Top 20 Influential AI-Based Technologies in Chemistry

Valentine P. Ananikov

Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences,
Leninsky pr. 47, Moscow, 119991, Russia; E-mail: val@ioc.ac.ru;
http://AnanikovLab.ru/

Abstract: The beginning and ripening of digital chemistry is analyzed focusing on the role of artificial intelligence (AI) in an expected leap in chemical sciences to bring this area to the next evolutionary level. The analytic description selects and highlights the top 20 AI-based technologies and 7 broader themes that are reshaping the field. It underscores the integration of digital tools such as machine learning, big data, digital twins, the Internet of Things (IoT), robotic platforms, smart control of chemical processes, virtual reality and blockchain, among many others, in enhancing research methods, educational approaches, and industrial practices in chemistry. The significance of this study lies in its focused overview of how these digital innovations foster a more efficient, sustainable, and innovative future in chemical sciences. This article not only illustrates the transformative impact of these technologies but also draws new pathways in chemistry, offering a broad appeal to researchers, educators, and industry professionals to embrace these advancements for addressing contemporary challenges in the field.

Contents

1. Introduction

The integration of digital technologies within the field of chemistry is forging unprecedented pathways in research methodologies, educational paradigms, and industrial practices.[1-5] The transformative impact is actively discussed in digital innovations, including machine learning, artificial intelligence, the Internet of Things (IoT), and blockchain, among others, in the chemical sciences.[6-9] The integration of machine learning (ML) approaches helps to develop new concepts in catalysis.[10-13] Researchers explore how automated laboratory platforms and high-throughput experimentation are redefining experimental protocols, enabling precise, reproducible research at an accelerated speed.[14, 15] Progress in the development of synthetic platforms may be further anticipated with the availability of ML-driven interpretation of analytic data.[16, 17] The emergence of digital twins and integrated data systems is highlighted for their ability to simulate and optimize chemical processes, enhancing efficiency and sustainability in chemical manufacturing.[18-21]

Furthermore, ongoing research has examined the role of deep learning in exploring complex structure–activity relationships, thereby revolutionizing drug discovery and material science by predicting molecular behaviors with reasonable accuracy. The application of natural language processing in mining the vast volume of chemical literature illuminates its potential for accelerating knowledge acquisition and fostering innovation. Additionally, the implementation of blockchain

technology is addressed for ensuring traceability and integrity within the chemical supply chain, marking a significant stride toward transparency and safety.

In this article, the top 20 highly influential AI technologies that have already transformed fundamental research and the industrial sector in chemistry or are expected to have profound influence in the future are summarized. The convergence of digital technologies and chemistry promises to address some of the most pressing challenges in the field, from the development of green chemistry practices to the rapid discovery of novel therapeutics. By offering an overview of these digital advancements, this article aims to inspire researchers, educators, and industry professionals to embrace these technologies, opening the way for a more efficient, sustainable, and innovative future in chemistry.

2. AI-Based Technologies in Chemistry

The influence of digital technologies has profoundly impacted the field of chemistry. Recent trends are empowered with artificial intelligence (AI), particularly machine learning (ML), deep learning (DL) and data analysis. Digital "wave" in chemistry is not limited to AI-related innovations and implements a number of other technologies. The following technologies highlight the key directions that mark the progress of science in chemistry within the trajectory of digital development:

1. **AI-driven Drug Discovery**: AI can be used to accelerate the identification of potential drug candidates through the prediction of biological activity and optimization of lead compounds.

2. **Big Data and Integrated Data**: Consolidation and harmonization of diverse chemical data sources, facilitating cross-disciplinary research and comprehensive data analysis.

3. **Automated Laboratory Platforms**: Development of automated laboratory systems for conducting experiments with minimal human intervention, improving precision and reproducibility.

4. **Integration of Laboratory Instruments and IoT**: Laboratory instrument connectivity to the Internet of Things (IoT) for real-time data collection, monitoring, and analysis.

5. **AI in Spectroscopy and Analytical Method Development**: AI applications improving complex spectroscopic data interpretation and developing novel analytical methods.

6. **Blockchain in Chemical Supply Chain**: Blockchain technology for secure, transparent chemical and material supply chain tracking.

7. **Digital Twins**: Creation of virtual replicas of chemical processes or systems for simulation, monitoring, and optimization purposes.

8. **Virtual Laboratories and Augmented Reality**: Digital platforms and simulation software enhancing teaching and providing virtual lab experiences.

9. **Natural Language Processing (NLP) in Chemical Space**: NLP tools for mining chemical information from scientific literature, patents, and databases for knowledge extraction.

10. **Predictive Toxicology**: Computational models predicting chemical toxicity to enhance environment protection and chemical manufacturing safety.

11. **AI in Environmental Chemistry and Sustainability**: Digital tools aiding environmental process analysis, pollution control, and green chemistry development.

12. **Machine Learning in Molecular Design**: Application of ML algorithms to predict molecular properties, enabling efficient design of new compounds and materials.

13. **Smart Control**: Implementation of intelligent control systems in chemical processes and equipment, enabling adaptive and optimized operations.

14. **Deep Learning in Structure-Activity Relationships (SAR)**: The utilization of deep learning models to decipher and predict the complex relationships between chemical structures and their biological activities enhances the efficiency of drug discovery processes.

15. **AI-driven High-throughput Experimentation (HTE)**: Robotics and AI integration for conducting and analyzing multiple parallel experiments, accelerating the research process.

16. **Digital Materials Design and Materials Informatics**: Data-driven approaches for discovering and designing new materials with desired properties and applications.

17. **Data-Driven Chemical Reaction Optimization**: Machine intelligence models predicting reaction outcomes, optimizing conditions, and discovering new reactivity.

18. **Automated Synthesis Planning**: Assistance from AI tools in planning and optimizing synthetic routes, reducing experimental trial time and resources.

19. **Chemoinformatics and Chemical Data Analysis**: Advanced techniques for managing and interpreting large chemical datasets, enhancing understanding of complex relationships in chemical structures and processes.

20. **AI in Quantum Chemistry and Simulations**: AI-enhanced accuracy and efficiency in quantum chemical calculations and molecular simulations.

These concepts collectively demonstrate the significant role that digital technologies play in transforming the field of chemistry, enhancing research capabilities, and fostering innovation. Below, each of the top 20 digital discovery transformative technologies are considered in view of their possible role in fundamental chemical science, applications and future development.

2.1 Comparative analysis of digital chemistry trajectories

AI-based innovations bring several highly promising opportunities in chemistry. However, for the considered digital technologies, different levels of development are required before possible practical implementations can be reached and tested (Table 1). Some of the listed AI-driven technologies introduce new concepts for the current stage of chemical development or do not have a strong prior background in chemical applications, thus more efforts and time would be required for real applications. However, other applications may rely on the existing background, or rapid integration may be expected due to possible connections with the existing research areas; therefore, a quicker application of practice may be expected.

Relative amount of research already done may be estimated from the number of currently published articles: higher amount for the items 1 – 5 (high), with decreased amount for the items 6 – 10 (medium to high), further lowering for the items 11 – 15 (medium) and items 16 – 20 (medium-small) (Table 1). The first-listed areas undergo a more intense development as compared to the next listed ones (Table 1), which is most likely connected with the number of researchers

and funding involved in each field as well as with a broad applicability vs. more specialized task been solved.

Table 1. Summary of the possible influence of AI-based technologies on chemical research, including their applications, their impact, the amount of research done.[a]

Item	Digital chemistry trajectory	Application	Impact	Relative amount of publications [b]
1	AI-driven drug discovery	AI to accelerate drug candidate identification	Speeds up the drug discovery process	High *****
2	Big Data and Integrated Data	Consolidation of diverse data sources	Facilitates comprehensive analysis	High *****
3	Automated laboratory platforms	Highly automated systems for experiments	Improves precision and reproducibility	High *****
4	Integration of Lab instruments and IoT	Real-time data collection and analysis	Enhances lab efficiency and data analysis	High *****
5	Spectroscopy and analytical method development	AI in spectroscopic analysis	Improves analytical methods	High *****
6	Blockchain in chemical supply chain	Secure, transparent tracking in	Increases supply chain integrity	Medium-High ****

		supply chain		
7	Digital twins	Virtual replicas for real-time simulation and optimization	Optimizes chemical processes and systems	Medium-High ****
8	Virtual Laboratories and Augmented Reality	Digital platforms for education	Enhances educational outcomes	Medium-High ****
9	NLP in chemical space	Extract information from scientific texts	Streamlines knowledge extraction	Medium-High ****
10	Predictive toxicology	Models predicting chemical toxicity	Enhances drug safety	Medium-High ****
11	Environmental chemistry and sustainability	Digital tools for green chemistry	Promotes sustainable practices	Medium ***
12	Machine learning in molecular design	ML for predicting molecular properties	Speeds up compound and material design	Medium ***
13	Smart control	Intelligent control systems for optimized operations	Improves process efficiency and sustainability	Medium ***
14	Deep learning in SAR	Predict complex relationships between	New DL-driven level in drug discovery	Medium ***

		chemical structures and activities		
15	High-throughput experimentation (HTE)	Automated devices and AI in parallel experiments	Speeds up research processes	Medium ***
16	Digital Materials Design and Materials Informatics	Data-driven approaches for material discovery	Accelerates material discovery and design	Medium -Small **
17	Data-driven chemical reaction optimization	ML models for reaction optimization	Improves reaction efficiency	Medium -Small **
18	Automated synthesis planning	AI for synthetic route optimization	Streamlines synthesis planning	Medium -Small **
19	Chemoinformatics and chemical data analysis	Managing large chemical datasets	Supports chemical research	Medium -Small **
20	Quantum chemistry and simulations	AI-enhanced molecular simulations	Advances theoretical understanding	Medium -Small **

[a] A tentative summary is provided, which may change depending on the particular field of chemistry or problem considered. [b] Relative amount of peer-reviewed publications concerning the corresponding item (in % related to the overall amount for all 20 items) according to bibliographic database search; does not necessarily correlate with impact or importance.

2.2. Detailed description of digital chemistry trajectories

1. AI-driven Drug Discovery

AI-driven drug discovery represents a cutting-edge approach in pharmaceutical research where artificial intelligence, particularly machine learning and deep learning techniques, are employed to enhance the process of finding new drugs [22-30]. This approach harnesses the power of AI to sift through vast datasets drawn from chemical, biological, and medical sources to unearth potential drug candidates with desirable properties. The core idea is to use AI algorithms to predict how different chemical structures interact with biological targets, such as proteins or DNA, which are implicated in diseases. This predictive capability is crucial for identifying compounds that could modulate these targets in beneficial ways, potentially leading to new treatments.

This research direction is closely related with the direction discussed above (Deep Learning in Structure-Activity Relationships). However, the overall field of AI-driven drug discovery is more diverse and in some aspects is less challenging as compared to the ultimate goal of unsupervised deep learning in revealing general SARs.

AI-driven drug discovery involves the integration of diverse data types, including genomic data that offer insights into the genetic foundations of diseases, proteomic data that elucidate the structure and function of proteins, and chemical data that describe the properties and behaviors of millions of potential drug molecules. By analyzing these datasets, AI models can identify patterns and relationships that might not be apparent to human researchers. For example, an AI model might predict that a certain molecular structure is likely to bind effectively to a protein that plays a key role in a specific type of cancer, suggesting that compounds with this structure could be promising candidates for new cancer drugs.

One of the significant advantages of AI in this field is its ability to optimize lead compounds. Once a potential drug candidate is identified, it is rarely perfect; it might need to be more potent, more selective, or safer. AI algorithms can suggest modifications to the chemical structure of the lead compound to enhance its properties, guiding chemists in synthesizing new variants that are more likely to succeed in clinical trials.

This approach is particularly valuable for addressing several problems inherent in traditional drug discovery processes. The traditional route is rather slow and expensive, often taking over a decade and costing significant funding to bring a single new drug to market. AI can accelerate this process

by rapidly identifying promising candidates and optimizing them, potentially saving years of work and significant financial investment. Moreover, the complexity of biological systems and the subtleties of drug-receptor interactions make drug discovery a highly challenging field. The ability of AI to analyze complex, multidimensional data can lead to a deeper understanding of these systems, revealing new drug targets or highlighting the unforeseen therapeutic potential of existing molecules.

The potential practical applications and expected results of AI-driven drug discovery are vast and promising. By accelerating the identification and optimization of drug candidates, AI has the potential to provide new treatments to patients much faster than traditional methods. This acceleration is particularly crucial for diseases that currently lack effective treatments, such as many rare diseases, or for rapidly evolving health threats, such as new viruses. Furthermore, the optimization of AI can lead to drugs that are not only effective but also have fewer side effects, improving patient outcomes and safety. In the area of personalized medicine, the ability of AI to incorporate genetic and other biomarker data into the drug discovery process can lead to the development of treatments that are tailored to individual patients, enhancing treatment efficacy and reducing adverse reactions.

It should be noted, that this area strongly benefits from the methodology development made within the “#14. Deep Learning in Structure-Activity Relationships (SAR)” research direction. The results achieved within this area may be also used to facilitate development of another related area – “#10. Predictive Toxicology” as described in the article.

