Global Journal of Chemistry
Vol. 1 No.1 2026 Page 29-39
E-ISSN: XXXX-XXXX, DOI: 10.XXXXX/XXXX

Strategies for Integrating Green Chemistry Principles into Data-Driven Synthesis of Functional Materials

Naufal Kurniawan[1]*, Meutia Firmansyah[2], Arif Setiawan[3]
[1,2] Department of Statistics, Faculty of Mathematics and Natural Sciences, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia; [3] Department of Information Systems, Faculty of Computer Science, Universitas Indonesia, Depok, Indonesia

Article Info

Received : 16 January 2026
Revised : 24 March 2026
Accepted : 19 April 2026

Keyword:

Green chemistry;
Data-driven synthesis;
Functional materials;
Materials informatics;
Sustainable laboratory practice.

ABSTRACT

This study aims to explore strategies for integrating *green chemistry* principles into the data-driven synthesis of functional materials, focusing on how researchers understand, negotiate, and apply sustainability considerations in laboratory practice. The study used a qualitative exploratory case study design involving researchers, lecturers, postgraduate students, laboratory analysts, and research managers in functional materials laboratories. Data were collected through semi-structured interviews, non-participant observation, and documentation of laboratory procedures, synthesis records, safety documents, and research reports. The data were analyzed using reflective thematic analysis to identify patterns of meaning related to green chemistry integration, data use, synthesis optimization, and laboratory decision-making. The findings reveal six main themes: the practical embedding of green chemistry in laboratory routines, the role of data-driven methods in reducing experimental uncertainty, the tension between material performance and environmental responsibility, the limited integration of sustainability indicators into predictive models, the influence of laboratory culture and institutional support, and the importance of documenting failed experiments for greener synthesis pathways. This study contributes to the literature by positioning data-driven synthesis as a socio-technical process that requires technical, institutional, and educational integration. The findings imply that laboratories need to connect performance data with safety, waste, solvent, and energy indicators to support more sustainable material discovery. Future research should develop measurable green chemistry indicators for machine learning models and compare implementation practices across different institutional contexts.

***) Corresponding Author:**

Naufal Kurniawan,
Department of Statistics, Faculty of Mathematics and Natural Sciences, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
Email: naufal.kurniawan068@gmail.com

1. Introduction

The introduction serves as a crucial opening for the research paper. It should capture the reader's attention and provide essential context for understanding the content of the research. The author(s) must clearly address the research problem, offer relevant background information, and establish their work within the existing body of research. Additionally, the introduction should justify the study's significance and outline the paper's structure and contributions.

The development of functional materials has become an important issue in chemistry, materials science, energy, environmental science, and biomedical research. Functional materials are widely needed for batteries, catalysts, sensors, membranes, pollutant adsorbents, optoelectronic devices, and health technologies. However, material synthesis processes still often depend on hazardous solvents, high temperatures, long reaction times, high energy consumption, and trial-and-error approaches that generate laboratory waste. This condition shows that material innovation should not only pursue performance but also consider safety, resource efficiency, and ecological impact. *Green chemistry* provides a basis for reducing waste, using safer solvents and raw materials, improving atom efficiency, lowering process energy, and designing safer products from the early stages. Digitalization and artificial intelligence have begun to be viewed as pathways to accelerate the transition toward sustainable chemistry because they can connect data, prediction, optimization, and research decision-making more systematically (Fantke et al., 2021). This direction also aligns with the development of *machine learning*-based materials research, which places data as the foundation for designing, selecting, and optimizing new materials (Wang et al., 2022).

In the global context, material synthesis research has shifted from conventional experimental approaches toward *data-driven synthesis*. This approach uses material databases, *machine learning* algorithms, *Bayesian optimization*, laboratory robotics, and *active learning* to predict structures, properties, synthesis routes, and the most promising reaction conditions. Hardian et al. (2020) show that artificial intelligence can support more sustainable material development through the combination of *design of experiments*, *machine learning*, and *desirability* functions to assess both performance targets and environmental aspects. Shields et al. (2021) also demonstrate that *Bayesian optimization* can accelerate the search for chemical synthesis conditions with a more controlled number of experiments. In addition, the development of *robotic chemists* and *autonomous laboratories* indicates that experiments can be conducted more rapidly, measurably, and repeatedly, opening opportunities to reduce the waste of materials, time, and energy in material discovery (Burger et al., 2020). Szymanski et al. (2023) further strengthen this direction by showing that autonomous laboratories can accelerate inorganic material synthesis through the integration of algorithms, robotics, and experimental validation. Nevertheless, the use of this technology does not automatically guarantee green practice because algorithms often prioritize accuracy, yield, or material performance, while toxicity indicators, energy consumption, carbon footprint, solvent selection, and process safety are not always treated as core variables in predictive models.

