REVIEW Open access

Autonomous and Semi-Autonomous Laboratories in Materials Science — Conceptual Foundations and Risks: A Review Study

Bruno Martins[1*], Lucas Pereira[1]

Abstract

Autonomous and semi-autonomous laboratories represent a transformative paradigm in materials science, integrating artificial intelligence, robotics, and high-throughput experimentation to accelerate discovery and optimization processes. This review examines the conceptual foundations of these systems, including closed-loop optimization, machine learning algorithms, and modular hardware architectures. We explore their applications in areas such as alloy development, perovskite synthesis, and nanoparticle engineering, highlighting successes that have reduced discovery timelines from years to days. However, we also critically assess associated risks, including data quality issues, algorithmic biases, ethical concerns in resource allocation, and potential safety hazards from unsupervised operations. Drawing on recent advances, we propose balanced implementation strategies that maximize innovation while mitigating risks. The review underscores the need for interdisciplinary collaboration to realize the full potential of these technologies in addressing global materials challenges.

Keywords Artificial intelligence, Materials discovery, Self-driving laboratories, Closed-loop optimization, Risk assessment, Ethical implications

*Correspondence:
Bruno Martins
bruno.martins@outlook.com
[1] Department of Materials Modeling and AI Systems, Faculty of Engineering, University of Minho, Braga, Portugal

Introduction

The field of materials science has traditionally advanced through iterative, labor-intensive experimentation, in which human expertise and intuition guide the design, synthesis, and characterization of new materials. While this paradigm has yielded numerous technological breakthroughs, it is inherently slow and resource-intensive. Conventional materials discovery often relies on trial-and-error approaches, incremental optimization, and heuristic-driven exploration of parameter spaces, resulting in development timelines that can span decades from initial concept to industrial deployment [1, 2]. These challenges are compounded by the increasing complexity of modern materials systems, which often involve high-dimensional compositional, structural, and processing variables that are difficult to exhaustively explore through manual experimentation alone.

At the same time, global demands for advanced materials are intensifying across multiple sectors, including renewable energy, micro- and nanoelectronics, aerospace, and biomedicine [3, 4]. Applications such as next-generation batteries, catalysts for carbon-neutral processes, flexible electronics, and personalized medical devices require materials with finely tuned and often competing properties. This growing urgency has exposed fundamental limitations in traditional research workflows, particularly in scalability, reproducibility, and the efficient use of experimental data. As a result, there has been

a paradigm shift toward data-driven, automation-enabled approaches that promise to accelerate materials discovery while reducing costs and human workload.

In this context, autonomous and semi-autonomous laboratories have emerged as a transformative concept in materials research. These systems integrate artificial intelligence (AI), machine learning (ML), robotic synthesis and characterization platforms, and automated data management to conduct experiments in a closed-loop manner with minimal human intervention [5, 6]. By continuously learning from experimental outcomes and adapting subsequent experimental decisions, such laboratories aim to maximize information gain and optimize target properties far more efficiently than conventional methods. This approach not only accelerates discovery but also enhances reproducibility by standardizing experimental execution and data acquisition.

Often referred to as “self-driving laboratories” (SDLs) or materials acceleration platforms (MAPs), autonomous laboratories represent a convergence of computational science, experimental materials science, and systems engineering [7, 8]. In a fully autonomous setting, AI-driven decision-making algorithms analyze data generated from prior experiments, generate hypotheses or candidate material formulations, and instruct robotic hardware to execute new experiments without direct human input [9, 10]. The resulting data are then automatically processed and fed back into the learning model, enabling iterative refinement until predefined optimization or discovery criteria are satisfied. Semi-autonomous laboratories, by contrast, retain human oversight at key decision points—such as defining objectives, validating model outputs, or intervening in unexpected experimental conditions—thereby combining machine efficiency with domain expertise and ethical judgment [11, 12].

The theoretical foundations of these systems are rooted in advanced optimization and learning frameworks, including Bayesian optimization (BO), active learning, reinforcement learning, and evolutionary or genetic algorithms [13, 14]. These methods are particularly well-suited for navigating the vast and often sparse parameter spaces characteristic of materials research, where experimental costs are high, and data are limited. By prioritizing experiments expected to yield the greatest improvement or uncertainty reduction, autonomous laboratories can achieve significant acceleration over brute-force or intuition-driven exploration.

Despite their promise, the deployment of autonomous and semi-autonomous laboratories also introduces a range of technical, ethical, and safety challenges. Issues such as model bias, data quality, algorithmic transparency, system robustness, and the potential for unsafe experimental actions raise important concerns that must be addressed before widespread adoption. Furthermore, the increasing delegation of experimental decision-making to AI systems prompts broader questions regarding accountability, trust, and the evolving role of human researchers in the scientific process.

This review aims to elucidate the conceptual foundations of autonomous and semi-autonomous laboratories in materials science while critically evaluating their associated risks. Specifically, the objectives are to: (1) delineate the core architectural components and operational workflows of these systems; (2) survey key applications and recent achievements across materials domains; (3) identify technical, ethical, and safety risks inherent to autonomous experimentation; and (4) propose practical mitigation strategies to ensure responsible and effective implementation. By synthesizing insights, this work provides a comprehensive and balanced perspective that highlights both the transformative potential and the critical limitations of autonomous laboratories, thereby guiding future research and development in this rapidly evolving field.

