively selected by the author of this paper. “Predicting Ground Surface Deformation Induced by Pressurized Fractures Using Conditional Generative Adversarial Networks” [36], “Advancing Digital Rock Imaging with Generative Adversarial Networks” [37], “On Raster Image Segmentation, Well Log Instance Detection, and Depth Information Extraction from Rasters Using Deep Learning Models” [38]. Subjects: Hydraulic Fracturing, Reservoir Characterization, Reservoir Simulation, Flow in porous media.

Figure 9 shows the term occurrence diagram for cluster #46 containing 847 documents and 559 terms obtained by GSDMM algorithm.

 doi:10.20944/preprints202501.1294.v1

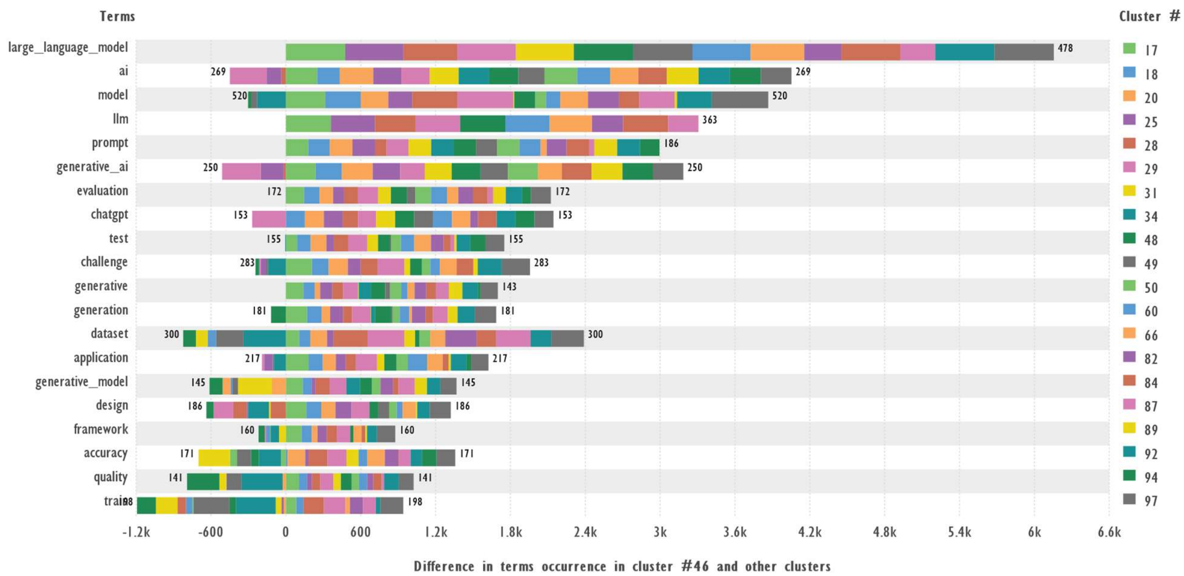


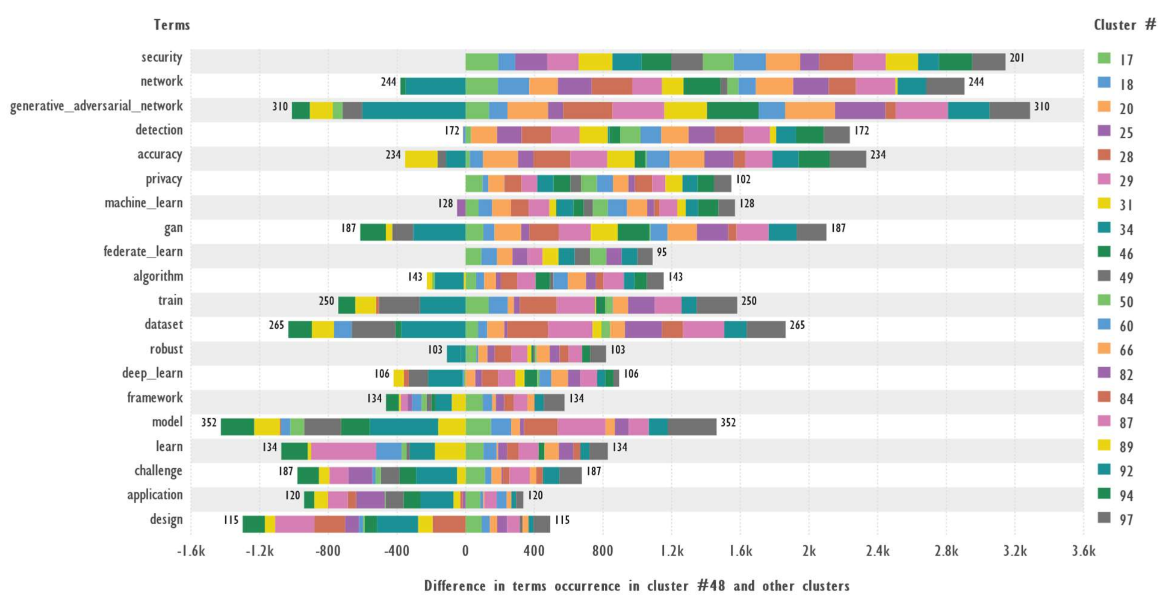
Figure 9. Stacked histogram of difference in term occurrence for cluster #46.

Examples of publications that contain the top terms of this chart and have a high score of belonging to this cluster.

"Using Generative Large Language Models for Hierarchical Relationship Prediction in Medical Ontologies" [39], "Detection of AI-Generated Text Using Large Language Model" [40].

According to the query to the OnePetro platform: '"large language model" prompt " generative ai" evaluation' — 119 Search Results obtained. The following works was most relevant to this cluster: "Empowering Drilling and Optimization with Generative AI" [41], "The Role of Personalized Generative AI in Advancing Petroleum Engineering and Energy Industry: A Roadmap to Secure and Cost-Efficient Knowledge Integration: A Case Study" [15], "Revolutionizing Drilling Operations: Next-gen LLM-AI for Real-time Support in Well Construction Control Rooms" [42]. Subjects: Professionalism, Training, and Education, Knowledge management, Data security.

Figure 10 shows the term occurrence diagram for cluster #48 containing 481 documents and 558 terms obtained by GSDMM algorithm.



Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 January 2025 doi:10.20944/preprints202501.1294.v1

Figure 10. Stacked histogram of difference in term occurrence for cluster #48.

Examples of publications that contain the top terms of this chart and have a high score of belonging to this cluster.

"SYN-GAN: A robust intrusion detection system using GAN-based synthetic data for IoT security" [43], "An Adaptive Federated Learning Intrusion Detection System Based on Generative Adversarial Networks under the Internet of Things" [44].

The query to the OnePetro platform: 'security "generative adversarial network" detection privacy "federate learn"' — yielded no results, the term "federate learn" is rarely used, when removing it from the query we get three results: the work [15] was previously cited, the other two works are: "Optimizing Pipeline Systems for Greater Precision, Efficiency & Safety Using Emerging Technologies" [45], "Detecting Anomalies in Water Quality Monitoring Using Deep Learning" [46]. Subjects: Pipelines, Flowlines and Risers, Water use, produced water discharge and disposal.

Figure 11 shows the term occurrence diagram for cluster #49 containing 809 documents and 620 terms obtained by GSDMM algorithm.

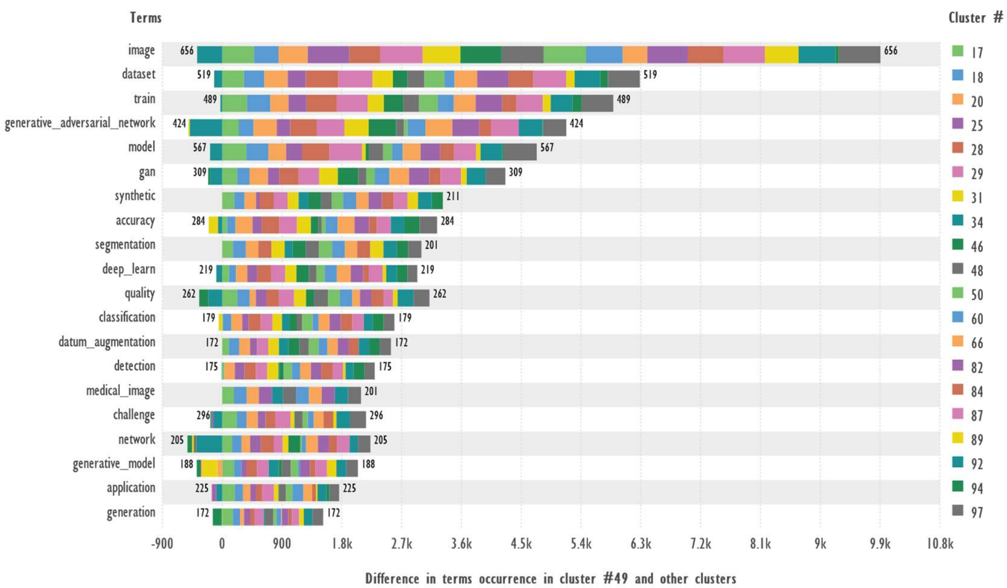


Figure 11. Stacked histogram of difference in term occurrence for cluster #49.

Examples of publications that contain the top terms of this chart and have a high score of belonging to this cluster.

