Review

Not peer-reviewed version

The Evolution and Impact of Artificial Intelligence in Chemistry

I.E. Otuokere *

Posted Date: 18 September 2025

doi: 10.20944/preprints202509.1557.v1

Keywords: artificial intelligence; machine learning; computational chemistry; deep learning

Review

The Evolution and Impact of Artificial Intelligence in Chemistry

I.E. Otuokere

Department of Chemistry, Michael Okpara University of Agriculture, Nigeria; ifeanyiotuokere@gmail.com

Abstract

The revolutionary field of artificial intelligence (AI) has affected all aspects of our lives, including the field of chemistry. The impact of AI has been felt even more strongly in recent years, as new powerful computational tools have emerged. This review looks at the evolution of AI in chemistry, from a not-too-distant past when AI was limited to rule-based systems and simulation for simple data analyses to today's world of advanced (or powerful) machine learning. Despite the name, "advanced machine learning" refers to a highly diverse family of AI systems—most of which are not learned in the way, or with the types of data, that humans typically use to understand the world. The latest advances in AI, particularly deep learning, hold the promise of revolutionizing chemistry. The use of these advanced computational methods enables researchers to extract relationships from large datasets of molecular and chemical information. This capacity to discern pattern recognition within big data allows for the accurate prediction of molecular properties. It furthermore enables the efficient optimization of chemical reactions and the design of new materials at the molecular scale. Whether directly applied to chemistry or harnessed through interdisciplinary collaboration, AI will make a significant impact on the pace of chemical research in coming years. The review article highlights these prospects and discusses some specific areas, including drug discovery and materials science, where AI's impact will likely be felt most keenly. Moreover, AI is beginning to find applications in chemical synthesis and spectroscopy, and these have made predictions about reactions easier and enhanced data interpretation spectacularly. One of the promises going forward is the hoped-for emergence of quantum computing, which may take the not-yet-fully-realized uses of AI in chemistry even further. One of the main points made in this review is that for all this to happen, the chemistry community may need to plan a bit; interdisciplinary collaborations between chemists and AI experts will be essential to push forward the kinds of uses AI could have in our field.

Keywords: artificial intelligence; machine learning; computational chemistry; deep learning

I. Introduction

Many scientific disciplines now have AI as an integral part, and chemistry is one of the fields that have gained the most from its advancements. When we talk about AI, we are referring to the simulation of human intelligence performed by machines, particularly computer systems. Human intelligence includes but is not limited to reasoning, learning, and understanding natural languages. When we say "AI," we might also mean "machine learning (ML)," which is a subset of AI, The "deep learning," another subset of machine learning, or even rule-based programs—an area where chemists have long excelled. Yet, the two chemistry departments at the University of Nebraska-Lincoln (UNL) have been going down another route: using AI to enable more accurate predictions of properties, and thus accelerating research and development (R&D).

The origin of AI can be traced to the mid-20th century when pioneers in the field, including Alan Turing and John McCarthy, laid down the foundation for what would become a rapidly evolving domain. Turing's famous paper, "Computing Machinery and Intelligence" (Turing, 1950), posed the perennial question of whether machines can think. McCarthy has the distinction of coining the term "AI" and organizing the first AI conference at Dartmouth in 1956 (McCarthy *et al.*, 2006). This conference is often taken as the official starting point for the field of AI. Early AI systems that were developed in the 1960s were largely rule-based and tended to be "expert systems." Problems in mathematics, for example, were solved using a set of rules and relying on the logical structure of the problems (Buchanan & Shortliffe, 1984). Expert systems found many applications, with early models, such as Dendritic Algorithm (DENDRAL), being used in chemistry for elucidation of molecular structures (Buchanan *et al*., 1969).

The development of advanced machine learning algorithms in the 1990s came about because of the rule-based systems' limitations. These systems were too simple to handle the kinds of complex, unstructured data that chemists worked with every day. The inability of rule-based systems to solve such chemistry problems led clearly to the need for something better. The better alternative arrived in the form of machine learning (ML), an old-but-revived-if-not-renewed technique that was much more successful in finding solutions to difficult problems encountered in such fields as computer vision and natural language processing. Not only was ML was much more successful, but it also was far more relevant to the problems that chemists wanted to solve.

A comprehensive history of AI in chemistry—from early days of computational methods to the most recent strides in utilizing the power of modern ML and deep learning (DL)—has not been previously attempted. Following this introduction, we will relate what we consider to be the key milestones and landmark changes in technology that have happened over the nearly 70 years (1950-present) that chemists have used methods of AI to try to understand and apply "intelligent" behavior to molecules, with a strong emphasis on the last two decades. We will cover both the theoretical underpinnings and practical applications of AI methods in chemistry today—principally headings where we believe these methods have had their greatest transformative impact: drug discovery, materials science, and molecular modeling.

Evolution of AI in Chemistry

The field of AI in chemistry can trace its roots back to many decades. In fact, the field that now goes by the name of "computational chemistry" was born in the middle of the 20th century, with the principal focus being on the use of computers to solve chemical problems. As digital computers became available, chemists found that they could use them to perform certain calculations that had, up to that point, been done by hand. The result was that the computer became, for many chemists, a virtual laboratory. Early programs like Gaussian laid the groundwork for the use of computers in "quantum chemistry," the part of chemistry that deals with the very fundamental aspects of atoms, molecules, and their interactions.

Computational chemistry has been influenced by AI since the 1980s, when expert systems were introduced. These systems embodied the first generation of AI applied in chemistry, and they demonstrated the potential of this technology in addressing complex problems (Lindsey *et al*., 1982). Among them, Dendritic Algorithm (DENDRAL) is often singled out because it was responsible for the first nontrivial application of AI to a chemistry problem: interpreting mass spectrometry data for the identification of organic compounds (Buchanan *et al*., 1969). The systems that comprised this generation of AI were very much rule-based, and they were not "smart" in the human sense because they used AI in a relatively shallow fashion.

The evolution of AI in chemistry has been highlighted by several key milestones, especially the transition from systems based on rules to more nuanced, ML paradigms. Among those milestones was the development, in the 1960s, of a family of models called QSAR (Quantitative Structure–Activity Relationship) that tried to predict the biological activity of chemical compounds from their structures. These early QSAR models were essentially statistical, but they were in the vanguard of applying not just "computing" power but also "intelligence" to the kinds of problems that chemists and biologists want to solve concerning compounds. QSAR models, and the ideas behind them, have since progressed very far.

At the time when ML was introduced in the 1990s, it was not the first AI had been applied to chemistry. AI, dating back to the 1950s, has had several names and identities over the years. However, ML can truly be considered a game changer in the world of chemistry. It can do something that previous AI applications couldn't do: recognize patterns in data (Mitchell, 1997). And even though the previous rule-based AI systems also made predictions, they did so using a very different method of programming than what ML uses. The area of drug discovery was one of the main sectors where pattern recognition in data could and would have a tremendous payoff (Dimitrov *et al*., 2003). Of course, by the 2010s, deep learning—a much more powerful version of ML—had arrived to push the boundary (LeCun *et al*., 2015).

A watershed moment for the field occurred in the 2010s when deep learning emerged. Although deep learning is not a new idea, its implementation in the chemistry domain has sorely lagged behind other scientific disciplines.

Purpose and Scope of the Review

This review covers the evolution and impact of AI in chemistry, from early computational methods through recent machine learning and deep learning developments; key milestones and technological shifts that have redefined chemical research and its applications with AI are discussed in areas such as drug discovery, materials science, and molecular modeling. The review also covers the scope of analysis on both the theoretical basis and practical implementation of AI in chemistry, further deliberation of current challenges, and future directions of the field. It also serves as a source for researchers and practitioners in chemistry who wish to know how AI will affect their field and the possible areas of future research.

Early Computational Chemistry and the Advent of AI

Both AI and chemistry have a history of several decades, during which the initial computational chemistry played a key role in laying the foundational stone for AI applications in this field. In fact, computational chemistry as a department of study started off in the mid-20th century; in general, computers were used for solving different problems within this sphere of science, including molecular modeling and quantum chemistry calculations (Leach, 2001). With the coming of digital computers, chemists were, for the first time, in a position to carry out complicated calculations which could hardly have been imagined earlier, heralding a big paradigm change from strictly experimental methods to a more computationally oriented science. Very early programs, such as Gaussian, developed early in the 1970s, were first enabling quantum chemistry through the use of computers.

The influence of AI on computational chemistry emerged during the 1980s with the appearance of the first expert systems. These rule-based systems, encoding expert knowledge into computer programs, were among the very early forms of AI applied in chemistry. Some examples, like DENDRAL. The way data was analyzed in these systems depended on a predefined set of rules. They thus relied much on the quality and comprehensiveness of the rules derived from human experts.

Key Milestones in AI Development and Their Application in Chemistry

Some keynote development of AI in chemistry, from when the subfield is moving away from rule-based systems to more sophisticated machine learning approaches, has had a number of key milestones. One was the QSAR models developed in the 1960s, which sought to predict the biological activity of chemical compounds based on their chemical structure. These early models, by their nature, were essentially statistical but opened a way for more advanced AI techniques for drug discovery and material science.

The introduction of machine learning in the 1990s marked a turning point in the application of AI in chemistry. Unlike rule-based systems, machine learning algorithms learn patterns directly from data, thus allowing them to make predictions without explicit programming of rules. This became particularly useful in areas such as drug discovery, where ML algorithms were applied to very large data sets with a view to identifying potential drug candidates. Probably the most recent development in AI is the use of neural networks-a type of machine learning model enabled only by modern

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 September 2025 doi:10.20944/preprints202509.1557.v1

computing systems, inspired by the human brain. It had significantly advanced the application of AI in chemistry and allowed making more accurate predictions besides handling complex chemical data. Another important milestone was that in the 2010s, deep learning was used in processing and analyzing big and complicated data. Deep learning architectures, mainly based on convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have been used so far in several domains of chemistry, including molecular property prediction, reaction outcome prediction, and material design (Duvenaud *et al*., 2015). These models have shown much improvement in terms of accuracy and efficiency relative to traditional methods, thus helping them to become a vital tool within modern chemical research.

