PERSPECTIVE `OPEN ACCESS`

The Alchemist, the Scientist, and the Robot: Exploring the Potential of Human-AI Symbiosis in Self-Driving Polymer Laboratories

Bahar Dadfar | Berna Alemdag | Gözde Kabay (ID)

Karlsruhe Institute of Technology (KIT), Institute of Functional Interfaces (IFG), Department For Bioengineering and Biosystems, Eggenstein-Leopoldshafen, Germany

Correspondence: Gözde Kabay (gozde.kabay@kit.edu)

ABSTRACT

Polymer chemistry research has progressed through three methodological eras: the alchemist's intuitive trial-and-error, the scientist's rule-based design, and the robot's algorithm-guided automation. While approaches combining combinatorial chemistry with statistical design of experiments offer a systematic approach to polymer discovery, they struggle with complex design spaces, avoid human biases, and scale up. In response, the discipline has adopted automation and artificial intelligence (AI), culminating in self-driving laboratories (SDLs), integrating high-throughput experimentation into closed-loop, AI-assisted design-build-test-learn cycles, enabling the rapid exploration of chemical spaces. However, while SDLs address throughput and complexity challenges, they introduce new forms of the original problems: algorithmic biases replace human biases, data sparsity creates constraints on design space navigation, and black-box AI models create transparency issues, complicating interpretation. These challenges emphasize a critical point: complete algorithmic autonomy is inadequate without human involvement. Human intuition, ethical judgment, and domain expertise are crucial for establishing research objectives, identifying anomalies, and ensuring adherence to ethical constraints. This perspective supports a hybrid model grounded in symbiotic autonomy, where adaptive collaboration between humans and AI enhances trust, creativity, and reproducibility. By incorporating human reasoning into adaptive AI-assisted SDL workflows, next-generation autonomous polymer discovery will be not only faster but also wiser.

1 | Introduction

Over the last century, polymer chemistry has evolved from an intuitive, trial-and-error approach to algorithmically driven, autonomous workflows. In 1920, Hermann Staudinger postulated that polymers consist of units of molecules linked together by the same covalent bonds found in smaller organic molecules. Following a decade-long controversy, Staudinger's theory, which posi-

tions polymers as macromolecules rather than colloidal aggregates, was supported by experimental evidence and established the theoretical groundwork for macromolecular science [1].

Early breakthroughs often arise from human intuition, creativity, and sometimes serendipity. For example, polytetrafluoroethylene (PTFE, also known as Teflon) was discovered by accident in 1938 when Plunkett was experimenting with tetrafluoroethylene as a

---

coolant for refrigerators and found a waxy solid clogging a gas line cylinder []. During the mid-20th-century polymer boom, successes were driven by heuristics: experts adopted a "cook-and-look" method, iteratively modifying recipes until they yielded transformative materials such as plastics, rubbers, and fibers that transformed society. However, this human-centered strategy had significant limitations in terms of scope and speed. The range of possible monomers, polymer configurations, and processing conditions is extensive; relying only on intuition and hands-on experimentation means that many potentially valuable polymers likely remain undiscovered.

By the late 20th century, scientists began adopting more systematic strategies to navigate chemical design space. Statistical design of experiments (DoE) [3–5] (e.g., factorial designs, response surface methodologies) and combinatorial chemistry [6, 7] (systematically generating and testing extensive libraries of chemical compounds under specified conditions) enable the structured exploration of the formulation variables and structure–property relationship (SPR) [7, 8]. While initially powerful, applying DoE-based statistical approaches in conjunction with manual experimentation faced scalability, throughput, and adaptability issues as the number of tunable variables (e.g., monomer composition, initiator type, solvent environment, and reaction temperature) and performance objectives (e.g., yield, flexibility, and thermal stability) increased. As a result, these rule-based methodologies became less practical.

This bottleneck motivated the adoption of high-throughput experimentation (HTE) platforms in the 90s, which offered automation, miniaturization, and online analytics. HTE allowed researchers to test hundreds of polymer compositions in parallel, significantly accelerating SPR mapping. Pioneered by Symyx Technologies, robotics and automated analysis have been integrated into polymer screening workflows by Borealis AG [9] and several academic laboratories [10]. Running multiple reactions in parallel enabled a more comprehensive SPR evaluation, thereby accelerating the discovery of various original polymers.

Over the last decade, a convergence of advances in HTE and data science has given rise to AI-assisted autonomous, or in other words, "self-driving" laboratories, capable of designing, executing, and learning from experiments with minimal human involvement [11, 12]. While these autonomous technologies have already enabled the discovery of novel polymer families with properties surpassing known materials, significant challenges remain. These include data sparsity, biased or overfitting of the model, limited transferability across chemical spaces, and difficulty interpreting SPR with solely AI decision-making. These issues suggest that algorithmic autonomy without human guidance is insufficient for advancing next-generation polymer science [11, 13].

This perspective presents the methodological evolution of polymer chemistry, from early heuristic and rule-based strategies to modern data-driven discovery. We thoroughly analyze the inherent strengths and weaknesses of each approach, highlighting how AI and machine learning (ML) empower SDLs to independently evaluate experimental data and improve polymer synthesis within closed-loop systems. It is also argued that true innovation in polymer science will not stem from full autonomy alone but from a symbiotic human–AI collaboration, where intuition, creative insights, and domain expertise are dynamically harmonized with algorithmic power to navigate and uncover the complexities of polymer design, with both learning from each other's strengths.

2 | A Brief Historical Evaluation of Polymer Chemistry

2.1 | Early Polymer Discoveries: From Myth to Method

The history of polymer chemistry is broad and multifaceted, starting with ancient practices guided by nature and intuition and evolving into modern experiments conducted by algorithmically guided robots. Ancient examples date back to 1600 B.C., when ancient Mesoamericans extracted latex from *Castilla elastica*, the rubber tree, then processed it with liquid from *Ipomoea alba* (a morning glory vine species) to manufacture rubber balls, hollow figurines, and other rubber artifacts [14]. In ancient Egypt, natural polymers were used in adhesives and embalming resins. Glues derived from animal collagen were widely used as adhesives, while tree resins such as myrrh and frankincense were used to formulate embalming balms for rituals. The Greeks and Romans utilized bitumen and pitch, hydrocarbon-rich substances obtained from pine tar or petroleum seeps, for sealing and waterproofing ships and amphorae [15]. Although processed without modern chemical understanding, these materials represented early human engagement with natural polymers and formed the empirical basis for centuries of polymer research and discovery.

The shift from alchemical practices to empirical methods began during the Enlightenment. European reformers increasingly challenged dogma in favor of scientific inquiry, systematic experimentation, and the practical application of scientific knowledge. One notable figure in this transition is Charles Marie de la Condamine, who introduced "caoutchouc" samples, harvested from rubber trees (*Hevea brasiliensis*) to the European world in the 1730s. This prompted a wave of scientific investigations only years later by figures such as Joseph Priestley, who noted caoutchouc's ability to erase pencil marks and coined the term "rubber", and Michael Faraday, in 1826, who explored its solubility features and analyzed its elemental composition [16]. Even though these early investigations were missing a theoretical foundation, they stimulated curiosity to understand materials better.

The 19th century witnessed significant breakthroughs, particularly in the development of polymeric materials. In the year 1839, Charles Goodyear made a significant discovery in the process of vulcanization by unintentionally subjecting rubber to heat in the presence of sulfur, which resulted in the formation of a material with enhanced durability and elasticity [17]. In 1861, British chemist Thomas Graham observed that certain organic compounds, such as cellulose, were dispersed in water; they did not crystallize and left sticky residues upon filtration. He hypothesized that these materials represented a different organization of matter, coining the term "colloids" from "kolla", meaning glue. In 1869, Hyatt introduced celluloid, a thermoplastic derived from camphor and cellulose nitrate, a versatile plastic used widely in goods ranging from hair combs to silent film rolls [18]. Later, in

1907, Leo Baekeland synthesized Bakelite by combining phenol and formaldehyde under applied heat and pressure in a sealed autoclave. Bakelite was the first synthetic plastic, specifically a thermosetting resin, that demonstrated commercial success, particularly in the electrical industry, establishing the foundation of the "Age of Plastics" [19]. Though groundbreaking, these breakthroughs were mainly driven by heuristics, which means that empirical observations, chemical intuition, and accumulated experience guided scientists in selecting which monomers to combine, how to adjust reaction conditions, or which additives could influence the thermal or mechanical properties of the materials without a molecular-level understanding of polymer structure and behavior.

A theoretical revolution in polymer chemistry began in the 1920s with Hermann Staudinger, who proposed that polymers were long chains of "covalently bonded macromolecules", denying colloidal aggregate theories. Staudinger's foundational work marked the era of macromolecular science and earned him the Nobel Prize in 1953 [20, 21]. At the same period, Wallace Carothers applied this macromolecular view to synthesize nylon-6,6 (1935) through the step-growth polycondensation of hexamethylenedi-amine by adipic acid [22, 23]. This work was guided by the mathematical principles that later became the Carothers equation, linking monomer conversion and polymer chain length, guiding the optimization of the molecular weight in linear condensation polymers [24]. Carothers also pioneered the development of neoprene and early polyesters, but abandoned this work due to the low molecular weight of the resulting polymer [25]. In 1938, Paul Schlack synthesized Nylon 6, polycaprolactam, by ring-opening polymerization of ε-caprolactam. Though initially encountered by reproducibility issues, the process was refined in the 1960s with acyl-lactam initiators.

2.2 | Shifting Emphasis from Intuition to Theory

In the decades following World War II, the polymer industry experienced rapid growth driven by the increasing demand for innovative and cost-effective materials. Advances in analytical characterization have led to the discovery of several new polymers and enhanced the understanding of the polymer structure–property relationship (SPR). Systematic experimentation, primarily grounded in empirical designs and early statistical methodologies, enabled numerous significant polymer discoveries, including polystyrene, poly(methyl methacrylate) (PMMA), polyvinyl chloride (PVC), polyethylene terephthalate (PET), polycarbonate (PC), and PTFE, and their industrialization which became foundational materials across multiple industries, from consumer products to advanced engineering applications [26].

As polymer chemistry matured theoretically, the field shifted from "cookbook" synthesis to systematic, rule-based experimentation. Consequently, researchers moved beyond purely intuitive judgment, formalizing chemical knowledge into collections of "if-then" rules or, in other words, rule-based heuristic systematics, grounded in accumulated domain knowledge and supported by statistical reasoning [27, 28]. This hybrid strategy aimed to structure chemical exploration, improve reproducibility, and facilitate the systematic optimization of polymer properties. Karl Ziegler's discovery (1953) of high-density polyethylene achieved

through a combination of titanium tetrachloride and an alkyl aluminum co-catalyst system enabled ethylene to polymerize under mild conditions [29]. The following year, Giulio Natta applied Ziegler's catalyst to propylene, synthesizing crystalline isotactic polypropylene. Their insight was that transition-metal complexes could facilitate the formation of polymers [30]. Their insight into transition-metal-catalyzed polymerization enabled the controlled, mild-condition synthesis of stereoregular polymers, earning them a Nobel Prize in 1963 [31].