Field problems can be seen in the gap between the predictive capability of digital technology and laboratory practices that remain constrained by data availability, research culture, instrumentation costs, researcher skills, and sustainability assessment standards. Wang et al. (2020) explain that the success of *machine learning* in materials research depends heavily on data quality, parameter traceability, feature representation, and proper model validation. In practice, synthesis data are often incomplete, inconsistent, and difficult to reuse because negative results, failed conditions, experimental parameter variations, and detailed laboratory processes are rarely reported openly. Yang et al. (2022) also emphasize that data-driven nanomaterial design still faces major challenges in integrating structural, property, and synthesis data. Hung et al. (2024) found that the research community does not always need full automation, but rather autonomous laboratory systems that are flexible, easy to integrate, and aligned with real experimental needs. This condition is relevant to the national context, including Indonesia, because materials research in universities and research institutions often faces limitations in facilities, chemicals, database access, and funding. Therefore, strategies for integrating *green chemistry* into data-driven synthesis need to be understood not only as technical issues but also as social and institutional processes involving researcher decisions, laboratory policies, resource availability, and awareness of environmental risks.

This issue is important to study because the development of functional materials is directly related to science education, industrial innovation, and environmental sustainability. *Machine learning* can accelerate the discovery of materials for sustainable energy, batteries, electrocatalysts, and energy conversion devices (Yao et al., 2023). However, *green synthesis* remains necessary so that the acceleration of material discovery does not overlook material safety, energy efficiency, and environmental impact. In this context, *deep eutectic solvents* have begun to be viewed as a green platform for the synthesis of functional materials because they offer potentially lower toxicity, broad component availability, and flexibility in synthesis processes (Ma et al., 2024). Deng et al. (2024) show that *deep eutectic solvents* can be used in the synthesis of functional powder materials for energy storage and conversion. Yadav et al. (2025) also emphasize that the integration of artificial intelligence and data science can support green chemistry through reaction optimization, new material discovery, process automation, and environmental monitoring. However, researchers' experiences in choosing green strategies are often complex. Researchers must balance material performance targets, material availability, predictive model validity, reproducibility, experimental costs, and publication demands. From the perspective of education and research culture, integrating *green chemistry* into *data-driven synthesis* requires a shift in thinking from searching for the best material toward designing materials that are functional, safe, efficient, and responsible.

Previous studies have widely discussed *machine learning* in material design, reaction optimization, nanomaterials, energy materials, and autonomous laboratories. Batra et al. (2021) explain that materials intelligence ecosystems develop through the integration of data, algorithms, prediction, and experimental validation. Moosavi et al. (2020) emphasize the role of *machine learning* in understanding and designing materials, especially when the material search space is too broad to explore manually. Lu et al. (2024) show that *machine learning* is increasingly used in two-dimensional materials research to predict properties and guide material discovery. However, most of these studies still focus on computational performance, prediction accuracy, discovery efficiency, and technical success. Studies that specifically examine how *green chemistry* principles are understood, negotiated, and applied by researchers in data-driven material synthesis remain limited. This gap is important because sustainability is not determined only by predictive models but also by human decisions in selecting data, determining optimization parameters, assessing risks, interpreting experimental results, and setting research priorities. Thus, a qualitative approach is needed to explore experiences, practices, and decision-making processes that are not always visible in quantitative reports or technical articles.

Based on this background, this study aims to explore strategies for integrating *green chemistry* principles into the data-driven synthesis of functional materials. The focus of the study is directed toward how researchers understand *green chemistry* principles, the strategies used to incorporate sustainability indicators into experimental design and data models, the barriers that emerge in laboratory practice, and the forms of adjustment made when material performance targets interact with safety and environmental considerations. Theoretically, this study contributes to the development of *green chemistry*, *materials informatics*, and socio-technical transition studies by positioning *data-driven synthesis* as a scientific practice shaped by technology, sustainability values, and institutional contexts. Practically, this study can provide a basis for laboratories, universities, and research groups to develop safer, more efficient, and evidence-based material synthesis guidelines. This study can also support the strengthening of chemistry and materials curricula so that students not only master algorithms and synthesis techniques but also understand ecological responsibility at every stage of material design. Therefore, the integration of *green chemistry* and *data-driven synthesis* can become a strategic pathway for producing innovative and sustainable functional materials.