AI-driven drug discovery is a demanding approach that leverages the analytical and predictive power of artificial intelligence to revolutionize how new medications can be found and developed. By enabling faster, more efficient identification and optimization of drug candidates, AI has the potential to significantly accelerate the pace of medical innovation, bringing much-needed therapies to patients more quickly and safely.

2. Big Data and Integrated Data

Transforming chemistry through Big Data analytics is an ongoing effort,[31-36] but desired ambitious goals are difficult to achieve using single-target/single-objective databases alone. Data fragmentation is a limiting factor that has to be overcome to improve the performance. Integrated

chemistry data involve the consolidation and harmonization of data from various sources within and related to the field, including experimental results, computational simulations, chemical databases, and literature. This approach aims to create a unified, accessible, and interoperable data ecosystem that can support more effective and comprehensive analysis, fostering cross-disciplinary research and innovation.

The process of integrating chemical data involves overcoming significant challenges, such as disparate data formats, inconsistent data standards, and varying levels of data quality and completeness. Data integration efforts typically require the development of sophisticated data management frameworks that can handle the complexity of chemical data, including the structural, spectral, and property data of chemical compounds, as well as the conditions and outcomes of chemical reactions. These frameworks often utilize advanced technologies such as semantic web tools, data ontologies, and machine learning algorithms to map relationships between datasets, ensuring that data from different sources can be effectively linked, queried, and analyzed together.

Integrated data can be applied to a wide range of problems in chemistry and related disciplines. For researchers working on drug discovery, integrated data allow for the aggregation of information on compound libraries, biological assays, and pharmacokinetic and pharmacodynamic (PK/PD) data, facilitating the identification of promising drug candidates and understanding their mechanisms of action. In materials science, the consolidation of data on material properties, synthesis methods, and application performance can accelerate the discovery of new materials with desired functionalities. Environmental chemists benefit from integrated data by being able to correlate pollution data with health outcomes and ecological impacts, supporting the development of more effective environmental protection strategies.

The practical applications and expected results of implementing integrated data in chemistry are profound. By breaking down data fragmentation and fostering interoperability among diverse data sources, integrated data enable a fresh view of research problems, allowing scientists to uncover insights that might be missed when data sources are considered in isolation. This comprehensive approach can lead to more robust scientific conclusions, drive innovation by uncovering new connections between seemingly unrelated pieces of information, and accelerate the speed of research by making relevant data more readily available to the scientific community.

Moreover, integrated data support the principles of open science and data sharing, encouraging collaboration among researchers across different fields and institutions. It can facilitate the development of predictive models and data-driven hypotheses, enhancing the efficiency of experimental research by guiding scientists toward the most promising research avenues. In the long term, integrated data initiatives can contribute to the creation of a more cohesive and collaborative scientific ecosystem, where data-driven insights lead to breakthroughs in understanding and technological advancements across the boundaries of traditional scientific disciplines.

The integration and harmonization of chemical data sources represent critical advancements in the management and utilization of scientific data. By enabling comprehensive data analysis and facilitating cross-disciplinary research, integrated data initiatives hold the promise of accelerating discovery and innovation in chemistry and beyond, driving forward our understanding of complex scientific phenomena and our ability to address pressing global challenges.

3. Automated Laboratory Platforms

Automated laboratory platforms refer to the integration of advanced robotics, computer systems, and software to create laboratory environments where experiments can be conducted with minimal to no direct human intervention.[37-42] These platforms may be equipped with robotic arms, automated pipetting systems, sensors, and other devices that can precisely handle liquids, solids, and gases; replicate experimental setups; and carry out a wide range of laboratory procedures, from simple mixing to more complex synthesis and analysis. Robotic process automation (RPA) is of course easier to implement for repetitive tasks but the ultimate goal is to achieve a universal laboratory platform.

The core of these automated systems lies in their sophisticated software, which allows researchers to design and program experiments, control laboratory instruments, and collect and analyze data, all within a unified interface. It is expected, that this software can be integrated with databases and computational tools, enabling automated platforms to not only execute experiments but also make data-driven decisions based on predefined criteria or real-time analysis. Synergy application with cobots technology (designed to work in a shared workspace alongside humans) may also be an option.

Automated laboratory platforms can be applied to a variety of problems across research and development in chemistry and related fields. They are particularly valuable in high-throughput screening, where thousands of samples need to be tested under uniform conditions to identify active compounds in drug discovery or to optimize catalysts in materials science. These platforms may be also superior in repetitive or time-consuming tasks that require high precision, such as the synthesis of many complex molecules or the preparation of sample libraries for analysis.

The implementation of automated laboratory platforms may result in numerous practical applications and benefits. One of the most significant advantages is the improvement in the precision and reproducibility of the experiments. Automation reduces the variability associated with manual handling, ensuring that each step of an experiment is performed consistently. This not only enhances the reliability of the results but also makes the findings more reproducible across different laboratories.

Moreover, automated systems can operate continuously without the need for breaks or shifts, significantly increasing the throughput of experiments and reducing the time from hypothesis to conclusion. This capability is especially critical in fast-paced research areas where speed is essential, such as in the development of new medications or materials.

Automated platforms also contribute to safer laboratory environments by handling hazardous materials and carrying out dangerous processes without exposing researchers to risks. Furthermore, they allow for more efficient use of resources, as precise control over reagents and conditions minimizes waste and optimizes experimental designs.

Automated laboratory platforms represent a transformative approach to conducting research in the chemical sciences and beyond. By leveraging robotics, software, and data analytics, these systems enable more efficient, precise, and safe experimentation, paving the way for accelerated discoveries and innovations in a wide range of scientific disciplines.

4. Integration of Laboratory Instruments and IoT

The integration of laboratory instruments and the Internet of Things (IoT) is a burgeoning area in scientific research that involves connecting various laboratory instruments to a network, allowing automated data collection, real-time monitoring, and efficient data analysis.[43-48] This integration harnesses IoT technology – a system of interrelated computing devices capable of transferring data across a network without the need of human-to-human or human-to-computer interactions.

As representative examples, laboratory instruments such as spectrometers, chromatographs, pH meters, and even temperature and pressure sensors can be equipped with IoT capabilities. These devices collect data continuously and transmit it to a centralized system or cloud-based platform. This setup enables researchers to monitor experiments remotely, adjust conditions in real time, and gather extensive data that can be used for in-depth analysis. The IoT in the laboratory also allows for the automation of routine tasks, such as reagent refilling or temperature control, enhancing the efficiency and accuracy of experimental work.

Not limited to already existing IoT devices, nano-/micro-chip integration - miniscule computer components – can be incorporated into tools and equipment widely used in the laboratories. Tiny electronic components are designed to speed up daily tasks in the lab, and this technology holds great promise to link many already existing laboratory devises with AI. Improving precession and reproducibility in often done operations is one of the expected outcomes.

The integration of laboratory instruments with IoT technology can be applied to a wide range of problems. In pharmaceutical research, it enables continuous monitoring of drug synthesis processes, ensuring quality and consistency. In environmental monitoring, IoT-equipped sensors can track pollutant levels or detect hazardous substances in air or water with high precision and in real time. In the field of materials science, the IoT can facilitate the detailed study of material properties under various conditions, accelerating the discovery of new materials.

The practical applications and expected results of integrating laboratory instruments with the IoT are substantial. This approach can significantly increase the efficiency of research laboratories by automating data collection and analysis, reducing manual errors, and allowing scientists to conduct more experiments simultaneously. Remote monitoring capabilities also mean that experiments can be managed outside of conventional lab hours, increasing productivity.

Furthermore, the vast amounts of data collected by IoT-enabled instruments can be used to uncover trends and patterns that might not be visible through manual analysis. This big data approach can lead to new insights and discoveries in various fields of science. In terms of safety, IoT integration can provide immediate alerts in hazardous conditions, such as chemical leaks or unsafe temperature fluctuations, thereby enhancing laboratory safety.

The integration of laboratory instruments with the IoT represents a significant improvement in scientific research methodology. By automating and streamlining data collection, monitoring, and

analysis processes, IoT technology not only boosts the efficiency and effectiveness of laboratory work but also opens up new possibilities for innovation and discovery across a range of scientific disciplines.

5. AI in Spectroscopy and Analytical Method Development

Spectroscopy and analytical method development, when combined with AI applications, represent a significant advancement in the way complex spectroscopic data are interpreted and how novel analytical methods are developed.[17, 49-53] This interdisciplinary approach demands artificial intelligence, particularly machine learning algorithms, to analyze and make sense of the vast and complex datasets generated by various spectroscopic techniques, such as NMR, IR, UV–Vis, mass spectrometry and others. The integration of AI helps to automate data analysis processes, enhance the accuracy of interpretations, and uncover patterns and insights that might be invisible to traditional analysis methods.

AI algorithms are trained on large datasets of spectroscopic measurements and their corresponding outcomes or interpretations. These algorithms learn to recognize the complex relationships between spectroscopic features and the chemical or physical properties they represent. This learning enables AI systems to predict properties, identify compounds, or elucidate structures on the basis of new spectroscopic data, significantly reducing the time and expertise required for data analysis.

The application of AI in spectroscopy and analytical method development spans a broad range of problems. For example, in pharmaceutical research, AI-enhanced spectroscopy can rapidly identify and quantify the presence of various compounds in drug formulations, ensuring quality and consistency. In environmental monitoring, AI can improve the detection and quantification of pollutants in air, water, and soil samples, even when those contaminants are present in trace amounts under complex backgrounds. In materials science, AI-driven spectroscopic analysis aids in characterizing materials and understanding their composition and structure at the molecular level.

The results and practical applications of integrating AI into spectroscopy and analytical method development are vast and manifold. For instance, in the field of drug discovery and development, this approach can lead to more efficient identification of potential drug candidates and monitoring of their stability and degradation products. In forensic science, AI-enhanced spectroscopy can

provide more rapid and accurate analyses of samples, aiding in criminal investigations. In the field of food safety, it can be used to detect adulterants and contaminants, ensuring the safety and integrity of food products.

Moreover, AI-driven spectroscopy opens up new avenues for developing novel analytical methods that are more sensitive, more selective, and more capable of handling complex mixtures or samples with minimal preparation. These advancements not only improve the throughput and reliability of spectroscopic analyses but also expand the scope of their applications, enabling new scientific discoveries and innovations.

The integration of AI applications in spectroscopy and analytical method development is reshaping the landscape of analytical chemistry. Achieving enhanced interpretability of complex spectroscopic data and facilitating the development of novel analytical methods, AI is setting new standards for efficiency, accuracy, and chemical analysis, with wide-ranging implications for science, industry, and environmental monitoring.

6. Blockchain in Chemical Supply Chains

Blockchain in the chemical supply chain refers to the application of blockchain technology to create a secure, immutable, and transparent record-keeping system for tracking the movement of chemicals and materials from production to end use.[54-57] Blockchain, a distributed ledger technology, enables multiple parties to have access to a common database, where transactions are recorded in a way that prevents tampering, alteration, or deletion, thereby ensuring data integrity and trust among participants.

In the context of the chemical industry, the supply chain involves numerous stakeholders, including retailers, raw material suppliers, distributors, manufacturers, as well as end users. Each of these transactions and handovers traditionally relies on paperwork and centralized databases, which can be prone to errors, inefficiencies and even fraud. By implementing blockchain technology, each transaction in the supply chain can be recorded as a block of data, which is then linked to the previous transaction, forming a chain. This decentralized and encrypted record ensures that all parties have real-time access to trustworthy data regarding the origin, movement, and handling of chemical products.

Blockchain technology can be applied to various problems within chemical supply chains. It addresses issues of counterfeiting and adulteration by providing an unforgeable record of product authenticity and provenance. It also enhances regulatory compliance by maintaining a transparent record of safety data, handling instructions, and environmental impact assessments, which can be readily audited by authorities. Moreover, blockchain facilitates more efficient recall processes and dispute resolutions by providing an indisputable record of transactions and product movements.

The results and practical applications of blockchain in chemical supply chains are manifold. Enhanced security and transparency lead to increased trust among supply chain participants, from suppliers to consumers, fostering stronger collaboration and partnerships. The technology's ability to provide real-time visibility to the supply chain helps companies manage inventory more effectively, reduce costs, and improve responsiveness to market demands. Additionally, blockchain's capacity for automating transactions through smart contracts can further optimize operations and reduce administrative overhead (including such possibilities as self-executing contracts, agreement terms directly incorporated into the code, etc.).

Furthermore, blockchain technology supports the sustainability goals of the chemical industry by enabling more effective tracking of materials' environmental footprints. Companies can monitor and verify sustainable sourcing practices, waste management, and recycling activities, contributing to more responsible and eco-friendly supply chains.

The integration of blockchain technology into chemical supply chains represents a significant improvement in ensuring the integrity, efficiency, and sustainability of chemical and material movements. By providing a secure and transparent mechanism for tracking and verifying transactions, blockchain technology not only enhances operational efficiency but also builds trust and collaboration among all stakeholders involved, paving the way for more resilient and responsible supply chains in the chemical industry.

7. Digital Twins

Digital Twins in the area of chemistry refer to the development of comprehensive virtual models that accurately replicate chemical reactions, processes or systems.[58-64] These digital models are designed to simulate the physical and chemical behaviors of their real-world counterparts in real time, integrating data from various sources, including sensors, operational data, and collected

performance records. The concept is based on the convergence of the physical and digital worlds, where every aspect of a chemical process or system is mirrored in a virtual model, allowing for in-depth analysis, prediction, and decision-making.