Despite the rapid growth of polymer science, discovery remained slow and labor-intensive, often struggling to meet industrial demands. Researchers mainly relied on one-factor-at-a-time experiments or small-scale factorial designs (early forms of Design of Experiments, DoE) to fine-tune polymer formulations. For instance, while optimizing Ziegler–Natta catalysts, metal halides and organoaluminium ratios were systematically varied in a pseudo-grid search to enhance activity and stereoselectivity [30, 32–34]. While such methods were interpretable and grounded in chemical intuition, they were inherently limited by human bias and a lack of prior knowledge. Exploration was often confined to familiar monomers and reaction types, minimizing risk but frequently neglecting unconventional or counterintuitive chemistries. A representative example is the design of photoresist polymers, where multiple performance metrics (such as etch resistance, resolution, and sensitivity) must be optimized simultaneously. Even expert chemists tended to test only incremental variations of known styrene-acrylate systems, following previous successful experimentations, and each design-test-analyze cycle could take weeks, further hindering the speed of innovation. Thus, heuristic methods achieved high success rates within narrow domains but lacked scalability. By the 1980s, the complexity of emerging polymers, from biomedical hydrogels to heat-resistant thermoplastics, had outgrown what purely human-guided discovery could handle.

One formalism that bridged intuition with a systematic exploration was the introduction of the design of experiments (DoE). Pioneered by Fisher [35] in the 1920s and later adapted to chemistry by Box and Wilson in 1951[4], DoE provided a statistical framework for efficiently navigating multivariable systems. Unlike one-factor-at-a-time trials, factorial designs could capture interactions between variables with minimal runs. In polymer science, these statistical methodologies, including factorial designs, fractional factorials, and response surface methodologies (i.e., central composite design and Box–Behnken design) [36], have been widely adopted to optimize complex formulations, such as polymer blends, flame-retardant additives, and catalyst systems [3]. DoE improved efficiency while enhancing interpretability, enabling researchers to identify key factor interactions (e.g., how initiator level and solvent polarity affect molecular weight) . For instance, the synthesis of styrene-butadiene rubber was streamlined in the late 1950s using a fractional factorial design, reducing hundreds of trials to 12 strategically selected experiments [5]. Rule-based decision systems utilizing "if-then" logic, such as SPR, also arose to encode known polymer property data into simple predictive rules. In parallel, factorial experimental designs (e.g., two-level factorial arrays) enabled the systematic optimization of key reaction conditions, such as monomer ratios, temperature, and curing conditions, with minimal trial requirements.

While many early statistical approaches focused on SPR evaluation, others targeted improving polymerization yield, molecular weight distribution, and refining polymer performance. For example, Taguchi's orthogonal array design, introduced in the 1960s, further refined process robustness in polymerization by systematically reducing variability. Bicerano's book, "Prediction of Polymer Properties", compiles extensive insights for estimating the SPR of polymers through their molecular structures, multiphase and interface modeling, and practical applications [37]. These early informatics tools were essentially human-crafted rules generalized from data, a step beyond pure intuition, yet still far from data-driven algorithms that characterize modern polymer informatics.

2.3 | High Throughput Technologies: Accelerating Polymer Discovery through Automation

The growing demand for scalable, innovative materials with advanced features has shifted emphasis from trial-and-error methods toward high-throughput experimentation (HTE). This shift led to the development of a family of high-throughput technologies (HTTs) by adopting combinatorial chemistry and automated screening, which miniaturized reactions to microliter or even nanoliter scales and ran them in parallel. Yet, early HTT implementations remained human-centered: experts chose variables based on their experience or literature, while automated equipment handled repetitive tasks, such as mixing in sealed parallel vessels and sampling, and offline analytics mapped SPR trends [7, 8]. For example, Kohn and coworkers pioneered a combinatorial HTE application in the polycondensation of 14 tyrosine-based diphenols and 8 diacids, generating a 112-member polyarylate library. This systematically developed library enabled them to elucidate how incremental structural changes influence polymer properties, such as hydrophobicity, glass transition temperature (T_g), and even cellular response, outpacing that of sequential synthesis [7, 8].

Despite the demonstrated success of early HTEs, manual intermediate steps and offline characterization limited scalability. Around the same time, Symyx Technologies introduced a *"fully integrated high-throughput workflow,"* initially demonstrated for the discovery of polyolefin catalysts. This platform utilized robotic liquid handlers to combinatorially mix metal precursors and ligands, followed by rapid screening assays to assess catalytic performance. End-to-end automation enabled the testing of dozens of catalyst candidates in the time it once took to run a single experiment, leading to the discovery of entirely new classes of olefin polymerization catalysts [38]. Polyolefin producers and academic laboratories soon adapted the same philosophy to polymer exploration, discovering chemistries that would have been impractical to probe manually.

As instrumentation matured, HTTs began to resemble miniature SDLs. Reactions were performed in multi-well plates or microfluidic chips, utilizing massive parallelization. Within a single automated run, discrete factors, such as catalysts and ligands, alongside continuous parameters, including temperature, residence time, and pressure, can be varied, yielding detailed readouts of monomer conversion, by-product profiles, solvent effects, and reproducibility online [39]. Modern HTE workflows increasingly embed statistical DoE into automation scripts, to ensure a balanced, information-rich set of experiments to be performed without unnecessary repetition, so that every additional experiment maximizes what could be gathered about the structure–property landscape. Without a systematic design plan, running hundreds of reactions would otherwise lead to redundant data of limited value.

In 2004, Schubert's group linked a Chemspeed (ASW 2000) robotic synthesizer with inline size exclusion chromatography (SEC) and automated matrix-assisted laser desorption/ionization time-of-flight (MALDI-ToF) mass spectrometry [40, 41]. This setup, housed in an "inert" glovebox environment, automatically performed three polymerization reactions: ATRP, cationic ring-opening polymerization (CROP), and emulsion polymerization, with immediate feedback on molecular-weight distribution [40, 41]. The products of fully automated runs matched the performance of manual experiments and reproduced SPR trends (e.g., initiator-dependent kinetics in CROP), at an accelerated pace. By the early 2010s, commercial platforms like Chemspeed's Accelerator and Swing could perform 16 or more parallel polymerizations, dispensing reagents with multi-head robots and injecting samples directly into GPC or HPLC for inline analytics [42]. This integration effectively merged synthesis and characterization workflows into a single closed-loop process, boosting discovery efficiency. Despite these advances, four main limitations remained: i) Fragmented workflow integration in many labs still separates synthesis, purification, and analysis steps, hindering overall throughput [43]; ii) the absence of unified, FAIR (Findable, Accessible, Interoperable, and Reusable) data repositories led to data silos and limited cross-lab data comparisons [44]; iii) miniaturized reactions in multi-well plates, impaired reproducibility mainly due to uneven mixing and heat transfer, affected data quality and cross-platform consistency [45]; iv) scaling up from micro-/nano-scale to industrial volumes was not always straightforward, as reaction kinetics and surface effects differ at larger scales [46]. Even with these caveats, HTTs fundamentally changed the culture of polymer discovery. Rather than relying on iterative "cook and look" approaches, researchers adopted data-rich exploration, which laid the physical and digital foundations for today's data-driven SDLs.

3 | The Data-Driven Revolution: Polymer Informatics and Algorithm-Guided Discovery

As computational power, automation capabilities, and data availability grew in the 2010s, polymer chemistry entered the era of polymer informatics: a field that utilizes ML and big-data analytics for polymer discovery. Polymers with their inherent combinatorial complexity (ranging from various monomer sequences and molecular weights to sophisticated architectures), and ML excel at detecting patterns in such high-dimensional data. Unlike heuristic "if-then" rules or simple linear quantitative models (e.g., a linear regression model predicting boiling point from molecular weight and polarity), modern ML models can capture nonlinear and high-dimensional relationships hidden in data, where multiple features jointly determine a property [10]. For instance, a polymer's T_g depends on backbone flexibility, intermolecular forces, chain-length distribution, and more features, typically too complex to decrypt by a single equation, yet

one that ML can handle to uncover. The pattern recognition capabilities of ML, when implemented in HTTs, sped up the discovery of next-generation polymers with properties surpassing the state of the art [10, 47, 48].

Crucial to this data-driven progress was the creation of publicly available, curated polymer databases. A notable example is the PoLyInfo database of the Japanese National Institute for Materials Science (NIMS), which has compiled an extensive archive of manually curated chemical structures and descriptors from the literature [49]. Similarly, the Massachusetts Institute of Technology (MIT)'s Polymer Genome project provides a platform to encode polymer structures (via specialized SMILES strings) and utilize pre-trained ML models to predict polymer properties [50]. With such informatics tools and open databases, polymer scientists gained access to vast amounts of data, enabling high-performing AI/ML models that no single lab's data could have ever produced. As an example, Tao and coworkers benchmarked various ML algorithms (e.g., deep and convolutional neural networks, random forests, support vector machines) on a feature-engineered dataset of approximately 7000 known polymers obtained from distinct polymer libraries. This process ultimately created a predictor capable of estimating the glass T_g of approximately 5700 hypothetical homopolymers [51]. Similarly, Zhang and Xu used Gaussian process regression (GPR) to predict the T_g of polyacrylamides utilizing quantum chemical descriptors derived from molecular structures. Even with only 37 training samples, they successfully captured SPR trends, demonstrating the predictive power of GPR despite the limited data [52]. Similar ML-driven efforts have tackled polymer crystallinity [53], mechanical properties like elastic moduli (e.g., Young's modulus [54]), and optical traits like refractive index [55], demonstrating broad applicability.

Beyond forward property prediction and SPR evaluation, data-driven approaches have enabled the inverse design of new polymers and other materials [56]. Generative models such as variational autoencoders (VAEs), generative adversarial networks can learn vast chemical design spaces and propose novel molecular structures that meet desired performance targets. For example, Zheng and coworkers combine a graph-based VAE with high-throughput molecular dynamics (MD) simulations to design vitrimers (recyclable polymers) [57]. Their model generated approximately 1 million hypothetical vitrimer chemistries and estimated T_g for 8424 candidates and proposed new candidates with target T_g values ($T_g \approx 323$ K), even beyond the training data range. One such vitrimer was found to have a T_g of approximately 315 K, along with desirable self-healing and flow properties, demonstrating how data-driven insights guided by expert knowledge can accelerate the discovery of sustainable polymers.

While inverse design enables the discovery of potential polymer candidates, another significant advantage of ML is the efficient optimization of experimental design spaces for predefined performance goals (e.g., maximizing T_g or yield). One such approach is Bayesian optimization (BO), a probabilistic and sample-efficient strategy for optimizing expensive or complex functions [58]. The BO typically works by building a surrogate model, often a Gaussian process, to predict the relationship between inputs (e.g., temperature, catalyst type) and the target properties (e.g., yield, molecular weight). An acquisition function then balances exploration versus exploitation by suggesting which experiment to run next, updating the surrogate as new data are fed. This iterative process can identify optimal reaction conditions in far fewer trials than brute-force or grid searches [58]. BO has been successfully applied to various experimental design optimizations, ranging from tuning stereoselective catalysts for ring-opening polymerization [59], to improving the copolymerization of high-density polyethylene [60], and to multi-objective optimization of additive manufacturing resin recipes [61].

For example, a recent study employed a single-objective BO application to minimize compositional drift in the free radical copolymerization of styrene and methyl methacrylate in a continuous-flow reactor [62]. The challenge was that differing monomer reactivities cause the copolymer composition to drift from the feed ratio. By employing BO with two different batch sizes per iteration (a "sparse" 4-candidate vs. an "exhaustive" 40-candidate strategy), the authors quickly converged on conditions that held the copolymer composition near the target. The sparse BO found an optimal formulation in just five iterations, while the exhaustive BO revealed there were multiple viable solutions (identifying solvent: monomer ratio as a crucial variable). This demonstrated BO's efficiency, validating that even with limited runs, BO could mitigate compositional drift. This study makes three key contributions: 1) validation of BO as an effective tool for minimizing compositional drift with limited experimental runs, 2) quantification of solvent-mediated kinetics in the copolymerization process, and 3) establishment of the infrastructure for multi-objective BO, aiming to optimize copolymer composition and physical properties simultaneously. These findings suggest that BO can be extended to other polymerization systems requiring precise microstructure control to tailor material properties.

