



# CASC Newsletter

Volume 17 | May 2026

## IN THIS ISSUE:

- [From the Director](#)
- [Collaborations: Spectral-Refiner for Turbulent Flow Simulation](#)
- [Lab Impact: Virtual Reality for Faster and Safer Decisions in Advanced Manufacturing](#)
- [Advancing the Discipline: Watching the Watchers: How to Trust the Tools Used to Establish Trust](#)
- [Machine Learning & Applications: Using AI to Turbocharge Chemistry](#)

---

## From the Director

Contact: [Kathryn Mohror](#)

It is hard to believe that it has already been a year since I stepped into the Division Lead role for CASC. It has been a busy and rewarding year, and I continue to be impressed by the talent, commitment, and collaborative spirit across the division. CASC has been home for me since I joined LLNL in 2010, and I am grateful to be part of this team and look forward to what we will do together in the coming year.

The research highlighted in this issue captures examples of how CASC is pushing science and engineering forward at LLNL, turning cutting-edge ideas into tools people can actually use. We start with Spectral-Refiner, which makes neural-operator turbulence prediction far more accurate and efficient, so high-fidelity flow modeling can be faster and more practical. Then we jump into immersive virtual reality (VR) for advanced manufacturing, where teams can meet inside a shared digital twin to inspect geometry, computed tomography (CT), and process data together, and even use the same platform for safer, hands-on training for high-consequence procedures like lockout-tagout (LOTO).

We also dig into the essential question of trust, building confidence in the software that checks other software through a verified G-code simulator and robust, uncertainty-aware analysis methods. Finally, we look at AI for chemistry and biology, from the BOOM benchmark that stresses models in realistic out-of-distribution settings to a

---

supercomputer-optimized pipeline that has already predicted structures for tens of millions of proteins. Together, these stories show CASC helping the Department of Energy (DOE) mission by speeding up discovery, strengthening reliability, and bringing powerful new capabilities to the Lab.

---

## Collaborations | Spectral-Refiner for Turbulent Flow Simulation

Contact: [Rui Peng Li](#)

Modeling the complex behavior of turbulent fluid flows is a long-standing challenge in science and engineering. Traditional numerical simulations can capture these flows, but they are computationally expensive and time consuming. Neural operators (NO) have become a powerful tool for approximating solutions to complex partial differential equations (PDEs), such as the Navier-Stokes equations.

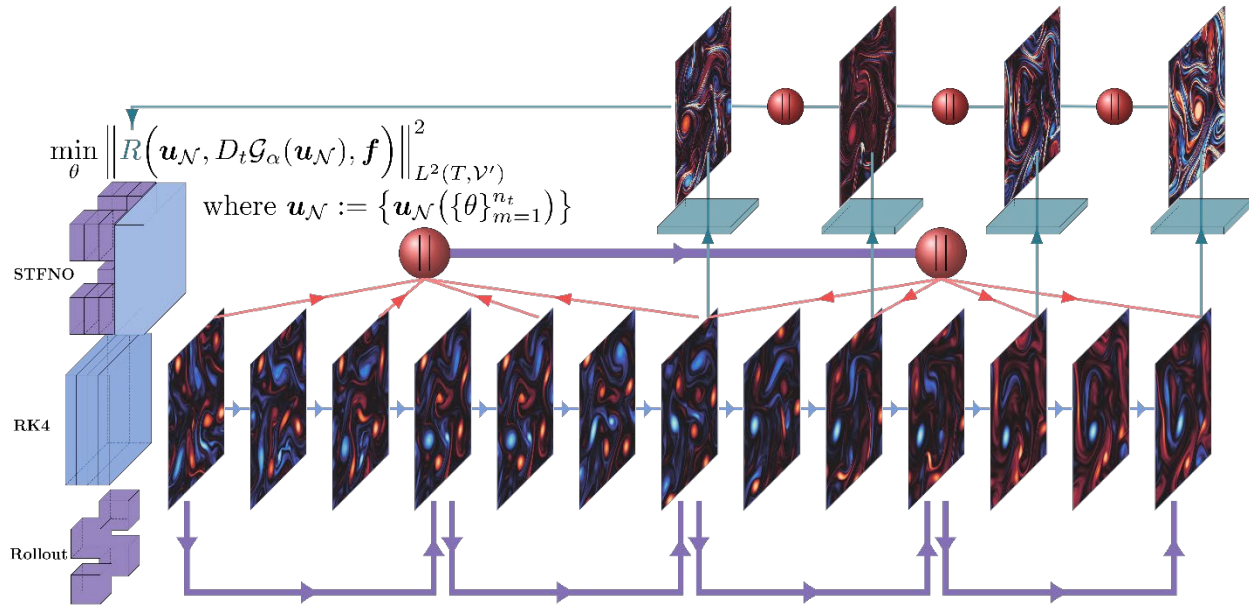
However, despite their promise, many existing operator learning approaches suffer from two major challenges: the need for extensive training to achieve modest accuracy, and the lack of theoretical reliability in their predictions over long time horizons. To address these limitations, CASC researcher Rui Peng Li and collaborators at University of Missouri-Kansas City and Emory University propose a novel training and fine-tuning framework that blends traditional numerical insights with modern neural architectures, enabling more accurate and computationally efficient learning for turbulent fluid dynamics [1].