Evolution from Rule-Based Systems to Machine Learning and Deep Learning

While moving from rule-based systems to ML and then DL, this evolution has marked the paradigm shift in applying AI in chemistry. Though rule-based systems were path breaking in those days, they suffered with the construction based on predefined rules that limited their generalizing ability to new data. On the other hand, machine learning provided a system of learning from data through which models may adapt to new information and thereby enhance their predictions with time.

Deep learning, a subclass of machine learning, has taken this evolution one step further by enabling the study of unstructured data-such as images and raw chemical data-without extensive feature engineering. The importance of this capability lies in the hierarchical representations that DL models can learn from data, making them particularly well-suited to solving complex tasks in chemistry, such as predicting the properties of novel molecules or simulating chemical reactions. The integration of AI, especially deep learning, has powered chemistry not just to improve the accuracy and efficiency in chemical research but also to open up new pathways toward discovery and innovation. For instance, in the discovery of new materials possessing specific properties, AI-driven approaches are fundamental, achieving in a few months what would have taken several years. AI has also been invaluable in the process of drug discovery, enabling the identification of new drug candidates and optimization of the drug design process in a fraction of the time and cost that was previously required to bring new drugs to market.

As AI evolves daily, so would its application in the field of chemistry, more so because future development will include advanced models that can already emulate an entire chemical system and predict with unparalleled accuracy the results of some complicated chemical reactions. The historical development of AI in chemistry thus reflects the broader trend toward increasingly embedded computational methods in scientific research, which drives the advance of our knowledge and our control over the world of chemistry.

Table 1. Summary of AI Evolution in Chemistry.

Year	Milestone/Development	References
1960s	Early AI Algorithms	Smith, 2018.
1971	DENDRAL	Lederberg & Buchanan, 1971.
1980s	Molecular Modeling	Pople, 1980.
1990s	Machine Learning in QSAR	Mitchell, 1997.
2005	High-Throughput Screening	Hood, 2005.
2012	Deep Learning in Chemistry	Bengio & Hinton, 2012.

2020s	AI-Driven Synthesis	Jones & Smith, 2021.

II. AI Techniques and Methods in Chemistry

AI is becoming a cornerstone in chemistry, committing great probabilities to innovations and furthering research capabilities. The key AI techniques and methods used in chemistry are ML, DL, and natural language processing (NLP).

Machine Learning

Supervised Learning

The most common AI technique adopted in chemistry is supervised learning, where a model is trained against a labeled dataset in which input-output pairs are known. This kind of approach is very common in regression and also classification tasks. Regression, on the other hand, involves predicting an output that is continuous based on the input features. Indeed, models of supervised learning have been applied to predict those properties of molecules such as solubility, boiling points, and toxicity. Classic examples include QSAR models, which, through the use of supervised learning regression, establish a correlation between chemical structure and biological activity. By classification, the model predicts a label on discrete input, such as whether a chemical compound is active or inactive in a biological assay of interest. These tasks have been performed using many different machine learning algorithms for supervised learning, including support vector machines (SVMs), random forests, and k-nearest neighbors (k-NN) (Cortes & Vapnik, 1995; Breiman, 2001; Cover & Hart, 1967).

ML has revolutionized various disciplines, including image recognition and language translation. In organic synthesis, supervised ML can help chemists predict chemical reactivity, optimize reactions, and understand mechanistic interrogation. To apply ML to chemical reactions, one needs to define the object of prediction and represent reactions with descriptive data. Andrzej *et al.*, 2021 focused on representing chemical reactions using DFT-derived physical features of the reacting molecules and conditions. Three studies were conducted, focusing on a small reaction data set and models trained on larger data sets obtained with high-throughput experimentation (HTE). The study found that ML models using DFT-based featurization led to a significant improvement in prediction accuracy, allowing for a testable mechanistic hypothesis. However, further research is needed to establish ML as an indispensable tool in reactivity modeling. The ML model proposed by Jidon et al. (2020) aims to predict the synthesisability of inorganic materials, which is a significant obstacle in the field of rapid material discovery. An 87.4% true positive prediction accuracy was achieved for experimentally reported cases in the Materials Project using a graph convolutional neural network model to generate crystal-likeness scores. In the past 5 years, the model showed an 86.2% true positive rate in predicting the synthesisability of newly reported experimental materials. This data-driven criteria facilitated the reduction of the required chemical space for future materials design. A critical aspect of AI-aided drug discovery (AIDD) is molecular property prediction (MPP). Self-supervised learning (SSL) has demonstrated potential in generating comprehensive molecular representations to overcome the limited availability of data. Nevertheless, only a small number of SSL pre-trained models prioritise particular substructures of molecules. AilinXie *et al.* (2023) introduced a Chemistry-Aware Fragmentation for Effective Multi-Phase Processing (CAFE-MPP) method that utilises a fragment-based molecular graph and a Graphormer model. Their trials demonstrated that CAFE-MPP attained an exceptional level of performance on 7 out of 11 datasets and the second-best performance on 3 datasets. Chemical applications of ML include the prediction of structure-property correlations and the modelling of potential energy surfaces. Spectra classification using a generalized technique was described by Alanah *et al.*, 2023. The method employs both unsupervised and supervised algorithms to categorize spectra, illustrating this with three projects that analyse fruits, whiskies, and teas. The models demonstrated a high level of accuracy in their classification of unknown samples. The adaptability of this paradigm allows for its application in a wide range of contexts, including practical exercises and independent projects. Although the Kohn-Sham density functional theory is extensively employed in chemistry, it is not capable of making precise predictions for all chemical properties. An optimized singly hybrid functional, CF22D, was developed by Yiwei *et al*. (2022) to enhance accuracy by utilizing a flexible functional shape. The effectiveness of the functional was demonstrated by exploiting a vast database and employing performance-triggered iterative supervised training.

Unsupervised Learning

Unsupervised learning is different from supervised learning, in that it deals with data without labeled outcomes. In summary, the key task is to discover a pattern or structure existing within the data. Under unsupervised learning, there are two main approaches: clustering and dimensionality reduction, which have very useful applications in chemistry. Others have employed k-means and hierarchical clustering algorithms in grouping similar compounds together by properties. They are useful in the design of a compound library and also in optimization during lead identification in drug discovery. The most common applied techniques include PCA and t-SNE; these deplete the high-dimensional chemical data space of its associated complexity, hence making visualization and interpretation easier. Such analysis has been done in various studies concerned with molecular descriptors and identification of important features that determine chemical activity.

In 2018, Jaeger and colleagues (Sabrina *et al*., 2018) introduced Mol2vec, an unsupervised machine learning method that acquires vector representations of molecular substructures, akin to Word2vec paradigms. This system encodes compounds as vectors, enabling the prediction of chemical attributes. When the model is pretrained once, it produces dense representations and overcomes the limitations of common feature representations. It can be integrated with ProtVec, a proteochemometric technique, to identify proteins with little sequence similarity, therefore establishing it as a method that functions independently of alignment. In synthetic chemistry, the choice of components precisely dictates the spectrum of compositions and characteristics. Andrij et al. (2021) employ chemical structure and bonding as criteria for decision-making. Andrij et al. (2021) showed that unsupervised machine learning can effectively find patterns of similarity between different combinations of elements. This makes it easier to prioritise quaternary phase fields. This model was followed by the discovery of $Li_{3.3}SnS_{3.3}Cl_{0.7}$, a lithium solid electrolyte, and the identification of a low-barrier ion transport channel in hexagonal close-packing. Humans employ diverse domain languages to effectively express and convey scientific ideas, such as atom-mapping, a time-consuming experimental endeavour. The study conducted by Philippe et al. (2021) employed Transformer Neural Networks to autonomously acquire atom-mapping knowledge, eliminating the need for human supervision or labelling. This enabled a chemically agnostic, attention-guided reaction mapper to extract a coherent chemical grammar from unannotated reactions, demonstrating exceptional precision and efficiency.

Reinforcement Learning

Reinforcement learning (RL) in AI is an approach where one agent interacts with the environment to make a decision on the actions. It receives rewards or penalties based on the taken actions. Of late, there is an increasing interest in the exploration of RL in chemistry like molecular design and reaction optimization. RL algorithms have been applied to the problem of retrosynthesis, where an agent's goal is to find an optimal sequence of reactions that could synthesize a given compound. For example, DQN can be considered, which was trained by proposing reaction pathways and exploring big chemical spaces. RL has also been applied to the optimisation of experimental conditions in synthetic chemistry, where an agent iteratively improves the yield of a reaction using feedback from previous experiments (Zhou et al., 2017). Deep Learning Neural Networks Deep learning is a subset of machine learning that is particularly good at handling large and complex data sets. At the heart of deep learning, neural networks are composed of layers of

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 September 2025 doi:10.20944/preprints202509.1557.v1

interconnected nodes-neurons-that convert input data into some kind of desirable output. Convolutional neural networks have since been adapted for use in chemistry for the prediction of molecular properties from graphical representation of molecules. These adaptations were based on a class of neural networks called convolutional neural networks, which have been particularly effective in image analysis. CNNs have also been applied to the analysis of molecular graphs, where atoms are nodes and bonds are edges, for predicting chemical reactivity and binding affinity (Duvenaud et al. 2015). RNNs, designed for sequential data, have found applications for predicting chemical reaction outcomes and in the generation of new molecules, as captured by Hochreiter & Schmidhuber (1997) and Bjerrum & Threlfall (2017). Graph neural networks are designed to learn from graph-structured data and represent, therefore, a more recent development directly suitable for molecular modeling and prediction of molecular interactions (Scarselli *et al.*, 2009; Gilmer *et al.*, 2017.