The creation of a digital twin involves mapping the physical attributes, operational dynamics, and interdependencies of a chemical process or system into a virtual model. This model is continuously updated with real-time data, enabling it to reflect the current state of the physical system more accurately. Advanced simulation techniques, coupled with data analytics and machine learning algorithms, allow the digital twin to predict future states, identify potential issues, and recommend optimizations.

Digital twins can be applied to a wide array of problems within the chemical engineering and processing industries. They are particularly useful in process optimization, where they can simulate the effects of changes in process parameters on efficiency, yield, and product quality without the risks and costs associated with physical experimentation. In safety and risk management, digital twins can predict the outcomes of hazardous scenarios, helping to develop more effective mitigation strategies. They also play a crucial role in maintenance and reliability, predicting equipment failures before they occur and suggesting preventative actions, thereby reducing downtime and extending the lifespan of physical assets. Stress testing is one of the areas where digital twins can take a leading role as practical stress testing could be undesirable, highly expensive or even impossible to perform.

The practical applications and expected results of employing digital twins in chemistry are manifold. First, they enable a more agile and informed decision-making process, as operators and engineers can visualize the impacts of their decisions in the virtual model before implementing them in the real world. This leads to more efficient, sustainable, and safe chemical processes, with optimized resource use and minimized waste production.

Moreover, digital twins facilitate the development of more innovative and complex chemical processes by allowing for the exploration of scenarios that are too costly, dangerous, or time-consuming to test physically. They support the transition toward Industry 4.0 in the chemical sector, promoting the integration of IoT, AI, and other digital technologies into traditional chemical manufacturing practices.

Digital twins represent a significant leap forward in the digitalization of the chemical industry, offering a powerful tool for simulation, monitoring, and optimization. By providing a dynamic, accurate, and real-time representation of chemical processes and systems, digital twins not only enhance operational efficiency and safety but also open new avenues for innovation and improvement in chemical engineering and process management.

8. Virtual Laboratories and Augmented Reality

Virtual reality (VR) and Augmented Reality (AR) represent a modern approach to teaching and learning in the field of chemistry, utilizing digital platforms and simulation software to enhance educational experiences and provide students with virtual laboratory opportunities.[65-71] This innovative teaching methodology addresses several challenges in traditional chemical education by offering interactive, safe, and accessible learning environments.

Already well-addressed opportunities, digital platforms for chemical education, encompass a wide range of tools, including online courses, interactive textbooks, and educational apps, that provide comprehensive content on various chemistry topics. These platforms often incorporate multimedia elements such as interactive surveys, animations, videos, and making the learning process more engaging and effective for students with diverse learning styles.

Virtual laboratories take this one step further by simulating lab experiments through software applications, allowing students to conduct experiments in a virtual environment. These simulations replicate the setup, procedures, and outcomes of real-life experiments without the need for physical laboratory equipment or materials. Students can change variables, conduct experiments, and observe outcomes in real time, just as they would in a physical laboratory, but with the added benefits of safety, cost-effectiveness, and the ability to repeat experiments multiple times with ease.

The application of digital platforms and virtual laboratories is particularly valuable in situations where access to physical lab facilities is limited, such as in schools with budget constraints, for distance learners, or during circumstances that prevent in-person lab work, such as the COVID-19 pandemic. Moreover, they provide an excellent platform for preliminary training and practice before students engage in real lab experiments, reducing the learning curve and potential safety risks associated with hands-on lab work.

The expected results and practical applications of integrating digital platforms and virtual laboratories into chemical education are manifold. Students gain a deeper understanding of complex chemical concepts through interactive and immersive learning experiences. The virtual labs allow for the exploration of a wider range of experiments, including those that might be too dangerous, expensive, or time-consuming to perform in a traditional laboratory setting. This approach also fosters critical thinking and problem-solving skills as students navigate through virtual experiments, analyze data, and make decisions based on their observations.

Furthermore, the use of digital tools in chemical education promotes inclusivity and accessibility, ensuring that students from diverse backgrounds and geographical locations have equal opportunities for high-quality education and lab experience. It also prepares students for the increasingly digital future of scientific research, where computational tools and simulations play a crucial role.

Augmented reality with artificial components is expected to play an important role. Unlike VR, AR does not replace the reality but enhances it by adding computer generated elements. For example, by adding virtual labels to advance learning and help navigation in the laboratory. Interactive hardware labeling in real time may help measurements and experiments, with the digital information smoothly aligned with physical reality.

Chemical education within the VR and AR landscape, powered by digital platforms and simulation software, represent a significant advancement in the way chemistry is taught and learned. Via providing engaging, safe, and accessible educational experiences, this approach not only enhances the quality of chemical education but also equips researchers with the skills and knowledge needed to succeed in the modern scientific world.

9. Natural Language Processing (NLP) in the Chemical Space

Natural language processing (NLP) in chemical space involves the application of computational techniques designed to understand, interpret, and output results in human language as it pertains to the field of chemistry.[5, 72-76] This technology is employed to mine vast quantities of textual data within scientific literature, patents, and chemical databases, extracting valuable chemical information and insights. NLP tools should be capable of recognizing and processing chemical terminology, structures, and reactions described in text, converting unstructured data into structured data that can be easily analyzed and utilized.

NLP techniques in chemistry involve tasks such as entity recognition, where chemical compounds, reactions, and other relevant entities are identified in text. Possible examples include relationship extraction, which involves determining the interactions and associations between identified entities; and text classification, which categorizes text segments according to their content, such as distinguishing between experimental sections and discussion sections in research papers.

NLP in chemical space can be applied to a range of problems. Researchers and chemists can rapidly review the existing knowledge, helping to identify previous work related to a specific compound or reaction, which is invaluable for avoiding redundant research efforts and for sparking new ideas. In the pharmaceutical industry, NLP can aid in drug discovery and development by extracting information on bioactive compounds, their targets, and therapeutic effects from a plethora of research articles and clinical study reports. Additionally, NLP tools can support regulatory compliance by automatically analyzing patents and regulatory documents to ensure that new chemical entities or processes do not infringe on existing intellectual property.
The results and practical applications of employing NLP in chemical space are significant. By automating the extraction of chemical information from text, NLP tools dramatically reduce the efforts and time needed for literature review and gathering data, allowing researchers to focus more on experimental work and less on manual information retrieval. This efficiency can accelerate research and development processes, leading to faster innovation cycles in chemical synthesis, materials science, and pharmaceuticals.

Moreover, NLP can uncover hidden connections and patterns within the scientific literature that might not be apparent through conventional reading, offering new perspectives and insights that can drive forward scientific discovery. For instance, NLP analysis might reveal an underexplored chemical reaction that could provide a new synthesis pathway for a valuable compound, or it might identify a potential drug candidate that has been overlooked in traditional reviews.

NLP in chemical space represents a powerful tool for modern chemists, harnessing the capabilities of artificial intelligence to mine the wealth of knowledge contained in scientific texts. Through transforming unstructured text into structured, actionable data, NLP opens up new avenues for research, innovation, and discovery in the chemical sciences, making the vast and ever-growing body of chemical space more accessible and actionable than ever before.

10. Predictive Toxicology

Predictive toxicology involves the use of computational models to assess the potential toxicity of chemical compounds, aiming to predict adverse effects on human health and the environment before they occur.[77-83] This field merges principles from toxicology, chemistry, biology, and computer science to forecast toxicological outcomes based on the molecular structure and properties of compounds. The essence of predictive toxicology is to shift the paradigm from traditional reactive approaches, which rely on experimental testing and observation of adverse effects, to proactive strategies that can anticipate risks and inform safer chemical design and usage. These efforts share some common points with "#14. Deep Learning in Structure-Activity Relationships (SAR)" and "#1. AI-driven Drug Discovery" and sometimes use common tools, however this topic has its own dedicated focus.

As key examples, the significance of predictive toxicology is particularly pronounced in the context of agrochemicals development and chemical manufacturing. Predictive toxicology models can also screen potential drug candidates early in the development process, identifying those that may pose significant toxicological risks and thus reducing the likelihood of late-stage failures, which are costly and time-consuming. This early assessment helps in prioritizing compounds with a higher safety profile for further development, thereby optimizing resource allocation and enhancing the overall efficiency of the drug discovery pipeline.

In chemical manufacturing, predictive toxicology plays a vital role in evaluating the safety of chemicals used in various consumer products, industrial processes, and environmental applications. By predicting the toxicological profile of new and existing chemical substances, manufacturers can make informed decisions about the handling, usage, and disposal of these chemicals to minimize occupational and environmental risks. This proactive approach not only safeguards human health and the environment but also helps companies comply with regulatory standards and avoid potential legal and financial repercussions associated with the production and use of hazardous substances.

The methodologies employed in predictive toxicology range from quantitative structure-activity relationship (QSAR) models, which correlate chemical structure with biological activity, to more complex systems biology models that simulate the interactions of chemicals with biological systems at various levels of organization. Advances in machine learning and artificial intelligence

have further enhanced the capabilities of predictive toxicology, enabling the analysis of large datasets to uncover subtle patterns and relationships that may indicate toxicological concerns.

The practical applications of predictive toxicology extend beyond safety assessments to include the design of safer chemicals known as "benign by design." This approach involves using predictive models to guide the synthesis of new compounds with reduced toxicity profiles, thereby inherently improving the safety of the design process. Additionally, predictive toxicology contributes to the reduction of animal testing by providing alternative in silico methods for toxicity assessment, aligning with ethical considerations and regulatory mandates aimed at minimizing animal use in research.

For environment/health protection and a number of related areas, predictive toxicology represents an important approach in chemical safety assessment, offering a powerful toolset for preemptively identifying potential toxicological risks associated with chemical compounds. Proposed integration into chemical manufacturing processes promises to enhance safety, efficiency, and sustainability, ultimately leading to safer pharmaceuticals and consumer products while protecting human health and the environment.

11. AI in Environmental Chemistry and Sustainability

Environmental chemistry and sustainability, supported by digital tools, encapsulates the innovative use of technology to analyze environmental processes, enhance pollution control measures, and foster the development of green chemistry practices.[84-89] This approach integrates computational models, data analytics, remote sensing technologies, and information systems to gain insights into environmental phenomena, assess the impact of pollutants, and design chemical processes that are benign to the environment.

Digital tools offer a comprehensive platform for monitoring environmental parameters in real time, allowing for the detailed tracking of air and water quality, soil composition, and the presence of hazardous substances. Advanced algorithms and data analytics can process vast amounts of environmental data collected from various sources, including satellite imagery, sensor networks, and historical databases, to identify trends, predict future conditions, and assess the effectiveness of pollution control strategies.

In the field of pollution control, digital tools enable the precise modeling of pollutant dispersion, the identification of pollution sources, and the assessment of potential environmental and health risks associated with exposure to various contaminants. This capability is crucial for developing targeted interventions and regulatory policies that can effectively mitigate the impact of pollutants on ecosystems and human health. For instance, predictive models can simulate the spread of an oil spill in marine environments, guiding cleanup efforts and minimizing environmental damage.

Furthermore, the principles of green chemistry, which are aimed to create chemical products and design processes that reduce or eliminate the use and generation of hazardous substances, are significantly supported by digital technologies. Computational chemistry and simulation tools can predict the environmental impact and toxicity of chemical compounds before they are synthesized, encouraging the development of safer alternatives. Life cycle assessment (LCA) software aids in evaluating the environmental footprint of chemical processes and products, from raw material extraction to end-of-life disposal, promoting more sustainable practices across the chemical industry.

The practical applications and expected results of employing digital tools in environmental chemistry and sustainability are diverse and impactful. Enhanced environmental monitoring and pollution control can lead to healthier ecosystems and improved public health by ensuring cleaner air, water, and soil. Digital tools also support the advancement of green chemistry by facilitating the discovery of environmentally friendly materials, energy-efficient processes, and waste-minimization techniques. In the broader context of sustainability, these technologies contribute to the achievement of global environmental goals, such as reducing greenhouse gas emissions, protecting biodiversity, and conserving natural resources.

The integration of digital tools in environmental chemistry and sustainability represents a forward-thinking approach that leverages the power of technology to address some of the most pressing environmental challenges. By providing a deeper understanding of environmental processes, enhancing pollution control, and driving the development of green chemistry, digital technologies play a pivotal role in promoting a more sustainable and environmentally conscious future.

12. Machine Learning in Molecular Design

ML in molecular design is a cutting-edge approach that employs advanced algorithms to predict the properties of molecules and thereby guide the synthesis of new compounds and materials with desired characteristics.[90-95] This technique leverages the vast amounts of chemical data available, including known molecular structures, their associated properties, and the outcomes of various chemical reactions, to train ML models. These models can then be used to analyze the relationships between molecular structures and their properties, making it possible to predict how changes in a molecule's structure might influence its behavior and effectiveness.

ML algorithms can digest and learn from the collected earlier data of chemical compounds, encompassing their successes and failures across various applications. By identifying patterns and correlations within these data, ML models can predict the properties of novel molecules that have not yet been synthesized. This capability is particularly beneficial in predicting the biological activity of new compounds aiming to drastically reduce the time and resources spent on experimental testing. Similarly, in materials science, ML can predict the physical, chemical, and mechanical properties of new materials, guiding researchers in creating substances with specific characteristics, such as high strength, conductivity, or durability.

