

DOI: https://doi.org/10.48009/4_iis_2026_131

Technological Trajectories and Future Roadmap of AI-Driven Drug Discovery: A Structural Topic Modeling Analysis of Global Patent Data

Anuj Sharma, *O.P. Jindal Global University*, anuj.sharma@jgu.edu.in

Alex Koohang, *Middle Georgia State University*, alex.koohang@mga.edu

Manoj Kumar Sharma, *Maharishi Markandeshwar University*, manojsharma198@gmail.com

Gopal Singh Charan, *SGT University*, pedslove@gmail.com

Rajesh Mahadeva, *Manipal Institute of Technology, Manipal Academy of Higher Education*, rajeshmahadeva15@gmail.com

Abstract

Artificial intelligence (AI) is radically transforming the pharmaceutical industry by converting protracted, trial-and-error drug discovery into rapid, precise, cost-effective, and data-driven approaches for molecular discovery, target identification, biomarker discovery, virtual screening, clinical trial optimization, and personalized therapeutics. Although research on AI-driven drug discovery has experienced substantial growth in recent years, the underlying innovation trajectories shaping this field remain insufficiently understood, as limited studies have mapped its technological structure and future evolutionary pathways. Moreover, the current literature lacks a future-oriented patent intelligence perspective on AI-driven drug discovery to uncover latent themes and emerging innovation hotspots. Addressing this gap, the present study leverages Structural Topic Modeling (STM) to identify and interpret the evolving technological landscape of AI-driven drug discovery from the patent dataset comprising 1078 patents from 2006 to 2025. The study identifies eight major themes, including deep learning for target identification and drug-target interaction modeling, AI-enabled in silico drug development and drug activity modeling, AI for clinical diagnostics and biomarker discovery, AI-assisted inhibitor discovery, AI for binding affinity prediction and molecular docking, precision medicine and AI-based treatment response prediction, intelligent drug screening, and generative AI for drug design. The analysis reveals the technological evolution from conventional computational drug discovery to generative AI-based design of high-potential compounds, target identification, virtual screening, precision therapeutics, and intelligent clinical trials. Further, the latent topics discovered using STM are classified into emerging, mature, and potentially stagnating themes to prepare a near-term, mid-term, and long-term technological roadmap of AI-driven drug discovery. By linking STM-derived topics with temporal dynamics, the study reveals how technological priorities evolve, identifies areas of technological convergence, and provides evidence of shifting innovation patterns within AI-driven drug discovery. These insights contribute to the understanding of innovation dynamics and technology evolution in emerging technological ecosystems. The theoretical, practical, and strategic implications for academic researchers, policymakers, and innovation strategists are highlighted.

Keywords: Artificial intelligence, drug discovery, drug design, patent text mining, topic modeling

Introduction

The pharmaceutical industry has historically faced significant challenges related to high research and development (R&D) costs, low productivity, scientific complexity and uncertainty, high failure rates, regulatory and compliance hurdles, long development timelines, and commercial pressures (Khanna, 2012; Munos, 2009; Paul et al., 2010). The traditional drug discovery process is often characterized by lengthy, resource-intensive, and complex trial-and-error-based scientific investigation and experimentation with substantial failure rates (Scannell et al., 2012). Hence, each successful therapeutic discovery has a prolonged commercialization timeline. Despite substantial technological innovations and digital transformation of drug discovery processes, the probability of delays, cost overruns, and failures across different phases of drug discovery remains significant due to inherent complexities and uncertainties associated with target identification, target validation, lead discovery, lead optimization, clinical testing, and regulatory approvals (Zheng et al., 2026). Moreover, the increasing complexity of multifactorial contemporary diseases such as cancer, Alzheimer's disease, autoimmune disorders, antibiotic resistance, rare genetic diseases, and viral pandemics has exceeded the capacity of traditional drug discovery methods (Fu & Chen, 2025). These stark realities have motivated intensive drug development efforts to focus on more efficient, cost-effective, adaptive, scalable, and precise approaches to drug discovery (Chen et al., 2023).

With significant advancements in recent years, artificial intelligence (AI) has emerged as a revolutionary technological paradigm capable of overcoming major challenges in pharmaceutical research and traditional drug discovery methods (Dharmasivam, Azad, et al., 2026; Jarallah et al., 2025). AI-driven drug discovery integrates state-of-the-art computational intelligence technologies, such as machine learning, deep learning, reinforcement learning, natural language processing, generative AI (GenAI), explainable AI (XAI), diffusion models, variational autoencoders, AI agents, etc., to accelerate various stages of the drug discovery and development processes (Luo et al., 2016; Pipada et al., 2026; Zheng et al., 2026). These transformative AI techniques used in modern drug discovery are considered superior to traditional approaches because they are capable of managing exponential biological data complexity, learning hidden patterns from massive datasets, automating drug discovery processes, generating entirely new molecular structures, and providing predictive insights about a drug's efficacy, target interactions, toxicity, and pharmacokinetic properties at a scale (Chen et al., 2023; Mustafa, 2026). Hence, modern AI-based systems have outperformed traditional drug discovery methods due to the inherent limitations of traditional methods in processing large multimodal biological datasets and in performing slow, computationally expensive molecular screening for therapeutic development (D. Paul et al., 2021; Singh, Gupta, et al., 2024).

The growing integration of AI in drug discovery is driven by the increasing complexity, cost, uncertainty, and data intensity of pharmaceutical innovation processes (Gupta et al., 2026; Jarallah et al., 2025). However, the strategic demand and competitive pressure for low cost, faster innovation cycles, more precise, scalable, and personalized therapeutic development in the pharmaceutical industry has further accelerated the adoption of AI as a strategically essential innovation capability (D. Paul et al., 2021; Pipada et al., 2026). This has resulted in a fundamental transformation of pharmaceutical drug discovery from traditional approaches that relied heavily on predefined rules, manual data processing, and hypothesis-driven trial-and-error-based experimentation to rapid, precise, productive, cost-effective, data-driven, adaptive, and autonomous therapeutic development (Dharmasivam, Kaya, et al., 2026).

Pharmaceutical and healthcare firms, biotechnology companies, startups, and technology corporations are increasingly investing in AI-driven drug discovery ecosystems to improve research productivity, reduce costs, shorten drug development cycles, and enhance therapeutic precision (Fu & Chen, 2025). The growing strategic significance of AI in drug discovery is evident from the rapidly increasing volume of global patents

filed in this domain (Sakaoka & Kano, 2025). These patents are a treasured source of technological intelligence as they detail commercially oriented innovations related to AI-driven drug discovery. As AI continues to rapidly evolve, mapping the landscape of AI-driven drug discovery from patent documents is vital because a vigilant exploration may reveal emerging technological paradigms, knowledge structures, and future-oriented innovation trajectories. Unlike scientific literature mapping, patent analysis can offer significant knowledge regarding technological convergence, industrial priorities, innovation trends, and future technology trajectories (Singh, Singh, et al., 2024).

Despite the growing importance of AI-driven drug discovery, the existing literature mostly focuses on experimental validation of AI algorithms, conceptual reviews of research articles exploring AI techniques in drug discovery, or case studies reporting specific AI platform-based drug discovery processes. Although these studies provide valuable insights related to the role of AI in drug discovery, a significant gap related to providing a holistic understanding of the evolving technological structure and innovation trajectories remains unfilled. Further, there is a significant requirement for research that advances existing patent analytics approaches that are based on keyword analysis and citation mapping (Sakaoka & Kano, 2025; Zhang et al., 2022). Another important limitation in the literature corresponds to the absence of future-oriented technological roadmapping and strategic foresight for AI-driven drug discovery. A vast number of existing studies primarily provide a retrospective overview of AI technologies in drug discovery without systematically examining which technological themes are emerging, which are maturing, and which may potentially stagnate in the future.