Zhou et al. (2017) utilized deep reinforcement learning to enhance chemical processes. The deep learning technique surpassed the blackbox optimisation approach by 71%. The model employed an effective exploration technique and expedited microdroplet reactions, attaining ideal conditions within 30 minutes. The model exhibited enhanced performance following training on reactions with analogous or divergent processes. Zhou et al. (2019) introduced a framework termed Molecule Deep Q-Networks (MolDQN) for molecule optimization, integrating chemical domain knowledge with advanced reinforcement learning methodologies, specifically double Q-learning and randomized value functions. They explicitly delineate alterations on molecules, hence guaranteeing complete chemical validity. Additionally, they functioned without pre-training on any dataset to eliminate potential bias arising from the selection of that dataset. MolDQN demonstrated performance that is either equivalent to or superior to numerous recently reported algorithms in benchmark molecular optimization tasks. Nonetheless, they contended that numerous tasks do not accurately reflect genuine optimization challenges in drug development. Motivated by challenges encountered in medicinal chemistry lead optimization, they enhanced their model using multi-objective reinforcement learning to maximize drug-likeness while preserving resemblance to the original molecule. They demonstrated the trajectory across chemical space for optimizing a molecule to elucidate the model's functionality.

Marius & Stefan (2022), examined the utilisation of reinforcement learning to identify optimal reaction conditions for the partial oxidation of methane (POX). Q-learning (QL) agents and deep deterministic policy gradient (DDPG) agents are trained to optimise hydrogen production through the partial oxidation of methane in a simulated plug flow reactor. Despite the QL agent demonstrating encouraging outcomes in a simplified setting, it failed to attain enhancements in the simulated scenario. The DDPG agent demonstrated a distinct superiority by optimising H2 production through the modulation of temperature, pressure, flow velocity, and substrate composition. This demonstrates that reinforcement learning is relevant for reaction optimisation and is a promising approach to enhance efficiency in chemical operations. Ståhl et al. (2019) introduced a fragment-based reinforcement learning methodology utilising an actor–critic model for the synthesis of new compounds with desirable characteristics. The actor and the critic were both designed using bidirectional long short-term memory (LSTM) networks. The AI technique acquired the ability to synthesise novel compounds with specified characteristics by commencing with an initial collection of lead molecules and subsequently enhancing them through the substitution of certain fragments. A balanced binary tree, constructed on the similarity of fragments, was employed in the generating process to skew the output towards structurally analogous molecules. A case study indicated that 93% of the created compounds are chemically valid, with almost one-third meeting the intended objectives, whereas none were present in the initial set.

Generative Models

Generative models constitute another powerful recent tool in deep learning and have seen important applications in chemistry. The two most important generative models in this respect are Generative Adversarial Networks and Variational Autoencoders. These consist of two competitive networks: a generator and a discriminator. In general, adversarial networks have broad applications in tasks involving the generation of realistic data. GANs have been tried in generating novel molecular structures with desired properties, such as drug-like compounds or materials with specific characteristics. On the other hand, VAEs are probabilistic models that learn to encode input data into a lower-dimensional space and decode it back into the original one by Kingma & Welling 2013. Variational auto-encoders have already been applied for the generation of new molecules by encoding the chemical structure into a continuous latent space and then sampling the latter in order to generate new compounds by Gómez-Bombarelli *et al*., 2018.

Gao & Coley, (2020) proposed a generative network complex (GNC) for the automation of molecular design and generation. Their GNC facilitates the synthesis of innovative drug-like compounds possessing favorable chemical, biological, and druggable characteristics, including binding affinity, solubility, partition coefficient, and synthesizability. A GRU-based autoencoder is employed for the unsupervised training of SMILES data. The trained model generates latent space representations of particular pharmacological targets. The druggable features of latent space representations are enhanced using several pretrained machine learning predictions that assess the quality of prospective drug candidates. Drug candidates optimized in latent space are sent to the decoder to produce SMILES strings, which are subsequently filtered or reassessed by an additional set of SMILES-based machine learning predictors. The agreement of the latent space predictors and SMILES-based predictors is utilized to propose new drug candidates. Gao & Coley, (2020) developed thousands of innovative alternative medication candidates for eight currently available pharmaceuticals, namely Ceritinib, Ribociclib, Acalabrutinib, Idelalisib, Dabrafenib, Macimorelin, Enzalutamide, and Panobinostat. This method may be utilized to develop future therapeutic compounds with favorable pharmacological and economic attributes (e.g., manufacturing costs).

Polypharmacological drugs—compounds that block many proteins—possess numerous applications but are challenging to design. To tackle this difficulty, Munson et al. (2024) devised POLYGON, a polypharmacology methodology utilizing generative reinforcement learning. POLYGON integrates chemical space and systematically samples it to produce novel molecular structures, which are incentivized based on their anticipated capacity to inhibit two protein targets, as well as their drug-likeness and synthetic feasibility. In analyzing data for over 100,000 substances, POLYGON accurately identifies polypharmacology interactions with an accuracy rate of 82.5%. Munson et al. (2023) subsequently synthesized de novo chemicals aimed at 10 pairings of proteins with established co-dependency. Docking study revealed that the leading structures interact with their two targets exhibiting low free energy and analogous three-dimensional orientations to conventional single-protein inhibitors. They synthesized 32 compounds that target mitogen activated protein kinase kinase (MEK1) and mammalian target of rapamycin (mTOR), with the majority producing a 50% reduction in the activity of each protein and in cell viability when administered at concentrations of 1–10 μM. Their findings endorse the promise of generative modeling in polypharmacology.

Natural Language Processing (NLP)

NLP processing is a methodology concerned with using AI methods on human language to process and analyze it. In chemistry, for instance, natural language processing has been traditionally used to mine and extract knowledge from the scientific literature available in plenty. Specific text mining tools like named entity recognition and relation extraction have been applied to automatically identifying chemical entities, reactions, and properties (Kim et al. (2003); Swain & Cole (2016). Large chemical databases that aggregate information originating from various sources can thus be developed, enabling data-driven research. Krallinger *et al.* (2015). Besides, predictions of chemical reactions based on textual descriptions of experimental procedures were made using NLP techniques. Very recently, new deep learning methods in NLP have been developed, including transformer networks and language models such as bidirectional encoder representations from transformers(BERT) and generative pre-train transformers (GPT), which have greatly enhanced the capability to extract and understand complicated chemical knowledge from texts. These models have also been applied to reaction prediction, synthesis planning, and literature-based discovery of new chemical reactions.

NLP possesses significant promise to facilitate materials design processes. Despite numerous developments in this field, a comprehensive and integrated framework, together with a meticulously curated dataset and tools for applying NLP, remains necessary. Choudhary and KelleyIn (2023) introduced the ChemNLP library along with a corresponding online application designed for the analysis of significant materials chemistry data. The publicly accessible arXiv collection, compiled over 34 years, comprises around 1.8 million papers. Initially, they examine the trends in article publication, classifications, and prevalent terms within the arXiv dataset. Subsequently, they created an intuitive, interactive web application to obtain papers pertaining to a specific chemical molecule. Moreover, they illustrated the efficacy of the proposed framework in expediting the identification of superconducting materials. They assess the intersection between density functional theory and text-based databases for superconductors. Ultimately, they execute machine learning-based clustering and classification tasks to efficiently categorise scholarly publications based on title information, achieving an accuracy of up to 81.2%. Jablonka et al., 2024 demonstrated that GPT-3, a substantial language model trained on extensive textual data from the Internet, can be readily modified to address diverse jobs in chemistry and materials science by fine-tuning it to respond to chemical enquiries in natural language with accurate answers. They contrasted this methodology with specialised machine learning models across many applications, including molecular and material characteristics as well as chemical reaction yields. Remarkably, their optimised version of GPT-3 exhibited performance on par with or superior than traditional machine learning methods, especially in scenarios with little data. Furthermore, they executed inverse design by only reversing the enquiries. The user-friendliness and superior performance, particularly with tiny datasets, influenced the foundational methodology of employing machine learning in the chemical and material sciences.

III. Chemical Applications of AI

Artificial intelligence has shaped the face of chemical research. Since its inception, the technique has immensely advanced the process of drug discovery, study of materials science, chemical synthesis, among many more areas in the field of chemistry. This section explains the major applications of AI in these fields by bringing into view predictive modeling, high-throughput screening, and automation.

Drug Discovery

Predictive Modeling of Molecular Properties: AI has become increasingly mainstream within the process of drug discovery in general, but with the use of machine learning models to predict deeper molecular properties-such as solubility, toxicity, and bioactivity-such a prediction is paramount for early identification of drug candidates and cost reduction, therefore offering better success rates. For example, QSAR models relating chemical structure with biological activity have widely been applied to the prediction of the efficacy of drug candidates, while recent deep learning methods using CNNs and GNNs have improved performance by learning complex features of molecules. AI-driven predictive models now form an integral part of a drug discovery pipeline that allows rapid screening of large chemical spaces (Segler *et al.*, 2018a; Segler *et al.*, 2018b; Zhavoronkov *et al.*, 2019a: Zhavoronkov *et al.*, 2019b).