Main Text

Architectural components and operational workflows

The core architecture of autonomous and semi-autonomous laboratories is typically organized around three tightly coupled and interdependent modules: (i) decision-making software, (ii) experimental hardware, and (iii) data management and infrastructure systems [1, 5]. The effectiveness of an autonomous laboratory hinges not on the sophistication of any

single component, but on the seamless integration and synchronization of all three within a closed-loop experimental framework.

Decision-making is primarily driven by machine learning (ML) and optimization algorithms that efficiently explore high-dimensional, often sparsely sampled parameter spaces. Bayesian optimization (BO) is among the most widely adopted approaches, owing to its ability to balance exploration and exploitation while minimizing experimental cost [9, 13]. BO frameworks typically rely on surrogate models—such as Gaussian processes—to approximate objective functions and guide experiment selection under uncertainty. Specialized variants have been developed to address domain-specific challenges. For example, GRYFFIN extends BO to handle categorical and discrete variables common in chemical synthesis, while explicitly incorporating expert priors to constrain and guide the search process [9]. Beyond single-objective optimization, multi-objective BO has been successfully employed to address competing performance metrics. In microfluidic synthesis of gold nanoparticles, for instance, multi-objective optimization enables the simultaneous control of particle size, morphology, and colloidal stability, revealing trade-offs that would be difficult to uncover manually [15].

Complementing the decision-making layer, experimental hardware provides the physical means to execute AI-generated experimental plans. Typical components include robotic arms, automated liquid-handling and dispensing systems, microfluidic reactors, and a range of in situ and ex situ characterization tools, all integrated into modular and reconfigurable platforms [7, 16]. High-precision digital pipettes and automated micropipetting systems are particularly critical in SDLs, enabling reproducible and high-throughput synthesis across large experimental spaces [16, 17]. In semi-autonomous configurations, computer vision and sensor-based monitoring systems are increasingly important. These systems enable real-time observation of experimental processes—such as color change, precipitation, or phase separation—and allow adaptive control strategies to modify experimental parameters on the fly [18, 19]. A notable example is an automated solubility screening platform that uses computer vision to detect phase transitions, improving data accuracy while reducing human subjectivity in interpretation [19].

Data management constitutes the third foundational pillar and is essential for sustaining closed-loop autonomy. Autonomous laboratories generate large volumes of heterogeneous data, including synthesis parameters, sensor outputs, images, and characterization results, which must be stored, curated, and made accessible to learning algorithms. Centralized and standardized databases—such as the Open Reaction Database—facilitate not only internal learning cycles but also data sharing and reproducibility across research groups [20]. In parallel, big-data approaches have gained prominence in materials genomics, particularly in porous and framework materials, where ML models leverage extensive datasets to predict properties and inform experimental design [2]. These data streams are continuously fed back into the decision-making layer, enabling iterative model refinement.

Operationally, autonomous laboratories typically follow a structured plan–execute–analyze workflow: AI models propose candidate experiments; robotic systems perform synthesis and characterization; sensors and analytical tools collect data; and ML algorithms update surrogate models or policies based on the results [6, 8]. This iterative loop has demonstrated significant acceleration in discovery and optimization tasks. In perovskite crystallization, for example, convolutional neural networks integrated into closed-loop workflows have been used to optimize processing conditions by learning directly from imaging data, rapidly converging on high-quality crystal growth regimes [21]. The modular architecture and the plan–execute–analyze feedback loop that operationalizes closed-loop autonomy are summarized in **Figure 1**.