The application of ML in molecular design addresses several challenges in chemistry and related fields. First, the traditional approach for discovering new molecules and materials is often slow and labor intensive and involves many trials and errors. ML can streamline this process, enabling rapid screening and evaluation of countless potential compounds. This approach is also invaluable in tackling complex problems where the relationships between molecular structures and their properties are not fully understood, offering insights that can lead to breakthroughs in understanding and innovation.

The results and practical applications of ML in molecular design are broad and impactful. In pharmaceuticals, this approach can lead to the discovery of new drugs with greater efficacy and fewer side effects, potentially revolutionizing treatments for various diseases. In the area of energy research, ML-designed materials can lead to more efficient solar cells, batteries, and catalysts, contributing to the advancement of renewable energy technologies. Moreover, in the field of electronics, ML can aid in the development of novel semiconductors, leading to faster and more efficient electronic devices.

Furthermore, the environmental sector can benefit from ML in molecular design through the development of materials that are more recyclable or better at capturing pollutants, thus contributing to sustainability efforts. Overall, the integration of ML in molecular design should accelerates the discovery and innovation in chemistry and materials science and also holds promise for addressing some of the most pressing challenges in healthcare, energy, electronics, and environmental protection, leading to advancements that could have a profound impact on society and the world at large.

13. Smart Control

In the area of chemical processes and equipment, smart control refers to the use of intelligent control systems that leverage advanced algorithms, sensors, and automation technologies to enhance the efficiency, safety, and reliability of chemical operations.[96-101] These systems are designed to monitor process parameters in real time, make data-driven decisions, and adjust operating conditions dynamically to achieve optimal performance. The concept also incorporates distance control since direct on-site interaction with human operator is not required.

At the heart of smart control systems are sophisticated algorithms, including the power of machine learning and artificial intelligence, that can predict process outcomes based on collected historical and real-time data. By analyzing data from a myriad of sensors installed throughout chemical reactors, separation units, and other equipment, these algorithms can identify patterns, anticipate potential issues, and optimize process conditions. For instance, smart control systems can adjust reaction temperatures, pressures, and feed rates to maintain ideal conditions, even in the face of fluctuating input qualities or unexpected disturbances.

Smart control can be applied to a broad range of problems in the chemical industry. Process optimization maximizes the efficiency and yield of chemical reactions, reducing waste and energy consumption. For safety and environmental compliance, intelligent control systems can detect and mitigate hazardous conditions, such as leaks or emissions, before they pose significant risks. In predictive maintenance, smart control systems monitor the health of equipment, predict failures and schedule maintenance to avoid costly downtime.

The practical applications and expected results of implementing smart control systems in chemical processes are significant. Enhanced process efficiency leads to cost savings, increased production

capacity, and reduced environmental impact. Improved safety measures protect workers and surrounding communities, while compliance with environmental regulations is ensured through more effective pollution control. Predictive maintenance capabilities extend the lifespan of equipment and reduce the likelihood of unexpected operational disruptions.

Furthermore, smart control systems facilitate the implementation of advanced manufacturing concepts such as process intensification and modular chemical processing. These systems support the development of more flexible and scalable manufacturing operations by enabling precise control over smaller, more efficient units. This adaptability is particularly valuable in the production of specialty chemicals and pharmaceuticals, where demand can be variable and product specifications stringent.

The implementation of intelligent control systems in chemical processes represents a transformative shift toward more adaptive, efficient, and safe operations. By harnessing the power of data and automation, smart control systems not only optimize individual processes but also contribute to the broader goals of sustainability and resilience in the chemical industry. As these technologies continue to evolve, they will undoubtedly play a pivotal role in shaping the future of chemical manufacturing, driving innovation and efficiency across the sector.

14. Deep Learning in Structure-Activity Relationships (SAR)

Deep learning in structure-activity relationships (SAR) involves the application of multilayer neural networks to understand and predict how the structure of chemical compounds affects their biological activity.[92, 102-105] This approach is particularly revolutionary in the field of drug discovery, where identifying compounds with desired biological effects is both critical and challenging.

In addition to their known biological activities, deep learning models, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), are trained on vast datasets comprising chemical structures, represented as molecular graphs or simplified molecular-input line-entry system (SMILES) strings. These models learn to recognize intricate patterns and features within the molecular structures that correlate with biological outcomes, going beyond traditional SAR analysis by considering not only simple molecular descriptors but also the entire structure in a holistic manner.

This deep learning approach can be applied to various problems within drug discovery and development. For instance, it can predict the efficacy of potential drug candidates against specific targets, thereby streamlining the initial screening process and focusing efforts on the most promising molecules. It can also be used to anticipate adverse drug reactions or off-target effects, enhancing the safety profile of new compounds. Furthermore, deep learning models can assist in the optimization of lead compounds, suggesting structural modifications that could improve efficacy, reduce toxicity, or enhance pharmacokinetic properties.

The practical applications and expected results of employing deep learning in SAR analysis are vast. By providing more accurate predictions of biological activity, deep learning models can significantly reduce the time and financial investment required to identify viable drug candidates, thereby accelerating the pace of drug development. This efficiency gain not only benefits pharmaceutical companies by speeding up the time-to-market for new drugs but also has broader implications for public health, potentially enabling quicker responses to emerging medical needs.

Moreover, deep learning models can uncover novel insights into the mechanisms of molecular interactions and biological pathways, contributing to a deeper understanding of disease processes and pharmacology. These insights can lead to the discovery of new therapeutic targets and the development of innovative treatment strategies. Methodology developments achieved in these efforts are very valuable for the “#1. AI-driven Drug Discovery” and “#10. Predictive Toxicology” discussed below.

The utilization of deep learning in understanding and predicting structure-activity relationships represents a paradigm shift in drug discovery. By harnessing the power of advanced computational models to analyze complex chemical and biological data, researchers can enhance the efficiency and effectiveness of the drug discovery process, paving the way for the development of new therapeutics and advancing our understanding of human health and disease.

15. AI-driven High-Throughput Experimentation (HTE)

High-throughput experimentation (HTE) is a methodological approach that leverages robotics, automation, and artificial intelligence (AI) to conduct a large number of chemical experiments simultaneously and rapidly.[106-111] This approach is useful for research processes, enabling scientists to explore vast chemical spaces and reaction conditions in a fraction of the time required by

traditional methods. HTE systems are designed to handle and analyze hundreds to thousands of samples per day, providing a level of efficiency and data generation that was previously unattainable.

At the core of HTE is the integration of robotic systems capable of performing repetitive laboratory tasks with precision and consistency. These tasks include pipetting, mixing reagents, controlling reaction conditions, and even analyzing the outcomes of reactions using various spectroscopic and chromatographic techniques. The automation of these processes reduces the potential for human error, increases throughput, and allows researchers to focus on more complex aspects of their work.

The role of AI in HTE is multifaceted. AI algorithms are employed to design experiments, predict outcomes, and analyze the outcome generated by high-throughput screens. For instance, machine learning models can suggest novel combinations of reaction conditions or materials to test based on patterns identified in previously collected data. After experiments are conducted, AI can also help in identifying successful outcomes, such as the synthesis of a desired compound or the discovery of a new catalytic activity, by analyzing complex data that might include spectral properties, reaction yields, and purity levels.

HTE is particularly valuable in fields where the exploration of a wide range of variables is crucial, such as in materials science, catalysis research, and drug candidates screening. In materials science, HTEs can accelerate the discovery of new materials with specific properties, such as high-strength alloys, superconductors, or photovoltaic materials, by systematically varying the composition and processing conditions. In catalysis research, HTE enables the rapid screening of catalyst libraries to identify candidates that facilitate reactions with high efficiency and selectivity and under milder conditions. In drug discovery, HTE is used to screen vast libraries of compounds for biological activity against various targets, vastly accelerating the identification of potential drug candidates.

The practical applications and results of HTEs are profound. By dramatically increasing the speed at which experiments can be conducted and analyzed, HTE has the potential to significantly shorten the research and development cycles for new chemicals, materials, and drugs. This not only leads to faster scientific advancements but also reduces the costs associated with R&D processes. Furthermore, the ability of HTE to generate large datasets provides a rich resource for further computational analysis and model building, enhancing the predictive capabilities of AI in science.

Active utilization of high-throughput experimentation may implement a paradigm shift in how regular laboratory research is conducted, characterized by a significant increase in efficiency and productivity. The integration of robotics and AI in HTE has not only accelerated the pace of discovery but also opened up new avenues for exploration that were previously considered impractical due to resource and time constraints. As such, the HTE is poised to continue playing a pivotal role in advancing research across a multitude of scientific disciplines.

16. Digital Materials Design and Materials Informatics

Digital materials design and materials informatics refers to the intersection of materials science and information technology, employing data-driven methodologies to understand, discover, and design new materials.[11, 112-117] This approach leverages large datasets of material properties, processing conditions, and performance metrics, along with advanced data analytics, machine learning algorithms, and computational modeling, to uncover patterns and relationships that can guide the development of materials with tailored properties. Development of new generation of high-performance catalytic materials is also included in this direction.

This field involves the collection and analysis of data from experimental studies, computational simulations, and the literature to construct predictive models. These models can forecast the properties and performance of materials based on structure, composition, and processing history. By systematically exploring the vast space of potential material compositions and configurations, digital materials design can identify promising candidates for specific applications much more rapidly than traditional trial-and-error methods.

Materials informatics can be applied to a wide array of problems in materials science and engineering. For instance, it is instrumental in the development of high-performance alloys for aerospace and automotive applications, where the specific strength, durability, and resistance to environmental degradation are critical. In the energy sector, materials informatics can accelerate the discovery of new battery materials with higher energy densities, faster charging rates, and longer lifespans, addressing the growing demand for efficient energy storage solutions. Additionally, in the electronics industry, this approach supports the development of novel semiconductors, dielectrics, and conductive materials essential for next-generation electronic and photonic devices.

The results and practical applications of digital materials design are profound and far-reaching. By enabling a more rational and efficient exploration of the material design space, it significantly shortens the development cycles for new materials, reducing costs and accelerating the pace of innovation. This will have a direct impact on various industries, leading to the creation of lighter, stronger, and more sustainable materials for transportation, more efficient and durable energy storage and conversion systems, and more advanced electronic devices with improved functionalities.

Moreover, digital technologies contribute to the sustainability of material development processes by identifying material compositions and processing routes that minimize the use of toxic or scarce elements and reduce energy consumption during manufacturing. It also plays an important role in advancing the fundamental understanding of material behavior, providing insights into the underlying mechanisms that govern material properties and performance.

Digital materials design and materials informatics contribute to development of a transformative approach in materials science, harnessing the power of data and computation to drive the discovery and design of new materials. Its integration into the material development process promises not only to enhance technological innovation across a broad spectrum of applications but also to contribute to the development of more sustainable material solutions for the future.

17. Data-Driven Chemical Reaction Optimization

Data-driven chemical reaction optimization, facilitated by machine learning models, represents a transformative approach in the field of chemistry, where algorithms are trained to predict the outcomes of chemical reactions, optimize reaction conditions, and uncover novel reactivity patterns.[118-123] This method leverages the vast amounts of reaction data accumulated in the scientific literature, databases, and experimental results to construct predictive models that can guide chemists in designing more efficient and effective chemical reactions. Bayesian optimization, reinforcement learning and a range of other methods are actively involved.

The process involves the collection and analysis of diverse data related to chemical reactions, including reactants, catalysts, solvents, temperatures, and yields. ML algorithms, such as neural networks or decision trees, are then trained on these datasets to identify patterns and correlations that can predict how changes in reaction conditions might affect the outcome. This predictive

capability enables chemists to simulate reactions in silico before conducting them in the laboratory, saving time and resources by focusing on the most promising conditions.

Data-driven chemical reaction optimization can be applied to a wide range of problems in synthetic chemistry, materials science, and pharmaceutical development. For instance, in the synthesis of complex organic molecules, optimizing reaction conditions is critical for achieving high yields and selectivities. ML models can predict the best combination of catalysts, solvents, and temperatures to achieve the desired results. In materials science, these models can help in the development of new synthesis routes for advanced materials with specific properties. In the pharmaceutical industry, optimizing reactions can lead to more efficient production processes for active pharmaceutical ingredients, reducing costs and environmental impact.

The expected results and practical applications of employing ML models for chemical reaction optimization are significant. With enhancement of the predictability of reaction outcomes, chemists can avoid costly and time-consuming experimental dead ends, accelerating the discovery and development of new compounds and materials. This approach also has the potential to reveal new chemical reactivities and mechanisms that have not been observed before, expanding the boundaries of chemical knowledge.

Moreover, data-driven chemical reaction optimization can contribute to the principles of green chemistry by enabling more sustainable chemical processes. Optimized reactions often require less energy, generate fewer byproducts, and use safer reagents, aligning with the goals of reducing the environmental footprint of chemical manufacturing.

The integration of ML models into the optimization of chemical reactions represents a significant advancement in the field, offering a more efficient, predictive, and exploratory approach to chemical synthesis. Harnessing the power of data and machine learning would help chemists to improve the efficiency of existing reactions and also discover new pathways and reactivities, driving innovation in chemistry and related disciplines.