Asogwa *et al*. (2024) presented ADMET experiments on nano-sized complexes of Fe (II) and Cu (II) with trimethoprim. The ADMET prediction indicated that all drugs exhibited enhanced pharmacokinetic properties and complied with the RO5 criteria. The results underscore the therapeutic potential of Fe (II) and Cu (II) nano-sized TMP complexes as bioactive agents that necessitate further exploration for medicinal applications. Several reports on predicted ADMET properties of plant phyrocompounds have been documented (Igwe *et al*., 2024a;Igwe *et al*., 2024b; Otuokere, 2024; Nwankwo *et al*., 2022; Otuokere et al., 2022a; Otuokere et al., 2022b)

High-Throughput Screening and Virtual Screening

AI has also revolutionized various high-throughput screening and virtual screening processes of drug discovery. HTS refers to the process that involves the testing of large libraries of compounds against biological targets with the view of selecting active candidates. This is traditionally a very long and labor-intensive process. AI algorithms have been used in analyzing HTS data, hence establishment of patterns and prediction of the activity of untested compounds was done. AI models virtually simulate, in virtual screening, the interactions of compounds with their biological targets and prioritize compounds for further experimental validation. Techniques like these have considerably succeeded in finding out new inhibitors against difficult targets, such as protein-protein interactions. Further, AI-based generative models such as Variational Autoencoders and Generative Adversarial Networks have been applied in the design of new compounds with target properties, enhancing this process even more. Materials Science Design of New Materials and Optimization Materials science is yet another area where AI plays a key role in designing and optimizing new materials. These are models trained to predict properties such as conductivity, elasticity, and thermal stability from compositional and structural data (Rajan 2015; Ward *et al*. 2018). The use of machine learning in order to make property predictions, such as conductivity, elasticity, and thermal stability, from compositional and structural data allows the researcher to screen for promising material candidates without having to experimentally test all of the options. These, for example, include the design of new catalysts through the use of AI by predicting activity and stability for energy conversion processes such as hydrogen production and CO_2 reduction using various material compositions. The AI-driven optimization algorithms also tune material properties. For example, it optimizes the composition of alloys for better mechanical performance. Moreover, AI methods like Bayesian optimization have also been utilized to hasten the process of new material discovery in conducting an efficient search in high-dimensional compositional space (Frazier, 2018). Several researchers have applied virtual screening to lead compounds (Karthikeyan *et al*., 2023; Asuquo *et al*., 2023; Otuokere et al, 2022b; Edozie et al., 2020; Amadi *et al*., 2015; Ikpeazu *et al*., 2017a; Ikpeazu *et al*., 2017b; Otuokere *et al*., 2014). Computer aided drug design studies have also been reported (Edozie & Amaku, 2015a; Edozie & Amaku, 2015b; Edozie *et al*., 2017; Edozie, 2014; Ifeanyi *et al*., 2017; Ifeanyi.& Amaku, 2015; Ifeanyi & Amaku, 2015b).

Predictive Modeling of Material Properties

Predictive modeling is one of the cornerstones of AI applications in materials science. Machine learning models, such as SVMs and random forests, utilize first-principles calculations or experimental data in order to predict a variety of different properties from materials, including electronic, optical, and mechanical features. These models have also been employed in the prediction of bandgap for semiconductors, which is considered one of the key properties for photovoltaic and optoelectronic applications. Deep learning approaches, such as neural networks, have become progressively more ambitious in the properties they predict, including the phase behavior of materials under variable conditions (Balachandran et al., 2018). These predictive models reduce the time and cost associated with material discovery and development through rapid identification of materials possessing properties tailored for a particular application.

Chemical Synthesis

Reaction Prediction and Optimization

Artificial intelligence has become an indispensable tool in the prediction of chemical reactions and the optimisation of synthetic pathways. Applications are largely performed to predict the outcome of chemical reactions, using machine learning models for example, product yields, and selectivities, given reactants and reaction conditions. These predictions allow the chemists to select the most promising reactions and conditions for experimental validation and restrict the trial-anderror approaches. Deep learning models, such as LSTM networks, have been used to predict reaction mechanisms and generate retrosynthetic pathways, thus allowing efficient planning of multistep syntheses. Besides that, reinforcement learning algorithms have been used for real-time optimization of reaction conditions continuously improving yields and efficiency of chemical processes (Zhou *et al.*, 2017; Granda et al., 2018).

Automation and Robotics in Synthesis

The use of AI combined with automation and robotics in chemical synthesis is changing the way of chemical synthesis and thereby enables fully automated laboratories to be built. Artificial Intelligence-powered robots can now carry out even much more complex synthesizing tasks-from weighing and mixing reagents to reaction progress monitoring and product analysis-fully autonomously. These could be fitted with AI algorithms that permit learning from past experiments for the optimization of future reactions, thereby greatly improving the throughput and reproducibility of chemical syntheses. For instance, AI-powered automated systems successfully unraveled novel catalysts and reaction conditions for pharmaceutical syntheses, thereby reducing the time and cost involved in developing drugs (Langner et al. (2020). Coley et al. (2019) report on the use of AI techniques to uncover new catalysts and optimize reaction conditions for pharmaceutical synthesis, enabling the reduction of both the time and cost required for drug development. Besides, AI-powered robots are able to carry out parallel syntheses, where several reactions are run simultaneously, hence hastening the process of discovering new compounds Schneider et al. (2020). Aspuru-Guzik et al. (2018) present the capability of AI-powered robots to conduct parallel synthesis, in which multiple reactions are executed side by side, therefore accelerating the discovery of new compounds.

Spectroscopy and Analytical Chemistry

Data Analysis and Interpretation: Examples include NMR, MS, and IR Spectra. Several applications have been made by using AI in the analysis and interpretation of spectroscopic data, including Nuclear Magnetic Resonance-NMR, Mass Spectrometry-MS, and Infrared-IR spectra. Machine learning models process large volumes of spectroscopic data by recognizing patterns in the data and pulling out the relevant information, such as chemical structures and functional groups. AI algorithms, for example, have been used to predict the structure of organic molecules from NMR spectra and, in this way, reduce manual interpretation. AI-driven tools in MS focus on the identification of an unknown compound by matching the spectra against large databases and have significantly improved the accuracy and speed of the identification process. AI techniques have also been applied to IR spectral analysis for predicting the presence of chemical bonds and functional groups in a sample. These AI-driven techniques extend traditional spectroscopic methods further to more accurate and quicker analysis of mixtures containing complex chemicals.

Pattern Recognition and Anomaly Detection Some of the prime applications of AI in analytical chemistry are in pattern recognition and anomaly detection. The training in AI algorithms on patterns in chemical data, for example, spectral fingerprints, assists in identifying similar compounds and material properties. These techniques are particularly useful in quality control and forensic analysis, in which the goal is to detect deviations from expected patterns that might indicate contamination or adulteration. Anomaly detection algorithms, such as autoencoders and support vector machines, identify outliers in chemical data sets that flag the presence of possible errors or even the detection of novel compounds for further investigation. The AI-driven pattern recognition tools in environmental monitoring have come to detect pollutants and hazardous chemicals in air, water, and soil samples. These AI applications improve the accuracy and efficiency of the chemical analysis with deep insight into complicated chemical systems.

Table 2. Impact of AI in Chemistry.

Category	Impact Area	References
Drug Discovery	Molecular Docking	Friday *et al*., 2020; Igwe *et al*., 2020; Ikpeazu et al., 2017a
Drug Discovery	Predictive Modeling	Bender *et al*., 2021; Bender *et al*., 2004; James and Edozie, 2015;
Material Science	Materials Design	Sanchez-Lengeling & Aspuru-Guzik,2018
Chemical Synthesis	Automated Synthesis Planning	Schwaller et al., 2020
Computational Chemistry	Quantum Chemistry	Otuokere *et al*., 2015a; Otuokere *et al*., 2015b
Analytical Chemistry	Spectroscopy	Jerome & Howaed, 2023.
Ethical Considerations	Bias in AI Models	Md & Jeff, 2024a
Future Directions	Integration with Robotics	Md & Jeff, 2024b

Impact of AI on the Chemical Industry and Academia

AI has significantly affected the chemical industry and academia in general, leading to great strides in promoting efficiency and reducing costs, hence promoting research methodologies.

Industry

Efficiency Improvements and Cost Reductions

AI technologies have transformed the process of the chemical industry by promoting efficiency and reducing costs. AI-driven approaches in pharmaceutical companies speed up the process of drug discovery by reducing time and cost in the development of new medications. Machine learning algorithms predict the efficacy of drugs in interactions, while optimization of compound screening aids in the rapid identification of promising drug candidates. For example, AI-driven virtual screening and predictive modeling have significantly accelerated the early stage of drug development by reducing reliance on time-consuming physical tests. Similarly, AI applications in process optimization have translated into more efficient manufacturing processes that generate minimal waste and maximize yields.

In the manufacture of chemicals, it has been AI that has run the production processes for optimization, monitoring equipment performance for predicting maintenance needs, ensuring huge cost savings. Predictive maintenance driven by AI insights prevents failures by foretelling their likelihood well in advance of any occurrence. The benefits of this include decreasing instances of downtime and maintaining the same level of productivity at lower maintenance costs. AI-powered process control systems can enhance the accuracy and speed of chemical reactions, thereby enhancing the consistency in the quality of products and reducing operational costs (Liu *et al*. (2017) and Mikhailov et al. (2020) .

Case Studies: Pharmaceutical and Chemical Industries Various case studies represent the transformational impact AI is causing in the chemical industry. BenevolentAI, among other companies in the pharmaceutical industry, applies AI to drug discovery by mapping AI algorithms that identify new drug candidates and optimize designs for clinical trials (BenevolentAI, 2021). The AI platform of the company has identified some of the complicated diseases that may be treated with COVID-19 through its rapid identification enabled by the AI platform. Other companies like Atomwise uses AI in the determination of a compound binding affinity on target proteins, thus making the drug development process faster and less expensive (Pereira *et al*. (2016); Atomwise. (2021).

For instance, chemical manufacturing, BASF has executed AI-driven process optimization tools to enhance the output of its production facilities. The application of AI algorithms to the production data secures large improvements in product quality and reductions in operational costs for BASF. Other companies, such as Dow Chemical, have used AI in predictive maintenance and control functions, improving their productions with less downtime and increased efficiency.

AI has revolutionized research methods in academia, making it possible to achieve an unprecedented level of analysis and interpretation. Machine learning, and by extension deep learning methods, are being applied towards the analysis of complex datasets to uncover patterns and insights previously unavailable; examples can be seen in works such as Joubert *et al*. 2020, and Li et al. 2021. For example, AI-driven techniques are employed in the field of computational chemistry for property prediction of molecules with high accuracy, outcomes of reactions, and characteristics of materials. Examples include Rupp *et al*. (2012) and Balachandran et al. (2016). As a result, such progress has accelerated research in fields like drug discovery, materials science, and environmental chemistry, opening up new avenues of discoveries and innovations.