Addressing these research gaps, the present study employs Structural Topic Modeling (STM) to investigate the evolving technological landscape of AI-driven drug discovery, leveraging a global patent dataset. Specifically, the current study aims to uncover the latent technological themes shaping AI-driven drug discovery, discover emerging themes, and develop a future-oriented technological roadmap. With these aims, the current study seeks to address the following research questions:

1. What are the key themes emerging from AI-driven drug discovery global patents?
2. Which themes represent emerging, mature, or potentially stagnating domains?
3. What future technological pathways are likely to shape the evolution of AI-driven drug discovery in the near-, medium-, and long-term future?

The study offers several important theoretical, methodological, and practical contributions. From a theoretical perspective, the study advances understanding of technological convergence and innovation evolution within AI-driven pharmaceutical ecosystems. The findings contribute to emerging literature on AI-enabled pharmaceutical innovation systems and strategic technology management by demonstrating how multiple technological paradigms under AI converge to shape future therapeutic innovation. From a methodological perspective, the study extends patent analytics research by integrating STM with technological roadmapping. STM enables more nuanced identification of latent technological structures. From a practical perspective, the study provides valuable insights for pharmaceutical firms, biotechnology companies, technology startups, policymakers, and research professionals. The identification of emerging technological domains and future innovation trajectories can support strategic investment decisions, R&D prioritization, and the formulation of effective technology commercialization strategies.

The remaining parts of this paper are organized as follows. The next section provides a review of the literature on AI-driven drug discovery and technology roadmapping. Then the research methodology, including the acquisition of patent data, preprocessing, and the topic modeling approach, is detailed. The following section elaborates on the identified latent themes and provides a discussion on emerging innovation trajectories. Finally, the study develops a future-oriented technological roadmap for AI-driven drug discovery and discusses theoretical and managerial implications.

Literature Review

Overview of AI-driven Drug Discovery

The drug discovery pipeline is a multi-phase procedure that starts with disease understanding, where researchers explore disease biology and mechanisms from information available from multiple sources (Zheng et al., 2026). AI techniques like NLP can summarize the voluminous data, and GenAI based on LLMs can present that information in an interactive way (Pipada et al., 2026). Further, biological targets (for example, proteins or genes) linked to disease are identified in the target identification phase, where AI enables the large-scale analysis of medical and pharmaceutical datasets (Chen et al., 2023). Subsequently, the identified disease-associated genes, proteins, or biomarkers that may serve as potential targets are validated in a target validation step where AI-based technologies such as machine learning, deep learning, GNNs, and NLP analyze complex datasets, map disease-target relationships, and predict drug-target interactions (Askr et al., 2023; Gupta et al., 2026). Hence, AI enhances the speed and precision of predicting target druggability (D. Paul et al., 2021). The next two important steps are hit discovery and lead discovery, where initial compounds that show promising activity against a target are discovered in hit discovery, and lead discovery steps refine these hits further into druggable molecules (Jarallah et al., 2025). AI supports hit discovery and lead discovery by modeling molecular structures and predicting molecular interactions using GNNs, Generative Adversarial Networks (GANs), and deep learning algorithms (Wang & Zhuang, 2025).

Further, the discovered lead compounds are chemically enhanced in the lead optimization phase by AI-supported specialized procedures like De Novo Design, Scaffold Hopping, and Pharmacokinetics prediction to reduce toxicity and enhance absorption (Fu & Chen, 2025). The next steps, named preclinical testing, involve assessing the safety, toxicity, and efficacy of the lead drug before human trials. AI-based simulations accelerate the preclinical testing processes, reducing reliance on animal testing and predicting compound safety and efficacy (Zheng et al., 2026). Techniques such as predictive modeling based on knowledge graphs, convolutional neural networks (CNNs), GenAI, in silico compound screening, etc., can predict absorption, distribution, metabolism, excretion, and toxicity (ADMET) of lead drug (Chang et al., 2023; D. Paul et al., 2021; Wang & Zhuang, 2025). Finally, in clinical trials of new drugs, AI supports patient stratification, trial recruitment, endpoint prediction, treatment response analysis, and adverse event forecasting through the integration of patient history, real-world evidence, and multimodal data (Probert et al., 2025). Additionally, AI contributes to drug repurposing by identifying new therapeutic applications for existing drugs through biomedical literature mining and network analysis (Mustafa, 2026).

AI-driven drug discovery adopts an ecosystem of multiple AI techniques applied across various phases of drug discovery and development. Different AI techniques address different scientific and operational challenges associated with key phases of drug discovery. For example, target identification and validation is one of the earliest and most critical phases in drug discovery, involving the identification of a biological molecule implicated in a disease, confirmation of its role in causing the disease, and validation that modifying the target can safely and effectively cure the disease (Chen et al., 2023). The key activities involved in this phase are literature analysis, laboratory experiments, protein interaction analysis, biomarker studies, clinical data assessment, genomic analysis, and a few others (Fu & Chen, 2025). This phase is significantly important, uncertain, slow, and resource-consuming because if an incorrect target is selected, then not only may the drug fail, but the toxicity may also increase (Du et al., 2026). Traditional approaches face several challenges in the target identification and validation phase due to biological complexity, huge genomic data, unknown molecular interactions, high experimental cost, slow discovery, and high uncertainty of results. However, AI drastically improves speed, cost efficiency, scalability, and prediction accuracy of biological insight generation. AI-based computer information systems can analyze massive biomedical datasets using natural language processing (NLP), detect latent disease-target relationships

using machine learning, improve biomarker discovery using deep learning, accelerate genomic analysis using GenAI and neural networks, and predict protein structure and function using Graph Neural Networks (GNNs), transformers, and large language models (LLMs) (Mustafa, 2026). Likewise, contemporary AI advancements such as GenAI models, comprising variational autoencoders, diffusion models, transformers, and LLMs, can support other phases of drug discovery with greater speed and precision than traditional experimental approaches.

Technology Roadmapping for AI-driven Drug Discovery

Technology Roadmapping is a data-driven strategic planning and forecasting method used to identify, visualize, and predict the evolution of emerging technologies, innovation trajectories, and future research priorities over time (Singh, Singh, et al., 2024). It helps researchers, practitioners, and policymakers understand how technologies are likely to evolve, converge, mature, or become obsolete while aligning technological developments with short-term strategic priorities and long-term strategic objectives (Ozcan et al., 2022). In the context of AI-driven drug discovery, technology roadmapping has become critically important because pharmaceutical innovation is undergoing rapid transformation due to advances in machine learning, deep learning, GNNs, GenAI, computer simulation, and autonomous design systems (Jarallah et al., 2025; Sun et al., 2023). AI-driven drug discovery as a field of research inquiry is highly dynamic, interdisciplinary, and data-intensive, making it difficult for pharmaceutical firms, researchers, and policymakers to identify which technological domains are emerging, which are maturing, and which may dominate future pharmaceutical ecosystems and innovation pathways. Hence, technology roadmapping of AI-driven drug discovery becomes crucial because it enables the systematic exploration of emerging innovation hotspots, future therapeutic paradigms, and evolving AI capabilities across the drug discovery pipeline, including target identification, target validation, lead discovery, virtual screening, novel molecular generation, precision medicine, and clinical intelligence. A technology roadmap can further support strategic R&D investment decisions, innovation prioritization, technology forecasting, collaborative ecosystem development, and competitive intelligence. Most importantly, technology roadmapping helps pharmaceutical firms transition from reactive innovation strategies toward proactive and future-oriented AI-enabled pharmaceutical ecosystems capable of accelerating therapeutic discovery, reducing R&D uncertainty, and improving long-term innovation competitiveness (Sakaoka & Kano, 2025).