"Synthetic MRI in action: A novel framework in data augmentation strategies for robust multi-modal brain tumor segmentation" [47], "Multi-modal brain tumor segmentation via conditional synthesis with Fourier domain adaptation" [48].

This cluster is similar to cluster 34 in its main terms, so we will focus on terms more typical for this cluster, but keeping the general context: 'image "generative adversarial network"' plus using distinctive terms: '"synthetic segmentation classification "data augmentation"'.

Examples of works that respond to such a synthetic query, found in OnePetro: "Seismic data augmentation for automatic faults picking using deep learning" [49] Subjects: Seismic processing and interpretation, "A Machine Learning-Based Data Augmentation Approach for Unconventional Reservoir Characterization Using Microseismic Data and EDFM" [50]. Subjects: Reservoir Characterization, Unconventional and Complex Reservoirs, Seismic processing and interpretation.

Figure 12 shows the term occurrence diagram for cluster #50 containing 619 documents and 553 terms obtained by GSDMM algorithm.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 January 2025 doi:10.20944/preprints202501.1294.v1

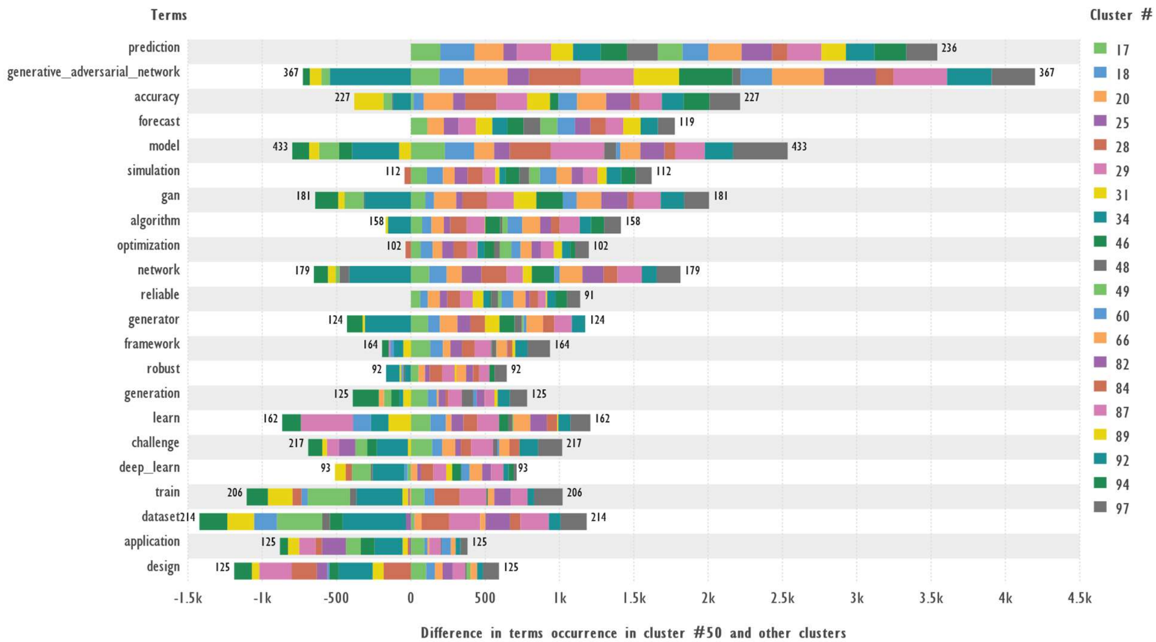


Figure 12. Stacked histogram of difference in term occurrence for cluster #50.

Examples of publications that contain the top terms of this chart and have a high score of belonging to this cluster.

“Networked Time-series Prediction with Incomplete Data via Generative Adversarial Network” [51], “Load forecasting based on dynamic adaptive and adversarial graph convolutional networks” [52].

According to the query to the OnePetro platform: ’prediction “generative adversarial network” forecast simulation reliable’ — 53 Search Results obtained. Subjective selection of publications relevant to the given query: “Towards Universal Production Forecasting via Adversarial Transfer Learning and Transformer with Application in the Shengli Oilfield, China” [53], Subjects: Unconventional and Complex Reservoirs, Information Management and Systems, Production forecasting. “Fracture Morphology Evaluation of Surrogate Model for Fast Prediction Using Machine Learning” [54],

Figure 13 shows the term occurrence diagram for cluster #60 containing 549 documents and 529 terms obtained by GSDMM algorithm.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 January 2025 doi:10.20944/preprints202501.1294.v1

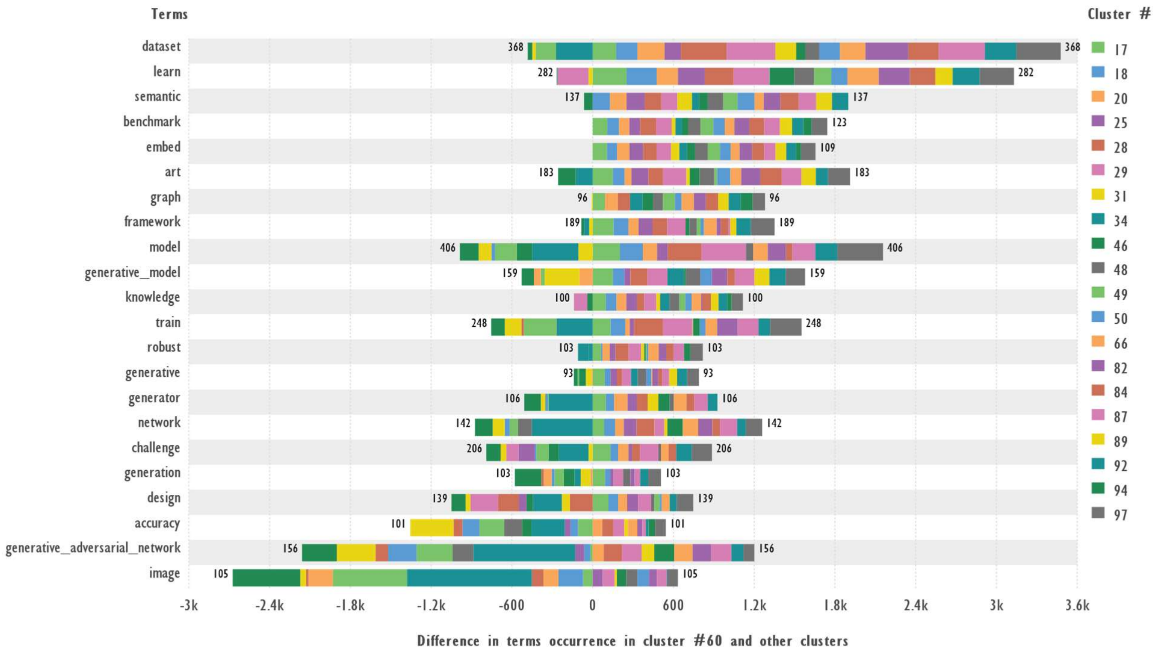


Figure 13. Stacked histogram of difference in term occurrence for cluster #60.

Examples of publications that contain the top terms of this chart and have a high score of belonging to this cluster.

“A semi-supervised learning approach for semantic parsing boosted by BERT word embedding” [55], “External Knowledge Enhancing Meta-learning Framework for Few-Shot Text Classification via Contrastive Learning and Adversarial Network” [56].

Using the terms of this cluster: ’learn semantic embed graph knowledge’, and context ‘“generative model”‘, on request to OnePetro we get a publication — “Using Generative AI to Build a Reservoir Simulation Assistant” [57], Subject: Reservoir Simulation.

Figure 14 shows the term occurrence diagram for cluster #66 containing 415 documents and 467 terms obtained by GSDMM algorithm.

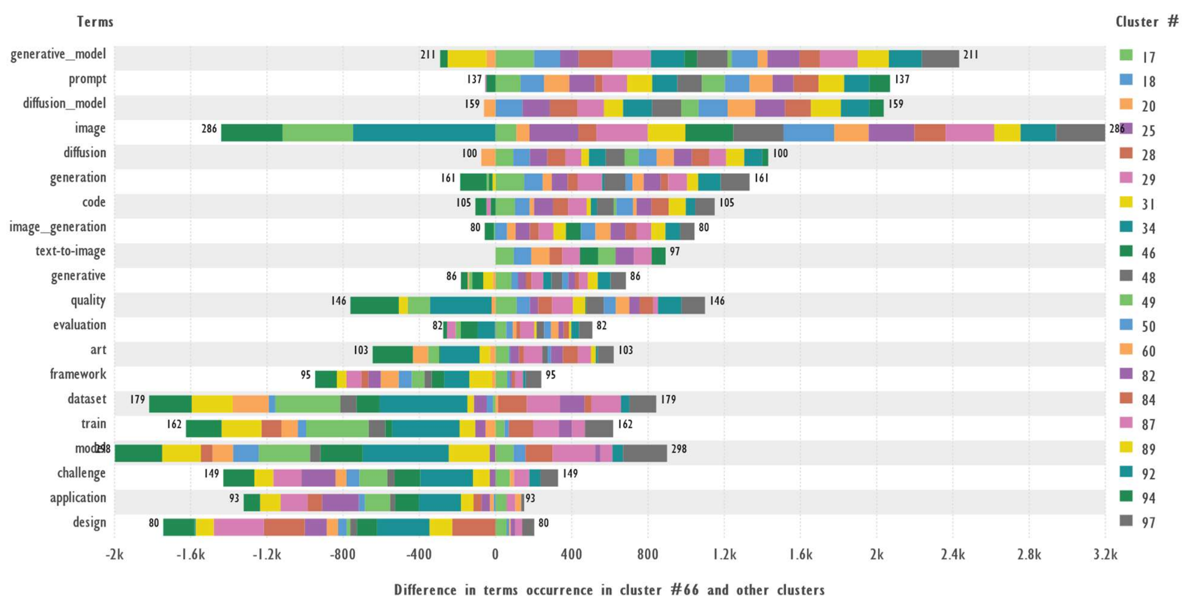


Figure 14. Stacked histogram of difference in term occurrence for cluster #66.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 January 2025 doi:10.20944/preprints202501.1294.v1

Examples of publications that contain the top terms of this chart and have a high score of belonging to this cluster.