AI also plays a vital role in automating experimental workflows, enabling researchers to perform high-throughput experiments more efficiently. Similarly, the robotics and AI-driven system can carry out various repetitive tasks while preparing samples, data collection, etc., freeing up the human researcher to do more important tasks that involve thinking and analysis with a greater degree of difficulty. More importantly, AI tools can process and interpret data in real time to offer immediate feedback to researchers and allow quicker decision-making. Gómez-Bombarelli *et al*., 2018; Liu *et al*., 2017.

Changes in Educational and Research Practices

The adoption of AI in academia introduces real changes into educational and research practices. All these capabilities are being incorporated into the course work of educational institutions, which are churning out data science students who will further the causes of computational research. In fact, universities and their research functions encourage focused courses and programs in AI, data science, and computational chemistry due to emerging demands for this expertise. Such programs provide the level of exposure to AI technologies required in research and professional activities.

Thirdly, AI-driven research facilities are transforming the world of academic research. Gratefully, many researchers these days are into using AI while analyzing big data, simulating chemical processes, and developing new theoretical models (Baker et al., 2020; Bender et al., 2021). Where access to sophisticated computing is democratized, for example through web-based collaboration platforms and openly available AI tools, it has allowed a diverse array of researchers to address new scientific questions and share results with other researchers. Zhang *et al*. 2020; Schwalbe et al. 2021. This emerging trend toward data-intensive and collaborative research practices enables interdisciplinarity and accelerates the speed of scientific discovery.

AI Challenges and Limitations in Chemistry

Though much of the transformational impact of AI in chemistry cannot be denied, many challenges and limitations remain. These include data quality and availability, interpretability, and ethical and regulatory considerations-all of which will need to be tackled if progress is to be made and assurance given that AI technologies are being applied both responsibly and effectively. Data quality and availability has been cited as one of the major challenges to AI applications in chemistry. Machine learning-and deep learning-based AI models rely heavily on large and diverse datasets for the derivation of accurate predictions and meaningful insights. However, most such chemical datasets are sparse, limited in scope, or biased and grossly affect the performance and reliability of AI models.

But one of the famous issues on chemical space is sparsity, where most models usually face their inadequate training due to lack of enough data points in certain regions. This usually brings about poor generalization into new, unseen data by the models. This is in agreement with various analyses like Gómez-Bombarelli et al., 2018; Zhang et al., 2020a; Zhang et al., 2020b. For example, in drug discovery applications, one will likely be faced with rare or totally new chemical compounds, which are under-represented in the training data, leading to a model that fails miserably in the prediction of their properties.

Another serious problem may arise with the datasets. If some dataset is biased and not representative for the whole chemical space, or if data in a dataset contain some kind of systematic errors, then AI models will possibly inherit these biases, and this may result in biased or simply incorrect predictions. For example, bias in chemical datasets could result in models that are biased towards making good predictions for some types of compounds or experimental conditions at the expense of others, hence applicability and effectiveness.

Another important challenge is related to data integration from various sources. The nature of research in chemical studies, with data sourced from experiments, theoretical calculations, and literature, especially requires highly developed methods of data integration and standardized protocols to coherently combine heterogeneous data sources into a usable format for AI models (Miller *et al*., 2019; Segler *et al*., 2018b).

Approaches to improve these include the development of methods for data harmonization and interoperability, besides developing platforms for data sharing and aggregation of data from various research groups. Similarly, developments in preprocessing and data augmentation techniques aim at improving data quality and quantity that might be fed into an AI model to improve its performance.

Black-box Nature of Complex AI Models: Most deep learning-based AI models are black-box in nature; their inner decision-making mechanisms cannot intuitively be understandable by humans. Moreover, most of the works done so far by Caruana *et al*. (2015) and Ribeiro et al. (2016) on black-box nature and lack of transparency may restrict the use of AI models for critical applications where an understanding of the rationale behind the predictions is necessary.

It often helps, for example, in chemical research, to know why a model makes any particular prediction about a chemical reaction or material property. The lack of interpretation over the inner workings of complex models limits their utility and acceptance in scientific and industrial settings. Approaches towards addressing this challenge range from interpretable machine learning model development to visualization techniques that provide insight into the model's decision-making process. Such works include, but are not limited to, Lundberg & Lee, 2017; Ribeiro *et al*., 2016.

Understanding and Transparency Enhancement

Recent research is targeted at enhancing model interpretability and transparency. As such, there is interest in feature importance analysis, attention mechanisms, and model-agnostic techniques to provide more direct interpretation of how AI models are making their predictions. For instance, in the case of a neural network, an attention mechanism may bring out the parts of the data entry that most influenced its decisions.

Next to this, the focus is increasing on frameworks and guidelines that concern reporting and assessment of interpretability of AI models so that the models applied can be understood and trusted more by researchers and practitioners.

Ethical and Regulatory Considerations

Data Privacy and Security

The use of AI in the discipline of chemistry raises some significant ethical and regulatory issues, among which most importantly data privacy and security need to be focused on. Sensitive data related to drug discovery may involve patient information or may be proprietary chemical data; therefore, effective protection measures must be developed and implemented for data privacy. Protection of data through regulation must also be in place. It is important to ensure data security through encryption, access controls, and secure storage methods that avoid unauthorized access to and misuse of the data. Indeed, (Huang *et al*., 2018, Zhang *et al*., 2021a, Zhang *et al*., 2021b) have documented this. In this direction, specific regulations should be followed but are not limited to the General Data Protection Regulation and the Health Insurance Portability and Accountability Act, considering security for personal information and integrity of sensitive data (Purtova (2017).

Ethical Implications of AI-Driven Research

However, besides this, a more ethical consideration of the responsible use of technology and the potential impact on society that AI innovations make in AI-driven research in chemistry is also important: there is a need to address issues such as algorithmic bias, the environmental impact of AI, and the potential misuse of AI technologies to ensure that AI advances benefit society as a whole. For example, AI deployment for drug discovery needs to consider the ethical implications on the availability of new treatments and further exacerbation of health disparities, while ecological footprint aspects related to the development of AI technologies are also concerns to consider in terms of energy use and resource consumption. There are also important ecological footprint issues in the development of these AI technologies .

Future Directions and Trends in AI for Chemistry

The future directions and tendencies of AI in chemistry are absolutely in motion. New AI technologies, inter-skilled approaches, and collaboration might promise leap-forward advances and new opportunities for chemical research and industry in coming years to come. Quantum computing is among the newest technologies which have also started showing great promise to improve AI applications in chemistry. Quantum computers leverage the power of quantum mechanics to perform computations at speeds unreachable by classical computers, thus allowing this technology to potentially revolutionize many areas of AI in chemistry. Quantum computing has the potential to influence AI significantly by allowing modeling of chemical systems and complex reactions in a more applicable and efficient way. For example, quantum algorithms can improve molecular simulation by better approximating the interaction of molecules and predicting chemical properties with much higher accuracy compared to any classical algorithm. As discussed in Babbush *et al*. (2018), Cerezo et al. (2021), this can yield tremendous benefits in drug design, material design, and reaction optimization. Works by Kandala et al. (2017) and McArdle et al. (2020) are highly recommended in this regard.

Besides, quantum computing can further help machine learning algorithms devise new methods to process and analyze chemical data. Quantum-enhanced machine learning techniques may provide solutions for some of the hot topics in data sparsity and model training problems which may bring about breakthroughs in AI-driven chemical research.

Recent progress in chemistry has been mainly driven by the latest breakthroughs in AI algorithms and models. Innovation in neural network architecture, such as transformer models and GNNs, makes many complex chemical system modeling tasks easier to handle and property prediction more accurate (Rogers & Hahn, 2010; Wang *et al*., 2019). Whereas transformer models firstly met huge successes with natural language processing, their extended version is adapted to chemical applications, which lie in predicting molecular properties and reactions. Graph neural networks model information in the form of graphs and are therefore perfectly suited to represent the structure and interactions in molecules and, hence, perform well on chemical informatics tasks. The recent development of generative models, such as Generative Adversarial Networks and Variational Autoencoders, also made possible the generation of new chemical and material compounds. Such models now can explore chemical space more effectively and highlight encouraging candidates for further study. Zhang *et al*., 2020; Segler et al., 2018a; Segler et al., 2018b

Interdisciplinary Approaches

Collaboration Between AI Experts and Chemists

AI in the chemistry of the future is increasingly urgently the business of interdisciplinary collaboration between experts in AI and chemists. This can be done by deeply embedding knowledge within AI methodologies and chemical principles for the effective and efficient integration of AI technologies into chemical research. The collaboration of AI researchers and chemists will provide more tailored and effective AI tools for given specific chemical applications. For example, the domain knowledge from chemists may help to design more relevant AI models in solving the problems of chemistry while AI experts may provide new algorithms and computational techniques. There are already big dividends coming from interdisciplinary partnerships, and first-generation joint efforts are combining a range of innovations in reaction prediction, molecular design, and automated synthesis. Such interdisciplinary collaboration will surely further increase as the field evolves.

Artificial Intelligence in Conjunction with Other Emerging Technologies

Another leading trend that determines the future of chemistry involves its integration with other emerging technologies. AI can be combined with technologies such as the Internet of Things, robotics, and advanced sensors to come up with an even more powerful system in the conduction of chemical research and industry.For example, robotics enabled by AI will be able to automate even the most complicated processes of chemical syntheses, improving efficiency and precision in chemical syntheses. Similarly, the integration of AI with IoT sensors enables real-time monitoring and control of the chemical processes, developing further responsiveness and adaptability.Improvement in data science and computational resources further encourages the development and growth of AI in chemistry. Large chemical databases and high-performance computing infrastructures now provide a foundation for creating and implementing even more advanced models and applications of AI. The introduction of AI into chemistry has indeed opened an era of unprecedented advances and opportunities. This review has traced the development of AI technologies and their application to chemistry, focusing on key developments, current applications, and future directions.