2. Materials and Methods

This study uses a qualitative approach with an exploratory case study design. This design was selected because the study focuses on an in-depth understanding of the strategies, experiences, considerations, and decision-making processes of researchers in integrating *green chemistry* principles into the data-driven synthesis of functional materials. A qualitative approach is appropriate for answering questions that emphasize reasons, processes, meanings, and the context of scientific practice, rather than testing relationships between variables statistically. Busetto et al. (2020) explain that qualitative research is suitable for understanding the nature of a phenomenon, the reasons why a practice occurs, and how an intervention can be improved in real contexts. The case study design is also relevant because the integration of *green chemistry* and *data-driven synthesis* is a complex phenomenon influenced by technology, laboratory culture, resources, research policy, and sustainability orientation. Priya (2021) states that case studies can be used to understand social or institutional phenomena contextually through multiple sources of data.

This study is designed to be conducted in three chemistry and functional materials laboratories at universities or research institutions in Indonesia. The locations are selected because they have material synthesis activities, the use of data analysis tools, or the development of predictive models in material experiments. The study is planned for three months, from February to April 2025, covering instrument preparation, data collection, data validation, thematic analysis, and report writing. The research subjects consist of principal researchers, lecturers, postgraduate students, laboratory analysts, and research managers involved in functional material synthesis. Informants are selected using *purposive sampling* because this study requires participants who have direct experience in material synthesis, experimental data processing, the use of predictive algorithms, solvent selection, reaction optimization, or the application of safety and sustainability principles. Campbell et al. (2020) explain that *purposive sampling* helps researchers select cases or participants who are most information-rich according to the research objectives. If the initial data do not sufficiently reach key actors, the researcher uses *snowball sampling* by asking early participants to recommend additional informants.

The informant criteria in this study include four aspects. First, informants must have at least one year of experience in functional material synthesis research, such as catalysts, nanomaterials, energy materials, sensors, membranes, or adsorbent materials. Second, informants must have used experimental data, computational tools, *machine learning*, *design of experiment*, or statistical optimization in the synthesis process. Third, informants must have knowledge or experience related to *green chemistry* principles, such as waste reduction, safer solvent selection, energy efficiency, chemical safety, and pollution prevention. Fourth, informants must be willing to participate in interviews, observation, and data confirmation processes. The target number of informants is 15 to 20 people, but the final number is determined by the depth of information and adequacy of meaning, not by a numerical quota. LaDonna et al. (2021) remind researchers that the quality of qualitative interview data should not be assessed only through claims of saturation, but also through the depth, relevance, variation, and strength of evidence obtained from participants.

Data are collected through semi-structured interviews, non-participant observation, documentation, and triangulation. Semi-structured interviews are used to explore informants' experiences regarding *green chemistry* integration strategies, the process of selecting synthesis parameters, the use of data in experimental design, technical barriers, institutional barriers, and ethical considerations in the use of chemicals and energy. Interview questions are designed openly so that informants can explain their experiences, assessments, and reflections in detail. Non-participant observation is conducted in laboratories to record work practices, experimental flows, material use, safety procedures, data recording, waste management, and the use of computational tools. Documentation is conducted by reviewing laboratory SOPs, experimental notes, synthesis protocols, chemical safety forms, research reports, publication articles, and laboratory policy documents related to sustainability. Mtisi (2022) explains that qualitative case studies require

multiple sources of evidence so that researchers can understand phenomena comprehensively and contextually. Therefore, interviews, observations, and documentation are used complementarily to reduce dependence on a single type of data.

Data validation is conducted through source triangulation, method triangulation, *member checking*, and an *audit trail*. Source triangulation is carried out by comparing information from principal researchers, students, laboratory analysts, and research documents. Method triangulation is carried out by comparing interview results, observation notes, and laboratory documents. *Member checking* is conducted by sending summaries of interview results or initial interpretations to informants to ensure that the meaning captured by the researcher is consistent with informants' experiences. An *audit trail* is conducted by storing research process notes, interview guidelines, analytical memos, transcripts, coding matrices, categorization decisions, and theme changes during the analysis. Johnson et al. (2020) emphasize that the quality of qualitative research must be demonstrated through credibility, process traceability, transparency of analytical decisions, and consistency between data and interpretation. Braun and Clarke (2022) also underline the importance of methodological awareness in thematic analysis so that themes do not merely summarize data but represent patterns of meaning relevant to the research question.

