c&en | WHITE PAPERS

PRACTICAL AI IN CHEMISTRY AND MATERIALS R&D

EXECUTIVE SUMMARY

Scientific advances, whether in energy, electronics, or manufacturing, depend on new materials, formulations, and chemical processes. Companies that undertake this innovation are under constant pressure to adhere to new regulations, improve safety, reduce environmental impact, lower costs, improve performance, accelerate time to market, and increase market share.

Chemical and materials informatics are emerging as essential tools in the modern product developer's toolbox. These applications use machine learning and data analytics to synthesize data, uncover hidden trends, and optimize product development for enhanced efficiency. Chemical and materials informatics can complement laboratory experiments with in silico experiments to screen more ideas than could be performed manually.

However, keeping pace with the breadth of emerging models, techniques, and applications is challenging for scientists whose primary responsibility is conducting laboratory experiments. Using a chemical and materials informatics platform to integrate these tools provides scientists with access to the best of modern machine learning and artificial intelligence. This white paper examines how NobleAI's platform empowers scientists to access the power of chemical and materials informatics for formulation and molecular design workflows.

INTRODUCTION

Chemical companies, as well as consumer and industrial product creators, are under mounting pressure to reformulate products in response to a mix of market demands, regulatory changes, and global supply chain disruptions.[1] In the personal care sector, for example, reformulation is increasingly driven by the consumer demand for ingredient transparency, natural alternatives, ethical sourcing, and improved sustainability profiles. The global market for

sustainable personal care products is estimated to reach $90.4 billion by 2032, which is approximately 1.7× the market size of this segment in 2024.[2]

Traditional approaches to reformulation are notoriously cumbersome and time-consuming, and there is a need for a better way to explore the vast formulation design space—one that remains scientifically grounded and data driven while being adaptable to ongoing uncertainty. Although mathematical modeling has been used for centuries to model physical phenomena and to inform better approaches to experimentation, many physical properties are not linear and do not conform to simple mathematical expressions.

Quantitative structure–activity relationship (QSAR)[3] models, for example, can correlate a bioconcentration factor against a handful of descriptors like the logarithm of the partition coefficient (logP), but it is rare that only a few descriptors need to be considered. Similarly, traditional design-of-experiment (DoE) approaches fit polynomial surfaces to outcomes like stability, viscosity, or cleaning efficiency, but they quickly break down as the number of features, interactions, and nonlinearities multiply.

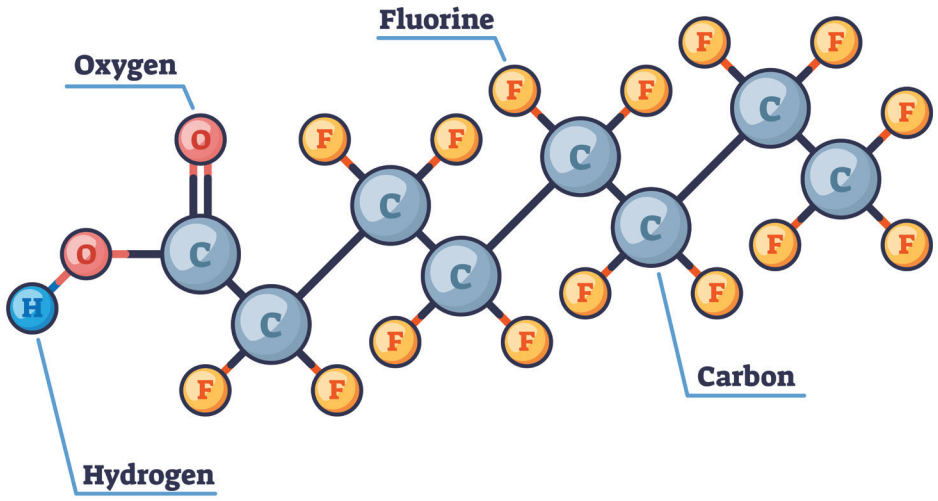
Beyond reformulation, chemists often design novel molecules to meet their needs when developing new materials or replacing ingredients in existing products. It is critical that chemists can predict the physicochemical properties of the individual compound as well as its impact when added to a complex formulation. However, these estimates often require considerable trial and error, leading to wasted time and resources.