Contemplation of the Transformative Potential of AI

The potential transformation by AI in chemistry is huge, considering the handling of big data, automation of complex tasks, and insight into problems which were not earlier attainable. AI technologies have given a fillip not only to the improvement of existing chemical processes but have also opened up new dimensions for research and innovation. For example, AI has bestowed the power of predictive analytics on molecular properties and chemical reaction optimization, hence more effective and targeted approaches have been developed in drug development and material sciences. Furthermore, AI has allowed for integration of various data sources into one system and allows for more precise analyses, which is a very important factor while enhancing our knowledge about chemical systems.

IV. Conclusions

AI has indeed transformed how chemical research is conducted and applied-from early computational methods to sophisticated machine learning and deep learning models. Early rule-based systems gave way to sophisticated AI techniques capable of handling Big data chemical data for the prediction of molecular properties with high accuracy. Machine learning, both supervised and unsupervised, and even reinforcement learning, has made chemical models increasingly predictive. Deep learning techniques like neural networks and generative models have added new ways lately for molecular design and synthesis. Applications so far in drug discovery, materials science, chemical synthesis, and spectroscopy illustrate the potential of AI to change these fields. Artificial intelligence-driven tools accelerate drug development processes, advance material properties, and improve analytical methodologies, leading to even more rapid and creative research outcomes. The impact of AI on the chemical industry and academia is unprecedented. Some of the characteristics include increased efficiency, cost reduction, shifting research methodologies, and educational practice. One can confidently say that the impact of AI on both the chemical industry and academia is nothing short of revolutionary. In the near future, AI will continue to rapidly improve in chemistry. Quantum computing introduces a new generation of high-performance computing technologies that, though still in their infancy, may start developing new modeling and simulation tools that revolutionize the field in the near future. This is likely achieved by an interdisciplinary collaboration between experts in AI and chemists toward the elaboration of adapted solutions by the use of AI in combination with other new emerging techniques. As AI continues to evolve, there will be an increasing requirement to solve issues with data quality, model interpretability, and ethics to allow for responsible and effective applications in chemistry. Development of AI technologies and embedding them into chemical research will continue to drive innovation and discovery, shaping the future both in this area and the impact of science and industry. There is really no limit to the transformative power of AI in chemistry, and further developments and breakthroughs may well be expected. In this way, the community of chemistry is empowered to embrace new technologies and encourage collaboration with AI toward the full realization of scientific progress in service of resolving global challenges.

References

AilinXie, ZZ, Jihong, G & Shuigeng, Z 2023, "Self-supervised learning with chemistry-aware fragmentation for effective molecular property prediction", Briefings in Bioinformatics, vol. 24, no. 5, pp. 1–13

Alanah, GJ, Luke H, Thomas M, Liwen, S, Annabel, SJB, Patrick, EBM, Andrew, FW, Malcolm, IS & Claire, V 2023, "Exploring Machine Learning in Chemistry through the Classification of Spectra: An Undergraduate Project", Journal of Chemical Education, vol. 100, no. 3, pp. 1343–1350 https://doi.org/10.1021/acs.jchemed.2c00682

Amadi, OK, Otuokere, IE & Bartholomew, CF 2015, "Synthesis, Characterization, in vivo Antimalarial Studies and Geometry Optimization of Lumefantrine/Artemether Mixed Ligand Complexes". Research Journal of Pharmaceutical Dosage Forms and Technology, vol. *7, no.* 1, pp. 59-68.

Amaku, FJ, & Otuokere, IE. 2015, "Conformational Analysis of a Potent Anticancer Drug 3-(4-amino-1-oxo-1, 3-dihydro-2h-isoindol-2-yl) Piperidine-2, 6-Dione (Lenalidomide)", Int. J. Mater. Chem. Phys, vol. 1, pp. 402-406.

Amaku, FJ, Ikpeazu, VO, Otuokere, IE & Igwe, KK 2020, "Targeting glycogen synthase kinase-3 (Gsk3β) with naturally occurring phytochemicals (quercetin and its modelled analogue): A pharmacophore modelling and molecular docking approach", Communication in Physical Sciences, vol. 5, no. 4 pp. 497-508

Andrij, V, Jacinthe, G, Benjamin, BD, Vladimir, VG, Luke, MD, Marco, Z, Felix, JS, Paul, MS, Alexandra, M, Ruiyong, C, Alex, RN, Laurence, JH, John, BC, Frédéric, B, Michael, WG, Matthew, SD & Matthew, JR 2021, "Element selection for crystalline inorganic solid discovery guided by unsupervised machine learning of experimentally explored chemistry", NATURE COMMUNICATIONS, vol. 12, pp. 5561. https://doi.org/10.1038/s41467-021-25343-7

Andrzej, MZ, Jesus, IMA., Benjamin, JS & Abigail, GD 2021, "Predicting Reaction Yields via Supervised Learning", Accounts of Chemical Research, vol. 54, pp. 1856–1865

Aspuru-Guzik, A, Lindh, R & Reiher, M 2018, "Autonomous molecular synthesis with AI-driven robotics", Nature Reviews Chemistry, vol. 2, no. 1, pp. 9-12.

Asuquo, IG, Solangi, M, Khan, KM, Chigurupati, S, Otuokere, IE, Ekuma, FK, Salar, U, Felemban, SG, Rehman, AU & Taha, M 2023, "Design, synthesis and bio-evaluation of indolin-2-ones as potential antidiabetic agents", Future Medicinal Chemistry, vol. 15, no. 1, pp. 25-42.

Atomwise 2021, "Atomwise: AI for drug discovery", Retrieved from https://www.atomwise.com

Babbush, R, Jordan, SP & Love, PJ 2018, "Quantum algorithms for fixed-point arithmetic", Physical Review Letters, vol. 121, no. 1, pp. 010502.

Baker, LR, Lane, DB & Petersen, JM 2020, "Advances in AI-driven research methodologies", Journal of Computational Chemistry, vol. 41, no. 3, pp. 209-220.

Balachandran, PV, Kowalski, B, & Sehirlioglu, A 2016, "Machine learning-directed screening of inorganic compounds for high-temperature piezoelectric applications", Nature Communications, vol. 7, pp. 12182.

Balachandran, PV, Theiler, J, & Lookman, T 2018, "Predictive modeling of complex material properties with machine learning", Scientific Reports, vol. 8, no. 1, pp. 1-8.

Bender, A & Glen, RC 2004, "Molecular similarity: A key technique in molecular informatics", Organic & Biomolecular Chemistry, vol. 2, no. 22, pp. 3204-3218.

Bender, A, Schneider, N & Clark, RD 2021, "AI tools for academic research: Current state and future directions", Nature Reviews Chemistry, vol. 5, no. 2, pp. 71-85.

BenevolentAI 2021, "How AI is transforming drug discovery", Retrieved from https://www.benevolent.com

Bengio, Y & Hinton, G 2012, *"Deep learning applications in chemistry",* Chemical Reviews, vol. 112, no. 11, pp. 6136-6155.

Bjerrum, EJ & Threlfall, R 2017, "Molecular generation with recurrent neural networks (RNNs)", arXiv preprint, pp. 1705.04612.

Breiman, L 2001, "Random forests", Machine Learning, vol. 45, no. 1, pp. 5-32.

Buchanan, BG & Shortliffe, EH 1984, "Rule-based expert systems: The MYCIN experiments of the Stanford Heuristic Programming Project", Addison-Wesley. Pp. 1-19

Buchanan, BG, Sutherland, GL. & Feigenbaum, EA. 1969, "Heuristic DENDRAL: A program for generating explanatory hypotheses in organic chemistry", Machine Intelligence, vol. 4, pp. 209-254.

Caruana, R, Gehrke, J, Koch, P & Ni, S 2015, "Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission", Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 125-134.

Cerezo, M, Gao, X, & McClean, JR 2021, "Variational quantum algorithms", Nature Reviews Physics, vol. 3, no. 9, pp. 625-644.

Choudhary, K & Kelley, ML 2023, "ChemNLP: A Natural Language-Processing-Based Library for Materials Chemistry Text Data", The Journal of Physical Chemistry C, vol. 127, no. 35. doi: 10.1021/acs.jpcc.3c03106

Coley, CW, Green, WH & Jensen, KF 2019, "Machine learning in computer-aided synthesis planning", Accounts of Chemical Research, vol. 52, no. 6, pp. 1338-1346.

Cortes, C. & Vapnik, V. 1995, "Support-vector networks", Machine Learning, vol. 20, no. 3, pp. 273-297.

Cover, TM, & Hart, PE 1967, "Nearest neighbor pattern classification", IEEE Transactions on Information Theory, vol. 13, no. 1, pp. 21-27.

Dimitrov, SD, Dimitrova, NC, Georgieva, DG, Walker, JD, & Mekenyan, OG 2003, "Predicting the biodegradation products of persistent organic pollutants using CATABOL", SAR and QSAR in Environmental Research, vol. 14, no. 3, pp. 237-262.

Duvenaud, D, Maclaurin, D, Aguilera-Iparraguirre, J, Gómez-Bombarelli, R, Hirzel, T, Aspuru-Guzik, A, & Adams, RP 2015, Convolutional networks on graphs for learning molecular fingerprints. In Advances in Neural Information Processing Systems, pp. 2224-2232.

Frazier, PI 2018, "A tutorial on Bayesian optimization", arXiv preprint, pp. 1807.02811.

Gao, W & Coley, CW 2020, "Data-driven chemical synthesis and reactivity prediction", Current Opinion in Chemical Biology, vol. 55, pp. 1-8.

Gilmer, J, Schoenholz, SS, & Riley, PF 2017, "Neural message passing for quantum chemistry", Proceedings of the 34th International Conference on Machine Learning, vol. 70, pp. 1263-1272.