At the core of this work is the spatiotemporal Fourier neural operator (ST-FNO), a new generalization of Fourier neural operators. Unlike earlier FNO variants that treat time as a series of fixed-length input-output pairs, the ST-FNO learns to map entire temporal trajectories, enabling it to perform *zero-shot super-resolution* across both space and time dimensions. This means the model can make long-range predictions over arbitrary time intervals with a single forward pass, rather than relying on autoregressive rollouts that suffer from accumulated error and stability issues.

What makes Spectral-Refiner especially effective is its hybrid fine-tuning strategy. Instead of training the full network end-to-end for hundreds of epochs, the model undergoes a short initial training phase (e.g., 10 epochs), after which most parameters are frozen. Only the final spatiotemporal spectral convolution layer—responsible for converting latent representations into high-resolution predictions—is further optimized.

This fine-tuning step is guided by a new loss function based on a negative Sobolev norm (computed via inverse fast Fourier transform), marking the first use of such a functional norm in NO learning. Remarkably, this error estimator does not require the true solution to

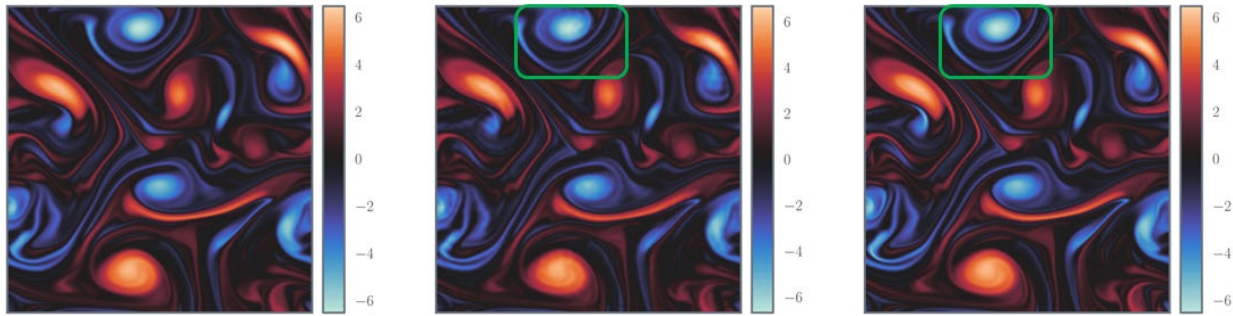
the PDE or even high-accuracy numerical simulations. Instead, it leverages *a posteriori* error estimation theory, using auto-differentiable solvers like Runge-Kutta, to refine the model output directly in the frequency domain.



**Figure 1:** Schematic differences between approaches. ST-FNO: spatiotemporal FNO with hybrid fine-tuning; large time steps; yielding arbitrary-length temporal prediction in a single evaluation; parallel-in-time optimization for spectral fine-tuning. RK4: fourth-order Runge-Kutta; small time steps bounded by the Courant-Friedrichs-Lewy condition; de-aliasing filter needed. Rollout: autoregressive NO rolling-out; using previous evaluation as input repetitively; large time steps; no stability guarantees.

