Global Journal of Chemistry
Vol. 1 No.1 2026 Page 20-28
E-ISSN: XXXX-XXXX, DOI: 10.XXXXX/XXXX

Development of a Data-Driven Chemistry Framework to Improve the Efficiency of Environmentally Friendly Material Synthesis

Nabila Hakim[1]*, Rudi Ramadhan[2], Fitri Fauzi[3]
[1,2] Department of Information Systems, Faculty of Computer Science, Universitas Indonesia, Depok, Indonesia
[3] Department of Informatics, School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Bandung, Indonesia

Article Info

Received : 20 January 2026
Revised : 11 March 2026
Accepted : 28 March 2026

Keyword:

Data-driven chemistry;
Environmentally friendly Synthesis;
Green chemistry;
Material synthesis.

ABSTRACT

This study aims to develop a data-driven chemistry framework to improve the efficiency of environmentally friendly material synthesis. The study responds to the problem of fragmented experimental records, repeated trial-and-error practices, and limited integration between synthesis data, green chemistry principles, and laboratory decision-making. A qualitative instrumental case study design was used to explore how researchers generate, document, interpret, and apply experimental data in material synthesis. Data were collected through semi-structured interviews, limited participant observation, and document analysis involving lecturers, material researchers, postgraduate students, laboratory analysts, instrument operators, and data-oriented researchers selected through purposive and snowball sampling. Thematic analysis identified five main themes: fragmented experimental data, tacit decision-making in synthesis design, practical meanings of green efficiency, cautious trust in data-driven recommendations, and the need for an adaptive laboratory framework. The findings show that data-driven chemistry can support greener synthesis when laboratory data are standardized, failed experiments are documented, characterization results are linked to synthesis parameters, and researcher interpretation remains central to decision-making. This study contributes theoretically by framing data-driven chemistry as a sociotechnical practice that connects human expertise, laboratory culture, data systems, and sustainability values. Practically, the proposed framework can help laboratories reduce unnecessary experiments, improve documentation quality, and strengthen environmentally responsible material synthesis. Future research should test the framework across different material systems and integrate qualitative findings with measurable environmental performance indicators.

***) Corresponding Author:**

Nabila Hakim,
Department of Information Systems, Faculty of Computer Science, Universitas Indonesia, Depok, Indonesia
Email: nabila.hakim069@gmail.com

1. Introduction

The development of environmentally friendly material synthesis has become a major concern in contemporary chemistry because material innovation now faces stronger demands for efficiency, safety, reproducibility, and sustainability. At the global level, material synthesis is no longer assessed only by the successful formation of the target product. It also needs to consider the

environmental consequences of solvent selection, precursor use, reaction temperature, synthesis duration, energy consumption, waste generation, and post-synthesis treatment. In this context, data-driven chemistry offers a relevant pathway because it enables researchers to use experimental data, computational modeling, and machine learning to support more informed decisions in designing synthesis routes. Recent studies show that artificial intelligence and machine learning can support sustainable chemistry by improving reaction optimization, reducing unnecessary experimental trials, and accelerating the discovery of safer and more efficient materials (Ghasemlou et al., 2025; Leonard & Hasan, 2021). This shift is also important for research laboratories in developing countries, including Indonesia, where limited research funds, equipment access, and reagent availability require more efficient experimental strategies.

In practice, many laboratories still rely on trial-and-error approaches, personal experience, and fragmented documentation when conducting material synthesis. Researchers often decide reaction parameters based on previous laboratory habits, supervisor recommendations, or partial literature reports rather than structured data integration. This condition makes synthesis processes difficult to reproduce and limits the opportunity to learn from failed or less optimal experiments. Data from laboratory notebooks, instrument outputs, synthesis logs, and characterization results often remain scattered across different formats. As a result, valuable experimental knowledge is not always transformed into reusable data for future synthesis planning. Wang et al. (2022) showed that structured datasets extracted from synthesis procedures can support data-driven material discovery, while Williamson and Brutchey (2023) emphasized that data-driven learning can help predict and control inorganic material synthesis outcomes. These findings indicate that the efficiency of green synthesis depends not only on advanced algorithms but also on the way laboratories record, organize, interpret, and reuse experimental data.