18. Automated Synthesis Planning

Automated synthesis planning harnesses AI tools to revolutionize the way chemists design and synthesize chemical compounds.[124-129] Traditionally, planning synthetic routes has been a complex

and time-consuming task, requiring chemists to draw upon their extensive knowledge of reaction mechanisms, reactivity, and functional group transformations. Automated synthesis planning systems aim to simplify the task by providing computer-generated pathways that predict the most efficient and practical routes for synthesizing desired molecules.

At the focus of automated synthesis planning is the use of ML algorithms to search through vast databases of chemical reactions, including both successful syntheses and experimental failures. These algorithms learn from the accumulated knowledge of chemical reactivity patterns, reaction conditions, yields, and scalability. By analyzing this information, the systems can propose multiple synthetic routes for a target molecule, ranking them based on factors such as the number of steps, the availability of starting materials, the likelihood of success, and the overall cost-effectiveness.

This approach can be particularly valuable in complex organic synthesis, such as the development of pharmaceuticals, where identifying the most efficient synthetic route can significantly accelerate drug discovery and development processes. This approach allows chemists to explore a broader array of synthetic possibilities, including routes that might not be immediately apparent through traditional methods. Moreover, automated synthesis planning can optimize existing synthetic pathways, suggesting modifications that reduce the number of steps, minimize hazardous reactions, or employ more readily available or sustainable reagents.

In practice, the application of automated synthesis planning can lead to substantial reductions in the time and resources required for experimental trials. By providing a ranked list of synthetic routes, chemists can prioritize their experimental work, focusing on the most promising approaches. This not only speeds up the laboratory workflow but also reduces the consumption of chemicals and energy, contributing to more sustainable research practices.

The practical applications of automated synthesis planning extend beyond academia to the pharmaceutical, agrochemical, and materials science industries. For instance, in drug development, it can be used to rapidly synthesize new drug candidates for testing or to find more efficient pathways for the production of approved drugs, thereby lowering manufacturing costs. Materials science can assist in the discovery of new materials with desirable properties by facilitating the synthesis of novel compounds.

As AI and ML technologies continue to advance, the capabilities of automated synthesis planning systems will expand, further enhancing their utility in chemical research. The integration of these

systems into the chemist toolkit represents a significant step forward in the digitalization of chemistry and is promising for accelerating scientific discovery and innovation across a wide range of chemical disciplines.

19. Chemoinformatics and Chemical Data Analysis

Chemoinformatics and chemical data analysis encompass a suite of advanced techniques focused on the management, analysis, and interpretation of large datasets in the chemical domain.[130-135] These methodologies leverage computational tools and statistical models to extract meaningful information from data related to chemical structures, properties, reactions, and biological activities. The essence of chemoinformatics lies in its ability to translate vast amounts of raw chemical data into actionable insights, thereby enhancing our understanding of chemical phenomena and facilitating the discovery of new chemical entities and materials.

At the core of chemoinformatics is the development of sophisticated algorithms and software tools capable of representing and adjusting chemical structures in silico. These representations enable the systematic analysis of molecular descriptors, which are numerical values that capture various aspects of a molecule's chemical and physical properties. By analyzing these descriptors, cheminformatics tools can identify patterns and relationships within chemical datasets that are not readily apparent through traditional analysis methods. This capability is invaluable in tasks such as virtual screening, where large libraries of compounds are evaluated for their potential biological activity based on their structural similarities to known active compounds.

Chemoinformatics techniques are particularly well suited for addressing problems where the complexity and scale of the data exceed the capacity for manual analysis. For example, in drug discovery, chemoinformatics can be used to analyze the structure-activity relationships of compound libraries, helping to identify promising drug candidates by predicting their activity against specific biological targets. Similarly, in materials science, chemoinformatics can facilitate the discovery of new materials with desired properties by analyzing the relationships between molecular structures and material characteristics.

The practical applications of chemoinformatics are wide-ranging and impactful. In the pharmaceutical industry, the identification of new drug candidates can be accelerated by enabling high-throughput screening of compound libraries against disease targets. In environmental

chemistry, cheminformatics tools can predict the toxicity and environmental fate of chemical pollutants, aiding in risk assessment and regulatory compliance. Additionally, in academia, cheminformatics supports fundamental research by providing insights into the mechanisms of chemical reactions and the principles governing molecular interactions.

As chemoinformatics methodologies continue to evolve, they are increasingly integrated with other computational disciplines, such as machine learning and data science, further enhancing their analytical capabilities. This integration allows for the development of predictive models that can forecast chemical properties and activities with high accuracy, opening new avenues for research and development in chemistry and related fields. The continued advancement of chemoinformatics promises to deepen our understanding of chemical systems, fasilitate the discovery process, and contribute to the development of new technologies and solutions across a broad spectrum of scientific disciplines.

20. AI in Quantum Chemistry and Simulations

Quantum chemistry and simulations involve the use of quantum mechanical principles to predict the properties and behaviors of atoms and molecules. This field is fundamental for understanding chemical reactions, material properties, and molecular structures at the most detailed level. However, quantum chemical calculations are notoriously complex and computationally demanding, often requiring significant resources and time, especially for large systems or highly accurate predictions.

The integration of artificial intelligence (AI) into quantum chemistry and molecular simulations represents a significant advancement, enhancing the accuracy and efficiency of these calculations.[28, 136-140] AI algorithms, particularly machine learning models, are trained on data generated from quantum chemical computations to predict the outcomes of similar calculations. This approach can significantly reduce the computational cost by providing accurate estimates of molecular properties without the need for full-scale quantum mechanical calculations for every new molecule or configuration.

AI-enhanced quantum chemistry can be applied to a wide range of problems where understanding the electronic structure and properties of molecules is crucial. This includes the design of new drugs, where accurate predictions of molecular interactions with biological targets are essential,

and the development of new materials with specific electronic, optical, or mechanical properties. In catalysis research, AI-enhanced simulations can predict the efficiency and selectivity of catalysts, guiding the design of better catalytic systems for chemical manufacturing processes.

The results and practical applications of AI-enhanced quantum chemistry and simulations are diverse and impactful. In drug discovery, it enables the rapid screening of vast libraries of compounds, identifying those most likely to bind effectively to disease targets with fewer experimental tests. Materials science can benefit from accelerating the discovery of novel materials with desirable properties, such as high-strength lightweight alloys for aerospace applications or efficient light-absorbing compounds for solar energy devices. The field of catalysis can facilitate the optimization of reaction conditions and the design of more efficient catalysts, contributing to more sustainable and cost-effective chemical processes.

Moreover, AI-enhanced quantum chemistry can lead to a deeper understanding of fundamental chemical phenomena, offering insights into reaction mechanisms and material behaviors that were previously inaccessible. This can open up new avenues in research, leading to innovations in various fields of chemistry and materials science.

The integration of AI into quantum chemistry and molecular simulations is changing the landscape of chemical and material research. By enhancing the accuracy and efficiency of quantum chemical calculations, AI enables scientists to address more complex systems and problems, accelerating the rate of discovery and innovation across a broad spectrum of applications.

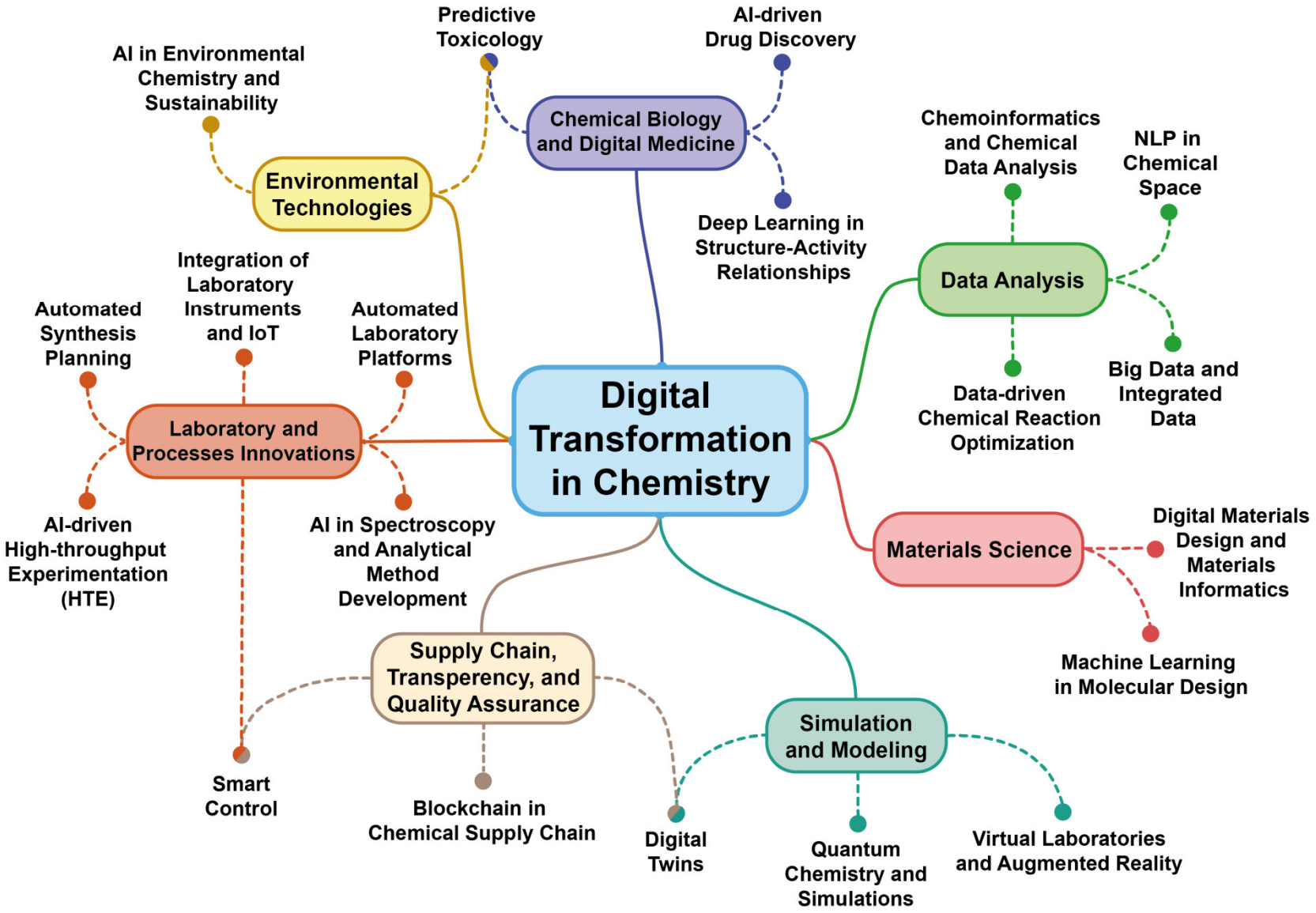

Figure 1. Connection of AI-based technologies with broader themes depending on their subject and applications.

2.3 Broader AI themes in chemistry

The defined 20 influential AI-based technologies in chemistry highlighted in the manuscript can be categorized into 7 broader themes that reflect their primary focus and application areas (Figure 1). These themes not only encapsulate the essence and application areas of the technologies but also offer a structured framework to understand their collective impact on the field of chemistry. Below is a description of these categories, the rationale for their formation, and their global importance in steering the evolution of chemistry into its digital era.

I. Chemical Biology and Digital Medicine

This category encompasses technologies focused on drug discovery, predictive toxicology, and understanding structure-activity relationships through deep learning (Figure 1). The grouping of these technologies underlines the critical role of AI in accelerating the development of therapeutics, predicting chemical safety, and elucidating complex biochemical interactions. In the global context, this theme is paramount for advancing personalized medicine, reducing drug development costs, and enhancing the safety profiles of chemical compounds, thereby significantly impacting healthcare outcomes worldwide.

II. Environmental Technologies

Application of AI in environmental chemistry and sustainability, along with predictive toxicology, constitutes this theme (Figure 1). It underscores the commitment to leveraging digital tools for environmental protection, green chemistry practices, and assessing chemical impacts on ecosystems. This thematic category highlights the global imperative to address environmental challenges, promote sustainability, and ensure chemical safety, reflecting a broader societal shift towards environmental stewardship.

III. Data Analysis

Integrating diverse data sources, managing large chemical datasets, extracting information through NLP, and optimizing chemical reactions are grouped here due to their collective focus on harnessing data for insightful chemical analysis (Figure 1). This theme showcases the transformative potential of AI in synthesizing knowledge from vast data pools, optimizing chemical processes, and driving discoveries. Globally, it represents a shift towards data-driven decision-making and research in chemistry, enhancing efficiency, innovation, and knowledge dissemination.

IV. Laboratory and Processes Innovations

This category brings together technologies that automate laboratory work, integrate IoT for data analysis, develop analytical methods, conduct high-throughput experimentation, plan synthesis, and implement smart control systems (Figure 1). The logic of grouping these technologies is based on their contribution to innovating laboratory practices and chemical processes, making research more precise, efficient, and scalable. On a global scale, this theme is critical for the future of research and manufacturing in chemistry, offering solutions to increase throughput, improve safety, and reduce costs.