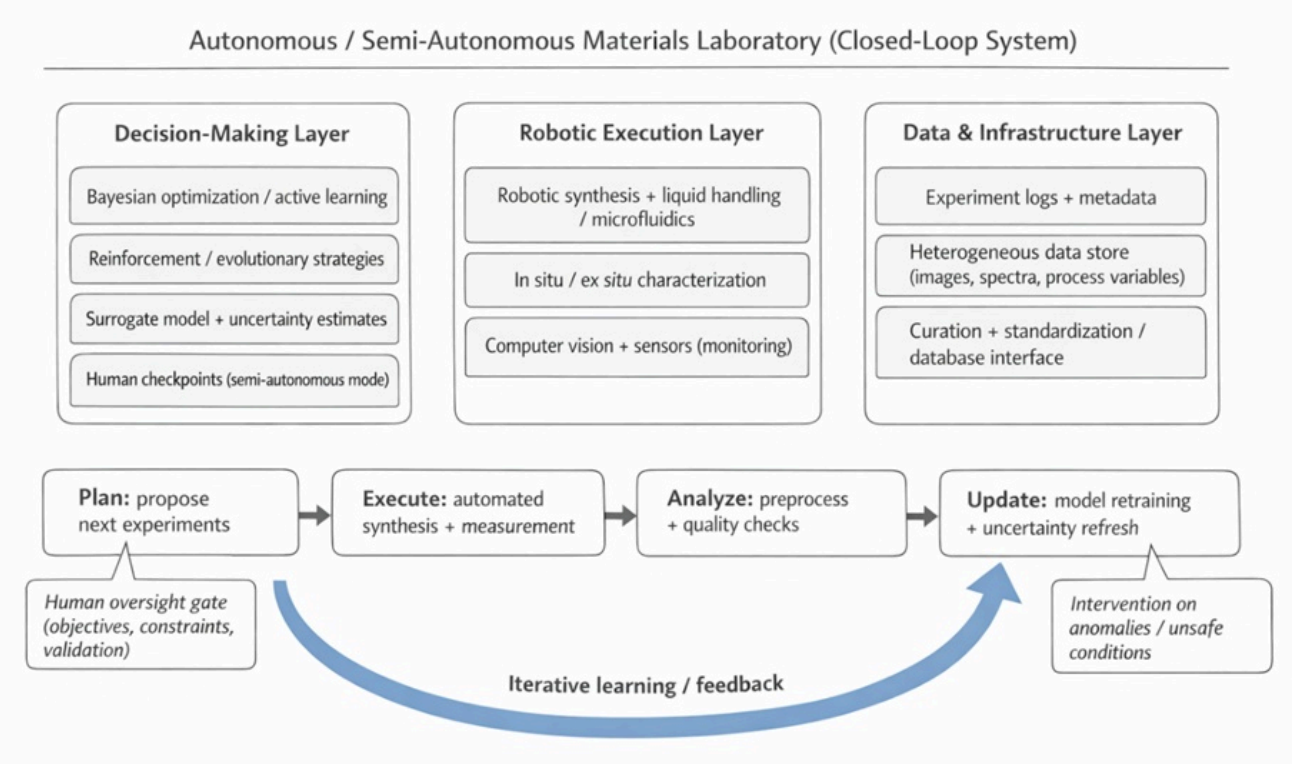

Figure 1. Modular architecture and plan–execute–analyze closed-loop workflow in autonomous materials laboratories

Despite these advances, substantial integration challenges remain. Interoperability between heterogeneous hardware and software components is often limited by proprietary interfaces and a lack of standardized communication protocols [10]. Additionally, experimental data are frequently noisy, incomplete, or biased, posing difficulties for ML models that assume well-behaved input distributions [12]. Addressing these challenges is critical for improving robustness and generalizability beyond tightly controlled laboratory settings.

Applications in materials discovery and optimization

Autonomous and semi-autonomous laboratories have demonstrated considerable effectiveness across a wide range of materials science applications, particularly in domains characterized by complex parameter spaces and costly experimentation. In alloy development, for example, high-throughput robotic synthesis and characterization platforms integrated with additive manufacturing workflows enable rapid evaluation of composition–microstructure–property relationships [22]. Autonomous scanning droplet cells further accelerate the discovery of corrosion-resistant alloys by automating electrochemical measurements across large compositional libraries, significantly reducing manual labor and experimental variability [23].

Perovskite materials for photovoltaic applications are among the most mature and extensively studied applications of SDLs. High-throughput synthesis platforms combined with automated photoluminescence and optical characterization have been used to rapidly screen compositional and processing variables, identifying optimal formulations for efficiency and stability [24]. More advanced systems integrate robotics-assisted crystallization with neural network-driven closed-loop optimization, enabling continuous refinement of growth conditions and structural properties [17, 21]. These approaches have been instrumental in addressing long-standing challenges related to perovskite degradation and reproducibility.

Nanoparticle engineering further highlights the strengths of autonomous experimentation, particularly when coupled with microfluidic platforms. Genetic algorithms implemented in liquid-handling systems have been used to optimize emission

properties of nanoparticles by iteratively refining synthesis conditions [25, 26]. Similarly, multi-objective BO frameworks deployed in microfluidic reactors enable precise control over gold nanoparticle characteristics for application-specific requirements [15]. Sonochemical SDLs have also demonstrated rapid synthesis and optimization of CdSe nanocrystals, illustrating the versatility of autonomous platforms across different synthesis modalities [5].

In the energy storage domain, autonomous laboratories have been applied to standardize and accelerate experimental workflows for non-aqueous redox flow batteries, improving reproducibility and comparability of performance metrics [27]. Robot-assisted platforms combined with ML have also been used to optimize organic photovoltaic (OPV) materials, revealing non-intuitive structure–property relationships that inform rational molecular design [4].

Beyond conventional materials systems, autonomous experimentation has been extended to unconventional domains such as concrete formulation. Here, adaptive experimentation guided by modern optimization algorithms has been employed to design mixtures that simultaneously maximize mechanical strength and minimize environmental impact, demonstrating the broader applicability of SDL concepts beyond nanoscale and electronic materials [28].

Collectively, these examples illustrate the potential of autonomous laboratories to dramatically compress discovery and optimization timelines, with reported acceleration factors ranging from one to two orders of magnitude compared to traditional workflows [1, 29]. However, these successes are often confined to well-defined parameter spaces with clearly measurable objectives. Scaling autonomous approaches to more open-ended discovery problems, poorly characterized systems, or multiscale phenomena remains an open challenge [14].

Integration of artificial intelligence and machine learning

Artificial intelligence (AI) and machine learning (ML) function as the intellectual core of autonomous and semi-autonomous laboratories, enabling predictive modeling, adaptive decision-making, and experimental automation across the materials discovery pipeline [1, 7]. Rather than serving merely as post hoc data analysis tools, modern ML models actively influence experimental design by forecasting outcomes, quantifying uncertainty, and recommending subsequent experiments in real time. This paradigm shift—from data-driven interpretation to data-driven experimentation—distinguishes autonomous laboratories from earlier high-throughput approaches.

Deep learning techniques, which have demonstrated transformative success in domains such as antibiotic and drug discovery, increasingly inspire analogous applications in materials science [30]. Neural networks trained on molecular structures, spectroscopy, or imaging data can predict physicochemical properties, stability, and performance metrics with high accuracy, thereby narrowing the experimental search space. Convolutional and graph-based neural networks are particularly effective in learning structure–property relationships, enabling rapid screening of candidate materials before physical synthesis.