The issue is important because environmentally friendly material synthesis is not only a technical matter. It also reflects the social and cultural practices of scientific work in the laboratory. The success of data-driven chemistry depends on how researchers understand data quality, trust model recommendations, document experimental failures, and connect sustainability indicators with chemical reasoning. In many research settings, successful experiments are more visible than failed trials, while failed synthesis data may remain undocumented even though they contain important information for reducing waste and avoiding repeated errors. This creates a gap between the potential of data-driven methods and their actual implementation in daily laboratory practice. Studies on self-driving laboratories and autonomous experimentation show that automation, machine learning, and closed-loop optimization can accelerate material discovery (Abolhasani & Kumacheva, 2023; Epps et al., 2021; Tom et al., 2024). However, these advances still require human interpretation, contextual judgment, and clear laboratory workflows to ensure that data-driven systems align with green chemistry principles.

Theoretically, this study is grounded in a sociotechnical perspective, which views data-driven chemistry as an interaction between researchers, instruments, data systems, algorithms, laboratory routines, and sustainability values. This perspective is relevant because a chemistry framework cannot be understood only as a computational tool. It also functions as a decision-making system shaped by human experience, institutional support, and laboratory culture. Stach et al. (2021) argued that autonomous experimentation systems require strong integration among data infrastructure, domain expertise, and stakeholder collaboration. In material synthesis, this integration becomes more complex because researchers must balance product performance, synthesis feasibility, environmental impact, and resource availability. Therefore, a qualitative approach is needed to explore how researchers give meaning to data, how they evaluate efficiency, and how they negotiate the use of predictive tools in real laboratory settings. This approach allows the study to capture processes that cannot be fully explained by numerical performance indicators alone.

Previous studies have made significant contributions to data-driven material synthesis, machine learning-assisted optimization, and autonomous experimentation. McDermott et al.

(2021) developed a graph-based model to predict reaction pathways in solid-state synthesis. Tang et al. (2020) demonstrated that machine learning can guide advanced inorganic material synthesis with fewer experimental trials. Szymanski et al. (2023) introduced an autonomous laboratory for accelerating the synthesis of novel materials. Other studies have reviewed machine learning applications in material property prediction, thermoelectric material synthesis, and self-driving laboratory design (Huang et al., 2023; Jain, 2024; Na, 2023; Rooney et al., 2022). Although these studies provide strong technical foundations, many of them focus on model performance, automation architecture, or prediction accuracy. Few studies explore how researchers experience, interpret, and adapt data-driven chemistry frameworks in environmentally friendly material synthesis, especially in laboratories with limited resources and diverse documentation practices.

Based on this gap, this study aims to develop a data-driven chemistry framework to improve the efficiency of environmentally friendly material synthesis through a qualitative exploration of laboratory practices, researcher experiences, and decision-making processes. The study focuses on how experimental data are generated, documented, interpreted, validated, and integrated with green chemistry principles during material synthesis. The theoretical contribution of this study lies in expanding the understanding of data-driven chemistry as a sociotechnical practice rather than merely a computational process. The practical contribution lies in providing a framework that can guide laboratories in designing structured documentation systems, reducing repetitive experiments, improving synthesis planning, and strengthening environmentally responsible research practices. The framework is expected to support more transparent, efficient, and sustainable material synthesis, especially in academic and research laboratories that need adaptable and realistic data-based decision tools.

2. Materials and Methods

This study used a qualitative approach with an instrumental case study design. The case study design was selected because the research aimed to examine a bounded phenomenon, namely the development and implementation of a data-driven chemistry framework in environmentally friendly material synthesis. This design allowed the researcher to explore laboratory practices, researcher experiences, data documentation routines, and decision-making processes in their natural context. A case study is suitable when the boundaries between the phenomenon and its context are closely connected, and when the researcher needs to understand how and why a process occurs in a specific setting (Priya, 2021). In this study, the case was not limited to the technical success of material synthesis. It also included how researchers interpret data quality, apply green chemistry principles, and use experimental evidence to improve synthesis efficiency.