V. Materials Science

Focused on the discovery and design of new materials, this theme encompasses technologies that apply AI to predict material properties and design compounds (Figure 1). It reflects the interdisciplinary nature of materials science and its pivotal role in developing next-generation materials for various applications, from energy storage to nanotechnology. Globally, this theme addresses urgent needs for aim-specific, high-performance materials, driving advancements in technology and industry.

VI. Supply Chain Transparency and Quality Assurance

Including technologies like blockchain for supply chain integrity and digital twins for process optimization, this category emphasizes the importance of transparency, traceability, and quality in chemical manufacturing (Figure 1). It highlights the role of AI in building trust, ensuring product quality, and optimizing production processes. Globally, this theme is fundamental for the ethical, safe, and efficient distribution of chemical products, aligning with demands for sustainability and consumer safety.

VII. Simulation and Modeling

This theme, encompassing digital twins, virtual laboratories, and quantum chemistry simulations, illustrates the power of AI-enhanced modeling and simulation in predicting chemical behaviors and facilitating education (Figure 1). It showcases the shift from traditional experimental methods to virtual experimentation and theoretical predictions. Globally, this theme advances our understanding of chemical phenomena, reduces experimental costs, and enables remote learning and research, democratizing access to chemistry education and innovation.

3. Conclusions

Transitioning into the modern era, digital chemistry represents a paradigm shift, leveraging computational power, machine learning algorithms, and vast datasets to accelerate and refine the discovery process. This digital approach offers unprecedented predictive capabilities, efficiency, and the ability to tackle complex problems that are beyond the reach of traditional methods alone. This may indicate a move toward more targeted, hypothesis-driven research, reducing the reliance on serendipity and making the discovery process more systematic and less resource intensive.

It should be noted that the some of the identified top 20 influential digital research trajectories have mutual interceptions and sometimes address similar targets using slightly different approaches. Not unexpectedly for such a new and actively evolving area, sometimes it is difficult to draw clear borders between the individual digital trajectories or distinguish clearly the particular segment in the chemistry research. Although some duplications in the methods used and targets stated may exist, the key trajectories identified upon analysis of current research data were described here. No attempts were made to manually merge or replace the trends captured. Further systematization and classification may be applied in the future when more trends and their impacts will be pronounced in details.

The future of chemical research lies in the integration of these diverse methodologies. The insights from classical chemistry continue to be invaluable, particularly in education and in areas where direct material interaction is crucial. Meanwhile, digital chemistry is set to expand, drive innovation and enable researchers to navigate the vast chemical space more effectively than ever before.

The categorization into themes (Figure 1) elucidates the multifaceted impact of AI on chemistry, highlighting areas of significant advancement and global challenges being addressed. These themes collectively underscore the shift from empirical, labor-intensive research towards a more predictive, efficient, and rapid approach facilitated by digital technologies. As chemistry transitions into its digital era, these themes play pivotal roles in shaping research priorities, industry practices, and educational paradigms, driving innovation, sustainability, and global collaboration. The integration of AI and digital tools across these themes not only transforms how chemistry is

conducted but also broadens its application, impact, and accessibility, marking a new frontier in the quest for scientific discovery and technological progress.

Despite an attempt at detailed analysis made in this study, plausible limitations should also be mentioned. The analysis within this article, while comprehensive, may have limitations in fully capturing or accurately predicting all trends within the rapidly evolving field of digital chemistry and AI applications. The dynamic nature of technological advancements and their interdisciplinary impacts make it challenging to foresee all future developments and their implications comprehensively. These constraints highlight the inherent difficulty in predicting the future trajectory and impact of such a rapidly changing domain, acknowledging that some emerging trends may be overlooked or misunderstood within the scope of this analysis.

As we move forward, the challenge and opportunity for the chemical research community will be to balance key approaches—joining the irreplaceable human intuition and expertise of classical methods—with the computational power of digital techniques. This balanced approach will not only enhance our understanding and capabilities within the field but also ensure that the journey of discovery remains as rich and unpredictable as the science of chemistry itself. The integration of these methodologies will prepare future chemists to navigate an increasingly complex and interdisciplinary landscape, driving forward innovations that address some of the most pressing challenges of our time.

Methodology

The present analysis is based on a multidisciplinary approach, incorporating insights from research articles, reviews, and trend analyses within the chemical sciences. Advanced data collection across major databases, including Google Scholar, Web of Science, Scopus, eLibrary, and PubChem, were employed to ensure a thorough representation of current and emerging trends.

AI technologies, including Large Language Models (LLM) and Generative Pre-trained Transformers (GPT), were utilized in this work. These tools were used to facilitate data processing, analysis, information clustering, trend distillation, and assessment of influential potential. Additionally, AI was utilized for text translation and preparation. Multiparametric comparison was used to enhance the overall efficiency and depth of the analysis. By integrating these methods, this article provides

a detailed understanding of the current landscape and future directions of AI in chemical sciences, highlighting the transformative potential and broad applicability of these technologies in the field.

Limitations. Despite the variety of tools employed, this study may still possess several limitations. The reliance on available literature and databases may not encompass all emerging technologies and innovations in AI and chemical sciences, potentially overlooking unpublished or proprietary research. Additionally, the rapid pace of technological advancement means that some findings may quickly become outdated. The interpretative nature of AI-based analysis also introduces a subjective element, which might affect the objectivity of trend predictions and significance assessments. These constraints underscore the need for ongoing research and validation in this dynamically evolving field.

Keywords: chemical science; artificial intelligence; digital technologies; research methodology; machine learning; chemical industry; digital twins; blockchain; large data

Acknowledgement. The author thanks Daniil Boyko, Darya Prima and Darya Arkhipova for manuscript reading and helpful discussion.

Cited literature

Uncategorized References

1. S. Back, A. Aspuru-Guzik, M. Ceriotti, G. Gryn'ova, B. Grzybowski, G. H. Gu, J. Hein, K. Hippalgaonkar, R. Hormazabal, Y. Jung, S. Kim, W. Y. Kim, S. M. Moosavi, J. Noh, C. Park, J. Schrier, P. Schwaller, K. Tsuda, T. Vegge, O. A. von Lilienfeld and A. Walsh, Accelerated chemical science with AI, *Digit Discov*, 2024, **3**, 23-33.
2. Z. Zheng, O. Zhang, C. Borgs, J. T. Chayes and O. M. Yaghi, ChatGPT Chemistry Assistant for Text Mining and the Prediction of MOF Synthesis, *J Am Chem Soc*, 2023, **145**, 18048-18062.
3. S. Matsubara, Digitization of Organic Synthesis — How Synthetic Organic Chemists Use AI Technology, *Chem Lett*, 2020, **50**, 475-481.
4. D. A. Boiko, R. MacKnight, B. Kline and G. Gomes, Autonomous chemical research with large language models, *Nature*, 2023, **624**, 570-578.

5. B. Mahjour, J. Hoffstadt and T. Cernak, Designing Chemical Reaction Arrays Using Phactor and ChatGPT, *Org. Process Res. Dev.*, 2023, **27**, 1510-1516.
6. J. B. Mitchell, Machine learning methods in chemoinformatics, *Wiley Interdiscip Rev Comput Mol Sci*, 2014, **4**, 468-481.
7. X. Zhou and M. Kraft, Blockchain Technology in the Chemical Industry, *Annu Rev Chem Biomol Eng*, 2022, **13**, 347-371.
8. J. V. Capella, A. Bonastre, J. C. Campelo, R. Ors and M. Peris, IoT & environmental analytical chemistry: Towards a profitable symbiosis, *Trends Environ. Anal. Chem.*, 2020, **27**, e00095.
9. V. Garrido-Momparler and M. Peris, Smart sensors in environmental/water quality monitoring using IoT and cloud services, *Trends in Environmental Analytical Chemistry*, 2022, **35**, e00173.
10. W. Yang, T. T. Fidelis and W. H. Sun, Machine Learning in Catalysis, From Proposal to Practicing, *ACS Omega*, 2020, **5**, 83-88.
11. A. S. Galushko, D. A. Boiko, E. O. Pentsak, D. B. Eremin and V. P. Ananikov, Time-Resolved Formation and Operation Maps of Pd Catalysts Suggest a Key Role of Single Atom Centers in Cross-Coupling, *J Am Chem Soc*, 2023, **145**, 9092-9103.
12. A. S. Galushko and V. P. Ananikov, 4D Catalysis Concept Enabled by Multilevel Data Collection and Machine Learning Analysis, *ACS Catal.*, 2024, **14**, 161-175.
13. T. Toyao, Z. Maeno, S. Takakusagi, T. Kamachi, I. Takigawa and K.-i. Shimizu, Machine Learning for Catalysis Informatics: Recent Applications and Prospects, *ACS Catal.*, 2020, **10**, 2260-2297.
14. M. H. S. Segler, M. Preuss and M. P. Waller, Planning chemical syntheses with deep neural networks and symbolic AI, *Nature*, 2018, **555**, 604-610.
15. T. Ha, D. Lee, Y. Kwon, M. S. Park, S. Lee, J. Jang, B. Choi, H. Jeon, J. Kim, H. Choi, H. T. Seo, W. Choi, W. Hong, Y. J. Park, J. Jang, J. Cho, B. Kim, H. Kwon, G. Kim, W. S. Oh, J. W. Kim, J. Choi, M. Min, A. Jeon, Y. Jung, E. Kim, H. Lee and Y. S. Choi, AI-driven robotic chemist for autonomous synthesis of organic molecules, *Sci Adv*, 2023, **9**, eadj0461.
16. K. McCardle, Accelerated mass spectra analysis, *Nat Comput Sci*, 2022, **2**, 556.
17. D. A. Boiko, K. S. Kozlov, J. V. Burykina, V. V. Ilyushenkova and V. P. Ananikov, Fully Automated Unconstrained Analysis of High-Resolution Mass Spectrometry Data with Machine Learning, *J. Am. Chem. Soc.*, 2022, **144**, 14590-14606.
18. M. Bortz, K. Dadhe and A. Mitsos, AI in Chemical Engineering – We Are Just at the Beginning, *Chem. Ing. Tech.*, 2021, **93**, 1875-1875.
19. A. Toniato, O. Schilter and T. Laino, The Role of AI in Driving the Sustainability of the Chemical Industry, *Chim.*, 2023, **77**, 144-149.
20. J. A. Keith, V. Vassilev-Galindo, B. Cheng, S. Chmiela, M. Gastegger, K. R. Muller and A. Tkatchenko, Combining Machine Learning and Computational

Chemistry for Predictive Insights Into Chemical Systems, *Chem Rev*, 2021, **121**, 9816-9872.
21. E. Cagno, A. Neri, M. Negri, C. A. Bassani and T. Lampertico, The Role of Digital Technologies in Operationalizing the Circular Economy Transition: A Systematic Literature Review, *Appl. Sci.*, 2021, **11**, 3328.
22. P. Schneider, W. P. Walters, A. T. Plowright, N. Sieroka, J. Listgarten, R. A. Goodnow, Jr., J. Fisher, J. M. Jansen, J. S. Duca, T. S. Rush, M. Zentgraf, J. E. Hill, E. Krutoholow, M. Kohler, J. Blaney, K. Funatsu, C. Luebkemann and G. Schneider, Rethinking drug design in the artificial intelligence era, *Nat. Rev. Drug. Discov.*, 2020, **19**, 353-364.
23. J. Vamathevan, D. Clark, P. Czodrowski, I. Dunham, E. Ferran, G. Lee, B. Li, A. Madabhushi, P. Shah, M. Spitzer and S. Zhao, Applications of machine learning in drug discovery and development, *Nat. Rev. Drug. Discov.*, 2019, **18**, 463-477.
24. F. Yu, N. D. Selva Kumar, D. Choudhury, L. C. Foo and S. H. Ng, Microfluidic platforms for modeling biological barriers in the circulatory system, *Drug Discov. Today*, 2018, **23**, 815-829.
25. A. Lavecchia, Machine-learning approaches in drug discovery: methods and applications, *Drug Discov. Today*, 2015, **20**, 318-331.
26. X. Yang, Y. Wang, R. Byrne, G. Schneider and S. Yang, Concepts of Artificial Intelligence for Computer-Assisted Drug Discovery, *Chem. Rev.*, 2019, **119**, 10520-10594.
27. D. Paul, G. Sanap, S. Shenoy, D. Kalyane, K. Kalia and R. K. Tekade, Artificial intelligence in drug discovery and development, *Drug Discov Today*, 2021, **26**, 80-93.
28. A. Sahu, J. Mishra and N. Kushwaha, Artificial Intelligence (AI) in Drugs and Pharmaceuticals, *Comb. Chem. High Throughput Screen.*, 2022, **25**, 1818-1837.
29. J. Jiménez-Luna, F. Grisoni and G. Schneider, Drug discovery with explainable artificial intelligence, *Nat. Mach. Intell.*, 2020, **2**, 573-584.
30. E. Gawehn, J. A. Hiss and G. Schneider, Deep Learning in Drug Discovery, *Mol. Inform.*, 2016, **35**, 3-14.
31. P. Schlexer Lamoureux, K. T. Winther, J. A. Garrido Torres, V. Streibel, M. Zhao, M. Bajdich, F. Abild-Pedersen and T. Bligaard, Machine Learning for Computational Heterogeneous Catalysis, *ChemCatChem*, 2019, **11**, 3581-3601.
32. L. Chiang, B. Lu and I. Castillo, Big Data Analytics in Chemical Engineering, *Annu. Rev. Chem. Biomol. Eng.*, 2017, **8**, 63-85.
33. R. Taylor and P. A. Wood, A Million Crystal Structures: The Whole Is Greater than the Sum of Its Parts, *Chem. Rev.*, 2019, **119**, 9427-9477.
34. E. Szymanska, Modern data science for analytical chemical data - A comprehensive review, *Anal. Chim. Acta*, 2018, **1028**, 1-10.