Recent advances also highlight the emerging role of large language models (LLMs) in laboratory automation. By interpreting natural language instructions, parsing experimental protocols, and interfacing with robotic control systems, LLMs offer a flexible layer of abstraction between human intent and machine execution [7]. In chemistry-oriented autonomous platforms, LLMs have been explored as tools for experiment planning, error diagnosis, and translation of textual knowledge into machine-readable workflows, potentially lowering the barrier to interaction with complex autonomous systems.

Among optimization strategies, Bayesian optimization (BO) remains the dominant framework for reaction and process optimization in autonomous laboratories. Its ability to operate effectively under limited data regimes and explicitly manage uncertainty makes it well-suited for expensive or time-consuming experiments [13]. BO has been successfully applied to computer-proposed multistep synthetic routes, which are subsequently executed on robotic flow chemistry platforms, demonstrating end-to-end automation from hypothesis generation to experimental realization [13]. In parallel, evolutionary and genetic algorithms have proven effective in domains such as nanoparticle synthesis, where parameter sets—such as

precursor concentrations, temperatures, and reaction times—are iteratively evolved toward desired optical or structural properties [26].

Active learning strategies further enhance autonomy by prioritizing experiments that maximize information gain rather than solely optimizing performance metrics. Multi-objective active learning platforms, including web-based interfaces for reaction tuning, allow simultaneous optimization of yield, selectivity, cost, or environmental impact [31]. In data-rich domains such as porous and framework materials, ML-driven “materials genomics” approaches extract latent design rules from large datasets, enabling predictive screening and inverse design within autonomous discovery loops [2].

Despite these successes, integrating AI and ML introduces significant algorithmic risks. Overfitting to limited or biased training datasets can cause models to converge prematurely on suboptimal regions of the search space, undermining exploratory discovery [9, 11]. Dataset biases—arising from historical experimental preferences or incomplete reporting—may be inadvertently amplified, leading to systematic exclusion of unconventional or novel materials classes [2, 20]. These limitations underscore the need for careful model validation, uncertainty quantification, and continual diversification of the dataset.

Safety and technical risks

While autonomous laboratories promise efficiency and acceleration, they also introduce nontrivial technical and safety risks, particularly when experiments are conducted with minimal human supervision. Fully autonomous operation increases the likelihood that equipment malfunctions, software errors, or unanticipated chemical behaviors may go undetected until damage or hazardous events occur [6, 10]. This risk is especially pronounced in laboratories handling volatile, reactive, or toxic substances, where improper reagent mixing, pressure buildup, or thermal runaway could result in fires or explosions [10]. In inorganic synthesis environments, for example, small deviations in stoichiometry or reaction conditions can have disproportionate safety consequences.

Data integrity represents another critical vulnerability. Autonomous decision-making relies heavily on sensor outputs, imaging systems, and automated characterization tools, all of which are susceptible to noise, drift, calibration errors, or data loss [18, 19]. Incomplete or corrupted datasets can mislead ML models, causing erroneous experimental decisions that propagate through closed-loop workflows. When combined with algorithmic biases, such data quality issues may systematically skew exploration and hinder the identification of genuinely novel materials [2, 9].

The increasing connectivity of autonomous laboratories further exposes them to cybersecurity risks. Networked control systems and cloud-based data storage create potential entry points for unauthorized access, intellectual property theft, or malicious interference with experimental operations [1, 8]. From a practical standpoint, scalability also remains a major challenge. High capital investment, specialized maintenance requirements, and limited robustness of robotic hardware constrain the deployment of SDLs beyond a small number of well-funded research institutions [7, 16].

Mitigation strategies emphasize layered safety architectures, including redundant hardware interlocks, real-time anomaly detection, and fail-safe shutdown mechanisms. Semi-autonomous operation modes that retain human-in-the-loop oversight at critical decision points can substantially reduce risk without negating the efficiency benefits of automation [12, 14]. Additionally, robust validation datasets, routine sensor calibration, and systematic stress-testing of ML models are essential for ensuring reliable and safe operation.

Ethical and societal risks

Beyond technical considerations, autonomous laboratories raise important ethical and societal questions that warrant scrutiny. One prominent concern is equitable access. The substantial financial and infrastructural requirements of SDLs risk exacerbating disparities between well-resourced institutions and those with limited funding, potentially concentrating discovery power within a small subset of organizations or regions [1, 29]. Such imbalances could shape research agendas in ways that favor commercial interests over broader societal needs.

Resource allocation decisions embedded within AI systems introduce further ethical complexity. Optimization objectives may implicitly prioritize metrics such as cost efficiency or performance, potentially sidelining considerations of sustainability, environmental impact, or social benefit unless explicitly encoded [4, 28]. Determining who defines these objectives—and whose values they reflect—remains an open question in the governance of autonomous research platforms.

Concerns regarding workforce displacement also persist. Increased automation may reduce demand for traditional laboratory technicians and experimental scientists, even as it creates new roles focused on AI supervision, data stewardship, and systems integration [5, 8]. Managing this transition requires proactive investment in interdisciplinary education and reskilling to ensure that human expertise remains central to scientific progress.