While traditional BO is typically applied to optimize a single objective, such as maximizing yields or minimizing dispersity, many real-world polymer systems involve multiple, often competing objectives, including achieving high molecular weight while maintaining low dispersity. To address this complexity, multi-objective BO (MOBO) balances trade-offs among multiple performance metrics by identifying Pareto-optimal solutions. This approach enables the simultaneous and data-efficient optimization of several objectives within a unified modeling framework, making it particularly valuable for guiding complex materials design tasks [58]. For example, a MOBO approach is applied to optimize the production of polyvinyl acetate (PVAc), aiming to achieve a balance between product quality and manufacturing cost [63]. Recognizing the intrinsic trade-offs in synthesizing long-chain branched polymers, the authors applied MOBO with a q-Expected Hyper-Volume Improvement (qEHVI) acquisition function to identify optimal reaction conditions. The approach drastically reduced computational demand by modeling the polymerization process with differential equations and leveraging the data efficiency of Bayesian inference. Such results demonstrate that MOBO can efficiently navigate trade-offs in complex polymer syntheses, identifying optimal conditions with orders of magnitude fewer trials than traditional methods, thereby highlighting the significant impact of incorporating multi-objective decision strategies on autonomous optimization.

4 | Self-Driving Labs: Toward Autonomous Polymer Discovery

An SDL is an extension of traditional HTE platforms, coupling robotic hardware, real-time feedback, and AI-based decision-making. These closed-loop platforms can autonomously design, execute, and analyze experiments without direct human intervention, iteratively improving their strategies with each cycle [11, 12]. Unlike standard automation, which operates static, predefined protocols, SDLs dynamically adapt their workflows based on real-time feedback, enabling adaptive exploration of synthesis and processing conditions [64].

A typical SDL is structured around hardware, software, and a data infrastructure layer, similar to the body, brain, and memory of a biological system [64]. In this configuration, the hardware mimics the physical body, performing operations traditionally performed by hand, such as dispensing reagents and transferring samples between the synthesis and characterization modules. It includes robotic components, such as reactors, mixers, heaters, and flow systems, for synthesis; as well as analytical instruments, including spectrometers, chromatographs, and microscopes, for characterization [11]. The software component acts as the brain, coordinating all hardware actions through custom scripting and orchestrating complex experimental workflows [11, 65, 66]. Above this, advanced algorithms, such as BO, serve as the cognitive cortex. These algorithms interpret incoming data in real-time and determine the next set of experimental conditions, such as optimizing conductivity or minimizing dispersity, in light of predetermined research objectives [64]. Ultimately, the data infrastructure serves as the memory backbone, supporting both real-time data acquisition and storage [12, 64]. This infrastructure also stores historical knowledge, such as literature data or prior experimental results, as well as metadata, sensor readings, and contextual annotations, all stored in centralized databases, providing the SDL with information to be utilized during decision-making by data-driven models. Comprehensive metadata standards enable traceability, reproducibility, and long-term sustainability [11, 67, 68]. Meanwhile, real-time dashboards provide both AI/ML systems and human operators with an interactive view of ongoing processes and results. Together, these components establish a self-reinforced "design-build-test-learn" cycle. AI algorithms design candidate experiments, the robotic hardware builds and tests them, instrumentation provides data for evaluation, and AI/ML models retrieve data from databases, actively learns, and continuously refine their predictive models, suggesting subsequent iterations. This closed loop can iterate quickly and even continuously, enabling SDLs to efficiently explore large, high-dimensional design spaces and find optimal solutions much faster than traditional methods [11].

As SDLs evolve, autonomy becomes a defining characteristic distinguishing them from conventional automated systems. With increasing sophistication, autonomy is no longer a binary concept but is instead classified across hardware and software axes [66]. These dimensions work in tandem; in other words, advances in hardware capabilities must be matched by corresponding software sophistication to progress between autonomy levels. Hardware autonomy refers to the degree of a system's control over physical tasks, ranging from simple actions, such as liquid dispensing, to complex ones, like inline sensing, sample transfers, and coordinating multiple instruments. Software autonomy refers to the AI's ability to plan, optimize, detect anomalies, and generate hypotheses. These two dimensions intersect to create discrete levels, where advancing hardware capabilities must be matched by corresponding software sophistication to achieve higher levels of autonomy. Combined, these axes yield a six-level autonomy framework, from Level 0 (manual operation) to Level 5 (full autonomy). According to Tom and his team's analysis, most SDLs operate at Levels 2 and 3, where algorithmic decision-making is combined with partial robotic task execution, with the long-term vision of reaching Level 5 [66]. At Level 5, while systems operate with full hardware autonomy, they still require human oversight for strategic oversight and workflow co-supervision. This persistent human involvement, even at the highest autonomy level, underscores the complex interplay between automated capabilities and human judgment, suggesting that advanced SDLs may be evolving toward sophisticated forms of human–AI partnership rather than pure autonomy. Figure 1 illustrates this six-level autonomy within the context of polymer chemistry, demonstrating how each level does not eliminate human expertise but instead signifies its evolution and harmonization with increasingly advanced AI capabilities. Using the example of Dr. Mehmet Mutlu's development of biodegradable polymers for food packaging, the evolving workflow illustrates how human intelligence adapts and becomes more specialized, rather than diminishing, as technology advances. This progression suggests that increasing laboratory automation reflects a growing collaboration between humans and AI, rather than a shift toward sole software autonomy, with human intelligence adapting to provide more sophisticated cognitive input. At the same time, AI supports, rather than replaces, vital human expertise in polymer discovery.

This vision is already becoming a reality through pioneering SDL implementations. One such example is Argonne National Laboratory's "Polybot," a modular platform that employed feature importance-guided BO to navigate a 7D parameter space and identify optimal parameters to synthesize conductive polymer films with conductivities exceeding 4500 S cm^{-1} [69]. Another example is Brabec's group's automated high-throughput platform for synthesizing multicomponent polymer blend films for photovoltaics. In this design, the SDL autonomously varied the ratios of donor and acceptor polymers and solvent mixtures to maximize solar cell efficiency. By coupling BO with automated film deposition and screening, the system rapidly mapped a 4D composition space and identified blend ratios that yielded significantly improved power conversion efficiency and stability [70].

Cloud connectivity is another emerging feature used to construct data infrastructure in SDL, as it enhances modularity, remote access, and cross-institutional collaboration. For example, Li and coworkers developed a remotely accessible microfluidic SDL for fully autonomous synthesis of perovskite nanocrystals [71]. The device combines robotic liquid handling with inline spectroscopy and a cloud-integrated data management layer, allowing users to remotely launch, set goals, and monitor experiments in real-
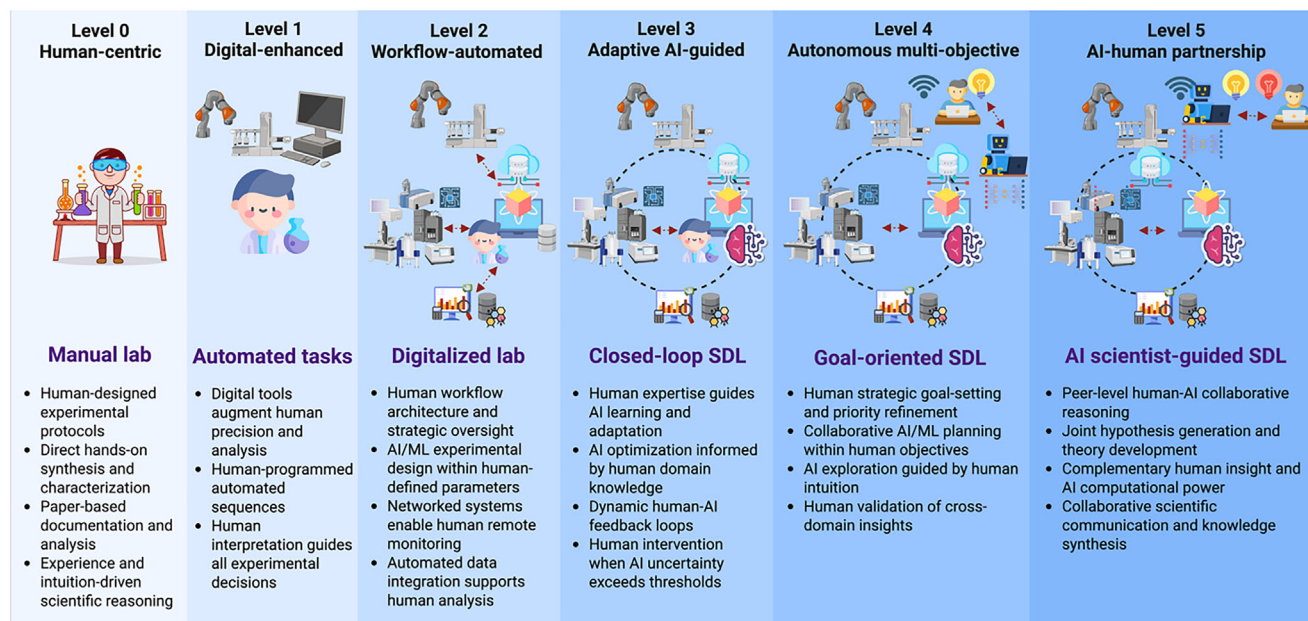
FIGURE 1 | The evolution of human–AI collaboration in laboratory automation is explained through an illustrative polymer synthesis workflow. This framework presents six distinct modes of cooperation between humans and AI-guided machines, which will be illustrated through a case study involving Dr. Mehmet Mutlu's work on biodegradable polymers in labs with varying autonomy levels. *Level 0* depicts traditional human-centered scientific practice where Dr. Mutlu manually designs experimental protocols, weighs lactide monomers on balances, monitors reaction temperatures personally, and records all polymerization data in lab notebooks. His decisions depend on experience and direct observation of reaction progress. *Level 1* demonstrates digital enhancement of human abilities, where Dr. Mutlu continues to design protocols, while digital balances automatically log measurements, and temperature probes provide continuous data. He uses this improved data to make better adjustments to heating and reaction conditions. *Level 2* demonstrates workflow automation with strategic human oversight, where Dr. Mutlu programs robotic synthesis platforms to perform multi-step polymerization sequences overnight. He designs automated protocols, reviews real-time molecular weight data, and adjusts workflow parameters based on his analysis of reaction kinetics. *Level 3* illustrates adaptive AI-guided collaboration, where Dr. Mutlu sets initial parameters and target properties, while AI systems monitor multiple reaction variables and make micro-adjustments to optimize synthesis pathways. When the AI detects unexpected reaction intermediates, it flags anomalies for user interpretation, creating dynamic feedback loops between human knowledge and AI pattern recognition. *Level 4* features a goal-oriented strategic partnership, where Dr. Mutlu defines overarching objectives for biodegradability, mechanical strength, and cost targets while collaborating with AI systems that design and run parallel experiments across multiple reactors. His research intuition guides resource decisions, while AI's computational power explores chemical space beyond his knowledge. *Level 5* exemplifies peer-level partnership, where Dr. Mutlu and an AI scientist jointly propose new hypotheses combining insights from marine biology and synthetic chemistry. Together, they design experiments that neither could conceive alone, with AI modeling molecular interactions and Dr. Mutlu contributing to the understanding of synthesis feasibility and practical applications, leading to co-developed research insights. The figure is created in BioRender under an academic subscription. (Kabay, G. (2025), https://BioRender.com/1gaa938).

time via a web interface. Knox and his team developed another SDL platform that integrates cloud-based machine learning and orthogonal online analytics for the synthesis of polymeric nanoparticles [72]. They successfully demonstrated the multi-objective optimization features of the ML algorithms, maximizing conversion and minimizing both molar mass dispersity and the particle polydispersity index, all while maintaining the particle size within a predetermined value.