"Prompt Optimizer of Text-to-Image Diffusion Models for Abstract Concept Understanding" [58], "Diffusion-Geo: A Two-Stage Controllable Text-To-Image Generative Model for Remote Sensing Scenarios" [59].

The query to the OnePetro platform: '"generative model" "diffusion model" "image generation" "text to mage"' not finding any publications. Excluding "text to image" yields 6 publications, examples of which are: "Enhancing the Resolution of Micro-CT Images of Rock Samples via Unsupervised Machine Learning based on a Diffusion Model" [60], Subject: Reservoir Characterization. "On-The-Fly History Matching of Simulation Models Using Generative Diffusive Learning (SimGDL)" [61], Subjects: Reservoir Simulation, History matching.

Figure 15 shows the term occurrence diagram for cluster #82 containing 799 documents and 539 terms obtained by GSDMM algorithm.

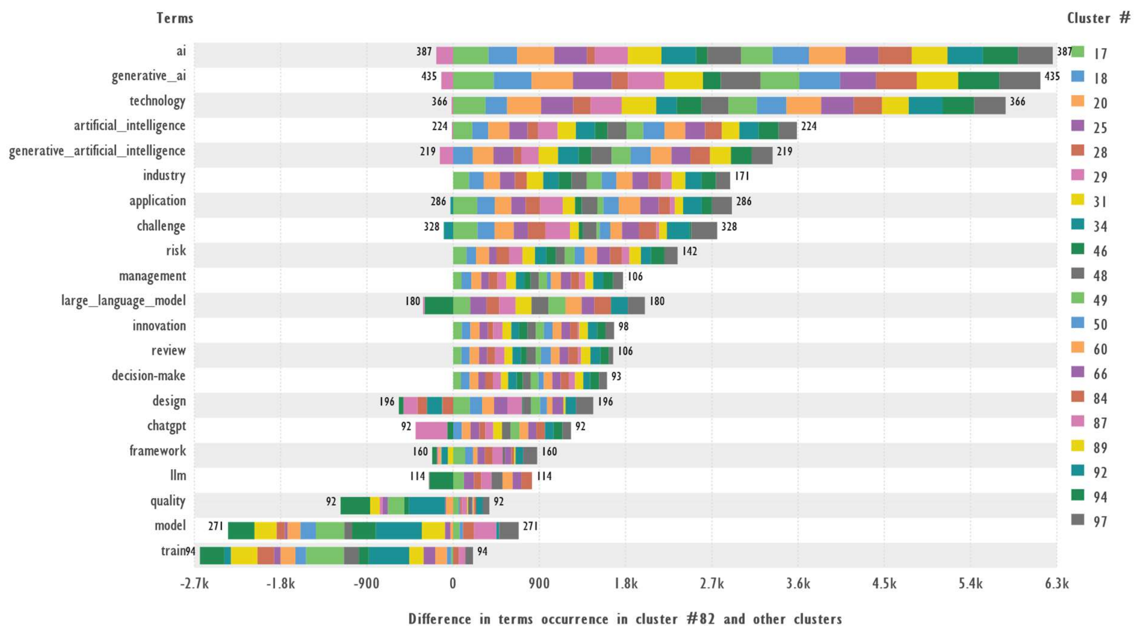


Figure 15. Stacked histogram of difference in term occurrence for cluster #82.

Examples of publications that contain the top terms of this chart and have a high score of belonging to this cluster.

"Generative AI for Industry 5.0: Analyzing the impact of ChatGPT, DALLE, and Other Models" [62], "Generative AI for Healthcare Engineering and Technology Challenges" [63].

A query to OnePetro: 'technology "generative ai" industry risk management innovation' returns 54 results, example publications are: "In-House Integrated Big Data Management Platform for Exploration and Production Operations Digitalization: From Data Gathering to Generative AI through Machine Learning Implementation Using Cost-Effective Open-Source Technologies - Experienced Mature Workflow" [64]. "Innovating Oil and Gas Forecasting: Developing a Trailblazing Generative AI Model" [65]. "Innovative Approach of Generative AI for Automating Technical Bid Evaluations in Oil Companies" [66]. All three works relate to Subject: Information Management and Systems.

Figure 16 shows the term occurrence diagram for cluster #84 containing 524 documents and 482 terms obtained by GSDMM algorithm.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 January 2025 doi:10.20944/preprints202501.1294.v1

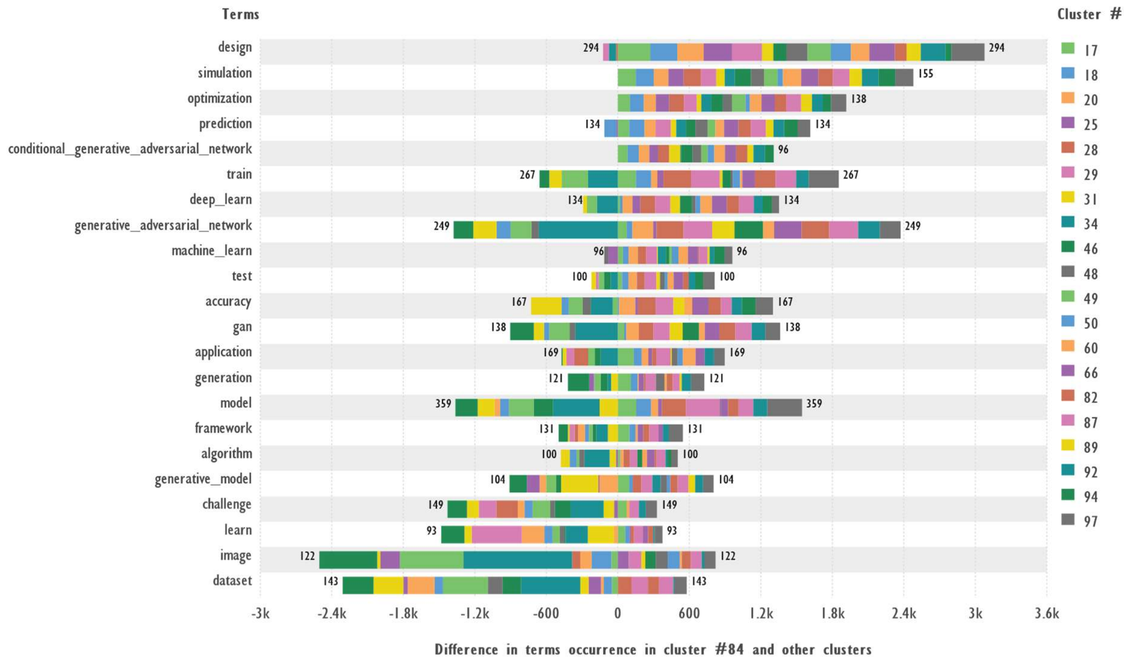


Figure 16. Stacked histogram of difference in term occurrence for cluster #84.

Examples of publications that contain the top terms of this chart and have a high score of belonging to this cluster.

“A conditional generative model for end-to-end stress field prediction of composite bolted joints” [67], “Conditional Generative Adversarial Network Enabled Localized Stress Recovery of Periodic Composites” [68].

Among the terms in this cluster: ’design simulation optimization “conditional generative adversarial network”‘, conditional_generative_adversarial_network is the most interesting term, the rest are those that are simply more frequent in this cluster compared to a number of other clusters. Samples of articles posted on OnePetro: “A new P-wave reconstruction method for VSP data using conditional generative adversarial network” [69] Subject: Seismic processing and interpretation. “Fracture Morphology Evaluation of Surrogate Model for Fast Prediction Using Machine Learning”. Subject: Hydraulic Fracturing. “Seismic data interpolation with conditional generative adversarial network in time and frequency domain” [70], Subject: Seismic processing and interpretation. “A novel deep learning-assisted reservoir fracture delineation with conditional generative adversarial networks” [71], Subjects: Reservoir Characterization, Faults and fracture characterization.