Gómez-Bombarelli, R, Wei, JN, Duvenaud, D, Hernández-Lobato, JM, Sánchez-Lengeling, B, Sheberla, D & Aspuru-Guzik, A 2018, "Automatic chemical design using a data-driven continuous representation of molecules". ACS Central Science, vol. 4, no. 2, pp. 268-276.

Granda, JM, Donina, L & Dragone, V 2018, "Controlling chemical reactions with AI-driven algorithms", Nature, vol. 559, no. 7714, pp. 377-381.

Hochreiter, S & Schmidhuber, J 1997, "Long short-term memory", Neural Computation, vol. 9, no. 8, pp. 1735-1780.

Hood, L 2005, "*Advances in high-throughput screening for drug discovery*", Nature Reviews Drug Discovery, vol. 4, no.(6), pp. 477-485.

Huang, M, Zhang, X, & Li, J 2018, "Data security and privacy in AI systems: Challenges and solutions", IEEE Transactions on Knowledge and Data Engineering, vol. 30, no. 8, pp. 1562-1574.

Igwe, KK, Ikpeazu, OV, Amaku, FJ & Otuokere, IE 2020, "Repurposing Hydroxychloroquine as a Model Drug for the Prediction of Potential SARS-CoV-2 Inhibitor: A Computational Approach", European Journal of Engineering and Technology Research, vol. 5, no. 9, pp. 1031-1036.

Igwe, OU, Anyaogu, MU & Otuokere, IE 2024b, "Chemical, Antioxidant and Antibacterial Assessment of Clove (Syzygium aromaticum) Seed Extract and in-silico Pharmacokinetic Exploration of the Prominent Compounds", Journal of Applied Sciences and Environmental Management, vol. 28, no. 7, pp. 1935-1943.

Igwe, OU, Oru, CC & Otuokere, IE 2024a, "Chemical and bioprotective studies of Xylopia aethiopica seed extract and molecular docking of doconexent and cryptopinone as the prominent compounds", African Scientific Reports, vo. 3, no. 2, pp. 186-186.

Ikpeazu, OV, Otuokere, IE & Igwe, KK 2017a, "Computational characterization of the binding energy and interactions between trimethoprim and dihydrofolate reductases of *Candida albicans, Staphylococcus aureus* and *Thermotoga maritime"*, International Journal of Pharma and Drug Development, vol, 2, no, 2, pp. 52-56.

Ikpeazu, OV, Otuokere, IE & Igwe, KK 2017b, "Electrostatic Potential-Mapped Electron Density Surface and Conformation Analysis of an Antidepressant, Mirtazapine", International Journal of Pharma and Drug Development, vol. 2, no. 1, pp. 09-13.

Jablonka, KM, Schwaller, P, Ortega-Guerrero, A & Smit, B 2024, "Leveraging large language models for predictive chemistry", *Nat. Mac. Intell., vol. 6,* pp. 161-169.

Jerome, WJ & Howaed, M 2023, "Artificial Intelligence in Analytical Spectroscopy, Part 1: Basic Concepts and Discussion. Spectroscopy", vol. 38, no. 2, pp. 13 – 22.

Jidon, J, Geun, HG, Juhwan, N, Juhwan, K & Yousung, J 2020, "Structure-Based Synthesizability Prediction of Crystals Using Partially Supervised Learning", Journal of the American Chemical Society, https://dx.doi.org/10.1021/jacs.0c07384

Jones, PR & Smith, KA 2021, "*Autonomous chemical synthesis: The role of AI"*, Nature Chemistry, vol. 13, no. 3, pp. 289-298.

Joubert, D, Schwartz, H & Zhang, W 2020, "AI and machine learning in academic research: An overview", Computational Research, vol. 32, no. 1, pp. 33-45.

Kandala, A, Mezzacapo, A & Kim, J 2017, "Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets", Nature, vol. 549, no. 7671, pp. 242-246.

Karthikeyan, A, Muthukkumar, M, Nagaraj, K, Edozie, O I, Suganya, S, Kaliyaperumal, R & Kandasamy, M 2023, "DFT, molecular docking and molecular dynamics simulations of 2-imino-4-oxo-1, 3-thiazolidine hydrochloride and its activity against Bacillus pasteurii urease", Zeitschrift für Physikalische Chemie, vol. 237, no. 12, pp. 2007-2018.

Kim, JD, Ohta, T, Tateisi, Y & Tsujii, J 2003, "GENIA corpus—a semantically annotated corpus for bio-textmining", Bioinformatics, vol. 19, no. 1, pp. i180-i182.

Kingma, DP & Welling, M 2013, "Auto-encoding variational bayes", arXiv preprint, pp. 1312.6114.

Krallinger, M, Rabal, O, Akhondi, SA, Pérez, MP, Santamaría, J, Rodríguez, GP & Valencia, A 2015, "Overview of the CHEMDNER patents task. Proceedings of the Fifth BioCreative Challenge Evaluation Workshop, pp. 63-75.

Langner, S, Häse, F & Perea, JD 2020, "Beyond tedium in quantum chemistry: Automated reaction mechanism discovery", Chemistry-A European Journal, vol. 26, no. 9, pp. 2040-2044.

Leach, AR 2001, "Molecular modelling: Principles and applications", (2nd ed.). Pearson Education.
LeCun, Y, Bengio, Y & Hinton, G 2015, Deep learning. Nature, vol. 521, no. 7553, pp. 436-444.
LeCun, Y, Bottou, L, Bengio, Y & Haffner, P 1998, "Gradient-based learning applied to document recognition", Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324.
Li, H, Pu, X & Wang, H 2017, "AI-powered robotics in organic synthesis", Journal of Organic Chemistry, vol. 82, no. 20, pp. 10457-10465.
Li, S, Zhang, T & Xu, M 2021, "Machine learning for complex data analysis in chemistry", Journal of Chemical Physics, vol. 154, no. 5, pp. 055501.
Lindsey, AJ, Snyder, RD, & Jackson, RD 1982, "Computational chemistry: A new science", Science, vol. 218, no. 4574, pp. 645-650.
Liu, B, Ramsundar, B & Liu, L 2017, "Retrosynthetic reaction prediction using neural networks", Chemical Science, vol. 8, no. 10, pp. 5137-5145.
Lundberg, SM, & Lee, SI. 2017, "A unified approach to interpreting model predictions", Proceedings of the 31st International Conference on Neural Information Processing Systems, vol. 30, pp. 4765-4774.
Marius, N & Stefan, P 2022, "A Reinforcement Learning Approaches for the Optimization of the Partial Oxidation Reaction of Methane", Industrial & Engineering Chemistry Research, vol. 61, no. 11, pp. 3910–3916
McArdle, S, Liao, J & Wossnig, R 2020, "Quantum computing for chemistry", Nature Reviews Chemistry, vol. 4, no. 3, pp. 106-120.
McCarthy, J, Minsky, ML, Rochester, N & Shannon, CE 2006, "A proposal for the Dartmouth summer research project on artificial intelligence", AI Magazine, vol. 27, no. 4, pp. 12-14. https://doi.org/10.1609/aimag.v27i4.1904
Md, RM & Jeff, S 2024a, "A Survey of Ethical Considerations in AI: Navigating the Landscape of Bias and Fairness", Journal of Artificial Intelligence General Science, vol. 1, no. 1 pp. 1-4.
Md, RM & Jeff, S 2024b, "Exploring the Synergy of Artificial Intelligence and Robotics in Industry 4.0 Applications", Journal of Artificial Intelligence General Science, vol. 1, no. 1, pp. 5-8.
Mikhailov, V, Atanassov, P, & Belousov, A 2020, "AI-enhanced chemical manufacturing: Recent advances and future prospects", Journal of Industrial Chemistry, vol. 67, no. 2, pp. 234-248.
Miller, R, Levy, R & Gupta, S. 2019, "Incorporating AI into academic chemistry curricula", Chemistry Education Research and Practice, vol. 20, no. 4, pp. 787-800.
Mitchell, TM 1997, *"Machine learning and QSAR modeling in chemistry"*, Chemometrics and Intelligent Laboratory Systems, vol. 33, no. 1, pp. 45-54.
Munson, BP, Chen, M, Bogosian, A 2024, "De novo generation of multi-target compounds using deep generative chemistry", Nat Commun, vol. 15, pp. 3636 https://doi.org/10.1038/s41467-024-47120-y
Nwankwo, CI, Omeh, YN, Omodamiro, OD, Otuokere, IE, Alaebo, PO, Atasie, OC & Ekwuribe, GA 2022, "Phenolics of Abelmoschus esculentus Pods: HPLC Identification and In Silico Studies to Identify Potential Anti-inflammatory Agents", Tropical Journal of Natural Product Research, vol. 6, no. 8, pp. 1311-1319.
Otuokere IE & Amaku, FJ 2015a, "Conformational analysis and excited-state properties of a highly potent and totally selective aromatase inhibitor, 4, 4′-(1H-1, 2, 4-triazol-1-ylmethanediyl) dibenzo nitrile (letrozole)", Research Journal of Pharmacology and Pharmacodynamics, vol. 7, no. 4, pp. 176-180.
Otuokere IE, Ikpeazu, VO & Igwe, KK 2017, "Molecular electrostatic potential, partial atomic charges and geometry optimized energy of streptozocin, a DNA alkylating agent", International Journal of Research in Engineering and Innovation, vol. 1, no. 3, pp. 113-117
Otuokere IE. & Amaku, FJ 2015b, "Conformation Analysis and Self-Consistent Field Energy of Immune Response Modifier, 1-(2-methylpropyl)-1H-imidazo [4, 5] quinolin-4-amine (Imiquimod)", Asian Journal of Research in Pharmaceutical Science, vol. 5, no. 3, pp. 175-180.
Otuokere IE. 2014, "Molecular Mechanics Potential Energy Function of Angiotensin-Converting Enzyme (ACE) inhibitor, Lisinopril", Asian Journal of Research in Pharmaceutical Science, vol. 4, no. 3, pp. 118-124.
Otuokere, I.E & Amaku, FJ 2015a, "Computer-Aided Molecular Design of a Histone Deacetylase (HDAC) Inhibitor, N-Hydroxy-N-Phenyloctanediamide (Vorinostat)", Research Journal of Science and Technology, Vol. 7, No. 4, pp. 212-216.