Intellectual property (IP) ownership presents another unresolved issue, particularly when discoveries emerge from AI-driven workflows with minimal direct human intervention. Questions regarding authorship, patent eligibility, and data ownership become increasingly complex in autonomous settings [20]. Additionally, the rapid acceleration of materials discovery carries the risk of unintended societal consequences, including the faster development of dual-use materials with potential military or surveillance applications [10, 30].

The major risk domains introduced by autonomous operation, along with the corresponding layered mitigation strategies, are shown in **Figure 2**.

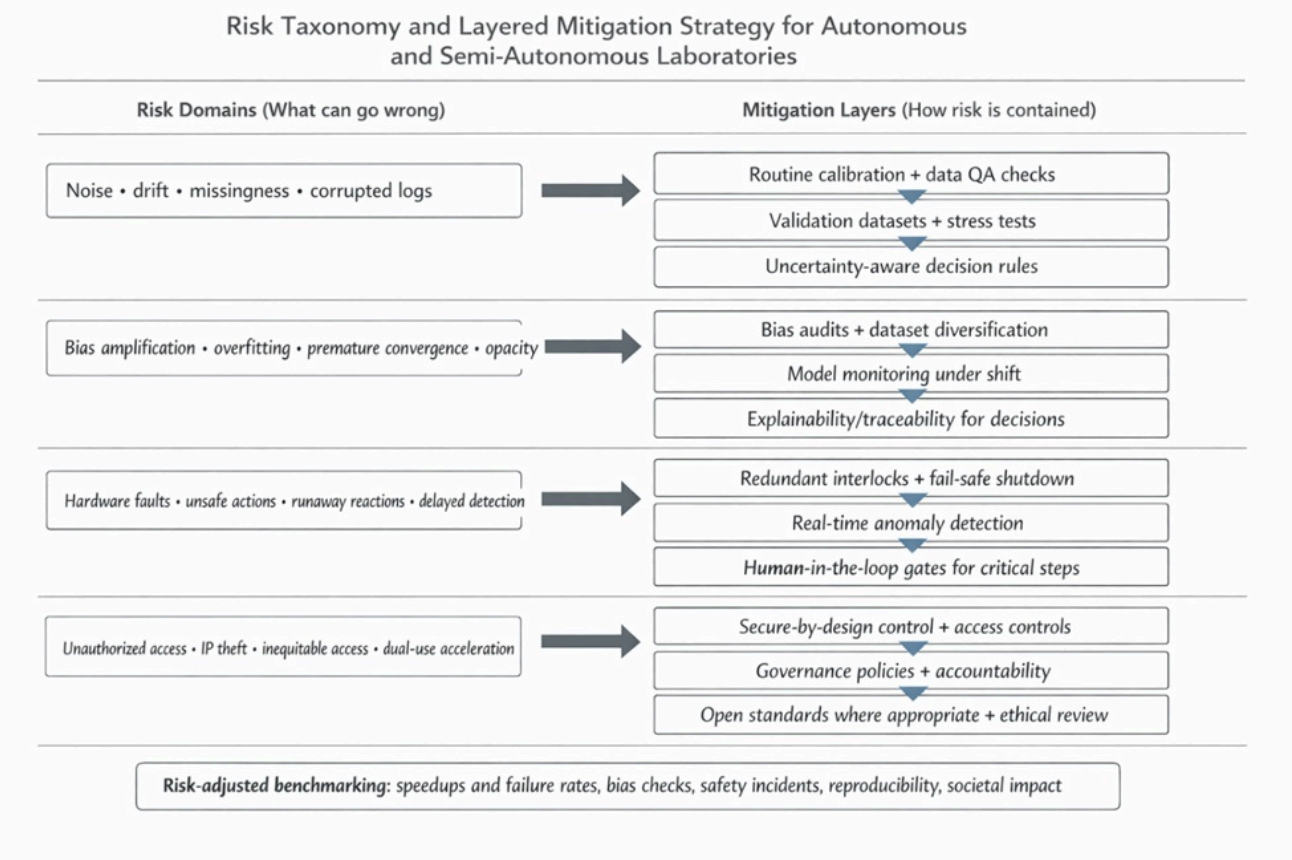

Figure 2. Risk taxonomy and layered mitigation strategy for autonomous and semi-autonomous laboratories

To address these challenges, proposed strategies include developing ethical frameworks and governance guidelines tailored to autonomous experimentation, increasing the adoption of open-source software and shared datasets to democratize access, and interdisciplinary training that integrates materials science, computer science, ethics, and policy

[1, 2]. Transparency and explainability in AI decision-making processes are particularly critical for building trust and ensuring accountability as autonomous laboratories become more prevalent [7, 11].

Results and Discussion

The conceptual foundations of autonomous and semi-autonomous laboratories in materials science, as delineated in the preceding sections, rest upon a synergistic integration of AI-driven decision-making, robotic execution, and iterative data feedback loops. This paradigm shifts materials discovery from serendipitous human-led efforts to systematic, data-informed processes capable of exploring chemical and structural spaces at unprecedented scales [1, 29]. Bayesian optimization has emerged as a cornerstone algorithm, enabling efficient sampling of both categorical and continuous variables while incorporating domain expertise, as demonstrated in GRYFFIN's application to chemical synthesis challenges [9]. Extensions to multi-objective frameworks further allow simultaneous optimization of conflicting properties, such as yield, purity, and stability in nanoparticle systems [15, 31]. To clarify design choices, **Table 1** contrasts autonomous versus semi-autonomous laboratory modes in terms of workflow control points, typical strengths, and dominant risk-control mechanisms.