Figure 17 shows the term occurrence diagram for cluster #87 containing 1064 documents and 465 terms obtained by GSDMM algorithm.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 January 2025 doi:10.20944/preprints202501.1294.v1

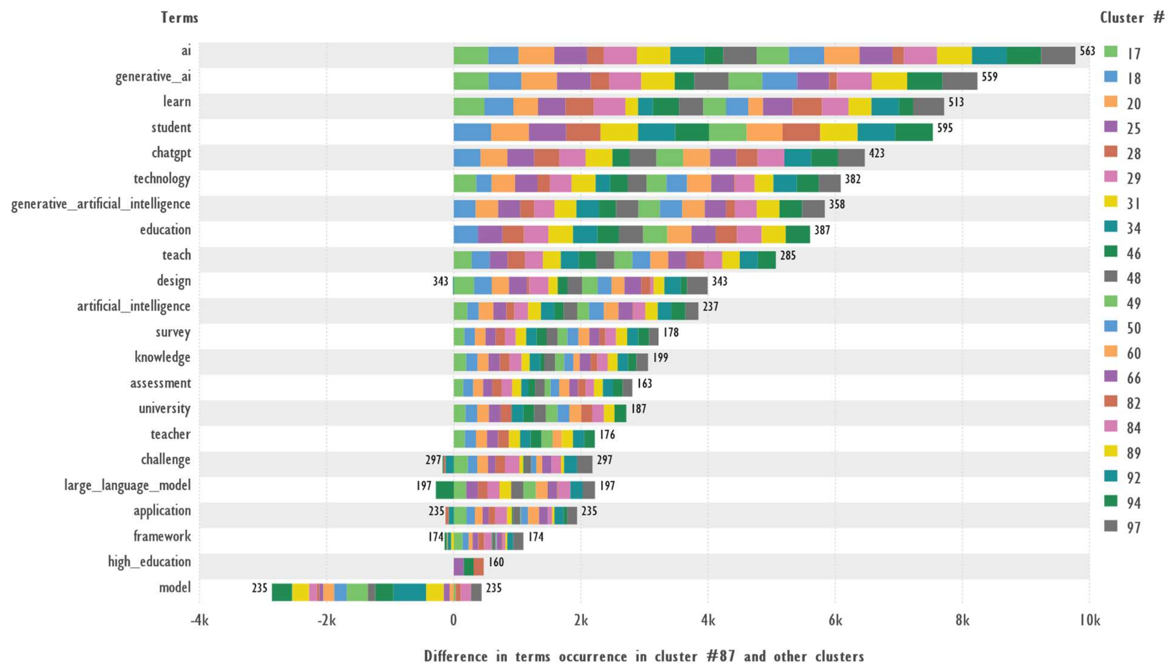


Figure 17. Stacked histogram of difference in term occurrence for cluster #87.

Examples of publications that contain the top terms of this chart and have a high score of belonging to this cluster.

"Student Perceptions of Generative Artificial Intelligence: Investigating Utilization, Benefits, and Challenges in Higher Education" [72], "The rapid rise of generative AI and its implications for academic integrity: Students' perceptions and use of chatbots for assistance with assessments" [73].

The terms: '"generative ai" student technology education teach' indicate that this cluster refers explicitly to the use of 'generative ai' in the teaching process. The term "technology" reveals the subject matter of teaching. Publications found in OnePetro: "Unleashing the Power of Generative AI and LLM for Training Evaluation" [74], Subjects: Information Management and Systems, Artificial intelligence. "State of the Art of Artificial Intelligence and Predictive Analytics in the E&P Industry: A Technology Survey" [75], Subjects: Artificial intelligence, Information Management and Systems. The latter study is not about learning, but rather about what topics are appropriate to focus on in the learning process.

Figure 18 shows the term occurrence diagram for cluster #89 containing 712 documents and 607 terms obtained by GSDMM algorithm.

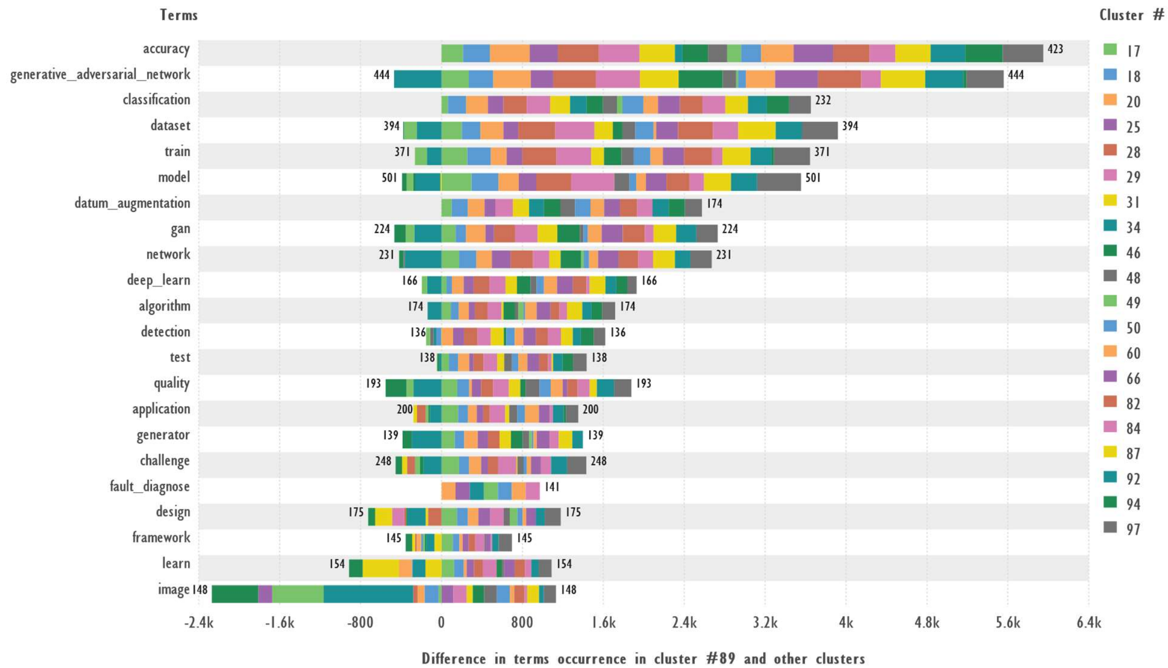


Figure 18. Stacked histogram of difference in term occurrence for cluster #89.

Examples of publications that contain the top terms of this chart and have a high score of belonging to this cluster.

“Generative AI with WGAN-GP for boosting seizure detection accuracy” [76], “Classification of multi-type bearing fault features based on semi-supervised generative adversarial network (GAN)” [77].

Terms: ’accuracy “generative adversarial network” classification “ data augmentation”‘ — highlight the importance of classification accuracy when using “generative adversarial network” — this is perhaps the most interesting aspect of this cluster. “Super-Resolution Reconstruction of Reservoir Saturation Map with Physical Constraints Using Generative Adversarial Network” [78], Subjects: Fluid Characterization, Reservoir Simulation. “ConGANergy: A Framework for Engineering Data Augmentation with Application to Solid Particle Erosion” [79], Subjects: Well & Reservoir Surveillance and Monitoring.

Figure 19 shows the term occurrence diagram for cluster #92 containing 347 documents and 442 terms obtained by GSDMM algorithm.

 doi:10.20944/preprints202501.1294.v1

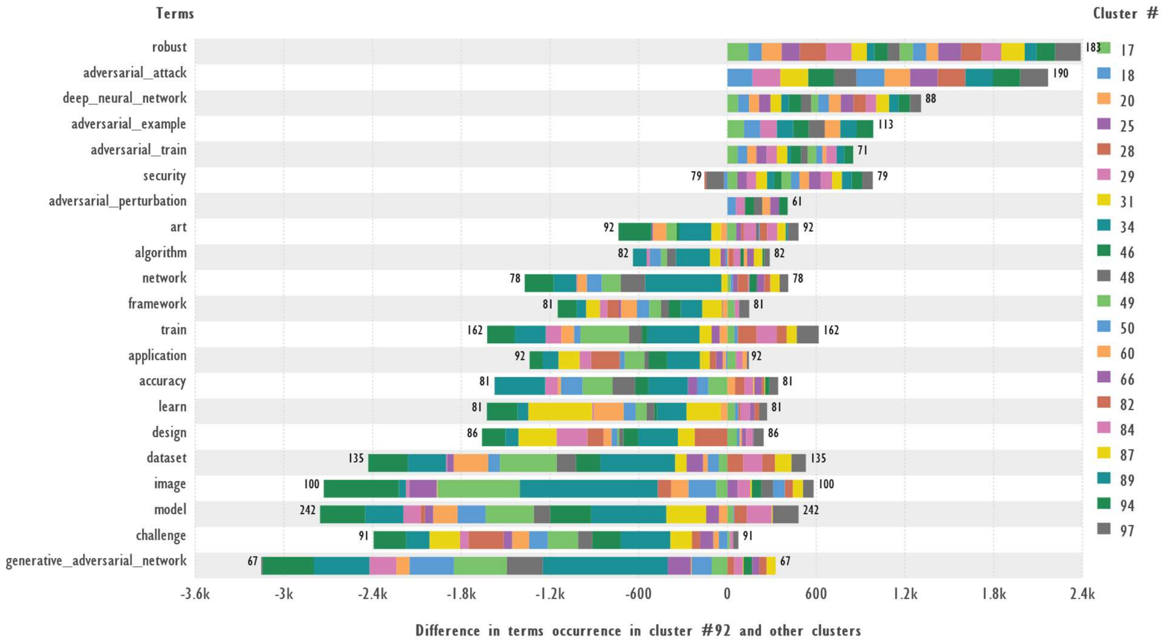


Figure 19. Stacked histogram of difference in term occurrence for cluster #92.

Examples of publications that contain the top terms of this chart and have a high score of belonging to this cluster.