This new approach yields impressive results. On standard turbulence benchmarks, including the Taylor-Green vortex and isotropic turbulence datasets, the Spectral-Refiner framework achieves up to  $10^5$  improvement in accuracy over baseline FNO models as well as significant gains in computational efficiency compared to traditional solvers. Furthermore, it generalizes well to out-of-distribution settings such as different Reynolds numbers or initial conditions, and it outperforms more complex methods like PDE-Refiner, all without needing to train auxiliary networks or engage in costly multi-step rollouts.

In addition to its technical contributions, the researchers provide a reproducible pipeline for the community. A PyTorch implementation is released on [GitHub](#), alongside curated datasets available at [Hugging Face](#). These tools are designed to facilitate further research in high-fidelity, data-driven simulations for scientific computing.



**Figure 2:** he ground truth versus the ST-FNO3d evaluation/fine-tuning results for the famous decaying turbulence (McWilliams) at  $Re=5,000$ . Left: the ground truth at  $t=5$  (prediction step 20) generated by IMEX RK4. Middle: ST-FNO3d prediction. Right: ST-FNO3d fine-tuned prediction. A region where the ST-FNO solution is more accurate, as can be visually observed, appears inside the green rectangles.

[1] S. Cao, F. Brarda, R.P. Li, and Y. Xi. “Spectral-Refiner: accurate fine-tuning of spatiotemporal Fourier neural operator for turbulent flows.” In *The Thirteenth International Conference on Learning Representations*, 2025. [doi.org/10.48550/arXiv.2405.17211](https://doi.org/10.48550/arXiv.2405.17211).

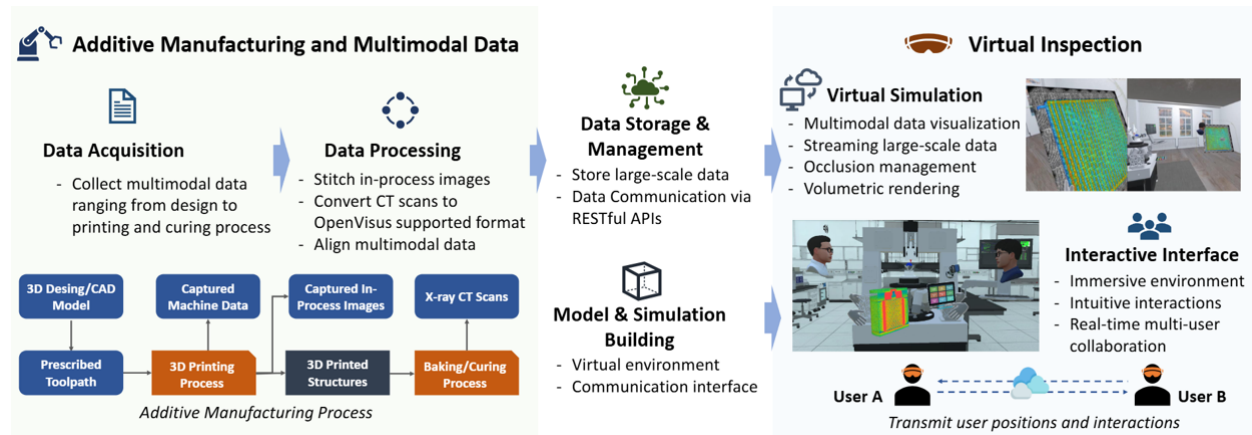
## Lab Impact | Virtual Reality for Faster and Safer Decisions in Advanced Manufacturing

Contact: [Haichao Miao](#)

Advanced manufacturing is a complex and geographically distributed process that relies on a data-driven procedure to certify the created parts. Inspecting complex parts and their build histories takes time, requires multiple experts, and often happens after problems become expensive. The LDRD-SI Digital Twins project, which CASC researcher Haichao Miao co-led, tackled this head-on by bringing digital twins of additively manufactured (AM) parts into a shared virtual environment. The result is a collaborative “war room” where materials scientists, nondestructive testing experts, and analysts can co-inspect geometry, CT data, and process telemetry together—whether they’re across the hall or across the country.