Table 1. Comparative characteristics of autonomous vs semi-autonomous materials laboratories

Dimension	Fully autonomous laboratory	Semi-autonomous laboratory (human-in-the-loop)
Decision authority	AI system proposes and executes experiment sequences end-to-end	AI proposes; humans approve/override at predefined gates (objectives, constraints, anomalies)
Typical optimization logic	Strong reliance on BO/active learning/RL to maximize information gain or objectives	Same algorithmic core, but constrained by human judgments on feasibility, safety, and values
Primary strength	Maximum throughput and iteration speed in well-defined objective spaces	Higher robustness in uncertain, open-ended, or safety-sensitive regimes
Main vulnerability	Undetected errors can propagate (sensor drift → model update → unsafe or unproductive actions)	Slower throughput; risk of inconsistent human interventions if oversight is not standardized
Data integrity exposure	High sensitivity to noise, drift, missingness, and automated mislabeling	Reduced exposure when humans validate critical measurements or interpret ambiguous signals
Safety profile	Requires stringent interlocks, anomaly detection, and fail-safe shutdown due to minimal supervision	Lower operational hazard when humans remain responsible for critical decisions and escalation
Bias and exploration risk	Greater chance of bias amplification and premature convergence if training data are skewed	Better opportunity to correct bias via oversight, constraints, and deliberate exploration policies
Governance and accountability	Accountability can be ambiguous when decisions are machine-generated end-to-end	Clearer accountability distribution (humans retain responsibility at gates and policy layers)
Best-fit use case	Narrow, measurable objectives; stable instrumentation; controlled parameter spaces	Complex chemistries, reactive hazards, shifting conditions, and value-laden objectives (e.g., sustainability)

In practice, these foundations manifest through closed-loop workflows that have yielded tangible accelerations in discovery. For carbon nanotube growth, Bayesian optimization closed the loop to maximize growth rates, achieving efficient parameter tuning in an otherwise intractable search space [32]. Parallel efforts in perovskite materials highlight the role of computer vision and neural networks in real-time monitoring and adaptive control, where convolutional neural networks refined crystallization conditions to enhance optoelectronic performance [14, 21]. Similarly, robotics-accelerated platforms for perovskite synthesis and photoluminescence analysis have automated high-throughput screening, enabling rapid identification of promising compositions [17, 24].

The integration of large language models represents a frontier in conceptual evolution, facilitating natural language interfaces for robotics in chemistry and materials experiments [7]. This enables semi-autonomous systems to interpret complex instructions, bridging the gap between human intent and machine execution. In alloy development, high-throughput robotic characterization and scanning droplet cells exemplify semi-autonomous risk mitigation, where human oversight complements automation to validate corrosion resistance and mechanical properties [22, 23]. Concrete formulation optimization further illustrates adaptability, employing modern algorithms to balance performance metrics in sustainable materials design [28].

Despite these advancements, the risks associated with autonomous and semi-autonomous laboratories warrant rigorous scrutiny. Technical risks primarily stem from data quality and algorithmic reliability. Noisy sensor data or incomplete datasets can propagate errors in ML models, leading to misguided experiment proposals [18, 19]. For instance, in solubility screening platforms, computer vision inaccuracies could misclassify phases, undermining the entire closed-loop process [19]. Overfitting in deep learning models trained on biased or limited training sets risks suboptimal exploration, as seen in early applications to porous materials genomics, where dataset diversity was critical for predictive accuracy [2, 30].

Safety hazards pose another layer of concern, particularly in unsupervised operations involving reactive chemicals or high-energy processes. In the synthesis of autonomous inorganic materials, uncontrolled parameters may trigger exothermic reactions or pressure buildup [10]. Semi-autonomous designs offer partial mitigation through human-in-the-loop interventions, but full autonomy demands advanced fail-safes, such as redundant monitoring systems and emergency shutdown protocols [6, 12]. Cybersecurity vulnerabilities in interconnected robotic platforms further amplify risks, potentially allowing remote interference that could compromise experiments or expose proprietary data [1, 8].

Ethical and societal risks are equally profound and underexplored. The high capital investment required for SDLs risks exacerbating inequalities in access to materials research, favoring resource-rich institutions and corporations [8, 29]. Algorithmic biases may inadvertently prioritize commercially viable materials over those addressing pressing societal needs, such as eco-friendly alternatives in energy storage or construction [4, 27, 28]. Job displacement remains a contentious issue; while automation handles repetitive tasks, it may devalue traditional lab skills unless reskilling programs are implemented [5, 11]. Intellectual property challenges arise when AI autonomously generates novel compositions, blurring lines of inventorship and raising questions about patentability [20].

Moreover, the potential for accelerated development of dual-use technologies—materials with both civilian and military applications—necessitates ethical oversight [10, 30]. Transparency in AI decision-making is paramount; black-box models hinder interpretability and accountability [7, 13]. Open reaction databases promote data sharing but introduce risks of misuse if not governed by ethical standards [20]. To address these, interdisciplinary frameworks incorporating ethicists, policymakers, and diverse stakeholders are essential [2, 14].