These limitations highlight why faster, more flexible approaches are required when addressing longstanding compounds that demand immediate replacement or when developing new compounds and products.

AI, a tool that can identify patterns in datasets, is combined with scientific expertise, proving highly effective for reengineering chemical products. It enables more efficient reformulation and molecular design by accurately predicting molecular properties, guiding experimental design, and uncovering novel candidates that will meet performance, safety, and supply requirements. Chemists can greatly improve their effectiveness using a digital experimentation approach.

A SENSE OF URGENCY TO CHANGE

The need for reformulation is especially important for previously accepted chemicals that have recently been deemed unsafe. Per- and polyfluoroalkyl substances (PFAS) are a class of thousands of synthesized chemicals that are found in everything from consumer to industrial products. These “forever chemicals” have been widely used in industry because of their stability in applications ranging from cookware coatings to firefighting foams. The characteristic carbon-fluorine bonds that make these chemicals so stable also makes them slow to degrade in the environment and in the human body.



Example molecular structure of PFAS.

Credit: Shutterstock

Regulatory pressure to eliminate PFAS is increasing worldwide, with bans and restrictions taking effect across Europe, North America, and parts of Asia.[4,5] Despite industry pushback, manufacturers must identify and design replacements for these substances while also ensuring that the alternatives are safer and more effective over the long term. "The impact of PFAS has been astronomical," says Kyle Snow, a solutions engineer at NobleAI. "There is the human health risk, but the expenses around remediation efforts have also been tremendous. There are legal penalties, supply chain and development costs, and lost revenue and credibility concerns for companies trying to pivot."

PREDICTIVE MODELING AND SCIENCE-BASED AI

Predictive modeling is an AI-based approach that uses algorithms to simulate the behavior of molecules under specific conditions. This technique enables researchers to predict key properties such as solubility, reactivity, and toxicity — all of which are critical aspects of product reformulation. By generating a database of physicochemical properties computationally, researchers can screen thousands of candidate molecules in silico before committing to time-consuming and costly lab experimentation.

For example, in replacing PFAS, predictive modeling can prioritize molecules that are structurally distinct from known hazardous compounds and that deliver key functional traits, such as hydrophobicity or thermal stability. AI can predict how new compounds will perform in a mixture or under various manufacturing conditions, which are essential considerations in real-world applications.

Chemists need to have confidence that the predictive models they are using for formulation and molecular design applications align with scientific knowledge. Modern science-based AI (SBAI) models improve on past models by using a variety of techniques to build more scientific information into prediction algorithms. This can take many forms, such as incorporating detailed chemical

descriptors, designing model architectures that respect physical constraints and symmetries, and employing hierarchical approaches to add characterization data (such as thermogravimetric curves) into models. When done correctly, this fusion of scientific information can provide a deeper mechanistic understanding of systems and improve accuracy.

Once scientists have a model trained with relevant data and principles, they need structured workflows to guide their reformulation or molecular design campaign. These workflows can range from simple queries (for example, "What if I substitute a methyl for a butyl group?") to constrained parameter sweeps, combinatorial searches across large design spaces, and iterative optimization cycles. However, implementing these workflows can require significant effort, including tracking data changes, training new models, crafting bespoke optimizations, and analyzing numerous experimental candidates.

No-code platforms such as NobleAI's NobleVIP help orchestrate these workflows, while graphs and other data visualization tools enable the identification of relationships that might otherwise remain hidden. This cycle of model training, candidate generation, visualization, and retraining creates a self-improving "data flywheel," explains Kevin Shen, an AI scientist at NobleAI.

If a candidate compound fails to meet performance specifications in the lab, these new data can be fed back to train updated models that suggest structural modifications to improve results. This iterative, hypothesis-driven capability yields additional insights, which are particularly valuable when researchers are operating in constrained or emerging data landscapes.