The researchers developed a pipeline that ingests part geometry and multimodal data (e.g., CT volumes, G-code, in-process images), aligns them in a single coordinate frame, and streams multiresolution views into a multiuser VR workspace. Participants can point, measure, slice volumes, and leave anchored annotations that persist back to the digital thread. Under the hood, they leverage GPU-accelerated rendering, level-of-detail volume streaming (supported by [OpenViSUS](#)), and advanced visualization and interaction

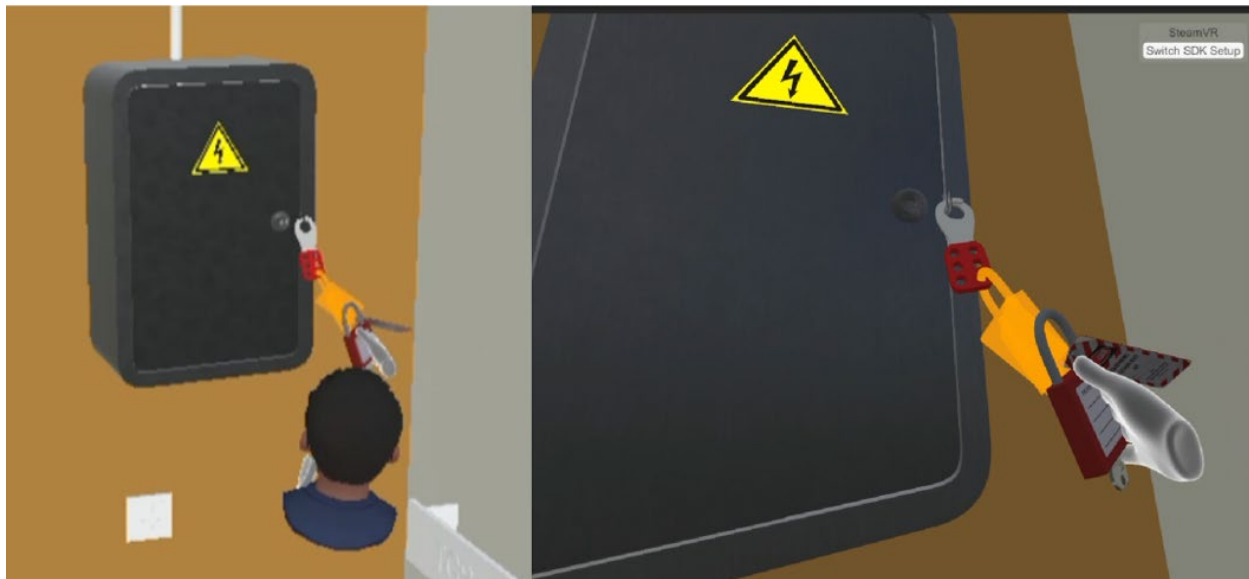
techniques so that users literally “sees the same thing” with low latency, even when datasets are large.



**Figure 3:** A virtual environment loads multimodal data from the AM process and allows distributed users to interact and inspect the digital twin. The main developer of this VR system is Vuthea Chheang, who is currently an assistant professor at San Jose State University.

The first and obvious advantage of VR is that it simulates a physical environment, where breaking simulated objects is much less expensive than breaking real ones. In addition, inspection and training in AM are inherently spatial as digital twins represent real-world objects. Hence, users must reason about complex 3D geometry, internal CT volumes, and process histories that are easy to misread on flat screens but are intuitive to do using VR with true scale, depth cues, and natural hand-eye interaction. In a shared virtual environment, remote and onsite users can co-locate around the same part, point, slice, measure, and leave anchored notes that persist to the digital thread, creating common ground and speeding consensus on defects and root causes.

Since VR is still an emergent technology, designing the interfaces requires careful consideration of human-computer interaction. The researchers collaborated with [Prof. Doug Bowman](#) and his 3D Interaction group at Virginia Tech to use a human-centric approach to develop and evaluate the visualization and interaction techniques. Beyond developing the actual virtual environment, they contributed methods for (1) co-registering large-scale AM data using progressive refinement [1,2]; (2) scalable rendering and interaction techniques for large CT volumes [3]; and (3) methods of remote collaboration to visualize multimodal AM data [4,5].



**Figure 4:** VR for safety-critical training lets trainees practice high-consequence procedures (LOTO) in a risk-free environment, while promising to increase performance and knowledge transfer.

An exciting offshoot of this work applies to the same immersive infrastructure to safety training, starting with LOTO procedures around high-energy equipment. VR lets trainees practice high-consequence procedures in a life-size, risk-free environment, building muscle memory without exposing people or equipment to actual danger. It also enables standardized, repeatable scenarios with consequence simulation and objective performance tracking, so teams can verify competence and close gaps before anyone touches a real machine. Trainees rehearse the exact steps to isolate energy sources, verify zero-energy state, and tag equipment in an interactive scenario that mirrors their facilities.