Data analysis uses reflective thematic analysis through systematic stages. The first stage involves transcribing all interview data and organizing observation notes. The second stage involves repeated reading to understand the context, initial patterns, and emerging issues. The third stage involves assigning initial codes to data segments related to *green chemistry* strategies, data use, material selection, synthesis optimization, barriers, negotiation of interests, and sustainability practices. The fourth stage involves grouping codes into broader categories, such as technical strategies, institutional strategies, educational strategies, data constraints, facility constraints, and safety considerations. The fifth stage involves developing main themes that explain the process of integrating *green chemistry* into *data-driven synthesis*. The sixth stage involves writing the findings narrative by connecting informant quotations, observation results, documents, and relevant theories. Lester et al. (2020) explain that qualitative data analysis must be conducted gradually, transparently, and reflectively so that readers can understand how researchers move from raw data to interpretation. Braun and Clarke (2020) also emphasize that thematic analysis allows researchers to identify patterns of meaning across cases, as long as the coding and theme development processes are explained methodologically. With these procedures, this study can be replicated in a limited way in laboratories or institutions with similar characteristics.

3. Results and Discussions

The findings show that the integration of green chemistry principles into data-driven synthesis of functional materials was understood by participants as a gradual and selective process rather than a fixed laboratory procedure. Most participants did not view green chemistry as a separate research agenda. Instead, they described it as a set of practical considerations that entered the synthesis process through solvent selection, waste reduction, energy efficiency, safety control, and the search for reproducible experimental conditions. One principal researcher explained, "We do not always start by saying that this is green chemistry, but we try to reduce solvent use, avoid highly toxic chemicals, and make the reaction more efficient." This statement reflects a pattern in which green chemistry appeared more as embedded laboratory judgment than as a formalized framework. Observations in the laboratories also showed that researchers often made small adjustments, such as reducing reaction volume, reusing safe solvents when possible, shortening reaction time, and comparing less hazardous precursors before selecting the final synthesis route.

The first major theme that emerged from the data was the use of data-driven methods to reduce experimental uncertainty. Participants reported that predictive models, previous experimental datasets, design of experiments, and optimization software helped them narrow the range of possible synthesis conditions before conducting laboratory trials. This practice reduced the

number of repeated experiments and helped researchers avoid unnecessary material use. A postgraduate researcher stated, “Before using data analysis, we repeated many trials without a clear direction. Now, we can estimate which temperature, ratio, or solvent system is more promising before starting.” This finding suggests that data-driven synthesis contributed not only to research efficiency but also to the practical reduction of chemical consumption. Documentation from laboratory notebooks supported this pattern. Several records showed that experimental conditions were selected based on prior datasets, published synthesis routes, and statistical comparison of parameters such as temperature, reaction time, solvent type, and precursor concentration.

The second theme was the tension between material performance and environmental responsibility. Participants consistently emphasized that the main expectation in functional material research remained performance, such as catalytic activity, conductivity, stability, adsorption capacity, sensitivity, or energy storage ability. Green considerations were often introduced after researchers identified the performance target. One participant noted, “The first question is still whether the material works. After that, we ask whether the process can be made safer or simpler.” This indicates that green chemistry had not fully become the starting point of synthesis design. Instead, it often functioned as a corrective principle during optimization. In several cases, researchers selected a less hazardous solvent only if the resulting material maintained acceptable performance. If the green alternative reduced the desired property too strongly, researchers tended to return to conventional synthesis routes. This pattern shows that sustainability was negotiated within the pressure to produce publishable and technically competitive materials.

The third theme concerned the limited integration of sustainability indicators into data models. Participants used data-driven tools mainly to predict material properties and optimize synthesis outcomes. However, toxicity, waste generation, energy input, solvent hazard, and carbon-related indicators were rarely included as formal variables in computational models. A laboratory analyst explained, “The model usually reads performance data, not waste data. We record waste separately, but it is not part of the prediction.” This finding shows a methodological gap between the logic of machine learning and the logic of green chemistry. While the former often focuses on predictive accuracy and target performance, the latter requires broader indicators that capture environmental and safety impacts. Observations confirmed that safety data sheets and waste logs were available in laboratories, but these documents were not systematically connected to experimental datasets used for modeling or optimization. As a result, green chemistry remained partly dependent on individual awareness rather than integrated data infrastructure.