The study was conducted in a chemistry material laboratory and a computational chemistry or data analysis unit at a university or research institution in Indonesia. The research was planned for four months, from February to May 2025. The research period covered preliminary observation, participant recruitment, semi-structured interviews, limited participant observation, document collection, data validation, thematic analysis, and framework development. The selected research site was considered relevant because it conducted material synthesis activities, used characterization instruments, produced experimental datasets, and had laboratory members who engaged with green chemistry or data-based decision-making. The site also represented a realistic laboratory context where data-driven chemistry could be developed gradually without assuming full automation or advanced self-driving laboratory infrastructure.

Participants were selected using purposive sampling. The selection focused on individuals who had direct experience with material synthesis, laboratory data documentation, characterization procedures, green chemistry principles, or data analysis. The participants included chemistry lecturers, material researchers, postgraduate students, laboratory analysts, instrument operators, and researchers familiar with machine learning or experimental data management. The inclusion criteria required participants to have at least one year of experience in synthesis, characterization, or laboratory data handling. They also needed to be willing to share their experiences and allow

the researcher to examine relevant non-confidential documents. The initial target was 12 to 18 participants. This number could increase through snowball sampling when key participants recommended other informants with specific expertise. The final number of participants was determined by data saturation, when additional interviews no longer produced new substantial themes (Hennink & Kaiser, 2022).

Data were collected through semi-structured interviews, limited participant observation, and document analysis. Semi-structured interviews were used to explore participants' experiences in selecting synthesis parameters, recording experimental data, evaluating failed experiments, assessing synthesis efficiency, and responding to data-driven recommendations. Each interview lasted around 45 to 90 minutes and was recorded with participant consent. Limited participant observation was conducted during laboratory routines, including material preparation, parameter setting, reaction monitoring, data recording, sample labeling, instrument use, and team discussions about synthesis results. Document analysis included laboratory notebooks, synthesis logs, standard operating procedures, characterization outputs, reagent records, safety documents, data spreadsheets, and research reports. These three data sources allowed the researcher to understand the relationship between laboratory practice, data culture, and environmentally friendly synthesis decisions.

Data validity was maintained through source triangulation, method triangulation, member checking, and an audit trail. Source triangulation was conducted by comparing information from lecturers, students, analysts, instrument operators, and data-oriented researchers. Method triangulation was conducted by comparing interview findings with observation notes and laboratory documents. Member checking was carried out by returning interview summaries and preliminary interpretations to selected participants to confirm whether the researcher's interpretation matched their intended meaning. McKim (2023) emphasized that member checking needs to be structured so that it strengthens interpretive validity rather than becoming a formal confirmation only. An audit trail was also maintained by documenting interview guides, consent records, coding decisions, theme revisions, analytical memos, and framework development notes. This strategy supported credibility, dependability, transferability, and confirmability in the qualitative research process (Ahmed, 2024; Carcary, 2020).

The data were analyzed using reflexive thematic analysis. The analysis followed several stages, namely familiarization with the data, initial coding, theme construction, theme review, theme definition, and narrative writing. Interview transcripts, observation notes, and documents were read repeatedly to identify patterns related to data documentation, synthesis decision-making, sustainability considerations, researcher trust in data, and barriers to framework implementation. Coding was conducted inductively so that themes could emerge from participants' experiences. However, the analysis was also informed by the concepts of data-driven chemistry, green chemistry, and sociotechnical systems. Braun and Clarke (2021) explained that reflexive thematic analysis requires active interpretation from the researcher, while Byrne (2022) showed that this approach can produce coherent themes when the analytical process is clearly documented. In this study, thematic analysis was used not only to describe participant experiences but also to construct the proposed data-driven chemistry framework.

The development of the framework followed the synthesis of empirical themes and theoretical concepts. First, themes from the field data were mapped into several components, including data generation, data documentation, data cleaning, synthesis parameter selection, green efficiency assessment, model-supported recommendation, researcher interpretation, and decision feedback. Second, these components were connected into a workflow that explains how experimental data can support more efficient and environmentally friendly material synthesis. Third, the initial framework was discussed with selected key informants to assess clarity, feasibility, and relevance to laboratory practice. The framework was revised based on this feedback. The final framework was designed as a practical guide for laboratories that want to move from fragmented experiment records toward more systematic and reusable synthesis data.