FORMULATION OPTIMIZATION FOR NONFLUORINATED BATTERY ELECTROLYTES

NobleAI challenged its SBAI modeling tools to design nonfluorinated electrolyte systems for improved performance and reduced environmental impact of lithium-metal batteries. They developed a model of coulombic efficiency trained on a combination of open and curated electrolyte data and applied a special model architecture to predict new formulations with new chemical ingredients.[7]

Figure 1 shows a visualization of the full performance landscape using a constrained variable sweep that adjusts the ratio of two solvents as a function of salt concentration. The heat map uses prior data to predict a performance landscape that would be challenging to intuitively predict.

Within this performance landscape, there are still too many possible combinations of candidate solvents, electrolytes, and additives to be studied manually, especially when considering novel chemicals or combinations. However, NobleAI's algorithm can predict the performance of new chemicals and their combinations, unlike traditional DoE models that allow researchers to search over combinations using only chemicals with existing formulation performance data. NobleAI's software also uses a proprietary formulation

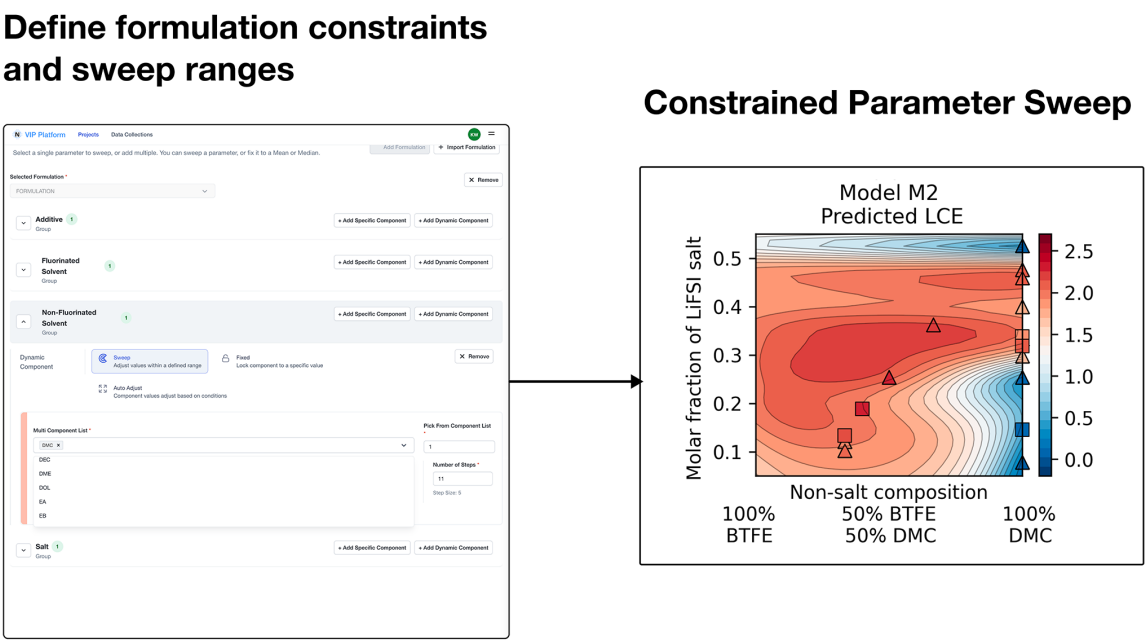
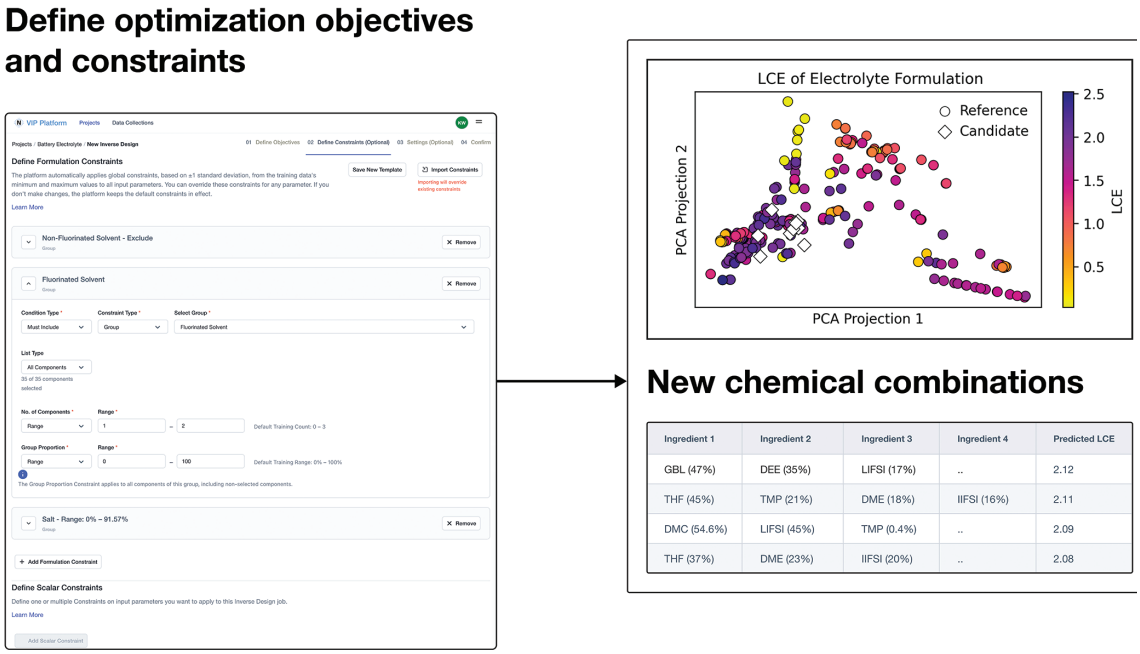