These examples collectively demonstrate how AI-assisted SDLs can handle various laborious tasks, from formulation to structural control, by navigating high-dimensional parameter spaces. While these examples support workflows with little or no direct human intervention, the goal is not to eliminate scientists but to reposition them as collaborators in a human–AI symbiosis. In the following section, we explore this adaptive model, where human expertise complements algorithmic reasoning to create more robust, interpretable, and ethical autonomous SDLs.

4.1 | Adaptive Autonomy: Symbiotic Human–AI Cooperation

As SDLs move toward increased autonomy, it is evident that full automation, while beneficial, is not always the optimal or most feasible choice, particularly in polymer discovery, where the chemical design spaces are vast and the available data can be limited or biased [73]. Rigid workflows may risk missing unexpected or creative acumens that often come from human intuition. To address this, hybrid autonomy models, which combine human intuition and machine intelligence, are evolving. These models combine AI's strengths in pattern recognition and computational efficiency with the scientist's ability to interpret context, apply domain-specific knowledge, and think creatively. Studies in active learning have shown that "human-in-the-loop" agents, where humans and machines collaborate to design and interpret experiments, outperform those performed by solely human or AI agents alone [74]. In this collaborative approach, AI excels at conducting in-depth searches, running large-scale experiments,

and detecting correlations across data, while humans bring domain-specific reasoning, anomaly detection, and hypothesis development. As a result, this synergistic strategy allows for more informed, adaptable, and innovative exploration of complex polymeric systems. This raises a critical question that one should ask: How can we determine when a machine should handle a task and when human intervention is essential?

Building on this hybrid strategy, the concept of adaptive autonomy arose to answer this critical question. Depending on the task context, adaptive autonomy dynamically allocates tasks between humans and machines in response to uncertainty, complexity, or the need for creativity [75]. For example, AI can autonomously execute routine SDL experiments, while human input is triggered when model uncertainty is high, a decision is risky, or expert opinion is required [76–78]. The ChemOS platform is a leading example of such adaptive autonomy integrated into SDL, enabling users to set experimental goals and parameters, then autonomously select reaction conditions and conduct experiments, pausing to seek human input when encountering unfamiliar situations or uncertain data [75]. Likewise, the Polybot system at Argonne National Laboratory links synthesis, processing, and characterization stations in a closed-loop pipeline, while allowing for human intervention as needed [69]. The Autonomous Formulation Laboratory (AFL) extends this strategy to an open-source and modular SDL architecture for complex soft matter formulations, combining ML with high-throughput, reproducible data generation and real-time feedback integration. The modular hardware–software architecture allows dynamic task allocation, demonstrating how human–AI teams can co-pilot high-throughput studies through adaptive autonomy [79]. By dynamically balancing machine automation and intelligence with human expertise and insights, these hybrid frameworks prevent blind spots, balancing reproducibility with creativity and safety [72].

4.2 | Critical Aspects of Human–AI Collaboration

4.2.1 | Transparency: Interpretable AI Models and Visualization Tools

As ML models become integral to polymer research, transparency, interpretability, and explainability are critical for building trust in human–AI teams. Researchers must understand why an algorithm suggests a particular polymer or predicts a specific property, especially when results contradict conventional wisdom. Various explainability methods, such as inherently interpretable models [74, 80–82], symbolic regression [83, 84], post hoc [85–88], attention mechanisms [89–91], and visualization tools (e.g., t-SNE/UMAP embeddings or uncertainty maps) [51, 80, 92, 93], are being integrated into AI workflows to enhance transparency. Also, inherently interpretable "glass-box" models can be used when applicable. For example, linear models and generalized additive models provide transparent relationships between inputs and outputs, allowing researchers to observe which molecular features define performance directly [74]. However, simplistic models, such as linear regression, shallow decision trees, or heuristic rule-based systems, often fail to capture complex, nonlinear SPR [94]. Symbolic regression offers a more interpretable alternative that can uncover equations

relating, for instance, the structural features of monomers to macroscopic polymer properties, providing clearer mechanistic insight than opaque black-box models [74]. Researchers have even proposed using symbolic regression piecewise in different "regimes" of the data (identified via clustering) to maintain interpretability across a complex domain [74]. On the other hand, for high-dimensional models that remain essentially black boxes (e.g., deep neural networks or graph neural networks), post hoc explanation techniques are increasingly employed to address interpretability issue by identifying the most influential features, verifying model reasoning against chemical knowledge, and guiding experimental design in polymer chemistry and materials science. Methods like SHAP and LIME can highlight which input features most influenced a specific prediction, albeit on a local case-by-case basis [74]. While not providing a complete mechanistic understanding, such techniques offer valuable insights and help to identify false correlations or domain-insensitive predictions.

Visualization tools bridge the interpretability gap by translating abstract model representations, such as attention weights or latent embeddings, into intuitive formats that enable human researchers to act upon them. For instance, attention-weight heatmaps in a transformer-based polymer predictor, such as TransPolymer [93] can be visualized to show which parts of a polymer sequence the model "attends" to most for a given property prediction, implying that it has learned SPR [92, 93, 95]. Dimensionality reduction plots (e.g., UMAP or t-SNE) of learned polymer embeddings can reveal clusters of chemically similar polymers in the model's latent space [96–98]. Likewise, uncertainty maps can be displayed to show specific regions of the design space where the model is uncertain, informing users where more data or caution is needed. Interactive dashboards are also handy, as users can adjust input parameters or constraints and immediately see the model's predictions or recommendations update, creating a two-way dialogue between the AI assistant and the user [99–101]. Such visual analytics builds user trust and encourages researchers to reintroduce their intuition into the loop, enabling users to guide, validate, and correct AI behavior if necessary, critical attributes in hybrid workflows where human judgment and algorithmic prediction must co-evolve.

For non-experts or students, simple explanations can be generated in natural language to summarize why the model made a specific choice. Indeed, an exciting approach is rationale generation, which produces human-like explanations for an AI's behavior [102, 103]. These rationales do not fully unveil the black box, but they translate the algorithm's decisions into narratives that a human can understand. Studies have demonstrated that human-like justifications, even if approximate, greatly improve users' trust and comprehension of autonomous systems. In one study, researchers trained a neural network to provide explanations, much like a human would justify a decision. While these explanations did not precisely reflect the AI's internal reasoning, they increased people's willingness to trust and collaborate with the AI [103]. This work highlights that explainability for non-experts is not just a nice addition, but a requisite, especially in interdisciplinary fields like polymer science, where end-users may not be AI experts. Multiple forms of interpretable feedback (e.g., visual, mathematical, linguistic) enable bidirectional knowledge transfer between humans and AI agents: the AI reveals SPRs,

generates hypotheses, and optimizes design parameters, while humans gain insights into the AI's decision-making process and grasp its limitations, guiding subsequent model refinement and prioritization [74].

4.2.2 | Explainability and Trust: Human–AI Co-Design

For hybrid human–AI systems to reach their full potential, explainability and trust must be explicitly engineered into every stage of the system lifecycle. It is insufficient for the AI to be powerful; its decisions and suggestions must be transparent, interpretable, and justifiable to the human collaborators [74, 103]. This requirement is underlined in the human–machine co-design concept, wherein humans assist in designing the AI's decision-making processes, constraints, and objectives that align with human reasoning patterns and domain-specific values.

Practically, co-design means incorporating domain knowledge and human-provided constraints into model architectures (for instance, encoding known chemical rules or safety limits prevents AI from proposing irrelevant or dangerous experiments). Additionally, it requires developing AI systems that can explain their rationale in user-friendly terms. When humans and AI collaboratively design experiments and models, the workflow reduces trust issues. Humans are not left uncertain about the machine's motives, and the machine is less prone to exploring irrational areas due to human guidance. For example, an AI agent may propose a polymer synthesis that is theoretically optimal but uses a toxic reagent or an impractically long reaction time. If so, a human expert can promptly reject these options and recognize the limitations of AI for further design adjustments.

Over time, reinforcement learning or user feedback loops can enable the AI system to infer implicit preferences, evolving toward context-aware optimization strategies that better reflect real-world laboratory conventions [103]. It is also important to mention that this co-design process is reciprocal, meaning that, similar to how AI systems learn from human constraints and corrections, humans must adapt by developing algorithmic literacy, refining their hypotheses based on machine-generated insights, and becoming comfortable with probabilistic or counterintuitive outputs. Through collaborative goal-setting and interpretive feedback, co-design stimulates mutual calibration: the AI adjusts to what the human considers valuable or valid, while the human learns to identify new opportunities, patterns, or design pathways that the AI reveals. This mutual trust is a critical component of enduring and symbiotic human–AI collaboration in SDLs.

4.2.3 | Uncertainty Quantification and Active Learning

A cornerstone of human–AI collaboration in experimental design is quantifying uncertainty. Rather than viewing uncertainty as a weakness, hybrid workflows treat it as an informative signal for guiding exploration. Understanding what the model does not know is crucial for deciding when to explore new experiments or seek human input [72].

Modern ML models can quantify uncertainty in their predictions. For example, Bayesian neural networks yield a confidence interval with each property prediction, distinguishing the design space regions where the model is extrapolating with sparse or low-quality data [76, 104, 105]. By identifying predictions with high epistemic uncertainty, which always stems from model limitations such as insufficient data or unfamiliar inputs, and high aleatoric uncertainty arising from inherent noise in the data, the system can flag those situations for closer inspection [72, 76–78]. In light of these predictions, an SDL might automatically suggest a new iteration in the region with the highest model uncertainty, assuming that new data in this region would reduce overall uncertainty. This forms the basis of active learning, where the AI proactively selects the most informative subsequent experiments to perform, instead of passively analyzing a predetermined dataset. Rather than relying on random or exhaustive searches, it directs experimental resources toward resolving uncertainties or marking a balance between exploration and exploitation in design space.

Active learning acts as a dynamic bridge between algorithmic exploration and human insights [106–110]. Instead of an exhaustive search, it prioritizes experimental resources toward resolving knowledge gaps. For instance, the Knox team recently demonstrated a cloud-connected SDL that enhanced polymeric nanoparticle synthesis by combining active learning with real-time analytics, effectively balancing various objectives such as particle size, shape, and yield [72]. Their system could autonomously navigate within the complex design space more efficiently than manual methods. Simultaneously, a remote team of scientists monitored and intervened as necessary via cloud interfaces, demonstrating successful utilization of adaptive autonomy.

To sum up, by quantifying models' confidence level or uncertainty, AI systems know when to ask for help or defer decisions, functioning as a primitive machine "self-awareness" that complements human judgment. The result creates a closed-loop feedback cycle (a cyclical process in which outputs from a system are used as inputs for future iterations), wherein the model's uncertainty prompts new experiments, those experiments generate data, and the human expert interprets outcomes, adjusts hypotheses, or corrects the system. In this continuous cycle, AI gains confidence and precision, while the human researcher maintains epistemic control and scientific direction.