"Comprehensive Botnet Detection by Mitigating Adversarial Attacks, Navigating the Subtleties of Perturbation Distances and Fortifying Predictions with Conformal Layers" [80], "Minimizing Adversarial Training Samples for Robust Image Classifiers: Analysis and Adversarial Example Generator Design" [81].

In OnePetro, for the query: '"adversarial attack'' ''adversarial example'' ''adversarial train'' ''adversarial perturbation''' there are no results, for the shorter '"adversarial attack'' ''adversarial perturbation''' there are no results either.

Technically, OnePetro produces 4 results for the query "adversarial attack", but none of them have the term in the title or abstract. For example, a review-type article "The Impact of Artificial Intelligence on Cybersecurity" [82] contains the term "adversarial attack" in the full text but not in the abstract. Unlike OnePetro, querying arxiv.org for ''adversarial attack'' in Abstract returns 4,833 results. Moreover, turning to the open and very high quality IEEE Xplore abstract database, even in the titles of publications for the query ''Document Title: ''Adversarial Attack''' we get 474 results — Conferences (324), Journals (137). The "adversarial attack" theme, widespread in other industries, has significant potential to investigate AI adoption in the petroleum sector and is recommended as a promising research topic due to its broad applicability.

General Note. There is now a tendency to create large "introductory" sections; they generate a large proportion of references in an article and contain a large variety of terms. On the one hand, this approach indicates that the authors have read a large number of publications before starting the paper; on the other hand, they lead to "bloated" citations. This reduces the value of citations as an assessment of the subject of the paper, since they are more relevant to the text of the Introduction section than to the Materials and Methods and Results and Discussion sections. The same can be said about the classification of publications by topic - many terms used in the classification are taken from the text of the Introduction, which blurs the classification used. Moreover, while the "bibliographic coupling" method previously yielded good results for similarity of research topics, the current exaggerated use of citations is likely to indicate similarity of the Introduction sections rather than similarity of the research topics themselves.

 doi:10.20944/preprints202501.1294.v1

Figure 20 shows the term occurrence diagram for cluster #94 containing 898 documents and 556 terms obtained by GSDMM algorithm.

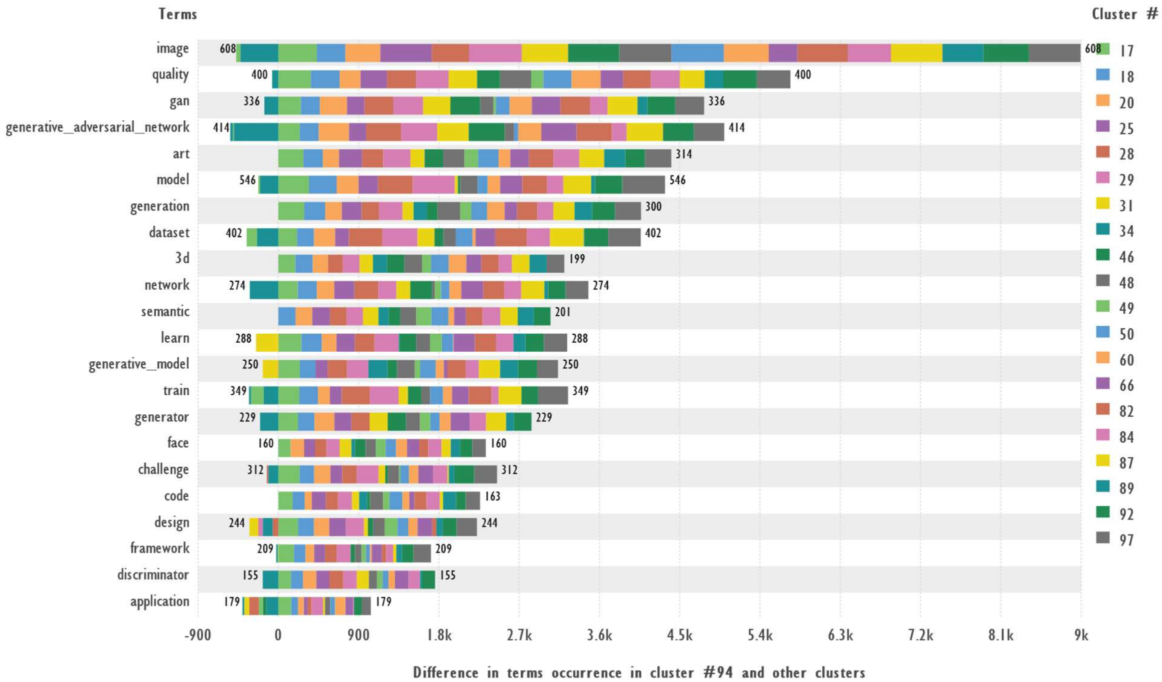


Figure 20. Stacked histogram of difference in term occurrence for cluster #94.

Examples of publications that contain the top terms of this chart and have a high score of belonging to this cluster.

“Gaussian Splatting Decoder for 3D-aware Generative Adversarial Networks” [83], “DepthGAN: GAN-based depth generation from semantic layouts” [84].

The terms ’image’ and ’generative adversarial network’ have already been presented in detail in clusters 34 and 49, so let’s focus on the distinguishing terms: ’3d semantic’. The most relevant publications from OnePetro: “Generating Geologically Realistic 3D Reservoir Facies Models Using Deep Learning of Sedimentary Architecture with Generative Adversarial Networks” [85], Subjects: Reservoir Characterization, Geologic modeling. “2D-to-3D reconstruction of carbonate digital rocks using Progressive Growing GAN” [86], Subject: Reservoir Characterization. Compared to ’semantic’, which only appears in the full text of articles but not in the title and abstract, ’3d’ is a dominant term.

Figure 21 shows the term occurrence diagram for cluster #97 containing 123 documents and 385 terms obtained by GSDMM algorithm.

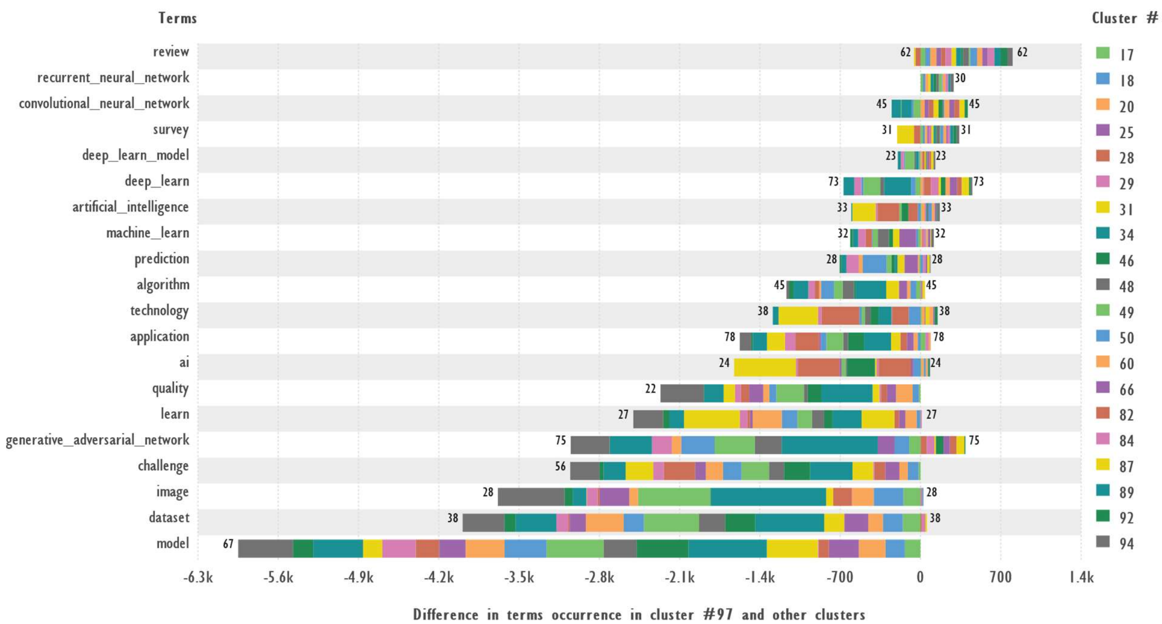


Figure 21. Stacked histogram of difference in term occurrence for cluster #97.

This cluster is related to 'review' and 'survey', the terms in this cluster are not unique, because in reviews there are terms typical for other clusters.

Examples of publications that contain the top terms of this chart and have a high score of belonging to this cluster.

"Exploring the Frontiers of Unsupervised Learning Techniques for Diagnosis of Cardiovascular Disorder: A Systematic Review" [87], "Comprehensive Literature Survey on Deep Learning Used in Image Memorability Prediction and Modification" [88].

Note: The total number of publications assigned to the 21 clusters was: 273+428+449+413+543+118+730+847+481+809+619+549+415+799+524+712+347+898+123=10077 out of 12424 total document list.

Note: There are interactive web pages in the attached archive for a better understanding of the results presented in the figures.

Conclusions

The clustering of the publications using the GSDMM algorithm applied to the titles and abstracts of the bibliometric records exported from Scopus resulted in 21 clusters.

The publications most frequently posted on OnePetro that are relevant to the topic under consideration are Subjects: information management and systems, reservoir characterization, seismic processing and interpretation, reservoir simulation, artificial intelligence, unconventional and complex reservoirs, well & reservoir surveillance and monitoring, data security, hydraulic fracturing.