Figure 1. Constrained formulation parameter sweep from NobleVIP. Left panel shows input into the programs and the right panel shows the output.



Ingredient 1	Ingredient 2	Ingredient 3	Ingredient 4	Predicted LCE
GBL (47%)	DEE (35%)	LIFSI (17%)	..	2.12
THF (45%)	TMP (21%)	DME (18%)	IIFSI (16%)	2.11
DMC (54.6%)	LIFSI (45%)	TMP (0.4%)	..	2.09
THF (37%)	DME (23%)	IIFSI (20%)	..	2.08

Figure 2. Formulation inverse design and optimization screens from the NobleVIP platform can cover a larger chemical space than can typically be considered in the laboratory.

optimizer that can efficiently explore this large chemical combination and formulation composition space.

Using this formulation optimization function, the NobleAI team was able to identify promising electrolyte systems that do not use fluorinated solvents yet achieve promising performance.

Figure 2 demonstrates how a user might define a search highlighting nonfluorinated solvents and shows the resulting exploration map of new formulations and new solvent combinations. The map selects for formulations and combinations that have not yet been reported in literature but are likely to show good performance.

MOLECULAR DESIGN

In addition to formulation predictions, researchers can use NobleVIP for molecular design. These algorithms must account for the discrete nature of molecules, which prevents continuous optimization and makes small property changes difficult to trace. Drug discovery experts have addressed these challenges with foundational modeling techniques. NobleAI has adapted these methods for the broader chemical industry, focusing on:

- Molecule generation seeded from domain-specific distributions
- Simpler, economically viable synthetic pathways
- Designs guided by multiple physicochemical properties