35. G. R. Schleder, A. C. M. Padilha, A. Reily Rocha, G. M. Dalpian and A. Fazzio, *Ab Initio* Simulations and Materials Chemistry in the Age of Big Data, *J. Chem. Inf. Model.*, 2020, **60**, 452-459.
36. K. S. Kozlov, D. A. Boiko, E. V. Detusheva, K. V. Detushev, E. O. Pentsak, A. N. Vereshchagin and V. P. Ananikov, Digital biology approach for macroscale studies of biofilm growth and biocide effects with electron microscopy, *Digit. Discov.*, 2023, **2**, 1522-1539.
37. C. W. Coley, N. S. Eyke and K. F. Jensen, Autonomous Discovery in the Chemical Sciences Part II: Outlook, *Angew. Chem. Int. Ed.* , 2020, **59**, 23414-23436.
38. M. Seifrid, R. Pollice, A. Aguilar-Granda, Z. Morgan Chan, K. Hotta, C. T. Ser, J. Vestfrid, T. C. Wu and A. Aspuru-Guzik, Autonomous Chemical Experiments: Challenges and Perspectives on Establishing a Self-Driving Lab, *Acc. Chem. Res.*, 2022, **55**, 2454-2466.
39. J. S. Manzano, W. Hou, S. S. Zalesskiy, P. Frei, H. Wang, P. J. Kitson and L. Cronin, An autonomous portable platform for universal chemical synthesis, *Nat. Chem.*, 2022, **14**, 1311-1318.
40. A. Vriza, H. Chan and J. Xu, Self-Driving Laboratory for Polymer Electronics, *Chem. Mater.*, 2023, **35**, 3046-3056.
41. Y. Li, L. Xia, Y. Fan, Q. Wang and M. Hu, Recent advances in autonomous synthesis of materials, *ChemPhysMater*, 2022, **1**, 77-85.
42. M. A. Soldatov, V. V. Butova, D. Pashkov, M. A. Butakova, P. V. Medvedev, A. V. Chernov and A. V. Soldatov, Self-Driving Laboratories for Development of New Functional Materials and Optimizing Known Reactions, *Nanomaterials*, 2021, **11**, 619.
43. J. V. Capella, A. Bonastre, J. C. Campelo, R. Ors and M. Peris, IoT & environmental analytical chemistry: Towards a profitable symbiosis, *Trends Environ. Anal. Chem.*, 2020, **27**.
44. J. C. Campelo, J. V. Capella, R. Ors, M. Peris and A. Bonastre, IoT Technologies in Chemical Analysis Systems: Application to Potassium Monitoring in Water, *Sensors*, 2022, **22**, 842.
45. J. Dai, O. Ogbeide, N. Macadam, Q. Sun, W. Yu, Y. Li, B. L. Su, T. Hasan, X. Huang and W. Huang, Printed gas sensors, *Chem. Soc. Rev.*, 2020, **49**, 1756-1789.
46. B. Horstkotte and P. Solich, The Automation Technique Lab-In-Syringe: A Practical Guide, *Molecules*, 2020, **25**, 1612.
47. J. A. Cardenas, J. B. Andrews, S. G. Noyce and A. D. Franklin, Carbon nanotube electronics for IoT sensors, *Nano Futures*, 2020, **4**.
48. P. Szymaszek, P. Fiedor, M. Tyszka-Czochara, M. Galek and J. Ortyl, An optical fluorescence sensor for IoT application in direct visualization of the curing process in polymer matrices, *Polym. Chem.*, 2024, DOI: 10.1039/d3py01326j, Advance Article.

49. M. Mousavizadegan, A. Firoozbakhtian, M. Hosseini and H. Ju, Machine learning in analytical chemistry: From synthesis of nanostructures to their applications in luminescence sensing, *TrAC Trends Anal. Chem.*, 2023, **167**, 117216.
50. F. Lussier, V. Thibault, B. Charron, G. Q. Wallace and J.-F. Masson, Deep learning and artificial intelligence methods for Raman and surface-enhanced Raman scattering, *TrAC Trends Anal. Chem.*, 2020, **124**, 115796.
51. D. Chen, Z. Wang, D. Guo, V. Orekhov and X. Qu, Review and Prospect: Deep Learning in Nuclear Magnetic Resonance Spectroscopy, *Chem. - Eur. J.*, 2020, **26**, 10391-10401.
52. N. Verbeeck, R. M. Caprioli and R. Van de Plas, Unsupervised machine learning for exploratory data analysis in imaging mass spectrometry, *Mass Spectrom. Rev.*, 2020, **39**, 245-291.
53. H. Zheng, X. Lu and K. He, *In situ* transmission electron microscopy and artificial intelligence enabled data analytics for energy materials, *J. Energy Chem.*, 2022, **68**, 454-493.
54. X. Zhou and M. Kraft, Blockchain Technology in the Chemical Industry, *Annu. Rev. Chem. Biomol. Eng.*, 2022, **13**, 347-371.
55. Z. Shahbazi and Y. C. Byun, Integration of Blockchain, IoT and Machine Learning for Multistage Quality Control and Enhancing Security in Smart Manufacturing, *Sensors*, 2021, **21**, 1467.
56. W. Deng, W. Fan, Z. Li, C. Cui, X. Ji and G. He, Security of Cyber-Physical Systems of Chemical Manufacturing Industries Based on Blockchain, *Processes*, 2023, **11**, 2707.
57. M. W. D. Hanson-Heine and A. P. Ashmore, Blockchain technology in quantum chemistry: A tutorial review for running simulations on a blockchain, *Int. J. Quantum Chem.*, 2022, **123**, e27035.
58. F. Tao, Q. Qi, L. Wang and A. Y. C. Nee, Digital Twins and Cyber–Physical Systems toward Smart Manufacturing and Industry 4.0: Correlation and Comparison, *Engineering*, 2019, **5**, 653-661.
59. J.-L. Kang, S. Mirzaei, Y.-C. Lee, Y.-C. Chuang, M. Frias, C.-H. Chou, S.-J. Wang, D. S. H. Wong and S.-S. Jang, in *Comput. Aided Chem. Eng.*, eds. M. Türkay and R. Gani, Elsevier, 2021, vol. 50, pp. 567-572.
60. Y. Chen, O. Yang, C. Sampat, P. Bhalode, R. Ramachandran and M. Ierapetritou, Digital Twins in Pharmaceutical and Biopharmaceutical Manufacturing: A Literature Review, *Processes*, 2020, **8**, 1088.
61. Z. Huang, Y. Shen, J. Li, M. Fey and C. Brecher, A Survey on AI-Driven Digital Twins in Industry 4.0: Smart Manufacturing and Advanced Robotics, *Sensors*, 2021, **21**, 6340.
62. M. Segovia and J. Garcia-Alfaro, Design, Modeling and Implementation of Digital Twins, *Sensors*, 2022, **22**, 5396.

63. W. Z. Bin Yu, Application exploration of digital twin technology in petrochemical industry, *Chem. Ind. Eng. Prog.*, 2019, **38**, 278–282.
64. D. I. Gerogiorgis and D. Castro-Rodriguez, in *Comput. Aided Chem. Eng.*, eds. M. Türkay and R. Gani, Elsevier, 2021, vol. 50, pp. 253-258.
65. A. Mazzuco, A. L. Krassmann, E. Reategui and R. S. Gomes, A systematic review of augmented reality in chemistry education, *Rev. Educ.*, 2022, **10**, e3325.
66. S. M. Reeves, K. J. Crippen and E. D. McCray, The varied experience of undergraduate students learning chemistry in virtual reality laboratories, *Comput. Educ.*, 2021, **175**, 104320.
67. Z. Merchant, E. T. Goetz, W. Keeney-Kennicutt, L. Cifuentes, O. Kwok and T. J. Davis, Exploring 3-D virtual reality technology for spatial ability and chemistry achievement, *J. Comput. Assist. Learn.*, 2013, **29**, 579-590.
68. J. B. Ferrell, J. P. Campbell, D. R. McCarthy, K. T. McKay, M. Hensinger, R. Srinivasan, X. Zhao, A. Wurthmann, J. Li and S. T. Schneebeli, Chemical Exploration with Virtual Reality in Organic Teaching Laboratories, *J. Chem. Educ.*, 2019, **96**, 1961-1966.
69. D. Manca, S. Brambilla and S. Colombo, Bridging between Virtual Reality and accident simulation for training of process-industry operators, *Adv. Eng. Softw.*, 2013, **55**, 1-9.
70. N. F. Saidin, N. D. Abd Halim and N. Yahaya, A Review of Research on Augmented Reality in Education: Advantages and Applications, *Int. Educ. Stud.*, 2015, **8**.
71. S. Cai, X. Wang and F.-K. Chiang, A case study of Augmented Reality simulation system application in a chemistry course, *Comput. Hum. Behav.*, 2014, **37**, 31-40.
72. H. Ozturk, A. Ozgur, P. Schwaller, T. Laino and E. Ozkirimli, Exploring chemical space using natural language processing methodologies for drug discovery, *Drug Discov. Today*, 2020, **25**, 689-705.
73. N. A. Lewinski and B. T. McInnes, Using natural language processing techniques to inform research on nanotechnology, *Beilstein J. Nanotechnol*, 2015, **6**, 1439-1449.
74. G. M. Hocky and A. D. White, Natural language processing models that automate programming will transform chemistry research and teaching, *Digit. Discov.*, 2022, **1**, 79-83.
75. Z. Zheng, O. Zhang, C. Borgs, J. T. Chayes and O. M. Yaghi, ChatGPT Chemistry Assistant for Text Mining and the Prediction of MOF Synthesis, *J. Am. Chem. Soc.*, 2023, **145**, 18048-18062.
76. X. Cai, H. Lai, X. Wang, L. Wang, W. Liu, Y. Wang, Z. Wang, D. Cao and X. Zeng, Comprehensive evaluation of molecule property prediction with ChatGPT, *Methods*, 2024, **222**, 133-141.

77. H. Zhu, J. Zhang, M. T. Kim, A. Boison, A. Sedykh and K. Moran, Big data in chemical toxicity research: the use of high-throughput screening assays to identify potential toxicants, *Chem. Res. Toxicol.*, 2014, **27**, 1643-1651.
78. T. Luechtefeld, D. Marsh, C. Rowlands and T. Hartung, Machine Learning of Toxicological Big Data Enables Read-Across Structure Activity Relationships (RASAR) Outperforming Animal Test Reproducibility, *Toxicol. Sci.*, 2018, **165**, 198-212.
79. H. L. Ciallella and H. Zhu, Advancing Computational Toxicology in the Big Data Era by Artificial Intelligence: Data-Driven and Mechanism-Driven Modeling for Chemical Toxicity, *Chem. Res. Toxicol.*, 2019, **32**, 536-547.
80. J. Hemmerich and G. F. Ecker, *In silico* toxicology: From structure-activity relationships towards deep learning and adverse outcome pathways, *WIREs Comput. Mol. Sci.*, 2020, **10**, e1475.
81. M. W. H. Wang, J. M. Goodman and T. E. H. Allen, Machine Learning in Predictive Toxicology: Recent Applications and Future Directions for Classification Models, *Chem. Res. Toxicol.*, 2021, **34**, 217-239.
82. M. Sabry Abdel-Messih and M. N. Kamel Boulos, ChatGPT in Clinical Toxicology, *JMIR Med. Educ.*, 2023, **9**, e46876.
83. I. Pantic, J. Paunovic, J. Cumic, S. Valjarevic, G. A. Petroianu and P. R. Corridon, Artificial neural networks in contemporary toxicology research, *Chem. Biol. Interact.*, 2023, **369**, 110269.
84. R. Vinuesa, H. Azizpour, I. Leite, M. Balaam, V. Dignum, S. Domisch, A. Fellander, S. D. Langhans, M. Tegmark and F. Fuso Nerini, The role of artificial intelligence in achieving the Sustainable Development Goals, *Nat. Commun.*, 2020, **11**, 233.
85. R. Nishant, M. Kennedy and J. Corbett, Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda, *Int. J. Inform. Manage.*, 2020, **53**, 102104.
86. D. Rangel-Martinez, K. D. P. Nigam and L. A. Ricardez-Sandoval, Machine learning on sustainable energy: A review and outlook on renewable energy systems, catalysis, smart grid and energy storage, *Chem. Eng. Res. Des.*, 2021, **174**, 414-441.
87. A. B. Culaba, A. P. Mayol, J. L. G. San Juan, C. L. Vinoya, R. S. Concepcion, II, A. A. Bandala, R. R. P. Vicerra, A. T. Ubando, W. H. Chen and J. S. Chang, Smart sustainable biorefineries for lignocellulosic biomass, *Bioresour. Technol.*, 2022, **344**, 126215.
88. P. Fantke, C. Cinquemani, P. Yaseneva, J. De Mello, H. Schwabe, B. Ebeling and A. A. Lapkin, Transition to sustainable chemistry through digitalization, *Chem*, 2021, **7**, 2866-2882.
89. C. R. Sagandira, S. Nqeketo, K. Mhlana, T. Sonti, S. Gaqa and P. Watts, Towards 4th industrial revolution efficient and sustainable continuous flow