Ethical procedures were applied throughout the study. Participants received information about the research objectives, data collection procedures, confidentiality protection, voluntary participation, and their right to withdraw from the study. Participant names, laboratory names, and sensitive experimental information were anonymized. Digital data were stored in password-protected files and accessed only by the research team. The study did not require participants to conduct unsafe experiments or change their standard laboratory procedures. The researcher only observed existing practices and discussed participants' experiences. These ethical considerations were important because laboratory data may include unpublished results, intellectual property concerns, and sensitive institutional information. Therefore, the study treated data protection, participant consent, and responsible reporting as essential parts of the research design.

3. Results and Discussions

The qualitative analysis identified five main themes that explain how a data-driven chemistry framework can improve the efficiency of environmentally friendly material synthesis. These themes emerged from semi-structured interviews, limited participant observation, and laboratory document analysis. The themes include fragmented experimental data, tacit decision-making in synthesis design, practical interpretations of green efficiency, cautious trust in data-driven recommendations, and the need for an adaptive laboratory framework. These findings show that data-driven chemistry is not only a technical innovation. It is also a change in laboratory culture, documentation habits, and scientific decision-making.

The first theme concerns fragmented experimental data. Most participants stated that synthesis data were available, but they were not always organized in a format that could support future analysis. Experimental records were often stored in laboratory notebooks, spreadsheet files, instrument folders, personal laptops, and informal communication channels. This condition made it difficult for researchers to trace previous synthesis conditions, compare failed and successful trials, or reuse historical data for new experiments. One participant explained, "We have many synthesis records, but they are not always in the same format. Sometimes the temperature is written, but the stirring speed or solvent volume is missing" (P4). Observation also showed that data from characterization instruments were often separated from synthesis notes. This separation limited the ability of researchers to connect synthesis parameters with material properties.

The second theme shows that synthesis decisions were strongly influenced by tacit knowledge. Researchers did not rely only on formal protocols or published procedures. They also used intuition, previous experience, and advice from senior researchers. This pattern helped researchers work efficiently in familiar experiments, but it created challenges for replication and systematic learning. A postgraduate student stated, "Usually we follow the method from previous students, but we adjust it based on what usually works in our lab" (P7). This finding indicates that laboratory experience remains important in material synthesis. However, tacit knowledge needs to be translated into explicit and structured data so that it can be shared, validated, and reused by other researchers.

The third theme relates to how participants understood green synthesis efficiency. Participants did not define efficiency only as a faster synthesis process. They also linked it to lower reagent consumption, safer solvent use, reduced energy demand, fewer failed trials, and simpler post-treatment procedures. A laboratory analyst noted, "For us, an efficient synthesis is not only about getting the product. It is better if we use less solvent, avoid repeated heating, and reduce waste from failed samples" (P2). This view shows that green efficiency has a practical meaning in laboratory work. It reflects the daily effort to balance material performance, cost, safety, and environmental responsibility.

The fourth theme concerns cautious trust in data-driven recommendations. Participants generally accepted the potential of data-driven chemistry, but they did not want algorithmic suggestions to replace chemical reasoning. They expected the framework to provide explainable

recommendations that could be checked against theory, laboratory constraints, and material characterization results. One participant said, "If the system suggests a condition, we still need to know why. We cannot just follow numbers without chemical logic" (P9). This statement reflects a critical attitude toward data-based tools. Researchers were willing to use data-driven recommendations when the system supported human judgment, not when it reduced the role of researchers as scientific decision-makers.

The fifth theme points to the need for an adaptive data-driven chemistry framework. Participants emphasized that the framework should not be too complex because many laboratories still face limited instruments, inconsistent documentation, and unequal data literacy. The framework should begin with simple but structured practices, such as standardized synthesis templates, complete parameter records, failed experiment logs, linked characterization data, and green efficiency indicators. Based on the data, the proposed framework consists of five stages: data capture, data standardization, green efficiency mapping, data-driven recommendation, and researcher validation. This framework allows laboratories to improve synthesis efficiency gradually while maintaining scientific control and contextual flexibility.