The fourth theme was the role of laboratory culture and institutional support. Participants indicated that green chemistry practices were easier to implement when laboratory leaders encouraged safe material use, careful documentation, and efficient experimentation. In laboratories where senior researchers emphasized sustainability, students were more likely to compare solvent hazards, record failed experiments, and discuss waste management during research meetings. One doctoral student stated, “Our supervisor always asks whether there is a safer way to do the synthesis. That habit makes us think before buying or using chemicals.” Conversely, participants from laboratories with limited funding or older instruments reported that green practices were harder to maintain. They often relied on available chemicals and established protocols because changing synthesis routes required additional validation, new reagents, and more time. This finding indicates that green chemistry integration depends not only on individual knowledge but also on institutional routines, leadership, infrastructure, and resource availability.

The fifth theme was the importance of negative results and failed experiments in building greener synthesis pathways. Participants explained that failed experiments often provided valuable information for avoiding unnecessary repetition. However, these results were rarely published or shared outside the research group. A participant said, “Failed synthesis actually teaches us which conditions waste materials, but usually we only report the successful route.”

This practice limits the development of reusable synthesis data. Observations of laboratory documentation showed that failed conditions were recorded inconsistently. Some researchers wrote detailed notes on failed trials, while others only marked them as unsuccessful without explaining possible causes. This creates a challenge for data-driven synthesis because incomplete records weaken model training and reduce the capacity to predict greener pathways. The finding suggests that sustainable material discovery requires stronger documentation cultures, including the systematic recording of failed trials, unsafe conditions, and inefficient routes.

The sixth theme involved education and researcher competence. Participants stated that many young researchers understood machine learning as a technical tool but had limited training in translating green chemistry principles into measurable variables. They could operate software for prediction or optimization, but they were less familiar with integrating solvent safety, waste minimization, or energy consumption into the data structure. One lecturer explained, "Students can run optimization models, but they still need guidance to decide which variables represent sustainability." This finding highlights the need for interdisciplinary competence. Researchers need knowledge of material chemistry, computational modeling, laboratory safety, and sustainability assessment. Without this combined competence, data-driven synthesis may accelerate material discovery without substantially improving environmental responsibility. Therefore, education plays a central role in transforming green chemistry from a general principle into an operational element of research design.

Discussion

The findings show that green chemistry integration in data-driven synthesis is not merely a technical adjustment. It is a socio-technical process shaped by research targets, laboratory culture, institutional resources, and the way researchers define success. This finding supports Fantke et al. (2021), who argue that digitalization can support sustainable chemistry when data, process design, and decision-making are connected to sustainability goals. However, the present study adds that digital tools do not automatically produce greener synthesis. Researchers still need to decide which variables matter, which risks should be prioritized, and how far performance targets can be balanced with environmental responsibility. This finding also expands the argument of Wang et al. (2022), who describe data-driven materials innovation as a route to accelerate discovery. Acceleration alone is not sufficient. Data-driven synthesis becomes more sustainable only when green chemistry indicators become part of the model, the documentation system, and the laboratory decision-making process.

The finding on experimental uncertainty is consistent with studies that highlight the value of artificial intelligence and optimization in reducing inefficient experimental work. Hardian et al. (2020) argue that artificial intelligence can support sustainable materials development by combining experimental design, machine learning, and desirability functions. Shields et al. (2021) also show that Bayesian optimization can reduce the number of experiments needed to identify effective chemical synthesis conditions. The present study confirms these arguments in a qualitative context. Participants experienced data-driven tools as a way to reduce trial-and-error practices, save materials, and narrow experimental choices. However, this study also shows that reduced experimentation does not always equal green chemistry. If the model only optimizes yield, performance, or reaction success, then the environmental dimension remains secondary. Therefore, the practical contribution of data-driven synthesis depends on how researchers design the optimization objective.

The tension between performance and sustainability reflects a central challenge in functional material research. Previous studies on materials informatics often emphasize prediction accuracy, material properties, and discovery efficiency (Batra et al., 2021; Moosavi et al., 2020). The findings of this study suggest that researchers still operate within a performance-oriented publication culture. Green chemistry becomes important, but it often enters the research process after the material target has been defined. This finding offers a critical perspective. It shows that

the slow integration of green chemistry does not result only from a lack of knowledge. It also results from incentive structures that reward novelty, high performance, and publication output more strongly than safer or lower-impact synthesis routes. In this sense, the integration of green chemistry requires changes in laboratory evaluation systems, journal reporting standards, and funding priorities.