Methodology

Patent Search

Developing a robust and precise search protocol is critically important when downloading patent data from proprietary patent databases to ensure comprehensive retrieval of relevant patents and reliability of subsequent patent analysis (Singh, Singh, et al., 2024). Patent databases are extremely large, multilingual, and technically complex as inventors use synonyms to represent technical terms, companies intentionally use broad or obscure language, and similar technologies are defined and classified differently (Chuprat et al., 2025). For example, AI-driven drug discovery patents may use different but semantically associated terms, such as ‘computational intelligence,’ ‘predictive modeling,’ ‘deep neural architecture,’ instead of using the direct term ‘AI.’ Previous studies have already evaluated the functionalities and capabilities of patent platforms and databases such as Google Patents, WIPO Patentscope, Espacenet, and other commercial platforms (Jürgens & Herrero-Solana, 2015; Mishra & Bonde, 2022) and have suggested selecting a database that combines large-scale global patent coverage, scholarly integration, open accessibility, and advanced analytics capabilities in a single ecosystem. This study found that the Lens.org platform offers open and free access to large-scale global patent data with advanced search capabilities. Hence, to develop a reproducible search protocol, an initial search was conducted using the seed words ‘AI’ and ‘drug discovery’ to collect patents from the Lens.org platform. These patents were thoroughly scanned to prepare two comprehensive lists of all the keywords and their synonyms that represent the AI domain

and the drug discovery domain. Two senior researchers and two practitioners from industry independently validated these lists, and after two iterations, the final search protocol was prepared using the Boolean operator ‘OR’ between the keywords (Table 1) in a list and the Boolean operator ‘AND’ to merge both lists. This study used the title and abstract fields of the Lens.org platform for retrieving relevant patent documents.

The search query was executed on March 2, 2026, which retrieved 1,218 patents. The main objective of the current study is to explore the technological landscape of AI-driven drug discovery for proposing a technology roadmap. Hence, the period from 2006 to 2025 is selected because this timeframe captures the modern technological evolution and commercialization of AI-driven drug discovery (Fu & Chen, 2025; Sakaoka & Kano, 2025). Limiting the analysis to the last two decades ensures that the final dataset reflects contemporary technological paradigms and current industrial practices. This also warrants an appropriate balance between historical depth and contemporary relevance. The final dataset represented 1,078 patents granted by the World Intellectual Property Organization (WIPO) and other patent offices from China, the United States, the Republic of Korea, Australia, Canada, and Japan, to name a few.

Table 1. List of Keywords for Search Protocol

Key Theme	Keywords
Artificial Intelligence	“artificial intelligence” OR “AI” OR “machine learning” OR “ML” OR “predictive modeling” OR “predictive analytics” OR “data mining” OR “expert system*” OR “knowledge-based system*” OR “cognitive computing” OR “robot*” OR “computer-aided drug design” OR “computational intelligence” OR “deep learning” OR “neural network*” OR “convolutional neural network*” OR “CNN” OR “recurrent neural network*” OR “RNN” OR “graph neural network*” OR “GNN” OR “generative adversarial network*” OR “GAN” OR “natural language processing” OR “NLP” OR “Explainable AI” OR “XAI” OR “generative AI” OR “GenAI” OR “large language model*” OR “LLM” OR “transformer” OR “diffusion model*” OR “autoencoder*” OR “attention mechanism” OR “reinforcement learning” OR “foundation model*” OR “Agentic AI” or “AI Agent*” OR “Multi-Modal AI” OR “foundation model*”
Drug Discovery	“drug discovery” OR “drug design” OR “drug development” OR “drug repositioning” OR “drug analysis” OR “drug screening” OR “drug repurposing” OR “pharmaceutical*” OR “medicinal chemistry” OR “lead identification” OR “lead optimization” OR “virtual screening” OR “molecular docking” OR “biomarker discovery” OR “compound screening” OR “therapeutic discovery” OR “precision medicine” OR “target identification” OR “target validation” OR “physicochemical property prediction” OR “drug–target interaction” OR “drug toxicity prediction” OR “Pharmacokinetics”

Note: * represents a wildcard character in the patent search protocol

Topic Modeling based on STM

Topic Modeling is particularly well-suited for analyzing patent data as it enables researchers to uncover the latent technological themes, temporal evolution, and contextual semantic relationships within large-scale patent corpora (Choi & Woo, 2022). Topic modeling in NLP is an unsupervised machine learning technique based on probabilistic mixture models that assumes that each document is a mixture of topics and each topic is a mixture of semantically associated terms (Sharma et al., 2021). In particular, STM provides a robust and scalable approach for uncovering latent themes from a patent dataset while simultaneously incorporating document-level metadata such as publication year or patent classifications for further analysis (Singh, Singh, et al., 2024). By taking publication year as a document-level covariate, STM can efficiently identify emerging themes and their evolution trajectories. Moreover, as AI-driven drug discovery is a temporally evolving convergence domain, STM as a text analytics method is superior to traditional patent analysis methods due to its ability to perform automated large-scale analysis with metadata integration to discover emerging and converging themes. The metadata integration enables STM to model topic prevalence and content variation across temporal covariates so that evolving AI applications in pharmaceutical research can be mapped.

STM uses each patent as one document in the dataset and then breaks the text into terms that appear together repeatedly across documents. For example, the terms like ‘molecule,’ ‘prediction,’ ‘neural network,’ and ‘drug candidate’ may often appear together. Hence, STM assumes that if some terms frequently appear together, then they can probably be mapped as one latent technological theme in AI-driven drug discovery. In this way, STM discovers latent topics automatically from patent documents by grouping frequent terms into semantically meaningful topics. Further, as one patent document may have several latent topics at the same time, STM creates a probabilistic topic model for each document by representing each patent as a mixture of all the topics, where the most prevalent topic may be used to visualize key innovation themes. Finally, STM uses a document-level covariate like publication year and associates it with the topic prevalence to show how these themes change over time.

Data Preprocessing and Model Building

The current study implements the key data preprocessing tasks and STM in the R programming environment and adapts the *tm* (Feinerer et al., 2008) and *stm* (Roberts et al., 2019) packages. The text corpus for the current study is prepared by merging the title and abstract of each patent document. Further, the most frequent bigrams and trigrams were extracted using an *n*-gram tokenizer, and subsequently, these were converted into unigrams to preserve their semantics. Further, the most common stop words, non-English characters, numbers, and special characters were removed. The current study experimented with different topic models with a total number of topics ranging from 5 to 30. Figure 1 (left map) shows that the final number of topics was selected as eight based on standard measures like held-out likelihood (which assesses the generalizability of the final model) and semantic coherence (which ensures that topics are interpretable and meaningful). To validate the topics further, the semantic coherence and exclusivity measures were checked, as shown in Figure 1 (right map), which confirms that no two topics have similar values for these measures. Finally, the topic correlation map (Figure 2) was created to confirm that the correlation among topics is not greater than 0.30, indicating that extracted topics are semantically distinct.