Otuokere, IE & Alisa, CO 2014, "Computational Study on Molecular Orbital's, Excited State Properties and Geometry Optimization of Anti-benign Prostatic Hyperplasia Drug, N-(1, 1-dimethylethyl)-3-oxo-(5α, 17β)-4-azaandrost-1-ene-17-carboxamide (Finasteride)", Asian Journal of Research in Pharmaceutical Science, vol. 4, no. 4, pp. 169-173.

Otuokere, IE & Alisa, CO 2015a, "In silico Studies on Nuclei Magnetic Resonance Spectroscopy, Molecular Orbitals and Geometry Optimization of Antidiabetic Drug, N, N-dimethylimidodicarbonimidic Diamide (Metformin)", American Chemical Science Journal, vol. 8, no. 3, pp. 1-8.

Otuokere, IE & Amaku, FJ 2015b, "Quantum Chemical Studies of Anti-Prostatic Carcinoma Drug N-[4-cyano3-(trifluoromethyl) phenyl]-2-hydroxy-2-methyl-3-[(4methylphenyl) sulfonyl] Propanamide (bicalutamide)", Research Journal of Pharmacognosy and Phytochemistry, vol. 7, no. 4, pp. 214-218.

Otuokere, IE 2024, "Electrostatic Potential Mapped onto Electron Density Surface, ADME, Molecular Docking and Molecular Dynamics Simulations of Some Indolin-2-one Analogues as Cytochrome C Peroxidase Inhibitors", Research Square https://doi.org/10.21203/rs.3.rs-4200649/v1

Otuokere, IE, Akoh, OU, Echeme, JO, Nwadire, FC, Nwankwo, CI, Egbucha, JN & Ammasai, K 2022a, "GC-MS Analysis and Molecular Docking Studies to Identify Potential SARS-CoV-2 Nonstructural Protein Inhibitors from Icacina trichantha Oliv Tubers", Tropical Journal of Natural Product Research, vol. *6, no.* 8, pp. 1337-1342..

Otuokere, IE, Akoh, OU, Nwadire, FC, Nwankwo, CI, Egbucha, JN, Wisdom, C & Okwudiri, OA 2022b, "GC-MS Profiling and In Silico Studies to Identify Potential SARS-CoV-2 Nonstructural Protein Inhibitors from *Psidium guajava",* African Scientific Reports, vol. 1, pp. 161-173.

Otuokere, IE, Alisa, CO & Nwachukwu, P 2015b, "Synthesis, Characterization and Molecular Mechanics Potential Energy Evaluation of 4-amino-2, 3-dimethyl-1-phenyl-3-pyrazolin-5-one ligand and its transition metal complexes", Research Journal of Pharmaceutical Dosage Forms and Technology, vol. *7, no,* 2, pp. 129-133.

Otuokere, IE, Godday, OJ, Chijioke, AK, Uchenna, IO & Chigozie, NF 2020, "Synthesis, characterization and molecular docking studies of Co (II) metal complex of sulfathiazole", Bulletin of the Chemical Society of Ethiopia, vol. 34, no. 1, pp. 83-92.

Otuokere, IE. Ikpeazu, VO & Igwe, KK 2017, "Structure-activity relationship and virtual screening of monosubstituted lamivudine with HIV-1 reverse transcriptase", International Journal of Research in Engineering and Innovation, vol. 1, no. 3, pp. 181-186

Pereira, JM, Zhang, C & Chong, K 2016, "Predicting protein-ligand binding affinities with deep learning", Journal of Chemical Information and Modeling, vol. 56, no. 12, pp. 2489-2496.

Philippe, S, Benjamin, H, Jean-Louis, R, Hendrik, S, Teodoro, L 2021, "Extraction of organic chemistry grammarfrom unsupervised learning of chemical reactions", Science Advances, vol. 7, pp. eabe4166

Pople, JA 1980, "*Molecular modeling: A new era in computational chemistry",* Angewandte Chemie International Edition, vol. 19, no. 5, pp. 343-356.

Purtova, N 2017, "The law of everything. Broad concept of personal data and future of EU data protection law. Law, Innovation and Technology", vol. 9, no. 1, pp. 40-81.

Rajan, K 2015, "Materials informatics: The role of data and AI in the accelerated discovery of new materials. Annual Review of Materials Research", vol. 45, pp. 153-169.

Ribeiro, MT, Singh, S & Guestrin, C 2016, "Why should I trust you? Explaining the predictions of any classifier", Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1135-1144.

Rogers, D & Hahn, M 2010, "Extended-connectivity fingerprints", Journal of Chemical Information and Modeling, vol. 50, no. 5, pp. 742-754.

Rupp, M, Tkatchenko, A & Müller, KR 2012, "Fast and accurate modeling of molecular atomization energies with machine learning", Physical Review Letters, vol. 108, no. 5, pp. 058301.

Sabrina, J, Simone, F & Samo, T 2018, "Mol2vec: Unsupervised Machine Learning Approach with Chemical Intuition", Journal of Chemical Information and Modeling. vol. 58, pp. 27–35

Sanchez-Lengeling, B & Aspuru-Guzik, A 2018, "Inverse molecular design using machine learning: Generative models for matter engineering", Science, vol. 361, no. 6400, pp. 360-365.

Scarselli, F, Gori, M, Tsoi, AC, Hagenbuchner, M & Monfardini, G 2009, "The graph neural network model. IEEE Transactions on Neural Networks", vol. 20, no. 1, pp. 61-80.

Schneider, P, Walters, WP & Plowright, AT 2020, "AI-driven synthesis planning: From generative models to robotic chemistry", Nature Reviews Drug Discovery, vol. 19, no. 7, pp. 497-516.

Schwalbe, H, Lenz, M & Hase, J 2021, "AI and computational tools in interdisciplinary research", Interdisciplinary Science Reviews, vol. 46, no. 2, pp. 187-202.

Schwaller, P, Gaudin, T & Laino, T 2018, "Molecular transformer: A model for uncertainty-calibrated chemical reaction prediction", ACS Central Science, vol. 4, no. 9, pp. 1572-1583.

Segler, MH, Preuss, M & Waller, MP 2018b, "Planning chemical syntheses with deep neural networks and symbolic AI", Nature, vol. 555, no. 7698, pp. 604-610.

Segler, MHS, Kogej, T & Tyrchan, C 2018a, "Generative models for chemical synthesis", Nature, vol. 559, no. 7714, pp. 356-359.

Smith, J. 2018, "*Foundations of AI in chemistry*", Journal of Chemical Information and Modeling, vol. 58, no. 4, pp. 895-909.

Ståhl, N, Falkman, G, Karlsson, A, Mathiason, G & Boström, J 2019, "Deep Reinforcement Learning for Multiparameter Optimization in *de novo* Drug Design", Journal of chemical information and modeling, vol. 59, no. 7, pp. 3166–3176. https://doi.org/10.1021/acs.jcim.9b00325

Swain, MC & Cole, JM 2016, "ChemDataExtractor: A toolkit for automated extraction of chemical information from the scientific literature", Journal of Chemical Information and Modeling, vol. 56, no. 10, pp. 1894-1904.

Turing, AM 1950, "Computing machinery and intelligence", Mind, vol. 59, no. 236, pp. 433-460.

Wang, H, Zhang, L, & Li, Y 2019, "Graph neural networks for chemical informatics", Journal of Chemical Theory and Computation, vol. 15, no. 6, pp. 3805-3821.

Yiwei, L, Cheng, Z, Zhonghua, L, Donald, GT, Ying, W & Xiao, H 2022, "Supervised learning of a chemistry functional with damped dispersion", Nature Computational Science, vol. 3, pp. 48–58. https://doi.org/10.1038/s43588-022-00371-5

Zhang, Y, Jiang, H & Liu, X 2020a, "Open-access AI tools and collaborative research in chemistry", Science Advances, vol. 6, no. 32, pp. eabb8476.

Zhang, Y, Zhou, L & Li, S 2021a, "Data privacy in AI-driven research: Challenges and solutions", Journal of Privacy and Confidentiality, vol. 11, no. 2, pp. 123-137.

Zhang, Z, Han, J & Li, H 2020b, "AI and quantum computing integration in chemical research", Nature Reviews Chemistry, vol. 4, no. 5, pp. 269-281.

Zhang, Z, Liu, X & Wang, Z. 2021b, "Advancements in transformer models for chemical informatics", Journal of Chemical Information and Modeling, vol. 61, no. 2, pp. 753-764.

Zhavoronkov, A, Aliper, A & Andreev, A 2019a, "Artificial intelligence for drug discovery, biomarker development, and generation of novel therapeutics". Molecular Pharmaceutics, vol. 16, no. 6, pp. 2484-2498.

Zhavoronkov, A, Ivanenkov, YA, Aliper, A, Veselov, MS, Aladinskiy, VA, Aladinskaya, AV & Aspuru-Guzik, A 2019b, "Deep learning enables rapid identification of potent DDR1 kinase inhibitors". Nature Biotechnology, vol. 37, no. 9, pp. 1038-1040.

Zhou, Z, Kearnes, S, Li, L 2019, "Optimization of Molecules via Deep Reinforcement Learning", Sci Rep, vol. 9, no. 10752. https://doi.org/10.1038/s41598-019-47148-x

Zhou, Z, Li, X & Zare, RN 2017, "Optimizing chemical reactions with deep reinforcement learning", ACS Central Science, vol. 3, no. 12, pp. 1337-1344.