Discussion

The findings show that the main challenge in developing data-driven chemistry is not the absence of data, but the weak organization and limited reuse of experimental data. This supports Wang et al. (2022), who argued that structured synthesis datasets are essential for data-driven material discovery. In this study, laboratory data existed in many forms, but they were not always ready for systematic analysis. This condition creates a gap between the potential of machine learning and the reality of laboratory documentation. Therefore, the first step toward data-driven chemistry should not always be advanced automation. It should begin with better data discipline, consistent documentation, and clear links between synthesis parameters and material properties.

The dominance of tacit knowledge also confirms that material synthesis remains a human-centered scientific practice. Researchers often depend on experience to adjust reaction parameters, solve unexpected problems, and interpret unclear results. This finding extends the sociotechnical perspective of Stach et al. (2021), which views autonomous experimentation as a system that requires integration between instruments, data, domain expertise, and human collaboration. In the context of this study, a data-driven framework must not eliminate tacit knowledge. Instead, it should help transform tacit knowledge into explicit data that can be discussed, tested, and improved across research teams.

The participants' interpretation of green efficiency offers an important contribution to the literature on sustainable chemistry. Green synthesis is often discussed through formal metrics, such as energy use, waste reduction, and solvent safety. However, the findings show that researchers in daily laboratory practice understand green efficiency through practical indicators, including fewer failed trials, lower reagent use, shorter heating duration, safer procedures, and easier waste handling. This finding supports Leonard and Hasan (2021), who emphasized that artificial intelligence and machine learning can accelerate sustainable chemistry when they are connected to real decision-making problems in chemical practice. In this study, sustainability becomes meaningful when it can guide practical choices in synthesis planning.

The cautious trust shown by participants also provides a critical perspective on the implementation of machine learning in chemistry. Previous studies have shown that machine learning can guide synthesis design, predict reaction pathways, and accelerate material discovery (McDermott et al., 2021; Tang et al., 2020). However, this study shows that researchers need interpretability before accepting data-driven recommendations. They do not reject algorithmic support, but they expect every recommendation to remain chemically reasonable, experimentally feasible, and environmentally justifiable. This finding is relevant to Jain (2024), who noted that interpretability and data quality remain key challenges in machine learning applications for materials research.

The proposed framework also responds to the gap between advanced self-driving laboratory models and the actual readiness of many research laboratories. Studies by Epps et al. (2021), Szymanski et al. (2023), and Tom et al. (2024) show that autonomous and self-driving laboratories can accelerate the discovery of materials and molecules. However, not all laboratories have the infrastructure to adopt full automation. The present study offers a more gradual approach. It positions data-driven chemistry as a scalable framework that can begin with structured documentation and researcher-guided decision-making before moving toward predictive modeling or laboratory automation.

Theoretically, this study contributes to the understanding of data-driven chemistry as a sociotechnical learning process. The framework is not only a technical workflow for managing data. It is also a mechanism for changing how researchers document experiments, evaluate failure, define efficiency, and justify synthesis decisions. This contribution is important because many studies on data-driven material synthesis focus on algorithmic performance, while fewer studies explore the meanings, practices, and experiences of researchers who generate and use the data. By using a qualitative approach, this study shows that the success of data-driven chemistry depends on the interaction between data infrastructure, researcher trust, laboratory routines, and sustainability values.

Practically, the findings suggest that laboratories need to build a stronger data culture before adopting advanced data-driven systems. This can be done by developing standardized synthesis forms, recording failed experiments, linking synthesis logs with characterization results, and including green indicators in every experimental record. Training is also needed so that researchers and students understand how to produce reusable data. Laboratory managers can use the proposed framework to reduce unnecessary repetition, improve resource efficiency, and support environmentally responsible research. For academic institutions, the framework can also strengthen research supervision because students and researchers can make synthesis decisions based on documented evidence rather than informal memory alone.