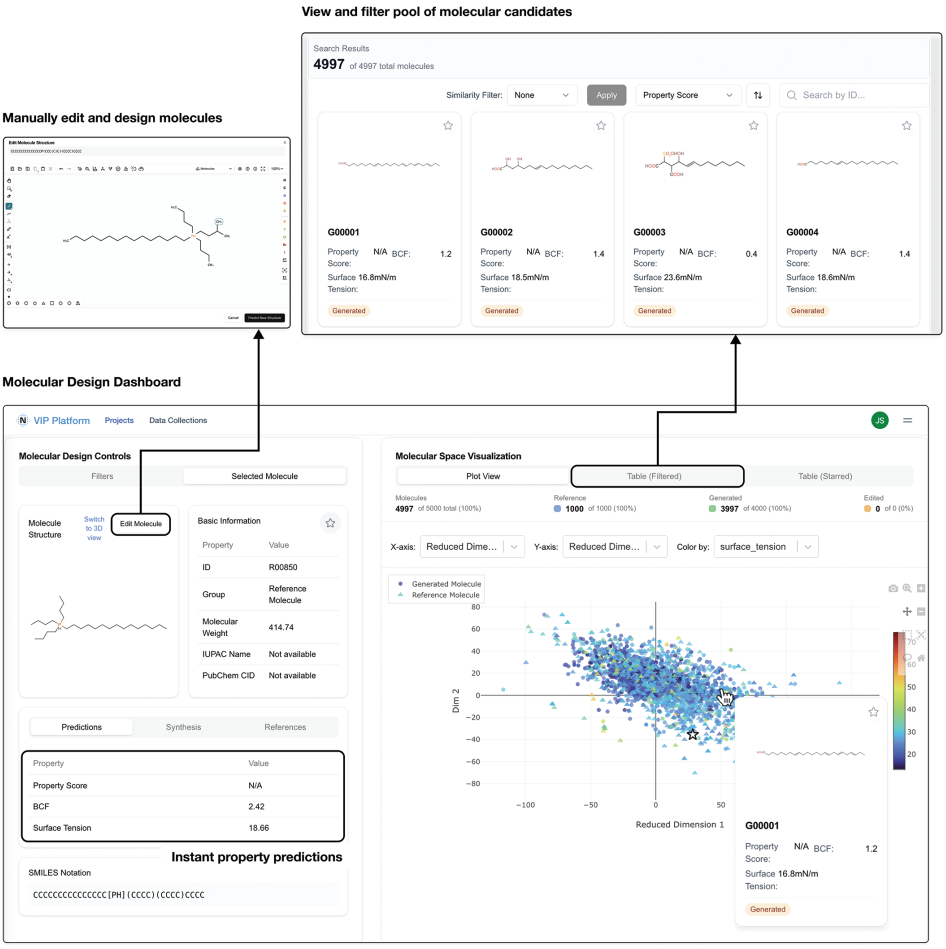


Figure 3. The NobleAI platform features molecular design tools, property predictions, and navigation tools that enable searching and filtering by chemical motifs, similarity, and property performance.

NobleVIP includes a customizable molecular design module. Users can make single-point predictions by editing or drawing molecules directly. When users' ideas prove insufficient, NobleVIP employs molecular generation algorithms guided by machine learning property models to propose promising candidates. It also offers tools to visualize and filter molecules—via chemical maps, tables, and filters based on functional groups, predicted properties, synthesizability, and chemical diversity.

CASE STUDY: ICL GROUP'S PARTNERSHIP WITH NOBLEAI

ICL Industrial Products, a global leader in agriculture, food, and engineered materials, partnered with NobleAI to accelerate the discovery and development of high-performing, environmentally sustainable next-generation flame retardants.

Developing replacement flame retardants that meet flammability standards and performance targets while reducing health and environmental risks is a complex, multidimensional challenge. ICL's goal was to develop novel flame-retardant chemistries that would not only comply with current and anticipated regulations but also offer improved sustainability and hazard profiles.

However, traditional approaches to materials discovery, which rely heavily on empirical testing, proved too slow and limiting for the ICL team. By working with NobleAI, ICL researchers aimed to extract knowledge from prior patents and published literature, thereby de-risking their search for new flame-retardant chemicals.

"Novel molecular design often has challenges with constraining its generation to 'reasonable' chemical structures, and the techniques are often not readily accessible to non-programmers or non-data scientists," says Shen. By integrating first-principles chemistry, custom tools can convert chemical data into a format that is usable for machine learning algorithms using flexible AI modeling architectures. This enabled ICL's scientists to explore thousands of chemical structures. The structures included ones that had not previously been evaluated for flame retardancy. The NobleAI molecular generation algorithm also proposed many new molecular ideas that the ICL team hadn't previously considered, yet found interesting enough to synthesize and test, thereby augmenting what their team would have explored otherwise.

For one of the molecular properties, ICL and NobleAI started with a prebuilt model based on data available in published literature. Through proprietary simulations, they improved the model's predictive accuracy on flame retardants by over 25%. As a result, the model could provide higher precision in identifying promising candidates during rapid screens of large candidate molecule pools. Overall, the NobleVIP platform streamlined molecular candidate identification and selection. The platform guided scientists toward more promising compound families, many of which are now undergoing further development and testing.