The limited use of sustainability indicators in data models confirms the challenges identified by Wang et al. (2020) and Yang et al. (2022). Data-driven materials research requires high-quality, traceable, and reusable data. Yet the present findings show that laboratory data often separate performance records from safety records, waste records, and solvent hazard information. This separation weakens the possibility of building models that can predict both material performance and environmental impact. The issue is not only technical but also organizational. Laboratories need data structures that connect synthesis parameters, material properties, failed experiments, chemical hazards, energy input, and waste outcomes. Without this integration, green chemistry remains difficult to operationalize in machine learning workflows. This study therefore suggests that sustainable data-driven synthesis requires a broader data culture, not only better algorithms.

The role of laboratory culture also supports qualitative perspectives on scientific practice. Busetto et al. (2020) explain that qualitative research is useful for understanding how practices occur in real contexts. In this study, laboratory leaders, supervisors, and institutional routines shaped how researchers applied green chemistry. When supervisors asked about safer alternatives, students became more attentive to solvent choice, waste reduction, and efficient experimentation. This shows that green chemistry is learned through everyday research interaction, not only through formal training. The finding also aligns with socio-technical transition thinking, which views technological change as a process shaped by institutions, actors, norms, and material resources. In the context of functional material synthesis, the transition toward greener practice depends on the interaction between digital tools, research culture, laboratory infrastructure, and sustainability values.

The finding on negative results provides an important contribution to discussions on data-driven synthesis. Burger et al. (2020) and Szymanski et al. (2023) show that autonomous and robotic laboratories can accelerate material discovery through repeated and controlled experimentation. However, the present study indicates that the value of automation and prediction depends on the quality of recorded data, including failed experiments. Negative results help researchers avoid wasteful synthesis routes, but they are often underreported. This limits the development of greener predictive models. If failed trials, unsafe routes, and inefficient conditions remain invisible, machine learning models may reproduce the same narrow logic that privileges successful outputs. Therefore, documenting failure should be treated as part of sustainable research practice. It can reduce repetition, improve model training, and support more responsible material discovery.

The findings also show that education is a key pathway for integrating green chemistry into data-driven synthesis. Researchers need to understand not only algorithms but also sustainability indicators. This supports the argument that materials intelligence ecosystems require the integration of data, domain knowledge, and experimental validation (Batra et al., 2021). The present study adds that ecological judgment must become part of this ecosystem. Students and early-career researchers need training that links machine learning, experimental design, solvent selection, waste minimization, energy efficiency, and risk assessment. Without this integration, data-driven synthesis may remain a powerful but incomplete tool. It can make discovery faster, but not necessarily safer or more sustainable.

Practically, the results suggest that laboratories should develop internal guidelines for greener data-driven synthesis. These guidelines can include the documentation of failed experiments, the use of solvent hazard scoring, the recording of energy input, the comparison of alternative

synthesis routes, and the integration of safety data into experimental datasets. Research groups can also develop decision matrices that balance performance, safety, cost, reproducibility, and environmental impact. Such tools can help researchers move from informal awareness toward systematic implementation. For universities, these findings imply the need to revise chemistry and materials curricula by combining materials informatics with green chemistry and laboratory sustainability. For research funders and journal editors, the findings suggest the importance of encouraging authors to report sustainability considerations in synthesis studies, including solvent choice, waste handling, energy conditions, and unsuccessful routes when relevant.

Theoretically, this study contributes to the literature by positioning the integration of green chemistry and data-driven synthesis as a contextual and negotiated process. Previous research has shown the technical promise of machine learning, Bayesian optimization, deep eutectic solvents, and autonomous laboratories in material discovery (Deng et al., 2024; Ma et al., 2024; Yao et al., 2023). This study extends that literature by showing how researchers interpret and apply these ideas in laboratory practice. The findings suggest that sustainable material discovery requires three forms of integration. The first is technical integration, which connects green indicators with predictive models. The second is institutional integration, which embeds sustainability into laboratory routines and documentation systems. The third is educational integration, which prepares researchers to think across chemistry, data science, safety, and environmental responsibility. These three forms of integration can strengthen the role of green chemistry in the future development of functional materials.