manufacturing of active pharmaceutical ingredients, *React. Chem. Eng.*, 2022, **7**, 214-244.
90. D. C. Elton, Z. Boukouvalas, M. D. Fuge and P. W. Chung, Deep learning for molecular design—a review of the state of the art, *Mol. Syst. Des. Eng.*, 2019, **4**, 828-849.
91. B. Sanchez-Lengeling and A. Aspuru-Guzik, Inverse molecular design using machine learning: Generative models for matter engineering, *Science*, 2018, **361**, 360-365.
92. D. Merk, L. Friedrich, F. Grisoni and G. Schneider, *De Novo* Design of Bioactive Small Molecules by Artificial Intelligence, *Mol. Inform.*, 2018, **37**, 1700153.
93. E. Putin, A. Asadulaev, Y. Ivanenkov, V. Aladinskiy, B. Sanchez-Lengeling, A. Aspuru-Guzik and A. Zhavoronkov, Reinforced Adversarial Neural Computer for de Novo Molecular Design, *J. Chem. Inf. Model.*, 2018, **58**, 1194-1204.
94. A. Mahmood and J.-L. Wang, Machine learning for high performance organic solar cells: current scenario and future prospects, *Energy Environ. Sci.*, 2021, **14**, 90-105.
95. W. Sun, Y. Zheng, K. Yang, Q. Zhang, A. A. Shah, Z. Wu, Y. Sun, L. Feng, D. Chen, Z. Xiao, S. Lu, Y. Li and K. Sun, Machine learning-assisted molecular design and efficiency prediction for high-performance organic photovoltaic materials, *Sci. Adv.*, 2019, **5**, eaay4275.
96. S. Mao, B. Wang, Y. Tang and F. Qian, Opportunities and Challenges of Artificial Intelligence for Green Manufacturing in the Process Industry, *Engineering*, 2019, **5**, 995-1002.
97. J. Chen, P. Hu, H. Zhou, J. Yang, J. Xie, Y. Jiang, Z. Gao and C. Zhang, Toward Intelligent Machine Tool, *Engineering*, 2019, **5**, 679-690.
98. J. Arents and M. Greitans, Smart Industrial Robot Control Trends, Challenges and Opportunities within Manufacturing, *Appl. Sci.*, 2022, **12**, 937.
99. W. Zhu, Y. Ma, M. G. Benton, J. A. Romagnoli and Y. Zhan, Deep learning for pyrolysis reactor monitoring: From thermal imaging toward smart monitoring system, *AIChE J.*, 2018, **65**, 582-591.
100. H. W. Song, W. Choi, T. Jeon and J. H. Oh, Recent Advances in Smart Organic Sensors for Environmental Monitoring Systems, *ACS Appl. Electron. Mater.*, 2022, **5**, 77-99.
101. A. B. Culaba, A. P. Mayol, J. L. G. San Juan, C. L. Vinoya, R. S. Concepcion, 2nd, A. A. Bandala, R. R. P. Vicerra, A. T. Ubando, W. H. Chen and J. S. Chang, Smart sustainable biorefineries for lignocellulosic biomass, *Bioresour Technol*, 2022, **344**, 126215.
102. J. Ma, R. P. Sheridan, A. Liaw, G. E. Dahl and V. Svetnik, Deep neural nets as a method for quantitative structure-activity relationships, *J. Chem. Inf. Model*, 2015, **55**, 263-274.

103. R. Gupta, D. Srivastava, M. Sahu, S. Tiwari, R. K. Ambasta and P. Kumar, Artificial intelligence to deep learning: machine intelligence approach for drug discovery, *Mol. Divers.*, 2021, **25**, 1315-1360.
104. W. P. Walters and R. Barzilay, Applications of Deep Learning in Molecule Generation and Molecular Property Prediction, *Acc. Chem. Res.*, 2021, **54**, 263-270.
105. G. Dorahy, J. Z. Chen and T. Balle, Computer-Aided Drug Design towards New Psychotropic and Neurological Drugs, *Molecules*, 2023, **28**, 1324.
106. N. S. Eyke, B. A. Koscher and K. F. Jensen, Toward Machine Learning-Enhanced High-Throughput Experimentation, *Trends Chem.*, 2021, **3**, 120-132.
107. Y. Gambo, S. Adamu, A. A. Abdulrasheed, R. A. Lucky, M. S. Ba-Shammakh and M. M. Hossain, Catalyst design and tuning for oxidative dehydrogenation of propane – A review, *Appl. Catal. A: Gen.*, 2021, **609**, 117914.
108. B. Mahjour, R. Zhang, Y. Shen, A. McGrath, R. Zhao, O. G. Mohamed, Y. Lin, Z. Zhang, J. L. Douthwaite, A. Tripathi and T. Cernak, Rapid planning and analysis of high-throughput experiment arrays for reaction discovery, *Nat. Commun.*, 2023, **14**, 3924.
109. R. A. Patel and M. A. Webb, Data-Driven Design of Polymer-Based Biomaterials: High-throughput Simulation, Experimentation, and Machine Learning, *ACS Appl. Bio Mater.*, 2024, **7**, 510-527.
110. Y. Xu, Y. Gao, L. Su, H. Wu, H. Tian, M. Zeng, C. Xu, X. Zhu and K. Liao, High-Throughput Experimentation and Machine Learning-Assisted Optimization of Iridium-Catalyzed Cross-Dimerization of Sulfoxonium Ylides, *Angew. Chem. Int. Ed.*, 2023, **62**, e202313638.
111. J. Yu, J. Liu, C. Li, J. Huang, Y. Zhu and H. You, Recent advances and applications in high-throughput continuous flow, *Chem. Commun.*, 2024, **60**, 3217-3225.
112. N. Kim, D. Lee and Y. Hong, Data-Efficient Deep Generative Model with Discrete Latent Representation for High-Fidelity Digital Materials, *ACS Materials Lett.*, 2023, **5**, 730-737.
113. J. Mańkowski and J. Lipnicki, Digital Materials – Evaluation of the Possibilities of using Selected Hyperelastic Models to Describe Constitutive Relations, *Int. J. Appl. Mech. Eng.*, 2017, **22**, 601-612.
114. R. Pollice, G. Dos Passos Gomes, M. Aldeghi, R. J. Hickman, M. Krenn, C. Lavigne, M. Lindner-D'Addario, A. Nigam, C. T. Ser, Z. Yao and A. Aspuru-Guzik, Data-Driven Strategies for Accelerated Materials Design, *Acc. Chem. Res.*, 2021, **54**, 849-860.
115. F. Sui, R. Guo, Z. Zhang, G. X. Gu and L. Lin, Deep Reinforcement Learning for Digital Materials Design, *ACS Materials Lett.*, 2021, **3**, 1433-1439.
116. P. Wu, Digital Materials Design: Computational Methodologies as a Discovery Tool, *MRS Bulletin*, 2011, **31**, 995-998.

117. J. Zhang, Y. Li, T. Zhao, Q. Zhang, L. Zuo and K. Zhang, Machine-learning based design of digital materials for elastic wave control, *Extreme Mech. Lett.*, 2021, **48**, 101372.
118. A. F. Almeida, F. A. P. Ataide, R. M. S. Loureiro, R. Moreira and T. Rodrigues, Augmenting Adaptive Machine Learning with Kinetic Modeling for Reaction Optimization, *J. Org. Chem.*, 2021, **86**, 14192-14198.
119. A. Filipa de Almeida and T. Rodrigues, in *Enabling Tools and Techniques for Organic Synthesis*, 2023, DOI: https://doi.org/10.1002/9781119855668.ch10, pp. 393-421.
120. P. Schwaller, A. C. Vaucher, R. Laplaza, C. Bunne, A. Krause, C. Corminboeuf and T. Laino, Machine intelligence for chemical reaction space, *WIREs Comput. Mol. Sci.* , 2022, **12**, e1604.
121. B. J. Shields, J. Stevens, J. Li, M. Parasram, F. Damani, J. I. M. Alvarado, J. M. Janey, R. P. Adams and A. G. Doyle, Bayesian reaction optimization as a tool for chemical synthesis, *Nature*, 2021, **590**, 89-96.
122. C. J. Taylor, K. C. Felton, D. Wigh, M. I. Jeraal, R. Grainger, G. Chessari, C. N. Johnson and A. A. Lapkin, Accelerated Chemical Reaction Optimization Using Multi-Task Learning, *ACS Cent. Sci.*, 2023, **9**, 957-968.
123. A. M. Zuranski, J. I. Martinez Alvarado, B. J. Shields and A. G. Doyle, Predicting Reaction Yields via Supervised Learning, *Acc. Chem. Res.*, 2021, **54**, 1856-1865.
124. A. D. Clayton, Recent Developments in Reactor Automation for Multistep Chemical Synthesis, *Chemistry–Methods*, 2023, **3**, e202300021.
125. M. K. Goshisht, Machine Learning and Deep Learning in Synthetic Biology: Key Architectures, Applications, and Challenges, *ACS Omega*, 2024, **9**, 9921-9945.
126. J. M. Gregoire, L. Zhou and J. A. Haber, Combinatorial synthesis for AI-driven materials discovery, *Nat. Synth.*, 2023, **2**, 493-504.
127. J. C. A. Oliveira, J. Frey, S.-Q. Zhang, L.-C. Xu, X. Li, S.-W. Li, X. Hong and L. Ackermann, When machine learning meets molecular synthesis, *Trends Chem.*, 2022, **4**, 863-885.
128. S. Szymkuc, E. P. Gajewska, T. Klucznik, K. Molga, P. Dittwald, M. Startek, M. Bajczyk and B. A. Grzybowski, Computer-Assisted Synthetic Planning: The End of the Beginning, *Angew. Chem. Int. Ed.* , 2016, **55**, 5904-5937.
129. J. Wu, X. Yang, Y. Pan, T. Zuo, Z. Ning, C. Li and Z. Zhang, Recent developments of automated flow chemistry in pharmaceutical compounds synthesis, *J. Flow. Chem.*, 2023, **13**, 385-404.
130. D. A. Boiko, E. O. Pentsak, V. A. Cherepanova, E. G. Gordeev and V. P. Ananikov, Deep neural network analysis of nanoparticle ordering to identify defects in layered carbon materials, *Chem. Sci.*, 2021, **12**, 7428-7441.
131. F. Di Palma, C. Abate, S. Decherchi and A. Cavalli, Ligandability and druggability assessment via machine learning, *WIREs Comput. Mol. Sci.*, 2023, **13**, e1676.

132. D. A. Olmedo, A. A. Durant-Archibold, J. L. Lopez-Perez and J. L. Medina-Franco, Design and Diversity Analysis of Chemical Libraries in Drug Discovery, *Comb. Chem. High Throughput Screen.*, 2024, **27**, 502 - 515.
133. M. A. Raslan, S. A. Raslan, E. M. Shehata, A. S. Mahmoud and N. A. Sabri, Advances in the Applications of Bioinformatics and Chemoinformatics, *Pharmaceuticals*, 2023, **16**, 1050.
134. R. Rodriguez-Perez, F. Miljkovic and J. Bajorath, Machine Learning in Chemoinformatics and Medicinal Chemistry, *Annu. Rev. Biomed. Data Sci.*, 2022, **5**, 43-65.
135. D. Zankov, T. Madzhidov, A. Varnek and P. Polishchuk, Chemical complexity challenge: Is multi-instance machine learning a solution?, *WIREs Comput. Mol. Sci.*, 2023, **14**, e1698.
136. Y. Han, I. Ali, Z. Wang, J. Cai, S. Wu, J. Tang, L. Zhang, J. Ren, R. Xiao, Q. Lu, L. Hang, H. Luo and J. Li, Machine learning accelerates quantum mechanics predictions of molecular crystals, *Phys. Rep.*, 2021, **934**, 1-71.
137. P. O. A. Navaux, A. F. Lorenzon and M. d. S. Serpa, Challenges in High-Performance Computing, *J. Braz. Comput. Soc.*, 2023, **29**, 51-62.
138. U. Raucci, A. Valentini, E. Pieri, H. Weir, S. Seritan and T. J. Martinez, Voice-controlled quantum chemistry, *Nat. Comput. Sci.*, 2021, **1**, 42-45.
139. A. S. Rosen, S. M. Iyer, D. Ray, Z. Yao, A. Aspuru-Guzik, L. Gagliardi, J. M. Notestein and R. Q. Snurr, Machine learning the quantum-chemical properties of metal–organic frameworks for accelerated materials discovery, *Matter*, 2021, **4**, 1578-1597.
140. J. Westermayr and P. Marquetand, Machine Learning for Electronically Excited States of Molecules, *Chem. Rev.*, 2021, **121**, 9873-9926.