This study has several limitations. The findings were developed from a qualitative case study, so they do not aim to produce statistical generalization. The framework also depends on the laboratory context, participant experience, and available documentation systems. Future research can test the framework in different types of material synthesis, such as biopolymers, catalysts, nanomaterials, or adsorbent materials. Further studies may also combine qualitative findings with experimental performance data to measure reductions in reagent use, synthesis time, energy consumption, and failed trials. A mixed-method design can strengthen the evidence on how data-driven chemistry improves both laboratory practice and environmental performance.

4. Conclusions

This study concludes that the development of a data-driven chemistry framework can improve the efficiency of environmentally friendly material synthesis by transforming fragmented laboratory records into structured, reusable, and decision-oriented data. The findings show that synthesis efficiency is not only determined by reaction success, but also by the quality of documentation, the reduction of repeated trials, the use of safer materials, lower energy consumption, and the ability to connect synthesis parameters with characterization results. The study also reveals that researchers still rely heavily on tacit knowledge, personal experience, and informal laboratory routines. Therefore, data-driven chemistry should not replace chemical reasoning, but should support researchers in making more transparent, systematic, and environmentally responsible synthesis decisions.

Theoretically, this study contributes to the understanding of data-driven chemistry as a sociotechnical practice that connects human expertise, experimental data, laboratory culture, and sustainability values. Practically, the proposed framework can guide laboratories in developing standardized synthesis records, failed-experiment logs, green efficiency indicators, and data-

based evaluation systems. From a policy perspective, research institutions and higher education laboratories need to strengthen data governance, laboratory documentation standards, and researcher training in sustainable data management. Future studies are recommended to test this framework in different types of material synthesis, such as catalysts, biopolymers, nanomaterials, or adsorbents, and to combine qualitative findings with experimental measurements of energy use, reagent reduction, waste minimization, and synthesis success rates.

References

Abolhasani, M., & Kumacheva, E. (2023). The rise of self-driving labs in chemical and materials sciences. *Nature Synthesis, 2*(6), 483–492. https://doi.org/10.1038/s44160-022-00231-0

Ahmed, S. K. (2024). The pillars of trustworthiness in qualitative research. *Journal of Medicine, Surgery, and Public Health, 2*, 100051. https://doi.org/10.1016/j.glmedi.2024.100051

Baird, S. G., & Sparks, T. D. (2022). What is a minimal working example for a self-driving laboratory? *Matter, 5*(12), 4170–4178. https://doi.org/10.1016/j.matt.2022.11.007

Braun, V., & Clarke, V. (2021). 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

Byrne, D. (2022). A worked example of Braun and Clarke's approach to reflexive thematic analysis. *Quality & Quantity, 56*, 1391–1412. https://doi.org/10.1007/s11135-021-01182-y

Carcary, M. (2020). The research audit trail: Methodological guidance for application in practice. *Electronic Journal of Business Research Methods, 18*(2), 166–177. https://doi.org/10.34190/JBRM.18.2.008

Epps, R. W., Volk, A. A., Ibrahim, M. Y. S., & Abolhasani, M. (2021). Universal self-driving laboratory for accelerated discovery of materials and molecules. *Chem, 7*(10), 2541–2545. https://doi.org/10.1016/j.chempr.2021.09.004

Ghasemlou, M., Nguyen, H. C., Talekar, S., Pfeffer, F. M., & Barrow, C. J. (2025). Artificial intelligence for more sustainable chemistry and a greener future. *ACS Sustainable Chemistry & Engineering, 13*(10), 3830–3833. https://doi.org/10.1021/acssuschemeng.5c00853

Hennink, M., & Kaiser, B. N. (2022). Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Social Science & Medicine, 292*, 114523. https://doi.org/10.1016/j.socscimed.2021.114523

Huang, G., Guo, Y., Chen, Y., & Nie, Z. (2023). Application of machine learning in material synthesis and property prediction. *Materials, 16*(17), 5977. https://doi.org/10.3390/ma16175977