Performance metrics for SDLs must evolve to encompass not only efficiency gains but also risk-adjusted indicators, such as failure rates, bias audits, and societal impact assessments [14]. Comparative analyses reveal that while autonomous labs achieve 10-100x speedups in targeted domains such as OPV materials and CdSe nanocrystals [4, 5], broader scalability is hampered by hardware modularity issues and the need for standardized protocols [16, 27]. Semi-

autonomous hybrids often outperform fully autonomous systems in complex, uncertain environments by leveraging human intuition for edge cases [11, 12].

The conceptual foundations also intersect with broader trends in big-data science and machine learning, where materials genomics approaches predict properties ahead of synthesis [2]. Genetic algorithms on microfluidic platforms for nanoparticle emission optimization demonstrate the efficacy of evolutionary strategies in navigating multimodal landscapes [25, 26]. However, these successes are domain-specific; extending to amorphous or multicomponent systems remains challenging due to the combinatorial explosion [3, 22].

Risk mitigation strategies should prioritize hybrid human-AI collaboration, robust validation pipelines, and open-source hardware, such as digital pipettes, to democratize access [7, 16]. Regulatory frameworks for SDL safety, analogous to those in autonomous vehicles, could standardize testing and deployment [6, 29]. Future conceptual developments may incorporate uncertainty quantification into BO to better handle experimental variability [9, 13].

In synthesizing these elements, the review reveals a field poised for exponential growth yet fraught with multifaceted risks that demand proactive governance. The balance between autonomy's efficiency and semi-autonomous safeguards will define sustainable progress [1, 8].

Conclusion

Autonomous and semi-autonomous laboratories embody a paradigm shift in materials science, grounded in AI, robotics, and closed-loop experimentation, and promise to reshape discovery. This review has outlined their conceptual foundations—from Bayesian and multi-objective optimization to integrated hardware-software architectures—and critically examined risks spanning technical failures, safety concerns, ethical dilemmas, and societal impacts.

Key achievements, including accelerated alloy characterization, perovskite optimization, and nanoparticle tailoring, underscore the transformative potential of this approach. Yet, risks such as data biases, operational hazards, and inequities necessitate vigilant management.

Future directions should focus on: (1) developing standardized, interoperable platforms with built-in risk assessments; (2) advancing interpretable AI for enhanced transparency; (3) fostering global collaborations to address access disparities; (4) integrating ethical AI principles into lab design; and (5) exploring hybrid models that scale to complex, real-world materials challenges. By addressing risks head-on, SDLs can fulfill their promise to tackle urgent material needs for sustainability, health, and technology.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 15 Aug 2025 Revised: 27 Aug 2025 Accepted: 08 Nov 2025
Published online: 18 January 2026

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Tom G, Schmid SP, Baird SG, Cao Y, Darvish K, Hao H, et al. Self-driving laboratories for chemistry and materials science. Chem Rev. 2024;124(16):9633-732.
https://doi.org/10.1021/acs.chemrev.4c00055.

Jablonka KM, Ongari D, Moosavi SM, Smit B. Big-data science in porous materials: Materials genomics and machine learning. Chem Rev. 2020;120(16):8066-129.
https://doi.org/10.1021/acs.chemrev.0c00004.

Vecchio KS, Dippo OF, Kaufmann KR, Liu X. High-throughput rapid experimental alloy development (HT-READ). Acta Mater. 2021;221:117352.
https://doi.org/10.1016/j.actamat.2021.117352.

Du X, Lüer L, Heumueller T, Wagner J, Berger C, Osterrieder T, et al. Elucidating the full potential of OPV materials utilizing a high-throughput robot-based platform and machine learning. Joule. 2021;5:495-506.
https://doi.org/10.1016/j.joule.2020.12.013.

Politi M, Baum F, Vaddi K, Antonio E, Vasquez JP, Bishop B, et al. A high-throughput workflow for the synthesis of CdSe nanocrystals using a sonochemical materials acceleration platform. Digit Discov. 2023;2:1042-57.
https://doi.org/10.1039/D3DD00033H.

Bayley O, Savino E, Slattery A, Noël T. Autonomous chemistry: Navigating self-driving labs in chemical and material sciences. Matter. 2024;7(7):2382-98.
https://doi.org/10.1016/j.matt.2024.05.022.

Yoshikawa N, Skreta M, Darvish K, Arellano-Rubach S, Ji Z, Bjørn Kristensen L, et al. Large language models for chemistry robotics. Auton Robots. 2023;47:1057.
https://doi.org/10.1007/s10514-023-10136-2.

Hung L, Yager JA, Potocek D, Baiocchi D, Kwon H-K, Sun S, et al. Autonomous laboratories for accelerated materials discovery: A community survey and practical insights. Digit Discov. 2024;3:1273-9.
https://doi.org/10.1039/D4DD00059E.

Häse F, Aldeghi M, Hickman RJ, Roch LM, Aspuru-Guzik A. GRYFFIN: An algorithm for Bayesian optimization of categorical variables informed by expert knowledge. Appl Phys Rev. 2021;8:031406.

https://doi.org/10.1063/5.0048164.

Szymanski NJ, Rendy B, Fei Y, Kumar RE, He T, Milsted D, et al. An autonomous laboratory for the accelerated synthesis of inorganic materials. Nature. 2023;624(7990):86-91.
https://doi.org/10.1038/s41586-023-06734-w.

Taylor CJ, Pomberger A, Felton KC, Grainger R, Barecka M, Chamberlain TW, et al. A brief introduction to chemical reaction optimization. Chem Rev. 2023;123(6):3089-126.
https://doi.org/10.1021/acs.chemrev.2c00798.