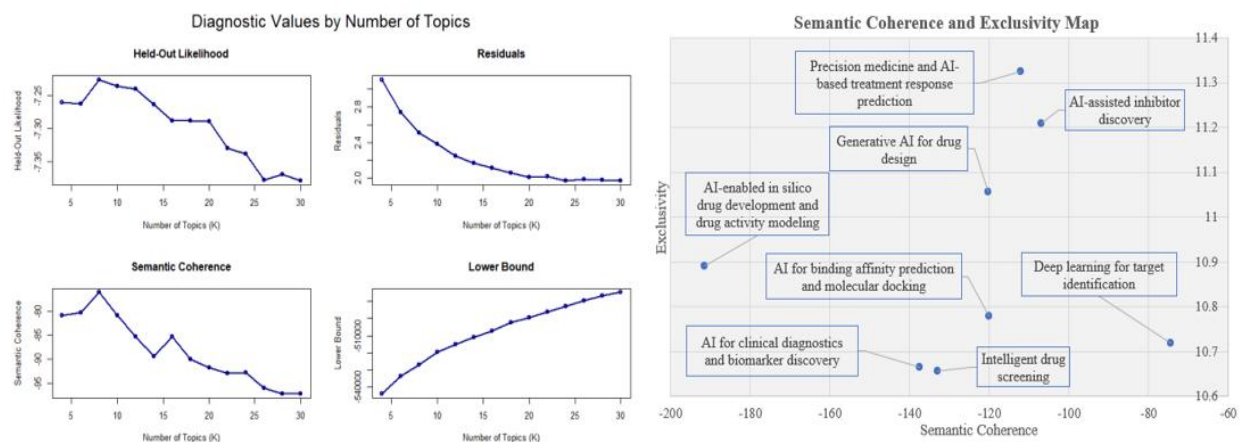


Figure 1. Selection of Number of Topics using Semantic Coherence, Held-out Likelihood, and Exclusivity

Results

STM maps patents to key topics that can be further represented using the top-most frequent terms. Table 2 reports the key topics, their definition, key terms, and representative patents that can be explored further to infer the underlying technological domain from the content of each topic. It is evident that the topics are reasonably differentiated and reflect distinct technological trajectories in AI-enabled drug discovery. STM doesn't generate topic titles automatically; the researchers manually inspect the semantics of each topic using the top words and the related top documents. This study, therefore, adopted a systematic interpretive

procedure based on prior works (Dwivedi et al., 2023; Nunkoo et al., 2025; Sharma et al., 2025) rather than a subjective guess to decide the title of each topic. Hence, the labeling process used contextual domain knowledge related to AI-driven drug discovery to label the patents. Since STM is an unsupervised machine learning technique, a subjective evaluation by humans is needed for the theme to be interpreted (Singh et al., 2023). The theme labels, therefore, were iteratively revised on the basis of domain-specific interpretation of the patents and keywords.

Topic 1: Deep Learning for Target Identification and Drug-Target Interaction Modeling

Deep learning has emerged as a transformative technique by accurately modeling molecular structures and predicting their interactions with binding targets during the drug development process (Askr et al., 2023). The advanced deep learning techniques, such as Convolutional Neural Networks (CNNs) and GNNs, can efficiently analyze molecular structures, chemical properties, and protein sequences and learn patterns from large biochemical datasets to predict which molecules are most likely to bind with the disease-related protein (Wang & Zhuang, 2025). Academic research literature also reports that deep learning techniques can effectively model the interactions between drug and target as well as between two drugs (Oliveira et al., 2025). Moreover, these techniques are also instrumental in assessing drug sensitivity and responsiveness, which is a resource-intensive process. Hence, Deep learning-based approaches are significantly shortening the time and reducing the overall cost of developing new drugs.

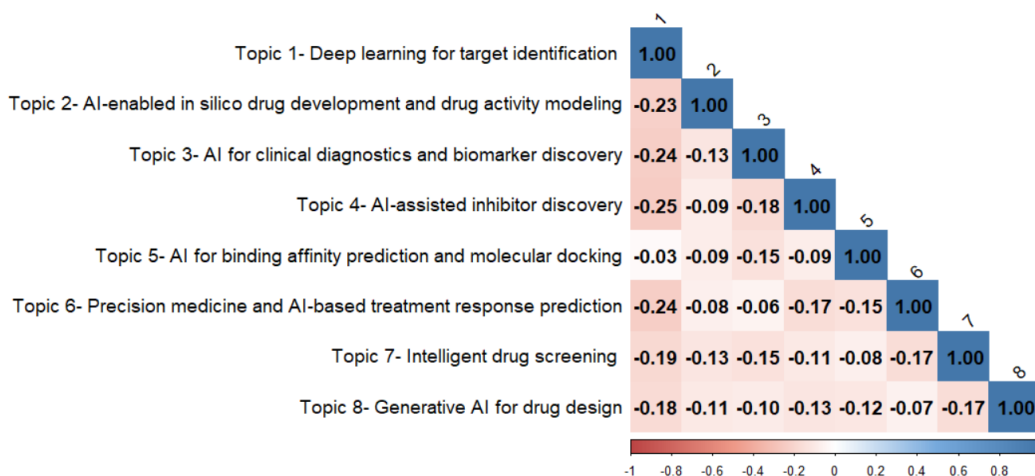


Figure 2. Correlation Analysis

Table 2. Summary of Extracted Topics and Sample Patents

Topic (Proportion in %)	Definition	Key Terms	Representative Patents
Deep learning for target identification and drug-target interaction modeling (26.8%)	Use of deep learning techniques to predict, understand, or simulate how biological molecules interact with each other during the drug discovery process	Target Identification, Graph, Connection, Molecular Structure, Candidate Drug Molecules, Molecular Property, Neural Network Model, Deep Learning, Interaction Prediction, Molecular Structure	1. "Drug-target interaction prediction method based on graph neural network block" (Yang Qing et al., 2025) 2. "Natural product-protein interaction prediction method based on deep learning" (Lei Zhou et al., 2025)
AI-enabled in silico drug development and drug activity modeling (8.1%)	Identify and design prospective drug candidates using AI-based techniques for optimizing the time and resources needed to find	Drug Development, Drug Discovery, Drug Activity, Pharmacokinetic, In Silico, Computational Modeling, Artificial Intelligence, Candidate, Invention, Estimate	1. "In silico drug development and determination of drug activity" (Packer Jonathan S et al., 2025) 2. "Synthesis and in silico studies of novel ANTI-SARS-COV Sulfonamides as potential inhibitors against COVID-19

Topic (Proportion in %)	Definition	Key Terms	Representative Patents
	and improve therapeutic candidates		protein target: SARS-COV-2 main protease” (M Pro) (Al-Sehemi Abdullah G et al., 2022)
AI for clinical diagnostics and biomarker discovery (12.2%)	Leveraging AI to identify biological indicators that can help diagnose diseases, predict drug response, monitor progression, or support personalized treatment	Diagnostics, Biomarker Discovery, Diagnosis, Platform, Risk, Prediction, Multimodal, Clinical Test, Clinical Trials, Machine Learning	1. “Artificial intelligence-based biomarker selection device and method” (Kim Joonghee & Cho Youngjin, 2025) 2. “Machine learning based predictive biomarker discovery and patient stratification using standard care data” (Probert Christopher et al., 2025)
AI-assisted inhibitor discovery (12.3%)	Use of AI techniques to identify, evaluate, and optimize potential inhibitor compounds for drug development	Inhibitor, Compound, Receptor, Screening, Effect, Inhibitory Activity, Protein, Cancer, Target Prediction, Virtual Screening	1. “Inhibitor drug screening method for targeting KRAS G12D mutation based on machine learning” (Zhou et al., 2025) 2. “Application of virtual screening compound in preparation of NLRP3 inflammasome inhibitor” (Yang Yanqiu et al., 2025)
AI for binding affinity prediction and molecular docking (7%)	AI-based techniques to predict the strength and stability of interactions between drug molecules and biological targets during drug discovery	Molecular Docking, Protein, Simulation, Affinity, Docking, Coronavirus, Binding Affinity Prediction, Protein Interaction, Absorption, Toxicity	1. “Molecular modification method for glycosyltransferase based on artificial intelligence” (Du Jun et al., 2025b) 2. “Enzyme modification method based on artificial intelligence” (Du Jun et al., 2025a)
Precision medicine and treatment response prediction (9.9%)	AI-based methods to predict personalized therapeutic responses and optimize patient-specific treatment strategies	Model, Disease, Response, Machine Learning, Data, Patient, Treatment, Precision Medicine, Algorithm, Disorder	1. “Pan-Cancer Platinum Response Predictor” (Jim Abraham et al., 2025) 2. “Dynamic and locally-faithful explanation of machine learning models for precision medicine” (Katuwal Gajendra Jung et al., 2024)
Intelligent drug screening (14.3%)	AI-based selection of potential drug candidates from a vast pool of compounds and evaluation of the efficacy and safety of new drugs in treating diseases	Screening, Model, Invention, Drug, Antiviral Drug Screening, Virus, Molecular Fingerprints, biological modeling, high-throughput screening	1. “Drug screening method and device based on neural network” (Zhang Yuezhou et al., 2024) 2. “Artificial intelligence expert system and method based on cell chip cancer drug screening” (Wan Zhitong, 2025)
Generative AI for drug design (9.4%)	Using AI for therapeutic molecule structure generation and molecular optimization to develop novel compounds with desired properties	Generative Models, Representation, Chemical Compound, Small Molecule, Autoencoder, ChatGPT, Protein, Biomolecular Generation, Antibody, Diffusion	1. “Generative AI methods and systems for small molecule structure generation” (Ning Xia & Chen Ziqi, 2025) 2. “Diffusion-based generative AI methods for protein and drug design” (Odaibo Stephen Gbejule, 2025)