Adversarial attacks pose a significant threat to generative models by exploiting vulnerabilities to manipulate input data and deceive models to produce incorrect predictions or classifications. This topic is underrepresented in petroleum publications, but it has significant research potential because it is well reported in publications from other subject areas. Analyzing the robustness of generative models to "adversarial attack" and "adversarial perturbation" can be recommended as a promising research topic.

The results of this study may help petroleum industry professionals to expand their search for publications on generative models to gain knowledge about them from a broader field of research.

References

[1] The state of AI in early 2024 | McKinsey n.d. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai (accessed January 8, 2025).
[2] Gen AI adoption: The next inflection point | McKinsey n.d. https://www.mckinsey.com/capabilities/people-and-organizational-performance/our-insights/gen-ais-next-inflection-point-from-employee-experimentation-to-organizational-transformation (accessed January 8, 2025).
[3] Zhang P, Shi J, Kamel Boulos MN. Generative AI in Medicine and Healthcare: Moving Beyond the 'Peak of Inflated Expectations.' Future Internet 2024;16:462. https://doi.org/10.3390/fi16120462.
[4] Fijačko N, Creber RM, Abella BS, Kocbek P, Metličar Š, Greif R, et al. Using generative artificial intelligence in bibliometric analysis: 10 years of research trends from the European Resuscitation Congresses. Resuscitation Plus 2024;18:100584. https://doi.org/10.1016/j.resplu.2024.100584.
[5] Van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics 2010;84:523–38. https://doi.org/10.1007/s11192-009-0146-3.
[6] Yin J, Wang J. A dirichlet multinomial mixture model-based approach for short text clustering. Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining, New York New York USA: ACM; 2014, p. 233–42. https://doi.org/10.1145/2623330.2623715.
[7] Li L. The Study on Food Safety of 15 'RCEP' Countries: Based on VOSviewer and Scimago Graphica. Science & Technology Libraries 2024;43:147–54. https://doi.org/10.1080/0194262X.2023.2237560.
[8] Krishnan SR, Sim EP, Varghese CM, Rajan B, S C, Thomas E. Detection of Diseases in Tomato Leaves Using Deep Learning Models: A Survey. 2024 1st International Conference on Trends in Engineering Systems and Technologies (ICTEST), Kochi, India: IEEE; 2024, p. 1–6. https://doi.org/10.1109/ICTEST60614.2024.10576182.
[9] Venkata Rami Reddy G, Niranjan A. War Snake Optimisation Algorithm with deep Q-Net for COVID-19 classification. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization 2024;11:2245925. https://doi.org/10.1080/21681163.2023.2245925.
[10] Yin C, Zhang K, Zhang L, Wang Z, Liu P, Zhang H, et al. Imbalanced Working States Recognition of Sucker Rod Well Dynamometer Cards Based on Data Generation and Diversity Augmentation. SPE Journal 2023;28:1925–44. https://doi.org/10.2118/214661-PA.
[11] Ayub M, Kaka SI. Automated Hyperparameter Optimization of Convolutional Neural Network (CNN) for First-Break (FB) Arrival Picking. Day 3 Wed, March 15, 2023, Dubai, UAE: SPE; 2023, p. D031S036R002. https://doi.org/10.2118/214253-MS.
[12] Barbadekar A, Sole S, Shekhavat A. Enhancing Social Media Security: LSTM-Based Deep Fake Video Detection. 2024 IEEE 9th International Conference for Convergence in Technology (I2CT), Pune, India: IEEE; 2024, p. 1–6. https://doi.org/10.1109/I2CT61223.2024.10543604.
[13] Kumar M, Rai PK, Kumar P. A Novel Approach for Detecting Deepfake Face Using Machine Learning Algorithms. 2024 2nd International Conference on Disruptive Technologies (ICDT), Greater Noida, India: IEEE; 2024, p. 1588–92. https://doi.org/10.1109/ICDT61202.2024.10489036.
[14] Aldossary RS, Almutairi MN, Alotaibi NM, Serkan D. Personal Protective Equipment Detection Using Computer Vision Techniques. Day 2 Tue, October 03, 2023, Abu Dhabi, UAE: SPE; 2023, p. D021S053R005. https://doi.org/10.2118/216253-MS.
[15] Gharieb A, Gabry MA, Soliman MY. The Role of Personalized Generative AI in Advancing Petroleum Engineering and Energy Industry: A Roadmap to Secure and Cost-Efficient Knowledge Integration: A Case Study. SPE Annual Technical Conference and Exhibition, New Orleans, Louisiana, USA: SPE; 2024, p. D011S007R002. https://doi.org/10.2118/220716-MS.
[16] Wang H, Zhai D, Zhou X, Jiang J, Liu X. Mix-DDPM: Enhancing Diffusion Models through Fitting Mixture Noise with Global Stochastic Offset. ACM Trans Multimedia Comput Commun Appl 2024;20:1–24. https://doi.org/10.1145/3672080.
[17] Heurtel-Depeiges D, Margossian CC, Ohana R, Blancard BR-S. Listening to the Noise: Blind Denoising with Gibbs Diffusion 2024. https://doi.org/10.48550/ARXIV.2402.19455.
[18] Wang L, Zhang L, Deng R, Wang H, Zhao X, Xu B. Integrating Multi-Source Experiment Data and Variational Diffusion Model for Intelligent Constructing Digital Core: A Case Study of Sandstone Reservoir

in the Turgay Basin Central South Kazakhstanc. SPE Caspian Technical Conference and Exhibition, Atyrau, Kazakhstan: SPE; 2024. https://doi.org/10.2118/223464-MS.
[19] Jiang F, Osypov K, Toms J. Implementation of denoising diffusion probability model for seismic interpretation. Third International Meeting for Applied Geoscience & Energy Expanded Abstracts, Houston, Texas: Society of Exploration Geophysicists and American Association of Petroleum Geologists; 2023, p. 1098–102. https://doi.org/10.1190/image2023-3907375.1.
[20] Meng H, Lin B, Zhang R, Jin Y. A Missing Well-Logs Imputation Method Based on Conditional Denoising Diffusion Probabilistic Models. SPE Journal 2024;29:2165–80. https://doi.org/10.2118/219452-PA.
[21] Yousuf MF, Mahmud MS. Generating Synthetic Time-Series Data on Edge Devices Using Generative Adversarial Networks. 2024 International Conference on Computing, Networking and Communications (ICNC), Big Island, HI, USA: IEEE; 2024, p. 441–5. https://doi.org/10.1109/ICNC59896.2024.10556140.
[22] Bhat KA, Sofi SA. From shallows to depths: unveiling hybrid synthetic data modeling for enhanced learning with privacy considerations in naturally imbalanced datasets. International Journal of Computers and Applications 2024;46:1088–103. https://doi.org/10.1080/1206212X.2024.2409989.
[23] Hassani H, Shahbazi A, Shahbalayev E, Hamdi Z, Behjat S, Bataee M. Machine Learning-Based CO2 Saturation Tracking in Saline Aquifers Using Bottomhole Pressure for Carbon Capture and Storage CCS Projects. Day 1 Wed, April 17, 2024, Bergen, Norway: SPE; 2024, p. D011S009R002. https://doi.org/10.2118/218445-MS.
[24] Huang R, Lin H, Chen C, Zhang K, Zeng W. PlantoGraphy: Incorporating Iterative Design Process into Generative Artificial Intelligence for Landscape Rendering. Proceedings of the CHI Conference on Human Factors in Computing Systems, Honolulu HI USA: ACM; 2024, p. 1–19. https://doi.org/10.1145/3613904.3642824.
[25] Liu X, Xu Y. Exploring the application path of AIGC technology in the styling design of traditional artifacts: a case study of Song dynasty lacquerware. In: Yin F, Zhan Z, editors. Third International Conference on Electronics Technology and Artificial Intelligence (ETAI 2024), Guangzhou, China: SPIE; 2024, p. 38. https://doi.org/10.1117/12.3045170.
[26] Sabbagh VB, Lima CBC, Xexéo G. Comparative Analysis of Single and Multiagent Large Language Model Architectures for Domain-Specific Tasks in Well Construction. SPE Journal 2024;29:6869–82. https://doi.org/10.2118/223612-PA.
[27] Akram J, Aamir M, Raut R, Anaissi A, Jhaveri RH, Akram A. AI-Generated Content-as-a-Service in IoMT-Based Smart Homes: Personalizing Patient Care With Human Digital Twins. IEEE Trans Consumer Electron 2024:1–1. https://doi.org/10.1109/TCE.2024.3409173.
[28] Liu G, Du H, Niyato D, Kang J, Xiong Z, Kim DI, et al. Semantic Communications for Artificial Intelligence Generated Content (AIGC) Toward Effective Content Creation. IEEE Network 2024;38:295–303. https://doi.org/10.1109/MNET.2024.3352917.
[29] Singh A, Jia T, Nalagatla V. Generative AI Enabled Conversational Chatbot for Drilling and Production Analytics. Day 2 Tue, October 03, 2023, Abu Dhabi, UAE: SPE; 2023, p. D021S065R002. https://doi.org/10.2118/216267-MS.
[30] Weller JA, Rohs R. Structure-Based Drug Design with a Deep Hierarchical Generative Model. J Chem Inf Model 2024;64:6450–63. https://doi.org/10.1021/acs.jcim.4c01193.
[31] Zhang H, Huang J, Xie J, Huang W, Yang Y, Xu M, et al. GRELinker: A Graph-Based Generative Model for Molecular Linker Design with Reinforcement and Curriculum Learning. J Chem Inf Model 2024;64:666–76. https://doi.org/10.1021/acs.jcim.3c01700.
[32] Kuzma HA, Arora NS, Farid K. Generative Models for Production Forecasting in Unconventional Oil and Gas Plays. Proceedings of the 2nd Unconventional Resources Technology Conference, Denver, Colorado, USA: American Association of Petroleum Geologists; 2014. https://doi.org/10.15530/urtec-2014-1928595.
[33] Perez Claro Y, Dal Santo N, Krishnan V, Kovscek A. Analyzing X-Ray CT Images from Unconventional Reservoirs Using Deep Generative Models. SPE Western Regional Meeting, Bakersfield, California, USA: SPE; 2022, p. D021S012R003. https://doi.org/10.2118/209280-MS.