Future research can deepen this topic in several directions. First, comparative studies can examine differences between laboratories that have strong green chemistry policies and laboratories that rely on informal practices. Second, future studies can explore how researchers translate specific green chemistry principles into measurable variables for machine learning models. Third, mixed-method research can combine qualitative interviews with quantitative assessment of solvent use, waste generation, energy consumption, and model efficiency. Fourth, further studies can examine the role of institutional policy, funding schemes, and journal requirements in encouraging sustainable synthesis reporting. These directions are important because the integration of green chemistry into data-driven synthesis is still developing. A deeper understanding of technical, cultural, and institutional factors can help ensure that functional material innovation becomes not only faster and more accurate but also safer, more efficient, and more environmentally responsible.

4. Conclusions

This study concludes that the integration of *green chemistry* principles into the data-driven synthesis of functional materials is not merely a technical process, but a socio-technical practice shaped by researcher awareness, laboratory culture, data quality, institutional support, and sustainability-oriented decision-making. The main findings show that researchers use data-driven tools to reduce experimental uncertainty, limit trial-and-error practices, and improve synthesis efficiency. However, sustainability indicators such as solvent hazard, waste generation, energy input, toxicity, and process safety are not yet fully integrated into predictive models or laboratory datasets. Green chemistry is often applied as a corrective consideration after performance targets are defined, rather than as a core principle from the beginning of synthesis design. This finding contributes to the literature by showing that sustainable material discovery requires a stronger connection between materials informatics, green chemistry, and laboratory decision-making.

Theoretically, this study strengthens the understanding of data-driven synthesis as a contextual and negotiated scientific practice, not only as a computational method for accelerating material discovery. Practically, the findings suggest that laboratories need clearer guidelines for documenting failed experiments, assessing solvent safety, recording energy use, and integrating environmental indicators into experimental datasets. From a policy perspective, universities, research institutions, funding bodies, and journals should encourage sustainability reporting in

material synthesis research. Future studies can expand this research through comparative case studies across different laboratory settings, mixed-method approaches that combine qualitative insights with quantitative environmental indicators, and deeper analysis of how specific green chemistry principles can be translated into measurable variables for machine learning models.

References

Batra, R., Song, L., & Ramprasad, R. (2021). Emerging materials intelligence ecosystems propelled by machine learning. *Nature Reviews Materials, 6*, 655–678. https://doi.org/10.1038/s41578-020-00255-y

Braun, V., & Clarke, V. (2020). One size fits all? What counts as quality practice in reflexive thematic analysis? *Qualitative Research in Psychology, 18*(3), 328–352. https://doi.org/10.1080/14780887.2020.1769238

Braun, V., & Clarke, V. (2022). Toward good practice in thematic analysis: Avoiding common problems and becoming a knowing researcher. *International Journal of Transgender Health, 24*(1), 1–6. https://doi.org/10.1080/26895269.2022.2129597

Burger, B., Maffettone, P. M., Gusev, V. V., Aitchison, C. M., Bai, Y., Wang, X., Li, X., Alston, B. M., Li, B., Clowes, R., Rankin, N., Harris, B., Sprick, R. S., & Cooper, A. I. (2020). A mobile robotic chemist. *Nature, 583*(7815), 237–241. https://doi.org/10.1038/s41586-020-2442-2

Busetto, L., Wick, W., & Gumbinger, C. (2020). How to use and assess qualitative research methods. *Neurological Research and Practice, 2*, Article 14. https://doi.org/10.1186/s42466-020-00059-z

Campbell, S., Greenwood, M., Prior, S., Shearer, T., Walkem, K., Young, S., Bywaters, D., & Walker, K. (2020). Purposive sampling: Complex or simple? Research case examples. *Journal of Research in Nursing, 25*(8), 652–661. https://doi.org/10.1177/1744987120927206

Deng, R., Gao, M., Zhang, B., & Zhang, Q. (2024). Solvent-mediated synthesis of functional powder materials from deep eutectic solvents for energy storage and conversion: A review. *Advanced Energy Materials, 14*(8), Article 2303707. https://doi.org/10.1002/aenm.202303707

Fantke, P., Cinquemani, C., Yaseneva, P., De Mello, J., Schwabe, H., Ebeling, B., & Lapkin, A. A. (2021). Transition to sustainable chemistry through digitalization. *Chem, 7*(11), 2866–2882. https://doi.org/10.1016/j.chempr.2021.09.012