Note: Topics are listed in the order of their discovery by STM algorithms. The sequence does not represent their importance, chronology, prevalence, or conceptual hierarchy.

Topic 2: AI-Driven In Silico Drug Development and Drug Activity Modeling

The associated keywords and patent documents related to this topic predominantly focus on computational approaches for drug discovery, pharmacological prediction, molecular activity analysis, and AI-assisted therapeutic modeling. In silico drug discovery involves the application of multiple AI techniques such as machine learning, NLP, virtual screening, computer simulation, etc., for identifying prospective drug candidates (Roney & Mohd Aluwi, 2024). For example, one of the most critical treatments for human immunodeficiency virus (HIV) adopts the medication raltegravir, which was designed using in silico drug discovery (Chang et al., 2023). In particular, AI techniques help analyze large biological and chemical datasets as well as biomedical literature, identify drug candidates, predict molecular properties and

behavior, model drug response, estimate drug-target interactions, and improve therapeutic decision-making (Hasnain et al., 2025).

Topic 3: AI for Clinical Diagnostics and Biomarker Discovery

Biomarkers are crucial in drug discovery because pharmaceutical researchers use these measurable biological indicators to identify disease mechanisms, select drug targets, predict treatment effectiveness, reduce clinical trial failure, and enable precision medicine (Kim Joonghee & Cho Youngjin, 2025). For instance, a new cerebrovascular disease medicine may only work for patients with a specific biomarker associated with chronic arterial plaques, termed large-artery atherosclerosis (LAA). However, novel and more efficient machine learning approaches may provide promising results for biomarker identification that can lead to diagnosis improvement, risk prediction, and better disease treatment (Sun et al., 2023).

Topic 4: AI-assisted Inhibitor Discovery

This topic focuses on the application of AI in inhibitor discovery and drug screening. An inhibitor is a substance or molecule that reduces, blocks, or stops the activity of a biological target (such as a protein or enzyme), to prevent harmful biological processes associated with diseases (Fu & Chen, 2025). The patents under this topic describe predictive systems for binding affinity estimation, toxicity assessment, and biological response analysis. The patents report that AI and machine learning techniques enhance virtual screening to identify potential inhibitor molecules by predicting drug-target interactions, estimating binding affinity, and prioritizing compounds with high inhibitory potential.

Topic 5: AI for Binding Affinity Prediction and Molecular Docking

This topic represents patents that describe AI innovations that improve molecular docking by making faster, more accurate, and more scalable predictions about how a drug molecule binds to a biological target such as a protein, enzyme, or receptor (Du Jun et al., 2025a). In drug discovery, molecular docking predicts the physical interaction pattern between a drug molecule and a target protein, while binding affinity prediction estimates the strength and stability of that interaction (Chen et al., 2023). AI-based binding affinity prediction estimates how strongly a molecule will bind to a protein target, which leads to prioritizing the most promising drug candidates (Jarallah et al., 2025).

Topic 6: Precision Medicine and AI-based Treatment Response Prediction

This topic focuses on the application of AI in precision medicine and in predicting drug responses to estimate how a biological system will respond to a specific drug (Zheng et al., 2026). The patents within this cluster emphasize AI-driven methods for analyzing genomic, clinical, pharmacological, and high-throughput molecular data to predict patient-specific treatment outcomes and drug responses (Jim Abraham et al., 2025). Machine learning and deep learning techniques are widely used to identify hidden biological patterns associated with treatment efficacy, adverse drug reactions, and disease progression (D. Paul et al., 2021). Some of the relevant technological subthemes in the present topic include molecular profiling, precision oncology, patient segmentation, risk prediction, and treatment optimization.

Topic 7: Intelligent Drug Screening

This topic focuses on drug screening, which is the broad process of testing large numbers of molecular compounds to identify molecules that show biological activity against a target or disease (Singh, Gupta, et al., 2024). The screening process results in the discovery of ‘hits,’ which are compounds that show promising activity. Subsequently, the hit compounds are further enhanced to become stronger candidate drugs termed as ‘leads’ (Jarallah et al., 2025). AI-based technologies help in virtual screening of compounds, activity prediction, candidate prioritization for hits, molecular optimization, and toxicity prediction for leads.

Topic 8: Generative AI for Drug Design

GenAI and related technologies help in generating new molecular structures, designing novel compounds, predicting molecular properties, simulating drug-target interactions, and optimizing drug candidates (Ning Xia & Chen Ziqi, 2025). Moreover, GenAI improves existing compounds by optimizing their efficacy, toxicity, solubility, stability, and binding affinity (D. Paul et al., 2021). Likewise, contemporary innovations including generative adversarial networks, variational autoencoders, and transformers, enable creating new molecules and bio-compounds, analyzing the target structure, generating candidate molecules, optimizing toxicity and stability, and testing new drugs experimentally, with greater speed and efficiency than traditional experimental approaches (Fu & Chen, 2025; Sakaoka & Kano, 2025).

Discussion

This study adopts patent publication year as a covariate in the topic modeling process to assess the temporal trends related to the prevalence of extracted topics. As evident from Figure 3, Topic 1 ('Deep Learning for Target Identification and Drug-Target Interaction Modeling'), Topic 3 ('AI for Clinical Diagnostics and Biomarker Discovery'), Topic 6 ('Precision Medicine and AI-based Treatment Response Prediction'), and Topic 8 ('Generative AI for Drug Design') are clearly showing an upward trends that confirms a clear evolution to autonomous and generative methods for drug discovery. Specifically, the trends in the last five years indicate current technological momentum due to the focus of industrial investment. These topics are hotspots for the future and are likely to dominate future innovation pathways in pharmaceutical ecosystems. Interestingly, similar trends are also reported in contemporary academic literature, which reports that diffusion models, transformers, LLMs, and end-to-end AI-powered platforms are likely to shape future advancements in drug discovery (Fu & Chen, 2025; Jarallah et al., 2025; Mustafa, 2026). On the contrary, Topic 2 ('AI-enabled In Silico Drug Development and Drug Activity Modeling') and Topic 5 ('AI for Binding Affinity Prediction and Molecular Docking') are showing a declining trend in topic proportions over the year, indicating these themes are becoming saturated or stabilizing due to technological maturity, methodological saturation, and the evolution of next-generation AI paradigms in drug discovery. In pharmaceutical ecosystems, methods such as in silico drug development, molecular docking, and binding affinity prediction are treated as foundational computational approaches under the umbrella of computer-assisted drug design (CADD) (Zheng et al., 2026). CADD has become extensively explored, leaving limited scope to attract future scholars' attention. Finally, Figure 3 reports that Topic 4 ('AI-assisted Inhibitor Discovery') and Topic 7 ('Intelligent Drug Screening') show a stable trend, which indicates that these topics remain contextual but may not become disruptive in the near future. Intelligent drug screening and AI-assisted inhibitor discovery remain strategically critical for pharmaceutical ecosystems; however, there is a possibility that future innovation on these themes may bring incremental improvement or may shift towards GenAI-based autonomous systems (Singh, Gupta, et al., 2024).