4.2.4 | Virtual Agents for Collaboration and Communication

One of the most game-changing developments for human–AI workflows is the rise of large language models (LLMs), such as GPT-4, and their domain-specific counterparts (e.g., ChemCrow in chemistry). These models act as intelligent assistants or scientific mediators in the lab [111, 112]. Instead of requiring a human to write code or manually query databases, an LLM-based system can understand a researcher's questions or instructions in

their natural language and translate them into actionable steps. For instance, a chemist who wants to perform polymer synthesis might ask in plain English, "What monomer should I try next to increase the T_g?". The LLM can interpret this query, combine knowledge from literature and the lab's databases, integrate them into the lab's knowledge base, and suggest an experiment or retrieve relevant data, complete with an explanation. Models like GPT-4 are already used to generate experimental procedures and optimize protocols from text inputs [111]. On top of that, LLMs enable conversational interfaces for complex AI systems: a researcher can interact with the model, refining questions or constraints in real time, much like collaborating with a virtual colleague. This vividly lowers the barrier for non-experts to engage with AI-driven platforms. Early applications, such as ChemCrow, have demonstrated that an LLM augmented with chemistry tools can assist in writing simulation code, controlling lab instruments via text commands, and even devising *de novo* molecular structures, all through a chat-based interaction [112]. By serving as a bridge between human language and machine-level execution, LLMs allow the unified integration of human expertise into autonomous workflows and vice versa. Humans can describe goals or constraints, and the AI interprets and acts while explaining its reasoning in understandable terms. This level of accessibility and transparency establishes trust, encouraging more people to utilize AI tools without requiring specialized training.

4.2.5 | Human–AI Synergy in SDL Practice

The polymer informatics field is rapidly evolving through the accumulation of examples, which are highlighting the potential of adaptive autonomy. On the AI side, polymer-savvy models such as TransPolymer and polyBERT treat polymer sequences as a language, utilizing transformer neural networks to learn rich representations of polymer chemistry [92, 93, 95]. TransPolymer utilizes self-attention and chemistry-aware tokenization of polymer chains to predict properties with high accuracy, leveraging pretraining on large unlabeled polymer datasets (similar to how GPT-4 is pre-trained on text) to capture chemical context. Meanwhile, polyBERT employs multitask learning that maps these learned polymer fingerprints to multiple property predictions, enabling the simultaneous prediction of performance metrics. This machine-learned fingerprint approach outperformed traditional, manually constructed descriptors by two orders of magnitude in speed while maintaining predictive accuracy [92]. Such models can be integrated into autonomous workflows to provide instant predictions of a candidate polymer's attributes, enabling researchers to screen possibilities or identify anomalies rapidly.

At the same time, advanced robotic platforms and high-throughput synthesis pipelines are expanding experimental capabilities in SDLs. A potential adaptive human–AI workflow may, for example, use polyBERT's fast property predictions to narrow down a vast library of potential polymer structures to a manageable subset, employ an active-learning-driven experimentation to synthesize and test the top candidates, and finally rely on the human expertise to interpret results, assess failure modes and refine the design objectives. In a recent demonstration, Polybot

autonomously performed hardware actions while an AI planner optimized a polymer formulation [69]. Yet, researchers remained "in the loop" to refine goals and enforce safety constraints, highlighting that neither humans nor machines alone can achieve the same speed and scientific insight as human–AI cooperation allows.

Notably, such hybrid workflows not only benefit from human input but also enable humans to learn from AI outputs. Models trained on vast chemical datasets can reveal unexpected patterns, suggest new candidates, or challenge common sense, stimulating new scientific questions, if taken seriously. This bidirectional flow of reasoning, which allows both humans and AI to learn from each other, forms the core of symbiotic autonomy-guided SDL.

Figure 2 schematically represents a conceptual human–AI symbiotic workflow for photoinduced electron/energy transfer–reversible addition-fragmentation chain transfer (PET-RAFT) polymerization of sustainable biomedical polymers, using a hypothetical AI-driven SDL assistant, "LabCoPilot Genie". While individual components shown (such as robotic platforms, inline analytics, and cloud connectivity) are based on existing SDL technologies, the complete integrated system with conversational AI interface represents our proposed vision for next-generation SDLs that operate with the highest level of hardware autonomy and human–AI symbiosis rather than describing any single existing implementation.

5 | Summary and Outlook

In this perspective, we provided a historical overview of polymer chemistry, from early alchemical practices to systematic rule-based methodologies and today's AI-driven autonomous SDLs, revealing how the limitations of each period have catalyzed subsequent innovation. The emergence of polymer informatics and ML has addressed historical constraints through data-driven pattern recognition and predictive modeling, with models such as polyBERT [92] and TransPolymer [93] demonstrating how AI can learn chemical language and predict properties with remarkable accuracy. SDLs exemplify symbiotic potential through adaptive autonomy frameworks, as shown by platforms like Polybot, which dynamically allocate tasks between human insight and algorithmic execution. Although there is an increasing demand to shift full software autonomy, we collectively believe that the actual breakthrough lies not in replacing human with AI, but in orchestrating partnerships where human creativity and AI computational power amplify each other's capabilities, transforming researchers into orchestrators of intelligent systems while maintaining essential roles in strategic oversight, ethical reasoning, and scientific interpretation.

Human–AI collaboration in self-driving polymer laboratories is restricted by fundamental challenges that must be addressed to realize its transformative potential. The primary obstacle is the limitations of the current training datasets, which are often small, biased toward well-explored materials, and lack the diversity needed for robust generalization. These constraints may create blind spots and lead to overfitting, ultimately limiting the generalizability, prediction performance, and restricting the real-world applicability of these models [113]. Digital twins,
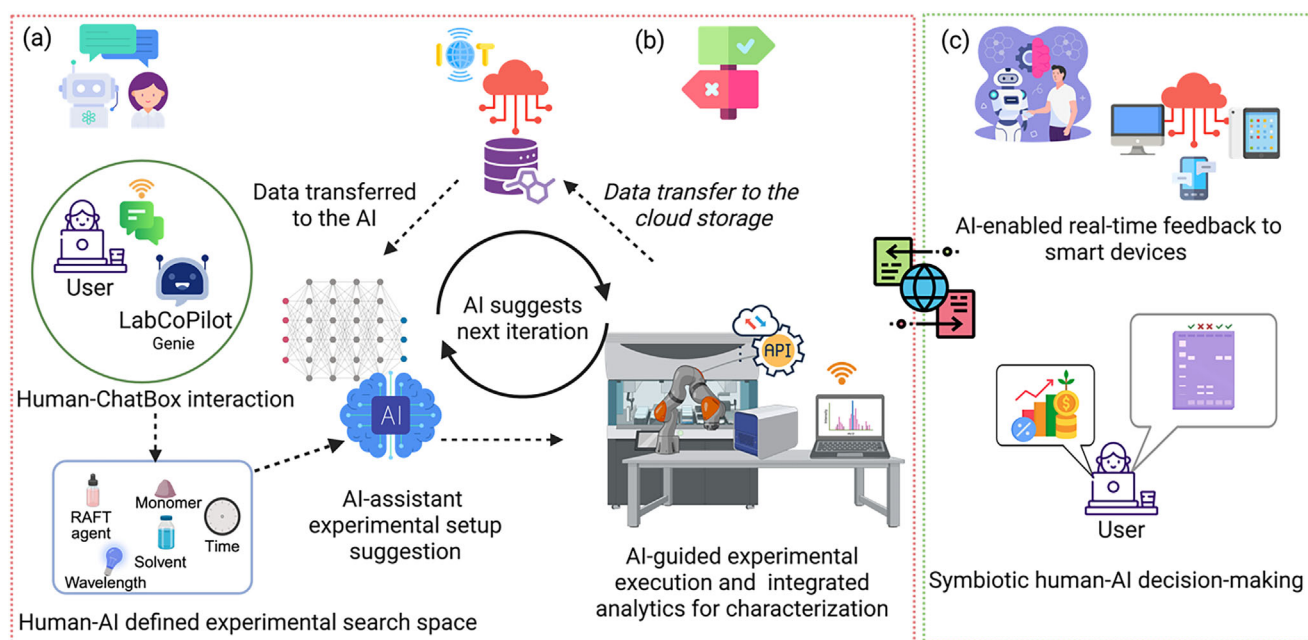
FIGURE 2 | Conceptual schematic overview of an adaptive SDL workflow enhanced by symbiotic human–AI collaboration for sustainable polymer synthesis via PET-RAFT polymerization. The user initiates the process by setting research goals and performance objectives through the ChatBox interface. LabCoPilot Genie searches chemical databases to determine the virtual design space, recommends HEMA, and proposes a set of reaction conditions for synthesis. Upon user confirmation, the robotic platform executes the experiments, while inline analytics continuously monitor key outputs such as monomer yield and molecular weight distribution. Experimental data are transferred to cloud storage, enabling real-time remote data access and monitoring, multi-user collaboration, and model retraining. As data accumulates, the AI learns and uncovers hidden SPR patterns using active learning strategies, suggesting subsequent iterations. Simultaneously, bidirectional feedback loops ensure continuous adaptation; the AI refines its suggestions based on evolving data and human feedback, while users can intervene to adjust goals, override AI decisions, interpret anomalies, or prioritize design trade-offs (e.g., favoring biocompatibility over maximum yield). In parallel, AI-generated insights are fed back to the cloud, connected smart devices, and user interfaces, facilitating remote monitoring, real-time alerts, and informed decision-making by users, as needed. This mutually adaptive architecture enables shared control, trust calibration, and the exchange of insights between human and AI agents. The platform emphasizes transparency, modularity, and openness through open-source software and cloud connectivity, supporting decentralized, collaborative SDL operations across diverse research teams. The figure is created in BioRender under an academic subscription. (Kabay, G. (2025) https://BioRender.com/tp96c6w).

virtual replicas of physical experimental systems, address this challenge by enabling safe, rapid, and cost-effectively simulations to investigate "what-if" scenarios before physical implementation. These virtual platforms integrate human intuition with AI predictions, providing interactive hypothesis testing while generating synthetic data to enhance training datasets [114–117]. ML strategies, such as multi-task learning (training models on multiple property prediction tasks simultaneously) and transfer learning (pre-training a model on big data, then fine-tuning it on specific tasks), further mitigate data-related limitations by leveraging broad chemical knowledge to improve model robustness and promote cross-task generalization [118]. These strategies also support community-wide knowledge transfer, where one lab's HTE improves another lab's AI model. That knowledge can even help to predict a niche property with only a few experimental iterations [118]. The critical need is to establish more initiatives for data sharing, measurement standardization, and real-time feedback from HTE systems to improve the dataset count and its relevance [119–121].

The persistent trade-off between model complexity and interpretability presents both technical and cultural challenges. While explainable AI tools help demystify black-box models, the inability of users to understand or question an AI's reasoning under-

mines their trust and adoption [122]. Integrating AI into SDL workflows requires a significant cultural shift; scientists must develop new skills to interpret uncertainty estimates, interact effectively with AI assistants, and critically assess machine-generated hypotheses. In parallel, AI systems must be trained to accommodate human preferences, for instance, by allowing users to provide clear rationales for suggestions, justify outputs in understandable terms, and enable manual human overrides. One of the most effective ways to bring AI and humans closer together as colleagues is to enhance their communication. This can be achieved through implementing user-centric interfaces and LLM-powered chatbots in the SDL workflow, as these tools can facilitate natural language interactions, thereby enhancing transparency in AI decision-making processes. Notably, special attention must be paid to fairness, bias, and the ethical use of AI to ensure diverse, transparent, and ethical recommendations, which are critical aspects of the successful implementation of human–AI symbiosis in SDL practice [103].

At the implementation level, human involvement in such AI-driven systems provides critical benefits in SDLs. Expert oversight adds a trust layer in algorithmic decision-making by validating AI-generated hypotheses and ensuring accountability against the black-box models [123]. By insisting on explainable, inter-

pretable outputs, scientists connect complex model predictions with chemical intuition, allowing limited experimental datasets to be utilized more effectively [124]. Indeed, human experts excel at generalizing from sparse data, infusing first-principles knowledge and symbolic reasoning so that AI systems can explain more observations with fewer experiments [125]. Equally important, human creativity and intuition inject innovative hypotheses and experimental directions that purely algorithmic searches might overlook [126]. The synergy of human insight and machine precision enhances the generalizability and reliability of SDL outcomes, enabling the identification and correction of model blind spots or biases while maintaining transparency, and ultimately building user trust in AI-guided autonomous workflows [125].