"ICL is in constant pursuit of new molecular formulations to help improve the performance and sustainability of our flame retardants," Yaniv Kabalek, president of the Industrial Products division at ICL, said in a recent press release. "By partnering with NobleAI, we can reduce wasted resources and speed time to market, while continuing to deliver the safest, highest performing products to our customers."

This collaboration between NobleAI and ICL accelerated ICL's molecular exploration and material innovation process and demonstrated that SBAI can enhance discovery, research, and development.

GETTING STARTED: ADOPTING SBAI FOR YOUR OWN REFORMULATION CHALLENGES

When considering chemical and informatics tools for their formulation and molecule- design workflows, one of the most common questions organizations ask is "What data will power the models?" NobleAI scientists suggest that the fastest path to gathering the data needed to train these models is to start now and start small. Best data practices will coevolve with modeling techniques.

Working with the right partner can accelerate modeling projects by providing understanding and insight — and helping scientists adopt digital experimentation practices. Partnering with NobleAI's expert AI scientists, chemists, and materials scientists and using the NobleVIP platform allows companies to:

- Accelerate collaboration and knowledge flow between digital and R&D teams
- Identify and launch the highest-impact projects to prove value quickly
- Define and align on key performance indicators that tie directly to business outcomes
- Build a data strategy that is practical, focused, and results driven

Connect with NobleAI to empower your scientists to derisk innovation, discover faster and more efficiently, and create new markets.

ABOUT NOBLEAI

NobleAI offers commercially-proven AI solutions for materials informatics powered by its unique Science-Based AI (SBAI) technology. SBAI models are developed quickly, securely and specifically for each customer and each specific use case. Via the cloud-based Visualizations, Insights, and Predictions (NobleVIP) platform, NobleAI technology delivers actionable insights to accelerate product development and reduce costs while improving product performance, sustainability, and reliability. NobleAI is supported by investments from world-class organizations such as Microsoft, Chevron, and Syensqo, and the company's solutions are already delivering real value in production deployments at leading chemical, materials, and energy companies around the globe.

REFERENCES:

1. Andreas Gocke et al., *Facing Unprecedented Pressures: Value Creation in Chemicals 2023* (Boston Consulting Group, Oct. 31, 2023), https://www.bcg.com/publications/2023/resilience-chemicals-industry.
2. Sustainable Personal Care Products Market Size, Share Market & Industry Analysis, (Fortune Business Insights, FBI110903, June 30, 2025), https://www.fortunebusinessinsights.com/sustainable-personal-care-products-market-110903.
3. Eugene N. Muratov et al., "QSAR Without Borders," *Chem. Soc. Rev.*, 49, no.11 (2020): 3525–64, https://doi.org/10.1039/D0CS00098A.

4. "World's Biggest Chemical Manufacturers Called On to Phase Out PFAS" (Ricardo, Feb. 7, 2023), https://www.ricardo.com/en/news-and-insights/industry-insights/world-s-biggest-chemical-manufacturers-called-on-to-phase-out-pfas.
5. Sheridan Wood, "2025 Outlook: Asia-Pacific Regulatory Landscape" (3E, Mar. 13, 2025), https://www.3eco.com/article/2025-asia-pacific-regulatory-landscape-3e/.
6. "Flame Retardants and Your Health," National Institute of Environmental Health Sciences, accessed April 2023, https://www.niehs.nih.gov/health/topics/agents/flame_retardants#:~:text=Brominated%20flame%20retardants%20(BFRs)%20contain,also%20show%20toxic%20endocrine%20effects.
7. Ying Shi Teh, "Electrolyte Formulation Design for Lithium Metal Batteries with AI" (presentation, Cambridge EnerTech's Battery Intelligence: Big Data, Machine Learning, and Artificial Intelligence Optimizing Battery Performance, 41st International Battery Seminar and Exhibit, Orlando, FL, March 12–15, 2024), https://www.internationalbatteryseminar.com/24/battery-intelligence.

AMERICAN CHEMICAL SOCIETY
1155 Sixteenth Street, NW,
Washington, DC 20036