Based on the patents registered in the last two years on the identified topics and related academic literature, this study develops a technology roadmap to chart future innovation pathways. Table 3 reports these pathways for near-, mid-, and distant-future innovations. It is evident that autonomous and generative intelligence is likely to become increasingly native to drug discovery. Techniques such as specialized and domain-oriented LLMs, AI-based simulations, and molecular design using GenAI may become prevalent in the near future. Progressing further, AI-guided systems with limited human intervention may lead the key drug discovery processes in the mid-future. Finally, this study predicts that future innovation may systematically shift toward autonomous AI ecosystems such as self-driving labs.

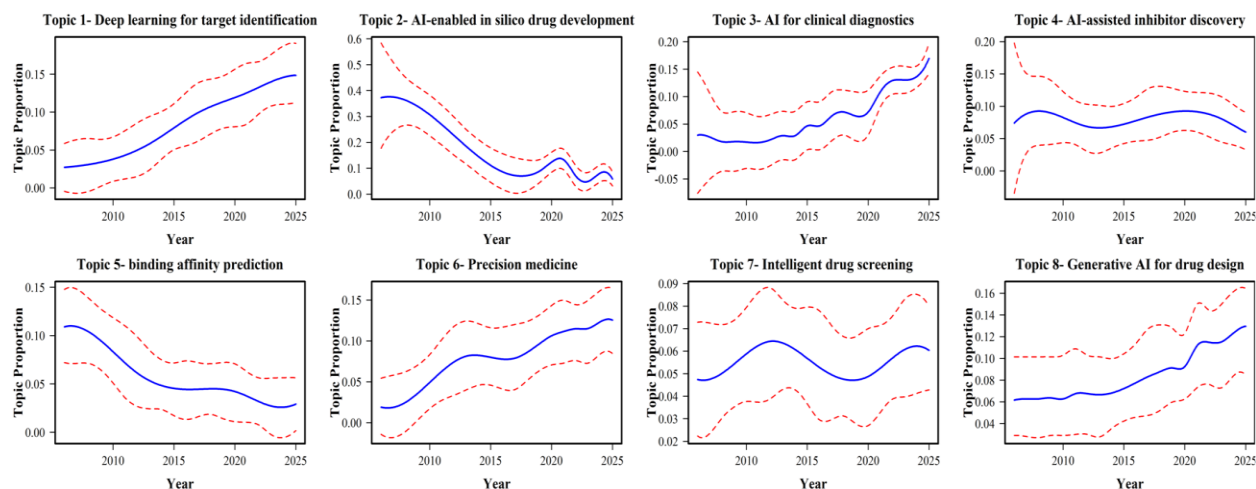


Figure 3. Temporal Topic Trends

Table 3. Summary of Topical Trends and Future Technological Pathways

S. No.	Topic	Current Status	Future Growth Potential	Near Future (0–2 Years)	Mid Future (2–5 Years)	Far Future (5–10 Years)
1	Deep Learning for Target Identification	Hot	Very high	Protein language models	Self-learning models	AI-generated disease models
2	AI-Enabled In Silico Drug Development and Drug Activity Modeling	Mature/stabilizing	Moderate	AI-enhanced simulations	Real-time AI simulation	Fully virtual drug discovery pipelines
3	AI For Clinical Diagnostics and Biomarker Discovery	Hot	High	Multimodal clinical AI	AI-guided precision diagnostics	Real-time remote biomonitoring
4	AI-Assisted Inhibitor Discovery	Stable but competitive	Moderate to high	AI-assisted virtual screening	Autonomous systems	Autonomous therapeutic engineering
5	AI For Binding Affinity Prediction and Molecular Docking	Becoming saturated	Moderate	Faster neural docking systems	AI-native interaction prediction	Simulation-free AI interaction prediction
6	Precision Medicine and AI-Based Treatment Response Prediction	Very hot	Extremely high	AI for clinical decision support	AI-guided precision medicine	Individualized AI-designed therapies
7	Intelligent Drug Screening	Stable topic	Moderate	AI automation	Robotic AI labs	Self-driving labs
8	Generative AI For Drug Design	Disruptive frontier	Extremely high	AI-based molecular generation	Autonomous systems with human-in-the-loop	Agentic systems

Implications, Limitations, and Conclusion

Theoretical Implications

Beyond mapping the technological landscape of AI-driven drug discovery, the findings provide insights into broader processes of technological evolution and innovation ecosystem development. The coexistence of emerging, mature, and stagnating technological themes suggests that AI-driven drug discovery is characterized by heterogeneous innovation trajectories, where different technological domains evolve at varying rates of maturity. This observation aligns with the notion that technological development is not linear but progresses through phases of emergence, growth, maturity, and decline. The identified topics further illustrate the increasing convergence of artificial intelligence and pharmaceutical sciences. Emerging themes such as generative AI, diffusion-based molecular design, and intelligent drug screening indicate that innovation is increasingly occurring at the intersection of previously distinct technological domains. Such convergence supports the view that breakthrough innovations often arise from the recombination of knowledge across disciplinary boundaries.

Moreover, the findings highlight the ecosystem-based nature of AI-driven drug discovery. The evolution of the identified technological themes reflects interactions among multiple actors, including pharmaceutical firms, biotechnology companies, AI developers, research institutions, and healthcare organizations. The emergence of new technological domains suggests that innovation outcomes are shaped not only by

advancements within individual organizations but also by knowledge flows and collaborative activities across the broader innovation ecosystem. Taken together, the study extends current understanding of technological evolution in AI-driven drug discovery by demonstrating how patent-based topic modeling can reveal patterns of technological convergence, ecosystem development, and shifting innovation priorities within a rapidly evolving technological domain. The study therefore suggests that AI-driven drug discovery should be viewed not merely as the adoption of AI tools within pharmaceutical research, but as an evolving innovation ecosystem characterized by technological convergence and the continuous recombination of knowledge assets across scientific and industrial domains

Implications for Practice

This study reports the technological evolution of AI-driven drug discovery based on global patent data. The results identify emerging and declining themes that can help potential scholars focus on research frontiers likely to have greater academic impact in the years to come. The technological roadmap presented in this study may assist professionals and practitioners in prioritizing strategic R&D activities and technology investment decisions. Furthermore, policymakers and regulatory bodies may leverage the insights and roadmap in designing policies and regulatory frameworks to facilitate responsible and sustainable AI-driven drug discovery ecosystems.

Limitations

Although the present study provides a comprehensive analysis of patents and develops key insights for researchers, professionals, and policymakers, the analysis is limited to patent data from a single source. Hence, future work may integrate patent data from multiple sources and incorporate other data, such as academic literature, industrial reports, and conference proceedings, for a holistic and multidimensional understanding of the domain. Moreover, as the current study uses a single data analysis method based on STM, future studies may complement the findings through stakeholder consultation and data or method triangulation.

Conclusion

This study employed STM to systematically examine patents related to AI-driven drug discovery and uncover the latent technological themes shaping the domain. The findings reveal a diverse innovation landscape characterized by both established and emerging technological trajectories, reflecting the dynamic evolution of AI applications across the drug discovery pipeline. By classifying technological themes according to their developmental trajectories, the study provides a data-driven roadmap that can support strategic decision-making by firms, investors, and policymakers. As AI technologies continue to mature, future research may extend this work through cross-database analyses, ecosystem-level investigations, and longitudinal assessments of emerging technological domains within drug discovery.