Polymer synthesis operates according to well-established physical and chemical principles (e.g., thermodynamics, reactivity, solubility) that purely data-driven methods may overlook. Determining when to impose human domain knowledge remains critical. Explicit constraints, such as toxicity avoidance, can be easily encoded, whereas nuanced insights, like feasibility, prove more challenging to formalize. Hybrid models that combine mechanistic reasoning with machine intelligence have yet to evolve, while knowledge graphs and symbolic engines may bridge the gap between chemical validity and data-driven creativity [118, 127, 128]. Despite these challenges, robotic platforms integrated with AI agents enable the synthesis and characterization of hundreds of polymer variants with unprecedented speed, creating opportunities for active learning approaches that focus resources on maximally informative experiments [94, 129, 130]. As costs fall and ease of use rises, more labs can deploy such material acceleration platforms, potentially networked via the cloud for global collaboration and simultaneously parallelized exploration.

Generative models represent a unique exploration capability by proposing novel polymer structures that extend beyond human intuition, as researchers evaluate their feasibility and accessibility [131, 132]. This co-creative partnership has already yielded breakthrough discoveries in recyclable materials and biomaterials [133, 134], demonstrating how human–AI collaboration expands the known boundaries of polymer discovery. The most profound opportunity lies in advancing fundamental understanding of polymer science through data-driven pattern recognition combined with human interpretation, enabling the discovery of new SPRs and theoretical principles that neither partner could achieve independently.

The next generation of SDLs will demand integration over replacement, with symbiotic intelligence reshaping how materials are discovered and knowledge is generated. This transformation redefines the human's role as an orchestrator of intelligent systems, setting research goals and performance targets, navigating algorithmic landscapes, and interpreting high-dimensional data within mechanistic frameworks. Looking ahead, several emerging directions will define the horizon of symbiotic human–AI-assisted SDL technologies for polymer discovery:

i. As polymer databases grow in volume and variety, next-generation models, such as polyBERT [92], TransPolymer [93], and graph neural networks, will continue to learn directly from chemical language and structure, capturing context-aware, transferable knowledge that improves across tasks and domains.

ii. With LLMs like GPT-4, domain-specific co-pilots such as ChemCrow, and dashboard-driven workflows, humans will interact with AI systems in their natural language, supported by visualization or tactile simulation, making collaboration fluid and conversational, rather than code-bound [131, 132].

iii. FAIR-compliant datasets, model repositories, and shared benchmarks will become standard. As sequence databases revolutionized bioinformatics, structured polymer data and interpretable models will catalyze community-driven discovery.

iv. Cloud-based orchestration will further enhance fully integrated robotic platforms, such as Polybot or NIST's AFL, facilitating distributed and parallelized experimentation on a global scale.

v. Decentralization is crucial for constructing durable and trustworthy networks of SDLs that securely share and verify experimental data, workflows, and discoveries across autonomous nodes, thereby eliminating the need for central authorities. Blockchain and smart contracts can support such infrastructure, ensuring data integrity, encouraging collaboration, and enabling the autonomous cooperation of tasks among distributed nodes [135–137].

vi. Future hybrid systems will not optimize solely for performance metrics. Instead, human-in-the-loop constraints will incorporate environmental impact, safety, and equity as primary objectives, guiding polymer discovery toward responsible innovation.

The evolution toward symbiotic intelligence transcends the capabilities of imaginary Level 5 automation, as AI tools enhance scientific imagination and reveal previously hidden connections, while human scientists contribute their expertise, real-world experience, and creativity. As the boundary between artificial and human intelligence blurs, polymer science will be characterized by adaptive autonomy and symbiotic intelligence: a blend of reasoning and data that is iterative, interpretable, and inclusive. This partnership will hence enable the exploration of previously inaccessible experimental territories, accelerating not only the speed of discovery but also the mechanistic understanding through rapidly tested and refined hypotheses. All in all, the upcoming age of polymer discovery, facilitating human–AI symbiotic partnerships in adaptive SDL workflows, promises to be not only faster but fundamentally wiser.

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. H. Staudinger, "Über Polymerization," *European Journal of Inorganic Chemistry* 53 (1920): 1073–1085.

2. R. Plunkett, "Tetrafluoroethylene Polymers," US2230654A (Kinetic Chemicals Inc, 1939), https://patents.google.com/patent/US2230654A/en.

3. N. Botha, H. M. Inglis, R. Coetzer, and F. J. W. J. Labuschagne, "Statistical Design of Experiments: An Introductory Case Study for Polymer Composites Manufacturing Applications," In *12th South African Conference on Computational and Applied Mechanics*, SACAM 2020, (EDP Sciences, 2020).

4. G. E. P. Box and K. B. Wilson, "On the Experimental Attainment of Optimum Conditions," *Journal of the Royal Statistical Society Series B: Statistical Methodology* 13 (1951): 1–38.

5. H. Scheffé, "Experiments With Mixtures," *Journal of the Royal Statistical Society Series B: Statistical Methodology* 20 (1958): 344–360.

6. P. M. Doyel, "Combinatorial Chemistry in the Discovery and Development of Drugs," *Journal of Chemical Technology and Biotechnology* (1995): 317–324.

7. S. Brocchini, K. James, V. Tangpasuthadol, and J. Kohn, "A Combinatorial Approach for Polymer Design," *Journal of the American Chemical Society* 119 (1997): 4553–4554.

8. S. Brocchini, K. James, V. Tangpasuthadol, and J. Kohn, "Structure-Property Correlations in a Combinatorial Library of Degradable Biomaterials," *Journal of Biomedical Materials Research* 42 (1998): 66–75.

9. Borealis. Annual Results Report: 2008 at a Glance, Vienna, Austria, 2008, https://www.borealisgroup.com/storage/Company/Investor-Relations/Financial-Reports/Annual-Report-of-Borealis-Group-2008-English.pdf (accessed 2025-06-23).

10. O. Queen, G. A. McCarver, S. Thatigotla, et al., "Polymer Graph Neural Networks for Multitask Property Learning," *npj Computational Materials* 9 90 (2023): 9.

11. M. Abolhasani and E. Kumacheva, "The Rise of Self-Driving Labs in Chemical and Materials Sciences," *Nature Synthesis* 2, no. 6 (2023): 483–492.

12. P. A. Beaucage, D. R. Sutherland, and T. B. Martin, "Automation and Machine Learning for Accelerated Polymer Characterization and Development: Past, Potential, and a Path Forward," *Macromolecules* 57, no. 18 (2024): 8661–8670.

13. S. Mavi, S. Kadian, P. K. Sarangi, et al., "Advancements in Machine Learning and Artificial Intelligence in Polymer Science: A Comprehensive Review," in *Macromolecular Symposia* (Wiley Online Library, 2025), 2400185.

14. D. Hosler, S. L. Burkett, and M. J. Tarkanian, "Prehistoric Polymers: Rubber Processing in Ancient Mesoamerica," *Science* 284 (1999): 1988–1991.

15. F. Dorrego, F. Carrera, and M. P. Luxán, "Investigations on Roman Amphorae Sealing Systems," *Materials and Structures* 37 (2004): 369–374.

16. J. P. Smith, "Plants and Civilization, An Introduction to the Interrelationships of Plants and People," *Plants, People, and Planet* (2006), https://digitalcommons.humboldt.edu/cgi/viewcontent.cgi?articl#x0003D;1008&context#x0003D;botany_jps_online_content.

17. C. Goodyear, *Gum-Elastic and Its Varieties: With a Detailed Account of Its Applications and Uses, and of the Discovery of Vulcanization* (Legare Street Press, 1853) 1-2.

18. J. W. .C. Hyatt, "Address of Acceptance," *Journal of Industrial and Engineering Chemistry* 6, no. 2 (1914): 158–161.

19. T. Bakelizer *A National Historic Chemical Landmark The "Old Faithful"* (National Museum of American History, 1993).

20. H. Staudinger, "A Source Book in Chemistry, 1900–1950," in *A Source Book in Chemistry* (Harvard University Press, 2013), 259–264.

21. V. Percec and Q. Xiao, "From Organic Chemistry to Chemical Biology via Macromolecules with Hermann Staudinger," *Giant* (2020): 100036.

22. C. W. Hume, "Synthetic Fiber," US2130948A (EIDP Inc, 1937) (accessed 2025-04-18), https://patents.google.com/patent/US2130948A/en.

23. W. H. Carothers, "An International Historic Chemical Landmark," *The Establishment of Modern Polymer Science* (2000): 1–8, https://www.acs.org/content/dam/acsorg/education/whatischemistry/landmarks/carotherspolymers/modern-polymer-science-wallace-h-carothers-historical-resource.pdf.

24. P. J. Flory, "Molecular Size Distribution in Linear Condensation Polymers 1," *Journal of the American Chemical Society* 58 (1936): 1877–1885.

25. W. H. Carothers, G. L. Dorough, and F. J. v. Natta, "Studies of Polymerization and Ring Formation. X. The Reversible Polymerization of Six-Membered Cyclic Esters," *Journal of the American Chemical Society* 54, no. 2 (1932): 761–772.

26. J. E. Mark, *Physical Properties of Polymers Handbook* (Springer, 2007), 1076.

27. L. M. Baker and K. Dunbar, "Experimental Design Heuristics for Scientific Discovery: The Use of "Baseline" and "Known Standard" Controls," *International Journal of Human-Computer Studies* 53 (2000): 335–349.

28. C. Schunn, "The Generality/Specificity of Expertise in Scientific Reasoning," *Cognitive Science* 23 (1999): 337–370.

29. K. Ziegler, H.-G., Gellert H. Martin, K. Nagel, and J. Schneider, "Metallorganische Verbindungen XIX: Reaktionen der Aluminium- Wasserstoff-Bindung mit Olefinen," *Justus Liebigs Annalen Der Chemie* 589 (1954): 91–121.

30. G. Natta, P. Pino, P. Corradini, F. Danusso, G. Moraglio, and E. Mantica, "Crystalline High Polymers of α-Olefins," *Journal of the American Chemical Society* 77 (1955): 1708–1710.

31. G. Natta, "Progress in Five Years of Research in Stereospecific Polymerization," *Stereoregular Polymers and Stereospecific Polymerizations* 2 (1967): 607–632.

32. G. Natta and P. Corradini, "The structure of crystalline 1,2-polybutadiene and of other "syndyotactic polymers"," *Journal of Polymer Science* 20, no. 95 (1956): 251–266.

33. K. Ziegler, R. Köster, H. Lehmkuhl, and K. Reinert, "Metallorganische Verbindungen XXX Neue Komplexverbindungen Der Aluminiumalkyle," *Justus Liebigs Ann Chem* 629 (1960): 33–49.

34. A. Vittoria, A. Meppelder, N. Friederichs, V. Busico, and R. D. Z.-N. C. Cipullo, "Demystifying Ziegler–Natta Catalysts: The Origin of Stereoselectivity," *ACS Catalysis* 7 (2017): 4509–4518.

35. F. Yates, "Sir Ronald Fisher and the Design of Experiments," *Biometrics* 20, no. 2 (1964): 307.

36. I. Elganidi, B. Elarbe, N. Ridzuan, and N. Abdullah, "Optimisation of Reaction Parameters for a Novel Polymeric Additives as Flow Improvers of Crude Oil Using Response Surface Methodology," *Journal of Petroleum Exploration and Production Technology* 12 (2022): 437–449.