Miracle DB, Thoma DJ. Autonomous research and development of structural materials: An introduction and vision. Curr Opin Solid State Mater Sci. 2024;33:101188.
https://doi.org/10.1016/j.cossms.2024.101188.

Nambiar AMK, Breen CP, Hart T, Kulesza T, Jamison TF, Jensen KF. Bayesian optimization of computer-proposed multistep synthetic routes on an automated robotic flow platform. ACS Cent Sci. 2022;8:825-36.
https://doi.org/10.1021/acscentsci.2c00207.

Volk AA, Abolhasani M. Performance metrics to unleash the power of self-driving labs in chemistry and materials science. Nat Commun. 2024;15:1378.
https://doi.org/10.1038/s41467-024-45569-5.

Tao A, Epps D, Abolhasani M, et al. Multi-objective Bayesian optimization for gold nanoparticle synthesis in microfluidics. ACS Nano. 2022;16(3):4321-30.
https://doi.org/10.1021/acsnano.1c09876.

Yoshikawa N, Darvish K, Vakili MG, Garg A, Aspuru-Guzik A. Digital pipette: Open hardware for liquid transfer in self-driving laboratories. Digit Discov. 2023;2:1745-51.
https://doi.org/10.1039/D3DD00115F.

Li Y, Kirman JR, Higgins MA, et al. Robotics-accelerated crystallization of perovskites using micropipetting systems. Chem Mater. 2021;33(14):5678-86.
https://doi.org/10.1021/acs.chemmater.1c01234.

El-khawaldeh R, Guy MA, Bork F, Taherimakhsousi N, Jones KN, Hawkins J, et al. Keeping an “eye” on the experiment: Computer vision for real-time monitoring and control. Chem Sci. 2024;15:1271.
https://doi.org/10.1039/D3SC05491H.

Shiri P, Lai V, Zepel T, Griffin D, Reifman J, Clark S, et al. Automated solubility screening platform using computer vision. iScience. 2021;24(3):102176.
https://doi.org/10.1016/j.isci.2021.102176.

Kearnes SM, Maser MR, Wleklinski M, Kast A, Doyle AG, Dreher SD, et al. The open reaction database. J Am Chem Soc. 2021;143:18820-6.
https://doi.org/10.1021/jacs.1c09820.

Kirman JR, Li Y, Higgins MA, et al. Closed-loop optimization of perovskite crystallization using convolutional neural networks. Nat Commun. 2020;11:4285.
https://doi.org/10.1038/s41467-020-18045-6.

Vecchio KJ, Sneddon J, Zhang Y, et al. High-throughput robotic characterization of additively manufactured alloys. Matter. 2021;4(4):1157-74.
https://doi.org/10.1016/j.matt.2021.01.019.

DeCost B, Sneddon J, Zhang Y, et al. Autonomous scanning droplet cell for accelerated discovery of corrosion-resistant alloys. ACS Cent Sci. 2022;8(5):612-23.
https://doi.org/10.1021/acscentsci.2c00123.

Higgins MA, Kirman JR, Sargent EH, et al. High-throughput synthesis and photoluminescence analysis of perovskites in a self-driving laboratory. J Am Chem Soc. 2020;142(48):20358-65.
https://doi.org/10.1021/jacs.0c09234.

Krishnadasan S, de Mello AJ, de Mello JC, et al. Microfluidics-based self-driving laboratory for nanoparticle emission optimization. Lab Chip. 2020;20(15):2789-98.
https://doi.org/10.1039/D0LC00456A.

Salley DJ, Krishnadasan S, de Mello AJ, et al. Seed-mediated synthesis of gold nanoparticles using genetic algorithms in a liquid-handling platform. Anal Chem. 2021;93(10):4567-75.
https://doi.org/10.1021/acs.analchem.0c04567.

Li M, Odom SA, Pancoast AR, Robertson LA, Vaid TP, Agarwal G, et al. Experimental protocols for studying organic non-aqueous redox flow batteries. ACS Energy Lett. 2021;6:3932-43.
https://doi.org/10.1021/acsenergylett.1c01675.

Völker A, Müller T, Oberacker R, et al. Adaptive experimentation for concrete formulation optimization using modern algorithms. Cem Concr Res. 2023;167:107189.
https://doi.org/10.1016/j.cemconres.2023.107189.

Abolhasani M, Kumacheva E. The rise of self-driving labs in chemical and materials sciences. Nat Synth. 2023;2:483-92.
https://doi.org/10.1038/s44160-023-00284-w.

Stokes JM, Yang K, Swanson K, Jin W, Cubillos-Ruiz A, Donghia NM, et al. A deep learning approach to antibiotic discovery. Cell. 2020;180:688-702.e13.
https://doi.org/10.1016/j.cell.2020.01.021.

Torres JAG, Lau SH, Anchuri P, Stevens JM, Tabora JE, Li J, et al. A multi-objective active learning platform and web app for reaction optimization. J Am Chem Soc. 2022;144:19999-20007.
https://doi.org/10.1021/jacs.2c08592.

Chang J, Nikolaev P, Carpena-Nunez J, Rao R, Decker K, Islam AE, et al. Efficient closed-loop maximization of carbon nanotube growth rate using Bayesian optimization. Sci Rep. 2020;10:9040.
https://doi.org/10.1038/s41598-020-64397-3.