References

- Al-Sehemi, A. G., Chinnam, S., Pannipara, M., Yallur, B. C., Chigurupati, S., Gowda, B. H. J., Paul, K., Shivanna, C. R., & Vardharajula, V. R. (2022). Synthesis and in silico studies of novel anti-SARS-CoV sulfonamides as potential inhibitors against COVID-19 protein target: SARS-CoV-2 main protease (Mpro) (Patent No. WO 2022/049567 A3).
- Askr, H., Elgeldawi, E., Aboul Ella, H., Elshaier, Y. A. M. M., Gomaa, M. M., & Hassanien, A. E. (2023). Deep learning in drug discovery: An integrative review and future challenges. *Artificial Intelligence Review*, 56(7), 5975–6037. <https://doi.org/10.1007/s10462-022-10306-1>

- Chang, Y., Hawkins, B. A., Du, J. J., Groundwater, P. W., Hibbs, D. E., & Lai, F. (2023). A Guide to In Silico Drug Design. *Pharmaceutics*, 15(1). <https://doi.org/10.3390/pharmaceutics15010049>
- Chen, W., Liu, X., Zhang, S., & Chen, S. (2023). Artificial intelligence for drug discovery: Resources, methods, and applications. *Molecular Therapy - Nucleic Acids*, 31, 691–702. <https://doi.org/10.1016/j.omtn.2023.02.019>
- Choi, H., & Woo, J. (2022). Investigating emerging hydrogen technology topics and comparing national level technological focus: Patent analysis using a structural topic model. *Applied Energy*, 313, 118898. <https://doi.org/10.1016/j.apenergy.2022.118898>
- Chuprat, S., Novianto, E. H. D., Matsuura, Y., Mahdzir, A. M., & Harun, A. N. (2025). A closer look on patent analytics through systematic literature review. *Management Review Quarterly*, 75(4), 3035–3065. <https://doi.org/10.1007/s11301-024-00452-x>
- Dharmasivam, M., Azad, M. G., Richardson, V., Kaya, B., & Richardson, D. R. (2026). Artificial intelligence–assisted drug discovery in 2025: Faster, but is it better? The robots are coming, look out! *Pharmacological Reviews*, 78(1), 100103. <https://doi.org/10.1016/j.pharmr.2025.100103>
- Dharmasivam, M., Kaya, B., Akinware, A., Azad, M. G., & Richardson, D. R. (2026). Leading artificial intelligence–driven drug discovery platforms: 2025 landscape and global outlook. *Pharmacological Reviews*, 78(1), 100102. <https://doi.org/10.1016/j.pharmr.2025.100102>
- Du Jun, Wang Xia, Xu Junrui, & He Shuche. (2025a). *Enzyme modification method based on artificial intelligence* (Patent No. CN 119649907 A). <https://lens.org/159-418-222-319-346>
- Du Jun, Wang Xia, Xu Junrui, & He Shuche. (2025b). *Molecular modification method for glycosyltransferase based on artificial intelligence* (Patent No. CN 119724342 A). <https://lens.org/172-893-692-334-560>
- Du, S., Hu, X., Li, P., Xu, S., Kim, M., Liu, X., & Zhan, P. (2026). Antiviral drug discovery and development: Challenges and future directions. *Signal Transduction and Targeted Therapy*, 11(1), 69. <https://doi.org/10.1038/s41392-025-02539-7>
- Dwivedi, Y. K., Sharma, A., Rana, N. P., Giannakis, M., Goel, P., & Dutot, V. (2023). Evolution of artificial intelligence research in *Technological Forecasting and Social Change*: Research topics, trends, and future directions. *Technological Forecasting and Social Change*, 192, 122579. <https://doi.org/10.1016/j.techfore.2023.122579>
- Feinerer, I., Hornik, K., & Meyer, D. (2008). Text mining infrastructure in R. *Journal of Statistical Software*, 25(5), 1–54. <https://doi.org/10.18637/jss.v025.i05>
- Fu, C., & Chen, Q. (2025). The future of pharmaceuticals: Artificial intelligence in drug discovery and development. *Journal of Pharmaceutical Analysis*, 15(8), 101248. <https://doi.org/10.1016/j.jpha.2025.101248>
- Gupta, S., Alphonse, A. S., & Kavitha, D. (2026). Uncertainty-aware hybrid deep generative framework for robust and explainable drug discovery. *BMC Bioinformatics*, 27(1), 84. <https://doi.org/10.1186/s12859-026-06405-3>

- Hasnain, A., Shabbir, M. A., Ullah, N., Kashif, A., & Shakeel, A. (2025). AI-driven in silico discovery, design, and evaluation of novel antiviral candidates targeting Marburg virus nucleoprotein for advanced therapeutic development. *Letters in Drug Design & Discovery*, 22(3), 100030. <https://doi.org/10.1016/j.lidd.2025.100030>
- Jarallah, S. J., Almughem, F. A., Alhumaid, N. K., Fayez, N. A., Alradwan, I., Alsulami, K. A., Tawfik, E. A., & Alshehri, A. A. (2025). Artificial intelligence revolution in drug discovery: A paradigm shift in pharmaceutical innovation. *International Journal of Pharmaceutics*, 680, 125789. <https://doi.org/10.1016/j.ijpharm.2025.125789>
- Jim Abraham, Wolfgang Michael Korn, & David Spetzler. (2025). *Pan-cancer platinum response predictor* (Patent No. JP 2025069209 A). <https://lens.org/195-177-928-524-377>
- Jürgens, B., & Herrero-Solana, V. (2015). Espacenet, Patentscope and Depatisnet: A comparison approach. *World Patent Information*, 42, 4–12. <https://doi.org/10.1016/j.wpi.2015.05.004>
- Katuwal Gajendra Jung, Conroy Bryan, & Rubin Jonathan. (2024). *Dynamic and locally-faithful explanation of machine learning models for precision medicine* (Patent No. US 12183465 B2). <https://lens.org/088-891-445-880-683>
- Khanna, I. (2012). Drug discovery in pharmaceutical industry: Productivity challenges and trends. *Drug Discovery Today*, 17(19), 1088–1102. <https://doi.org/10.1016/j.drudis.2012.05.007>
- Kim Joonghee & Cho Youngjin. (2025). *Artificial intelligence-based biomarker selection device and method* (Patent No. US 2025/0384985 A1). <https://lens.org/025-136-599-481-779>
- Lei Zhou, Li Jiabo, Shen Wenfeng, & Chen Si. (2025). *Natural product-protein interaction prediction method based on deep learning* (Patent No. CN 120375972 A).
- Luo, J., Wu, M., Gopukumar, D., & Zhao, Y. (2016). Big Data Application in Biomedical Research and Health Care: A Literature Review. *Biomedical Informatics Insights*, 8, BII.S31559. <https://doi.org/10.4137/BII.S31559>
- Mishra, B., & Bonde, G. V. (2022). Patent Searching. In V. A. Saharan (Ed.), *Computer Aided Pharmaceutics and Drug Delivery: An Application Guide for Students and Researchers of Pharmaceutical Sciences* (pp. 473–503). Springer Nature Singapore. https://doi.org/10.1007/978-981-16-5180-9_15
- Munos, B. (2009). Lessons from 60 years of pharmaceutical innovation. *Nature Reviews Drug Discovery*, 8(12), 959–968. <https://doi.org/10.1038/nrd2961>
- Mustafa, Y. F. (2026). Artificial intelligence-driven discovery of coumarin-based therapeutics: Revolutionizing target identification and validation. *Pharmaceutical Science Advances*, 4, 100119. <https://doi.org/10.1016/j.pscia.2026.100119>
- Ning Xia & Chen Ziqi. (2025). *Generative ai methods and systems for small molecule structure generation* (Patent No. WO 2025/038843 A1). <https://lens.org/044-827-984-269-290>
- Nunkoo, R., Sharma, A., So, K. K. F., Hu, H., & Alrasheedi, A. F. (2025). Two decades of research on customer satisfaction: Future research agenda and questions. *International Journal of*