37. J. Bicerano, *Prediction of Polymer Properties* (CRC Press, 2002).

38. V. Murphy, X. Bei, and T. R. Boussie, "High-Throughput Approaches for the Discovery and Optimization of New Olefin Polymerization Catalysts," *The Chemical Record* 2 (2002): 278–289.

39. H. Zhang, V. Marin, M. W. M. Fijten, and U. S. Schubert, "High-throughput experimentation in atom transfer radical polymerization: A general approach Toward a directed design and understanding of optimal catalytic systems," *Journal of Polymer Science Part A: Polymer Chemistry* 42 (2004): 1876–1885.

40. H. Zhang, R. Hoogenboom, M. A. R. Meier, and U. S. Schubert, "High-Throughput Experimentation in Polymer Chemistry," *Transactions of the Materials Research Society of Japan* 29 (2004): 319–324, http://www.schubert-group.com.

41. R. Hoogenboom and U. S. Schubert, "The fast and the curious: High-throughput experimentation in synthetic polymer chemistry," *Journal of Polymer Science Part A: Polymer Chemistry* 41 (2003): 2425–2434.

42. R. Upadhya, S. Kosuri, M. Tamasi, T. A. Meyer, S. Atta, M. A. Webb, and A. J. Gormley, "Automation and Data-Driven Design of Polymer Therapeutics," *Advanced Drug Delivery Reviews* 171 (2021): 1–28.

43. A. Weber, E. von Roedern, and H. U. Stilz, "SynCar: An Approach to Automated Synthesis," *Journal of Combinatorial Chemistry* 7, no. 2 (2005): 178–184.

44. J. D. Sutherland, N. P. Tu, T. A. Nemcek, et al., "An Automated Synthesis–Purification–Sample-Management Platform for the Accelerated Generation of Pharmaceutical Candidates," *SLAS Technology* 19, no. 2 (2014): 176–182.

45. M. Baumann, "Integrating Continuous Flow Synthesis With in-Line Analysis and Data Generation," *Organic and Biomolecular Chemistry* 16, no. 33 (2018): 5946–5954.

46. R. Rojas, N. K. Harris, K. Piotrowska, and J. Kohn, "Evaluation of automated synthesis for chain and step-growth polymerizations: Can robots replace the chemists?," *Journal of Polymer Science Part A: Polymer Chemistry* 47 (2009): 49–58.

47. D. C. Struble, B. G. Lamb, and B. Ma, "A Prospective on Machine Learning Challenges, Progress, and Potential in Polymer Science," *MRS Communications* 14, no. 5 (2024): 752–770.

48. S. Wu, H. Yamada, Y. Hayashi, M. Zamengo, and R. Yoshida, "Potentials and Challenges of Polymer Informatics: Exploiting Machine Learning for Polymer Design," *arXiv:2010.07683* (2020).

49. M. Ishii, T. Ito, H. Sado, and I. Kuwajima, "NIMS polymer database PoLyInfo (I): An overarching view of half a million data points," *Science and Technology of Advanced Materials: Methods* 4 (2024): 2354649.

50. Y. Qi, A. Reyes, S. E. Johnstone, M. J. Aryee, B. E. Bernstein, and B. Zhang, "Data-Driven Polymer Model for Mechanistic Exploration of Diploid Genome Organization," *Biophysical Journal* 119, no. 9 (2020): 1905–1916.

51. L. Tao, V. Varshney, and Y. Li, "Benchmarking Machine Learning Models for Polymer Informatics: An Example of Glass Transition Temperature," *Journal of Chemical Information and Modeling* 61, no. 11 (2021): 5395–5413.

52. Y. Zhang and X. Xu, "Machine Learning Glass Transition Temperature of Polyacrylamides Using Quantum Chemical Descriptors," *Polymer Chemistry* 12, no. 6 (2021): 843–851.

53. S. Venkatram, R. Batra, L. Chen, C. Kim, M. Shelton, and R. Ramprasad, "Predicting Crystallization Tendency of Polymers Using Multifidelity Information Fusion and Machine Learning," *The Journal of Physical Chemistry B* 124 (2020): 6046–6054.

54. J. A. Pugar, C. Gang, C. Huang, K. W. Haider, and N. R. Washburn, "Predicting Young's Modulus of Linear Polyurethane and Polyurethane–Polyurea Elastomers: Bridging Length Scales With Physicochemical Modeling and Machine Learning," *ACS Applied Materials and Interfaces* 14 (2022): 16568–16581.

55. J. Najeeb, S. S. A. Shah, M. H. Tahir, A. I. Hanafy, S. M. El-Bahy, and M. El-Bahy, "Machine Learning Assisted Designing of Polymers and Refractive Index Prediction: Easy and Fast Screening of Polymers From Chemical Space," *Materials Chemistry and Physics* 324 (2024): 129685.

56. B. Alemdag, G. Saygili, M. Franzreb, and G. Kabay, "Toward High-Performance Electrochemical Energy Storage Systems: A Case Study on Predicting Electrochemical Properties and Inverse Material Design of MXene-Based Electrode Materials With Automated Machine Learning (AutoML)," *Advanced Electronic Materials* (2025): 2400818.

57. Y. Zheng, P. Thakolkaran, A. K. Biswal, et al., "AI-Guided Inverse Design and Discovery of Recyclable Vitrimeric Polymers," *Advanced Science* 12, no. 6 (2025): 2411385.

58. E. Brochu, V. M. Cora, and N. A. de Freitas, "Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning," *arXiv:1012.2599* (2010).

59. X. Wang, Y. Huang, X. Xie, et al., "Bayesian-Optimization-Assisted Discovery of Stereoselective Aluminum Complexes for Ring-Opening Polymerization of Racemic Lactide," *Nature Communications* 14, no. 1, 4 (2023): 3647.

60. S. Takasuka, S. Ito, S. Oikawa, Y. Harashima, T. Takayama, A. Nag, and M. Fujii, "Methods Bayesian optimization of radical polymerization reactions in a flow synthesis system,", *Science and Technology of Advanced Materials*: (2024): 4, no. (1), https://doi.org/10.1080/27660400.2024.2425178.

61. K. Sattari, Y. Wu, Z. Chen, A. Mahjoubnia, C. Su, and J. Lin, "Physics-Constrained Multi-Objective Bayesian Optimization to Accelerate 3d Printing of Thermoplastics," *Additive Manufacturing* (2024): 86.

62. S. Takasuka, S. Ito, S. Oikawa, et al., "Bayesian Optimization of Radical Polymerization Reactions in a Flow Synthesis System," *Science and Technology of Advanced Materials: Methods* 4 (2024): 2425178.

63. A. Manoj, S. S. Miriyala, and K. Mitra, "Multi-Objective Bayesian Optimization for Computationally Expensive Reaction Network Models," in *2022 8th Indian Control Conference, ICC 2022 – Proceedings*, (Institute of Electrical and Electronics Engineers Inc., 2022), pp. 428–433.

64. A. Vriza, H. Chan, and J. Xu, "Self-Driving Laboratory for Polymer Electronics," *Chemistry of Materials* 35, no. 8 (2023): 3046–3056.

65. S. Sadeghi, R. B. Canty, N. Mukhin, J. Xu, F. Delgado-Licona, and M. Abolhasani, "Engineering a Sustainable Future: Harnessing Automation, Robotics, and Artificial Intelligence with Self-Driving Laboratories," *ACS Sustainable Chemistry and Engineering* 12 (2024): 12695–12707.

66. G. Tom, S. P. Schmid, S. G. Baird, et al., "Self-Driving Laboratories for Chemistry and Materials Science," *Chemical Reviews* 124, no. 16 (2024): 9633–9732.

67. M. Seifrid, R. Pollice, A. Aguilar-Granda, et al., "Autonomous Chemical Experiments: Challenges and Perspectives on Establishing a Self-Driving Lab," *Accounts of Chemical Research* 55, no. 17 (2022): 2454–2466.

68. S. Lo, S. G. Baird, J. Schrier, et al., "Review of low-cost self-driving laboratories in chemistry and materials science: The "frugal twin" concept," *Digital Discovery* 3, no. 5 (2024): 842–868.

69. J. Xu and P. Darancet, Polybot, https://www.anl.gov/cnm/polybot.

70. S. Langner, F. Häse, J. D. Perea, et al., "Beyond Ternary OPV: High-Throughput Experimentation and Self-Driving Laboratories Optimize Multicomponent Systems," *Advanced Materials* 32, no. 14 (2020): 1907801.

71. J. Li, J. Li, R. Liu, et al., "Autonomous Discovery of Optically Active Chiral Inorganic Perovskite Nanocrystals Through an Intelligent Cloud Lab," *Nature Communications* 11, no. 1 (2020): 2046.

72. S. T. Knox, K. E. Wu, N. Islam, et al., "Self-Driving Laboratory Platform for Many-Objective Self-Optimisation of Polymer Nanoparticle Synthesis With Cloud-Integrated Machine Learning and Orthogonal Online Analytics," *Polymer Chemistry* 16, no. 12 (2025): 1355–1364.

73. D. P. Tabor, L. M. Roch, S. K. Saikin, et al., "Accelerating the Discovery of Materials for Clean Energy in the Era of Smart Automation," *Nature Reviews Materials* 3, no. 5 (2018): 5–20.

74. T. B. Martin and D. J. Audus, "Emerging Trends in Machine Learning: A Polymer Perspective," *ACS Polymers Au* 3, no. 3 (2023): 239–258.

75. L. M. Roch, F. Häse, and C. Kreisbeck, "ChemOS: Orchestrating autonomous experimentation," *Science Robotics* 3, no. 19 (2018): aat5559.

76. S. Ryu, Y. Kwon, and W. Y. Kim, "Uncertainty Quantification of Molecular Property Prediction with Bayesian Neural Networks," *arXiv:1903.08375* (2019).

77. Y. Gal and Z. Ghahramani, "Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning," *PMLR* 48 (2016): 1050–1059.

78. Y. Zhang, "Bayesian Semi-Supervised Learning for Uncertainty-Calibrated Prediction of Molecular Properties and Active Learning," *Chemical Science* 10, no. 35 (2019): 8154–8163.

79. P. Bearucage and T. Martin Autonomous Formulation Lab, https://www.nist.gov/ncnr/ncnr-facility-upgrades/autonomous-formulationlab-afl.

80. Y. Lu, D. Yalcin, P. J. Pigram, and L. D. Blackman, "Interpretable Machine Learning Models for Phase Prediction in Polymerization-Induced Self-Assembly," *Journal of Chemical Information and Modeling* 63, no. 11 (2023): 3288–3306.

81. A. Menon, C. Gupta, K. M. Perkins, et al., "Elucidating Multi-Physics Interactions in Suspensions for the Design of Polymeric Dispersants: A Hierarchical Machine Learning Approach," *Molecular Systems Design and Engineering* 2, no. 3 (2017): 263–273.

82. D. J. Audus, A. McDannald, and B. DeCost, "Leveraging Theory for Enhanced Machine Learning," *ACS Macro Letters* 11, no. 9 (2022): 1117–1122.

83. W. Li, Z. Wang, M. Zhao, et al., "Symbolic Regression Accelerates the Discovery of Quantitative Relationships in Rubber Material Aging," *Journal of Materials Informatics* 5, no. 3 (2025): 29.

84. Z. Guo, S. Hu, Z.-K. Han, and R. Ouyang, "Improving Symbolic Regression for Predicting Materials Properties With Iterative Variable Selection," *Journal of Chemical Theory and Computation* 18, no. 8 (2022): 4945–4951.