[34] Gao M, Li Z, Wang Q, Fan W. DAE-GAN: Underwater Image Super-Resolution Based on Symmetric Degradation Attention Enhanced Generative Adversarial Network. Symmetry 2024;16:588. https://doi.org/10.3390/sym16050588.

[35] Lin Z, Lin B, Ye W, Liu Y. Trans-CNN GAN: Self-Attention Generative Adversarial Networkd for Remote Sensing Image Super-Resolution. 2024 4th International Conference on Neural Networks, Information and Communication (NNICE), Guangzhou, China: IEEE; 2024, p. 456–9. https://doi.org/10.1109/NNICE61279.2024.10498876.

[36] Salimzadeh S, Kasperczyk D, Kadeethum T. Predicting Ground Surface Deformation Induced by Pressurized Fractures Using Conditional Generative Adversarial Networks. 57th U.S. Rock Mechanics/Geomechanics Symposium, Atlanta, Georgia, USA: ARMA; 2023, p. ARMA-2023-0218. https://doi.org/10.56952/ARMA-2023-0218.

[37] Temizel C, Odi U, Cetin C, Pamukcu Y. Advancing Digital Rock Imaging with Generative Adversarial Networks. SPE Western Regional Meeting, Palo Alto, California, USA: SPE; 2024, p. D011S001R003. https://doi.org/10.2118/218833-MS.

[38] Gune OA, Lokhande AR, Mangsuli PR. On Raster Image Segmentation, Well Log Instance Detection, and Depth Information Extraction from Rasters Using Deep Learning Models. Day 2 Tue, October 03, 2023, Abu Dhabi, UAE: SPE; 2023, p. D021S053R004. https://doi.org/10.2118/216206-MS.

[39] Liu H, Zhou S, Chen Z, Perl Y, Wang J. Using Generative Large Language Models for Hierarchical Relationship Prediction in Medical Ontologies. 2024 IEEE 12th International Conference on Healthcare Informatics (ICHI), Orlando, FL, USA: IEEE; 2024, p. 248–56. https://doi.org/10.1109/ICHI61247.2024.00040.

[40] Prajapati M, Baliarsingh SK, Dora C, Bhoi A, Hota J, Mohanty JP. Detection of AI-Generated Text Using Large Language Model. 2024 International Conference on Emerging Systems and Intelligent Computing (ESIC), Bhubaneswar, India: IEEE; 2024, p. 735–40. https://doi.org/10.1109/ESIC60604.2024.10481602.

[41] Amour M, Rachmat BK, Rementeria AS, Guillot V, Millan E. Empowering Drilling and Optimization with Generative AI. ADIPEC, Abu Dhabi, UAE: SPE; 2024, p. D011S008R005. https://doi.org/10.2118/221862-MS.

[42] Ferrigno E, Rodriguez M, Davidsson E. Revolutionizing Drilling Operations: Next-gen Llm-AI for Real-time Support in Well Construction Control Rooms. SPE Annual Technical Conference and Exhibition, New Orleans, Louisiana, USA: SPE; 2024, p. D021S017R007. https://doi.org/10.2118/220798-MS.

[43] Rahman S, Pal S, Mittal S, Chawla T, Karmakar C. SYN-GAN: A robust intrusion detection system using GAN-based synthetic data for IoT security. Internet of Things 2024;26:101212. https://doi.org/10.1016/j.iot.2024.101212.

[44] Zhao H, Liu L, Fan F, Zhang H, Ma Y. An Adaptive Federated Learning Intrusion Detection System Based on Generative Adversarial Networks under the Internet of Things. Proceedings of the 2024 3rd Asia Conference on Algorithms, Computing and Machine Learning, Shanghai China: ACM; 2024, p. 1–6. https://doi.org/10.1145/3654823.3654824.

[45] Dickerson P, Worthen J. Optimizing Pipeline Systems for Greater Precision, Efficiency & Safety Using Emerging Technologies 2024.

[46] Tharayil SM, Alomari NK, Bubshait DK. Detecting Anomalies in Water Quality Monitoring Using Deep Learning. Day 2 Wed, March 06, 2024, Abu Dhabi, UAE: SPE; 2024, p. D021S010R004. https://doi.org/10.2118/219049-MS.

[47] Pani K, Chawla I. Synthetic MRI in action: A novel framework in data augmentation strategies for robust multi-modal brain tumor segmentation. Computers in Biology and Medicine 2024;183:109273. https://doi.org/10.1016/j.compbiomed.2024.109273.

[48] Khalil YA, Ayaz A, Lorenz C, Weese J, Pluim J, Breeuwer M. Multi-modal brain tumor segmentation via conditional synthesis with Fourier domain adaptation. Computerized Medical Imaging and Graphics 2024;112:102332. https://doi.org/10.1016/j.compmedimag.2024.102332.

[49] Pham N, Fomel S. Seismic data augmentation for automatic faults picking using deep learning. Second International Meeting for Applied Geoscience & Energy, Houston, Texas: Society of Exploration Geophysicists and American Association of Petroleum Geologists; 2022, p. 1719–24. https://doi.org/10.1190/image2022-3745790.1.

[50] Leines-Artieda JA, Fiallos-Torres MX, Alharthi A, Mahmoud SED, Al Hashmi A, Alqaydi M, et al. A Machine Learning-Based Data Augmentation Approach for Unconventional Reservoir Characterization Using Microseismic Data and EDFM. Day 2 Tue, November 01, 2022, Abu Dhabi, UAE: SPE; 2022, p. D021S062R004. https://doi.org/10.2118/210989-MS.

[51] Zhu Y, Jiang B, Jin H, Zhang M, Gao F, Huang J, et al. Networked Time-series Prediction with Incomplete Data via Generative Adversarial Network. ACM Trans Knowl Discov Data 2024;18:1–25. https://doi.org/10.1145/3643822.

[52] Liu J, Dong S, Zhang P, Li T, Peng C, Hu Z. Load forecasting based on dynamic adaptive and adversarial graph convolutional networks. Energy and Buildings 2024;312:114206. https://doi.org/10.1016/j.enbuild.2024.114206.

[53] Meng J, Xiao Y, Wang H, Ye T, Chang J, Zhang D. Towards Universal Production Forecasting via Adversarial Transfer Learning and Transformer with Application in the Shengli Oilfield, China. The Unconventional Resources Technology Conference, Houston, TX USA: American Association of Petroleum Geologists; 2024. https://doi.org/10.15530/urtec-2024-4032318.

[54] Hu Z, Sheng M, Wei Z, Tian S, Zhou J, Hu S. Fracture Morphology Evaluation of Surrogate Model for Fast Prediction Using Machine Learning. International Geomechanics Symposium, Al Khobar, Saudi Arabia: ARMA; 2023, p. ARMA-IGS-2023-0199. https://doi.org/10.56952/IGS-2023-0199.

[55] Bu Y, Chen T, Duan H, Liu M, Xue Y. A semi-supervised learning approach for semantic parsing boosted by BERT word embedding. IFS 2024;46:6577–88. https://doi.org/10.3233/JIFS-233212.

[56] Sun X, Yang Y, Liu Y. External Knowledge Enhancing Meta-learning Framework for Few-Shot Text Classification via Contrastive Learning and Adversarial Network. In: Zhang W, Tung A, Zheng Z, Yang Z, Wang X, Guo H, editors. Web and Big Data, vol. 14961, Singapore: Springer Nature Singapore; 2024, p. 46–58. https://doi.org/10.1007/978-981-97-7232-2_4.

[57] Wiegand K, Bedewi M, Mukundakrishnan K, Tishechkin D, Ananthan V, Kahn D. Using Generative AI to Build a Reservoir Simulation Assistant. ADIPEC, Abu Dhabi, UAE: SPE; 2024, p. D011S020R001. https://doi.org/10.2118/221987-MS.

[58] Fan Z, Li X, Nag K, Fang C, Biswas T, Xu J, et al. Prompt Optimizer of Text-to-Image Diffusion Models for Abstract Concept Understanding. Companion Proceedings of the ACM Web Conference 2024, Singapore Singapore: ACM; 2024, p. 1530–7. https://doi.org/10.1145/3589335.3651927.

[59] Cai M, Zhang W, Zhang T, Zhuang Y, Chen H, Chen L, et al. Diffusion-Geo: A Two-Stage Controllable Text-To-Image Generative Model for Remote Sensing Scenarios. IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium, Athens, Greece: IEEE; 2024, p. 7003–6. https://doi.org/10.1109/IGARSS53475.2024.10641523.