Contemporary Hospitality Management, ahead-of-print(ahead-of-print).
<https://doi.org/10.1108/IJCHM-03-2024-0436>

- Odaibo Stephen Gbejule. (2025). *Diffusion-based generative ai methods for protein and drug design* (Patent No. US 2025/0210130 A1). <https://lens.org/084-210-820-048-353>
- Oliveira, R. I., Pereira, T. O., Abbasi, M., Salvador, J. A. R., & Arrais, J. P. (2025). Deep learning for small-molecule drug discovery: From molecular design to clinical translation. *Journal of Pharmaceutical Analysis*, 101533. <https://doi.org/10.1016/j.jpha.2025.101533>
- Ozcan, S., Homayounfard, A., Simms, C., & Wasim, J. (2022). Technology Roadmapping Using Text Mining: A Foresight Study for the Retail Industry. *IEEE Transactions on Engineering Management*, 69(1), 228–244. <https://doi.org/10.1109/TEM.2021.3068310>
- Packer Jonathan S, Emdin Connor, Dewey Frederick, & Black Mary Helen. (2025). *In Silico Drug Development And Determination Of Drug Activity* (Patent No. AU 2023/379456 A1).
- Paul, D., Sanap, G., Shenoy, S., Kalyane, D., Kalia, K., & Tekade, R. K. (2021). Artificial intelligence in drug discovery and development. *Drug Discovery Today*, 26(1), 80–93. <https://doi.org/10.1016/j.drudis.2020.10.010>
- Paul, S. M., Mytelka, D. S., Dunwiddie, C. T., Persinger, C. C., Munos, B. H., Lindborg, S. R., & Schacht, A. L. (2010). How to improve R&D productivity: The pharmaceutical industry's grand challenge. *Nature Reviews Drug Discovery*, 9(3), 203–214. <https://doi.org/10.1038/nrd3078>
- Pipada, V. S., Bikkina, D. J. B., Joshi, S. K., Tharigoppala, N. R., Zade, S., & Vooturi, R. (2026). Artificial intelligence accelerates drug discovery and enhances commercialization efficiency. *Discover Artificial Intelligence*, 6(1), 173. <https://doi.org/10.1007/s44163-026-00859-3>
- Probert Christopher, Zachary Ryan Mccoy, Koller Daphne, & Anna Shcherbina. (2025). *Machine learning based predictive biomarker discovery and patient stratification using standard care data* (Patent No. CN 120958527 A). <https://lens.org/195-301-948-930-95X>
- Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). stm: An R Package for Structural Topic Models. *Journal of Statistical Software*, 91(2), 1–40. <https://doi.org/10.18637/jss.v091.i02>
- Roney, M., & Mohd Aluwi, M. F. F. (2024). The importance of in-silico studies in drug discovery. *Intelligent Pharmacy*, 2(4), 578–579. <https://doi.org/10.1016/j.ipha.2024.01.010>
- Sakaoka, S., & Kano, S. (2025). Quantitative insights on artificial intelligence in the pharmaceutical industry: A patent-basis analysis of technological trends and key players. *World Patent Information*, 82, 102381. <https://doi.org/10.1016/j.wpi.2025.102381>
- Scannell, J. W., Blanckley, A., Boldon, H., & Warrington, B. (2012). Diagnosing the decline in pharmaceutical R&D efficiency. *Nature Reviews Drug Discovery*, 11(3), 191–200. <https://doi.org/10.1038/nrd3681>
- Sharma, A., Koohang, A., & Singh, S. P. (2025). Information Security Policy Compliance: A Structured Review Using Scientometric Analysis and Topic Modeling. *Journal of Global Information Management*, 33(1). <https://doi.org/10.4018/JGIM.389715>

- Sharma, A., Rana, N. P., & Nunkoo, R. (2021). Fifty years of information management research: A conceptual structure analysis using structural topic modeling. *International Journal of Information Management*, 58, 102316. <https://doi.org/10.1016/j.ijinfomgt.2021.102316>
- Singh, S., Gupta, H., Sharma, P., & Sahi, S. (2024). Advances in Artificial Intelligence (AI)-assisted approaches in drug screening. *Artificial Intelligence Chemistry*, 2(1), 100039. <https://doi.org/10.1016/j.aichem.2023.100039>
- Singh, S., Singh, S., Koohang, A., Sharma, A., & Dhir, S. (2023). Soft computing in business: Exploring current research and outlining future research directions. *Industrial Management & Data Systems*, 123(8), 2079–2127. <https://doi.org/10.1108/IMDS-02-2023-0126>
- Singh, S., Singh, S., Kraus, S., Sharma, A., & Dhir, S. (2024). Characterizing generative artificial intelligence applications: Text-mining-enabled technology roadmapping. *Journal of Innovation & Knowledge*, 9(3), 100531. <https://doi.org/10.1016/j.jik.2024.100531>
- Sun, T.-H., Wang, C.-C., Wu, Y.-L., Hsu, K.-C., & Lee, T.-H. (2023). Machine learning approaches for biomarker discovery to predict large-artery atherosclerosis. *Scientific Reports*, 13(1), 15139. <https://doi.org/10.1038/s41598-023-42338-0>
- Wan Zhitong. (2025). *Artificial intelligence expert system and method based on cell chip cancer drug screening* (Patent No. CN 120877937 A). <https://lens.org/035-851-826-504-842>
- Wang, R., & Zhuang, C. (2025). Graph neural networks driven acceleration in drug discovery. *Acta Pharmaceutica Sinica B*, 15(12), 6163–6177. <https://doi.org/10.1016/j.apsb.2025.10.011>
- Yang Qing, Shi Changsheng, Ge Ruiquan, Deng Guojian, Lin Yiben, & Wang Changmiao. (2025). *Drug-target interaction prediction method based on graph neural network block* (Patent No. CN 119517226 A). <https://lens.org/070-995-428-787-367>
- Yang Yanqiu, Zhou Yi, Wang Zhengsong, Ma Zhenhe, & Dong Leilei. (2025). *Application of virtual screening compound in preparation of NLRP3 inflammasome inhibitor* (Patent No. CN 121154643 A). <https://lens.org/056-076-165-354-812>
- Zhang, T., Chen, J., Lu, Y., Yang, X., & Ouyang, Z. (2022). Identification of technology frontiers of artificial intelligence-assisted pathology based on patent citation network. *PLOS ONE*, 17(8), e0273355. <https://doi.org/10.1371/journal.pone.0273355>
- Zhang Yuezhou, Sun Shaohua, Liang Fei, Sun Xichen, & Zhang Tinghao. (2024). *Drug screening method and device based on neural network* (Patent No. CN 117708679 A).
- Zheng, L., Cao, J., Jing, L., Kang, D., Wang, Z., & Liu, X. (2026). Convergence of computer-aided drug discovery and artificial intelligence: Towards next-generation therapeutics. *Pharmaceutical Science Advances*, 4, 100100. <https://doi.org/10.1016/j.pscia.2025.100100>
- Zhou Chenxi, Zhu Yanni, Ye Yuchen, Yang Chenrui, Gao Yu, Lu Jianyang, & Ming Dengming. (2025). *Inhibitor drug screening method for targeting KRAS G12D mutation based on machine learning* (Patent No. CN 121191580 A). <https://lens.org/063-054-318-633-502>