Hardian, R., Liang, Z., Zhang, X., & Szekely, G. (2020). Artificial intelligence: The silver bullet for sustainable materials development. *Green Chemistry, 22*(21), 7521–7528. https://doi.org/10.1039/D0GC02956D

Hung, L., Yager, J. A., Monteverde, D., Baiocchi, D., Kwon, H., Sun, S., & Suram, S. (2024). Autonomous laboratories for accelerated materials discovery: A community survey and practical insights. *Digital Discovery, 3*(7), 1273–1279. https://doi.org/10.1039/D4DD00059E

Johnson, J. L., Adkins, D., & Chauvin, S. (2020). A review of the quality indicators of rigor in qualitative research. *American Journal of Pharmaceutical Education, 84*(1), Article 7120. https://doi.org/10.5688/ajpe7120

LaDonna, K. A., Artino, A. R., Jr., & Balmer, D. F. (2021). Beyond the guise of saturation: Rigor and qualitative interview data. *Journal of Graduate Medical Education, 13*(5), 607–611. https://doi.org/10.4300/JGME-D-21-00752.1

Lester, J. N., Cho, Y., & Lochmiller, C. R. (2020). Learning to do qualitative data analysis: A starting point. *Human Resource Development Review, 19*(1), 94–106. https://doi.org/10.1177/1534484320903890

Lu, B., Xia, Y., Ren, Y., Xie, M., Zhou, L., Vinai, G., Morton, S. A., Wee, A. T. S., van der Wiel, W. G., Zhang, W., & Wong, P. K. J. (2024). When machine learning meets 2D materials: A review. *Advanced Science, 11*(13), Article 2305277. https://doi.org/10.1002/advs.202305277

Ma, Y., Yang, Y., Li, T., Hussain, S., & Zhu, M. (2024). Deep eutectic solvents as an emerging green platform for the synthesis of functional materials. *Green Chemistry, 26*(7), 3627–3669. https://doi.org/10.1039/D3GC04289H

Moosavi, S. M., Jablonka, K. M., & Smit, B. (2020). The role of machine learning in the understanding and design of materials. *Journal of the American Chemical Society, 142*(48), 20273–20287. https://doi.org/10.1021/jacs.0c09105

Mtisi, S. (2022). The qualitative case study research strategy as applied on a rural enterprise development doctoral research project. *International Journal of Qualitative Methods, 21*, 1–13. https://doi.org/10.1177/16094069221145849

Priya, A. (2021). Case study methodology of qualitative research: Key attributes and navigating the conundrums in its application. *Sociological Bulletin, 70*(1), 94–110. https://doi.org/10.1177/0038022920970318

Shields, B. J., Stevens, J., Li, J., Parasram, M., Damani, F., Alvarado, J. I. M., Janey, J. M., Adams, R. P., & Doyle, A. G. (2021). Bayesian reaction optimization as a tool for chemical synthesis. *Nature, 590*(7844), 89–96. https://doi.org/10.1038/s41586-021-03213-y

Szymanski, N. J., Rendy, B., Fei, Y., Kumar, R. E., He, T., Milsted, D., McDermott, M. J., Gallant, M., Cubuk, E. D., Merchant, A., Kim, H., Jain, A., Bartel, C. J., Persson, K. A., Zeng, Y., & Ceder, G. (2023). An autonomous laboratory for the accelerated synthesis of inorganic materials. *Nature, 624*(7990), 86–91. https://doi.org/10.1038/s41586-023-06734-w

Wang, A. Y. T., Murdock, R. J., Kauwe, S. K., Oliynyk, A. O., Gurlo, A., Brgoch, J., Persson, K. A., & Sparks, T. D. (2020). Machine learning for materials scientists: An introductory guide toward best practices. *Chemistry of Materials, 32*(12), 4954–4965. https://doi.org/10.1021/acs.chemmater.0c01907

Wang, Z., Sun, Z., Yin, H., Liu, X., Wang, J., Zhao, H., Pang, C. H., Wu, T., Li, S., Yin, Z., & Yu, X. F. (2022). Data-driven materials innovation and applications. *Advanced Materials, 34*(36), Article 2104113. https://doi.org/10.1002/adma.202104113

Yao, Z., Lum, Y., Johnston, A., Mejia-Mendoza, L. M., Zhou, X., Wen, Y., Aspuru-Guzik, A., Sargent, E. H., & Seh, Z. W. (2023). Machine learning for a sustainable energy future. *Nature Reviews Materials, 8*(3), 202–215. https://doi.org/10.1038/s41578-022-00490-5