[60] Ma Z, Sun S, Yan B, Kwak H, Gao J. Enhancing the Resolution of Micro-CT Images of Rock Samples via Unsupervised Machine Learning based on a Diffusion Model. SPE Annual Technical Conference and Exhibition, San Antonio, Texas, USA: SPE; 2023, p. D021S028R005. https://doi.org/10.2118/214883-MS.

[61] Alnasser MH, Maucec M. On-The-Fly History Matching of Simulation Models Using Generative Diffusive Learning (SimGDL). All Days, Dhahran, Saudi Arabia: IPTC; 2024, p. IPTC-23666-EA. https://doi.org/10.2523/IPTC-23666-EA.

[62] Sai S, Sai R, Chamola V. Generative AI for Industry 5.0: Analyzing the impact of ChatGPT, DALLE, and Other Models. IEEE Open J Commun Soc 2024:1–1. https://doi.org/10.1109/OJCOMS.2024.3400161.

[63] Jain S, Subzwari SWA, Subzwari SAA. Generative AI for Healthcare Engineering and Technology Challenges. In: Sharma SK, Dwivedi YK, Metri B, Lal B, Elbanna A, editors. Transfer, Diffusion and Adoption of Next-Generation Digital Technologies, vol. 697, Cham: Springer Nature Switzerland; 2024, p. 68–80. https://doi.org/10.1007/978-3-031-50188-3_7.

[64] Gharieb A, Gabry MA, Elsawy M, Edries T, Mahmoud W, Algarhy A, et al. In-House Integrated Big Data Management Platform for Exploration and Production Operations Digitalization: From Data Gathering to Generative AI through Machine Learning Implementation Using Cost-Effective Open-Source Technologies - Experienced Mature Workflow. Day 2 Tue, April 23, 2024, Muscat, Oman: SPE; 2024, p. D011S011R004. https://doi.org/10.2118/218560-MS.

[65] Sidharth Manikandan S, Sharma A, Pellegrino S. Innovating Oil and Gas Forecasting: Developing a Trailblazing Generative AI Model. The Unconventional Resources Technology Conference, Houston, TX USA: American Association of Petroleum Geologists; 2024. https://doi.org/10.15530/urtec-2024-4044635.
[66] Abdelaziem OE, Khafagy AN, Yehia TA. Innovative Approach of Generative AI for Automating Technical Bid Evaluations in Oil Companies. Mediterranean Offshore Conference, Alexandria, Egypt: SPE; 2024, p. D021S012R003. https://doi.org/10.2118/223359-MS.
[67] Zhao Y, Liu Y, Lin Q, Pan W, Yu W, Ren Y, et al. A conditional generative model for end-to-end stress field prediction of composite bolted joints. Engineering Applications of Artificial Intelligence 2024;134:108692. https://doi.org/10.1016/j.engappai.2024.108692.
[68] Xu C, Wang X, Li Y, Wang G, Zhang H. Conditional Generative Adversarial Network Enabled Localized Stress Recovery of Periodic Composites. CMES 2024;140:957–74. https://doi.org/10.32604/cmes.2024.047327.
[69] Wei Y, Fu H, Li YE, Yang J. A new P-wave reconstruction method for VSP data using conditional generative adversarial network. SEG Technical Program Expanded Abstracts 2019, San Antonio, Texas: Society of Exploration Geophysicists; 2019, p. 2528–32. https://doi.org/10.1190/segam2019-3206719.1.
[70] Chang DK, Yang WY, Yong XS, Li HS. Seismic data interpolation with conditional generative adversarial network in time and frequency domain. SEG Technical Program Expanded Abstracts 2019, San Antonio, Texas: Society of Exploration Geophysicists; 2019, p. 2589–93. https://doi.org/10.1190/segam2019-3210118.1.
[71] Koeshidayatullah A, Ferreira I. A novel deep learning-assisted reservoir fracture delineation with conditional generative adversarial networks. Second International Meeting for Applied Geoscience & Energy, Houston, Texas: Society of Exploration Geophysicists and American Association of Petroleum Geologists; 2022, p. 3234–7. https://doi.org/10.1190/image2022-3751545.1.
[72] Almassaad A, Alajlan H, Alebaikan R. Student Perceptions of Generative Artificial Intelligence: Investigating Utilization, Benefits, and Challenges in Higher Education. Systems 2024;12:385. https://doi.org/10.3390/systems12100385.
[73] Gruenhagen JH, Sinclair PM, Carroll J-A, Baker PRA, Wilson A, Demant D. The rapid rise of generative AI and its implications for academic integrity: Students' perceptions and use of chatbots for assistance with assessments. Computers and Education: Artificial Intelligence 2024;7:100273. https://doi.org/10.1016/j.caeai.2024.100273.
[74] Tharayil SM, Idris MA, Alfaifi OM, Alghafis MS. Unleashing the Power of Generative AI and LLM for Training Evaluation. ADIPEC, Abu Dhabi, UAE: SPE; 2024, p. D021S056R006. https://doi.org/10.2118/222374-MS.
[75] Bravo C, Saputelli L, Rivas F, Pérez AG, Nikolaou M, Zangl G, et al. State of the Art of Artificial Intelligence and Predictive Analytics in the E&P Industry: A Technology Survey. SPE Journal 2014;19:547–63. https://doi.org/10.2118/150314-PA.
[76] Abou-Abbas L, Henni K, Jemal I, Mezghani N. Generative AI with WGAN-GP for boosting seizure detection accuracy. Front Artif Intell 2024;7:1437315. https://doi.org/10.3389/frai.2024.1437315.
[77] Li X, Zhang F-L. Classification of multi-type bearing fault features based on semi-supervised generative adversarial network (GAN). Meas Sci Technol 2024;35:025107. https://doi.org/10.1088/1361-6501/ad068e.
[78] Doloi N, Ghosh S, Phirani J. Super-Resolution Reconstruction of Reservoir Saturation Map with Physical Constraints Using Generative Adversarial Network. Day 1 Tue, January 24, 2023, Abu Dhabi, UAE: SPE; 2023, p. D011S001R006. https://doi.org/10.2118/212611-MS.
[79] Zhang J, Li Y, Pei W, Shirazi S. ConGANergy: A Framework for Engineering Data Augmentation with Application to Solid Particle Erosion. SPE Annual Technical Conference and Exhibition, New Orleans, Louisiana, USA: SPE; 2024, p. D031S032R009. https://doi.org/10.2118/220954-MS.
[80] Yumlembam R, Issac B, Jacob SM, Yang L. Comprehensive Botnet Detection by Mitigating Adversarial Attacks, Navigating the Subtleties of Perturbation Distances and Fortifying Predictions with Conformal Layers. Information Fusion 2024;111:102529. https://doi.org/10.1016/j.inffus.2024.102529.
[81] Wang Y, Sun T, Yuan X, Li S, Ni W. Minimizing Adversarial Training Samples for Robust Image Classifiers: Analysis and Adversarial Example Generator Design. IEEE TransInformForensic Secur 2024;19:9613–28. https://doi.org/10.1109/TIFS.2024.3474973.

 doi:10.20944/preprints202501.1294.v1

[82] Basu A. The Impact of Artificial Intelligence on Cybersecurity. ADIPEC, Abu Dhabi, UAE: SPE; 2024, p. D021S077R001. https://doi.org/10.2118/222493-MS.

[83] Barthel F, Beckmann A, Morgenstern W, Hilsmann A, Eisert P. Gaussian Splatting Decoder for 3D-aware Generative Adversarial Networks. 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA: IEEE; 2024, p. 7963–72. https://doi.org/10.1109/CVPRW63382.2024.00794.

[84] Li Y, Xiao J, Wang Y, Lu Z. DepthGAN: GAN-based depth generation from semantic layouts. Comp Visual Media 2024;10:505–22. https://doi.org/10.1007/s41095-023-0350-8.

[85] Zhang T, Tilke P, Dupont E, Zhu L, Liang L, Bailey W. Generating Geologically Realistic 3D Reservoir Facies Models Using Deep Learning of Sedimentary Architecture with Generative Adversarial Networks. Day 3 Thu, March 28, 2019, Beijing, China: IPTC; 2019, p. D031S067R001. https://doi.org/10.2523/IPTC-19454-MS.

[86] You N, Elita Li Y, Cheng A. 2D-to-3D reconstruction of carbonate digital rocks using Progressive Growing GAN. First International Meeting for Applied Geoscience & Energy Expanded Abstracts, Denver, CO and Virtual: Society of Exploration Geophysicists; 2021, p. 1490–4. https://doi.org/10.1190/segam2021-3592148.1.

[87] Priyadarshi R, Ranjan R, Kumar Vishwakarma A, Yang T, Singh Rathore R. Exploring the Frontiers of Unsupervised Learning Techniques for Diagnosis of Cardiovascular Disorder: A Systematic Review. IEEE Access 2024;12:139253–72. https://doi.org/10.1109/ACCESS.2024.3468163.

[88] Sadana A, Thakur N, Poria N, Anand A, Seeja KR. Comprehensive Literature Survey on Deep Learning Used in Image Memorability Prediction and Modification. In: Hassanien AE, Castillo O, Anand S, Jaiswal A, editors. International Conference on Innovative Computing and Communications, vol. 731, Singapore: Springer Nature Singapore; 2024, p. 113–23. https://doi.org/10.1007/978-981-99-4071-4_10.