85. J. G. Ethier, D. J. Audus, D. C. Ryan, and R. A. Vaia, "Integrating Theory with Machine Learning for Predicting Polymer Solution Phase Behavior," *Giant* 15 (2023): 100171.

86. E. R. Antoniuk, P. Li, B. Kailkhura, and A. M. Hiszpanski, "Representing Polymers as Periodic Graphs With Learned Descriptors for Accurate Polymer Property Predictions," *Journal of Chemical Information and Modeling* 62, no. 22 (2022): 5435–5445.

87. M. J. Uddin and J. Fan, "Interpretable Machine Learning Framework to Predict the Glass Transition Temperature of Polymers," *Polymers* 16, no. 8 (2024): 1049.

88. H. A. Gandhi and A. D. White, "Explaining Molecular Properties with Natural Language," *Theoretical and Computational Chemistry* (2022).

89. B. Tang, S. T. Kramer, M. Fang, Y. Qiu, Z. Wu, and D. Xu, "A Self-Attention Based Message Passing Neural Network for Predicting Molecular Lipophilicity and Aqueous Solubility," *Journal of Cheminformatics* 12 (2020): 1–9.

90. A. Y.-T. Wang, S. K. Kauwe, R. J. Murdock, and T. D. Sparks, "Compositionally Restricted Attention-Based Network for Materials Property Predictions," *Npj Computational Materials* 7, no. 1 (2021): 77.

91. D. Li, Y. Ru, and J. G. Liu, "Mining Graph Attention Networks-Based Important Substructures of Polymers for a Better Property Prediction," *Materials Today Communications* 38 (2024): 107577.

92. C. Kuenneth and R. P. B. Ramprasad *A Chemical Language Model to Enable Fully Machine-Driven Ultrafast Polymer Informatics* 14 (2023): 4099.

93. C. Xu, Y. Wang, and A. Barati Farimani, "TransPolymer: A Transformer-Based Language Model for Polymer Property Predictions," *Npj Computational Materials* 9, no. 1 (2023): 64.

94. P. K. Sarangi, N. Goel, A. K. Sahoo, and L. Rani, "Machine Learning in Polymer Science: Emerging Trends and Future Directions," *Macromolecular Symposia* 414, (2025): 2400101.

95. W. Ge, R. De Silva, Y. Fan, S. A. Sisson, and M. H. Stenzel, "Advanced Materials," *Machine Learning in Polymer Research* 37, (2025): 2413695.

96. L. Van Der Maaten and G. Hinton, "Visualizing Data Using T-SNE," *Journal of Machine Learning Research* 9 (2008): 2579–2605.

97. J. Cook, I. Sutskever, A. Mnih, and G. Hinton, "Visualizing Similarity Data With a Mixture of Maps," in *Artificial Intelligence and Statistics*, 2, (2007), 67–74.

98. L. McInnes, J. Healy, and J. Melville, "Umap: Uniform Manifold Approximation and Projection for Dimension Reduction," *arXiv:1802.03426* (2018).

99. R. Gurnani, C. Kuenneth, A. Toland, and R. Ramprasad, "Polymer Informatics at Scale With Multitask Graph Neural Networks," *Chemistry of Materials* 35, no. 4 (2023): 1560–1567.

100. A. Chandrasekaran, C. Kim, S. Venkatram, and R. Ramprasad, "A Deep Learning Solvent-Selection Paradigm Powered by a Massive Solvent/Nonsolvent Database for Polymers," *Macromolecules* 53, no. 12 (2020): 4764–4769.

101. R. Gómez-Bombarelli, J. N. Wei, D. Duvenaud, et al., "Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules," *ACS Central Science* 4, no. 2 (2018): 268–276.

102. U. Ehsan, P. Tambwekar, L. Chan, B. Harrison, and M. O. Riedl, "Automated Rationale Generation: A Technique for Explainable AI and Its Effects on Human Perceptions," In *Proceedings of the 24th International Conference on Intelligent User Interfaces*, (PMLR, 2019), pp. 263–274.

103. M. O. Riedl, "Human-centered artificial intelligence and machine learning," *Human Behavior and Emerging Technologies* 1, no. 1 (2019): 33–36.

104. S. Ryu, Y. Kwon, and W. Y. Kim, "A Bayesian Graph Convolutional Network for Reliable Prediction of Molecular Properties With Uncertainty Quantification," *Chemical Science* 10, no. 36 (2019): 8438–8446.

105. C.-I. Yang and Y.-P. Li, "Explainable Uncertainty Quantifications for Deep Learning-Based Molecular Property Prediction," *Journal of Cheminformatics* 15, no. 1 (2023): 13.

106. V. Duros, J. Grizou, A. Sharma, et al., "Intuition-Enabled Machine Learning Beats the Competition When Joint Human-Robot Teams Perform Inorganic Chemical Experiments," *Journal of Chemical Information and Modeling* 59, no. 6 (2019): 2664–2671.

107. C. Kim, A. Chandrasekaran, A. Jha, and R. Ramprasad, "Active-Learning and Materials Design: The Example of High Glass Transition Temperature Polymers," *MRS Communications* 9, no. 3 (2019): 860–866.

108. H. Abroshan, H. S. Kwak, Y. An, et al., "Active Learning Accelerates Design and Optimization of Hole-Transporting Materials for Organic Electronics," *Frontiers in Chemistry* 9 (2022): 800371.

109. A. G. Kusne, H. Yu, C. Wu, et al., "On-the-Fly Closed-Loop Materials Discovery via Bayesian Active Learning," *Nature Communications* 11, no. 1 (2020): 5966.

110. J. Xu and T. Luo, "Unlocking Enhanced Thermal Conductivity in Polymer Blends Through Active Learning," *Npj Computational Materials* 10, no. 1 (2024): 74.

111. A. M. Bran, S. Cox, O. Schilter, C. Baldassari, A. D. White, and P. Schwaller, "Augmenting Large Language Models With Chemistry Tools," *Nature Machine Intelligence* 6, no. 5 (2024): 525–535.

112. A. M. Bran, S. Cox, O. Schilter, C. Baldassari, A. D. White, and P. .C. Schwaller, "Augmenting Large-Language Models with Chemistry Tools," *arXiv:2304.05376* (2023).

113. P. Xu, X. Ji, M. Li, and W. Lu, "Small Data Machine Learning in Materials Science," *Npj Computational Materials* 9, no. 1 (2023): 42.

114. S. D. Rihm, J. Bai, A. Kondinski, S. Mosbach, J. Akroyd, and M. Kraft, "Transforming Research Laboratories With Connected Digital Twins," *Nexus* 1, no. 1 (2024): 100004.

115. E. F. Sheka, "Virtual Free-Radical Polymerization of Vinyl Monomers in View of Digital Twins," *Polymers* 15, no. 14 (2023): 2999.

116. M. V. Taylor, Z. Muwaffak, M. R. Penny, et al., "Optimising Digital Twin Laboratories with Conversational AIs: Enhancing Immersive Training and Simulation Through Virtual Reality," *Digital Discovery* 4 (2025): 1134–1141.

117. X. Liu and I. A. I. David, "Simulation By Digital Twins: Systematic Survey, Reference Framework, and Mapping to a Standardized Architecture," *arXiv:2506.06580* (2025).

118. C. Kuenneth, A. C. Rajan, H. Tran, L. Chen, C. Kim, and R. Ramprasad, "Polymer Informatics With Multi-Task Learning," *Patterns* 2, no. 4 (2021): 100238.

119. R. Ferreira da Silva, R. Moore II, B. Mintz, et al., *Shaping the Future of Self-Driving Autonomous Laboratories Workshop* (Oak Ridge National Laboratory (ORNL), 2024).

120. S. M. McDonald, E. K. Augustine, Q. Lanners, C. Rudin, L. Catherine Brinson, and M. L. Becker, "Applied Machine Learning as a Driver for Polymeric Biomaterials Design," *Nature Communications* 14, no. 1 (2023): 4838.

121. C. Ed-Driouch, F. Mars, P.-A. Gourraud, and C. Dumas, "Addressing the Challenges and Barriers to the Integration of Machine Learning Into Clinical Practice: An Innovative Method to Hybrid Human–Machine Intelligence," *Sensors* 22, no. 21 (2022): 8313.

122. M. Karuppusamy, R. Thirumalaisamy, S. Palanisamy, S. Nagamalai, E. E. S. Massoud, and N. Ayrilmis, "A Review of Machine Learning Applications in Polymer Composites: Advancements, Challenges, and Future Prospects," *Journal of Materials Chemistry A* 13 (2025): 16290–16308.

123. .G. Mengaldo, "Explainable AI and the Scientific Method: Interpretability-Guided Knowledge Discovery," *arXiv:2406.10557v5* (2025).

124. W. Ge, R. Silva, Y. Fan, S. Sisson, and M. Stenzel, "Machine Learning in Polymer Research," *Advanced Materials* 37 (2025): 2413695.

125. R. King and H. Zenil, "Artificial Intelligence in Scientific Discovery: Challenges and Opportunities," *Artificial Intelligence in Science* (2023): 181, https://doi.org/10.1787/a8d820bd-en.

126. H. Hysmith, E. Foadian, S. P. Padhy, et al., "The Future of Self-Driving Laboratories: From Human in the Loop Interactive AI to Gamification," *Digital Discovery* 3, no. 4 (2024): 621–636.

127. L. Rajulapati, S. Chinta, B. Shyamala, and R. Rengaswamy, "Integration of Machine Learning and First Principles Models," *AIChE Journal* 68, no. 6 (2022): 17715.

128. N. Sharma and Y. A. Liu, "A Hybrid Science-Guided Machine Learning Approach for Modeling and Optimizing Chemical Processes," *arXiv:2112.01475* (2021).

129. L. Gao, J. Lin, L. Wang, and L. Du, "Machine Learning-Assisted Design of Advanced Polymeric Materials," *Accounts of Materials Research* 5, no. 5 (2024): 571–584.

130. A. Jayaraman and B. Olsen, "Convergence of Artificial Intelligence," in *Machine Learning, Cheminformatics, and Polymer Science in Macromolecules*, 57, no. 16 (2024), 7685–7688.

131. T. Yue, L. Tao, V. Varshney, and Y. Li, "Benchmarking Study of Deep Generative Models for Inverse Polymer Design," *Digital Discovery* 4 (2025): 910–926.

132. H. Qiu and Z.-Y. Sun, "On-Demand Reverse Design of Polymers With PolyTAO," *Npj Computational Materials* 10, no. 1 (2024): 273.

133. G. Chen, Z. Shen, A. Iyer, et al., "Machine-Learning-Assisted De Novo Design of Organic Molecules and Polymers: Opportunities and Challenges," *Polymers* 12, no. 1 (2020): 163.

134. A. Khajeh, X. Lei, W. Ye, Z. Yang, D. Schweigert, and H.-K. Kwon, "A Self-Improvable Polymer Discovery Framework Based on Conditional Generative Model," *arXiv:2312.04013* (2023).

135. C. Laroiya, M. K. Bhatia, S. Madan, and C. Komalavalli, "IoT and Blockchain-Based Method for Device Identity Verification," In *International Conference on Innovative Computing and Communications: Proceedings of ICICC 2022*, 1, Springer, 2022, pp. 269–280.

136. M. Pilkington, *Blockchain Technology: Principles and Applications* (Edward Elgar Publishing, 2016), 225–253.

137. Y. Chen, Y. Lu, L. Bulysheva, and M. Y. Kataev, "Applications of Blockchain in Industry 4.0: A Review," *Information Systems Frontiers* 26, no. 5 (2024): 1715–1729.