Jain, A. (2024). Machine learning in materials research: Developments over the last decade and challenges for the future. *Current Opinion in Solid State and Materials Science, 33*, 101189. https://doi.org/10.1016/j.cossms.2024.101189

Leonard, K. C., & Hasan, M. F. (2021). Can artificial intelligence and machine learning be used to accelerate sustainable chemistry and engineering? *ACS Sustainable Chemistry & Engineering, 9*(18), 6126–6129. https://doi.org/10.1021/acssuschemeng.1c02741

McDermott, M. J., Dwaraknath, S. S., & Persson, K. A. (2021). A graph-based network for predicting chemical reaction pathways in solid-state materials synthesis. *Nature Communications, 12*, 3097. https://doi.org/10.1038/s41467-021-23339-x

McKim, C. (2023). Meaningful member-checking: A structured approach to member-checking. *American Journal of Qualitative Research, 7*(2), 41–52. https://doi.org/10.29333/ajqr/12973

Na, G. S. (2023). Artificial intelligence for learning material synthesis processes of thermoelectric materials. *Chemistry of Materials, 35*(19), 8272–8280. https://doi.org/10.1021/acs.chemmater.3c01834

Naeem, M., Ozuem, W., Howell, K., & Ranfagni, S. (2023). A step-by-step process of thematic analysis to develop a conceptual model in qualitative research. *International Journal of Qualitative Methods, 22*, 16094069231205789. https://doi.org/10.1177/16094069231205789

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

Rooney, M. B., MacLeod, B. P., Oldford, R., Thompson, Z. J., White, K. L., Tungjunyatham, J., Stankiewicz, B. J., & Berlinguette, C. P. (2022). A self-driving laboratory designed to accelerate the discovery of adhesive materials. *Digital Discovery, 1*(4), 382–389. https://doi.org/10.1039/D2DD00029F

Stach, E., DeCost, B., Kusne, A. G., Hattrick-Simpers, J., Brown, K. A., Reyes, K. G., Schrier, J., Billinge, S. J. L., Buonassisi, T., Foster, I., Gomes, C. P., Gregoire, J. M., Mehta, A., Montoya, J., Olivetti, E., Park, C., Rotenberg, E., Saikin, S. K., Smullin, S., & Maruyama, B. (2021). Autonomous experimentation systems for materials development: A community perspective. *Matter, 4*(9), 2702–2726. https://doi.org/10.1016/j.matt.2021.06.036

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 novel materials. *Nature, 624*(7990), 86–91. https://doi.org/10.1038/s41586-023-06734-w

Tang, B., Lu, Y., Zhou, J., Chouhan, T., Wang, H., Golani, P., Xu, M., Xu, Q., Guan, C., & Liu, Z. (2020). Machine learning-guided synthesis of advanced inorganic materials. *Materials Today, 41*, 72–80. https://doi.org/10.1016/j.mattod.2020.06.010

Tom, G., Schmid, S. P., Baird, S. G., Cao, Y., Darvish, K., Hao, H., Lo, S., Pablo-García, S., Rajaonson, E. M., Skreta, M., Yoshikawa, N., Corapi, S., Akkoc, G. D., Strieth-Kalthoff, F., Seifrid, M., & Aspuru-Guzik, A. (2024). Self-driving laboratories for chemistry and materials science. *Chemical Reviews, 124*(16), 9633–9732. https://doi.org/10.1021/acs.chemrev.4c00055

Wang, Z., Kononova, O., Cruse, K., He, T., Huo, H., Fei, Y., Zeng, Y., Sun, Y., Cai, Z., Sun, W., & Ceder, G. (2022). Dataset of solution-based inorganic materials synthesis procedures extracted from the scientific literature. *Scientific Data, 9*, 231. https://doi.org/10.1038/s41597-022-01317-2

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), 2104113. https://doi.org/10.1002/adma.202104113

Williamson, E. M., & Brutchey, R. L. (2023). Using data-driven learning to predict and control the outcomes of inorganic materials synthesis. *Inorganic Chemistry, 62*(40), 16251–16262. https://doi.org/10.1021/acs.inorgchem.3c02697