The early-stage effort in collaboration with experts in LLNL’s Environment, Safety & Health organization will support critical training with built-in assessment to make sure personnel have been sufficiently trained for these high-consequence scenarios. Besides the technical objectives to improve procedural recall and cut instructor time per trainee, the ultimate goal is to help protect personnel from injuries.

[1] L. Pavanatto, A. Giovannelli, B. Giera, P.-T. Bremer, H. Miao, and D. Bowman. “Exploring multiscale navigation of homogeneous and dense objects with progressive refinement in virtual reality.” *IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, pp. 228–237, 2025. [doi.org/10.1109/VR59515.2025.00047](https://doi.org/10.1109/VR59515.2025.00047).

[2] F. Rodrigues, A. Giovannelli, L. Pavanatto, H. Miao, J. de Oliveira, and D. Bowman. “AMP-IT and WISDOM: improving 3D manipulation for high-precision tasks in virtual reality.” *IEEE International*

---

*Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 303–311, 2023.  
[doi.org/10.1109/ISMAR59233.2023.00045](https://doi.org/10.1109/ISMAR59233.2023.00045).

[3] V. Chheang, B. Weston, R. Cerda, B. Au, B. Giera, P.-T. Bremer, and H. Miao. “A virtual environment for collaborative inspection in additive manufacturing.” *CHI EA '24: Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, pp. 1–7, 2024.  
[doi.org/10.1145/3613905.3650730](https://doi.org/10.1145/3613905.3650730).

[4] V. Chheang, S. Narain, G. Hooten, R. Cerda, B. Au, B. Giera, P.-T. Bremer, and H. Miao. “Enabling additive manufacturing part inspection of digital twins via collaborative virtual reality.” *Nature Scientific Reports*, 14:29783, 2024. [doi.org/10.1038/s41598-024-80541-9](https://doi.org/10.1038/s41598-024-80541-9).

[5] A. Giovannelli, L. Pavanatto, S. Davari, H. Miao, V. Chheang, B. Giera, P.-T. Bremer, and D. Bowman. “Investigating the influence of playback interactivity during guided tours for asynchronous collaboration in virtual reality.” *IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, pp. 23–33, 2025.  
[doi.org/10.1109/VR59515.2025.00027](https://doi.org/10.1109/VR59515.2025.00027).

---

## Advancing the Discipline | Watching the Watchers: How to Trust the Tools Used to Establish Trust

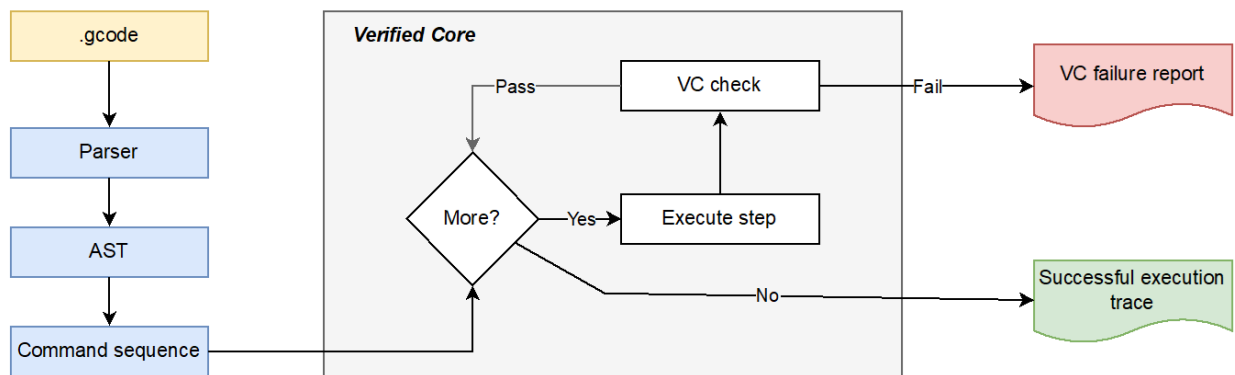
Contact: [Matthew Sottile](#)

Software often is used to check that hardware, (other) software, and data meet a set of requirements that define what it means to be “correct.” When these correctness checkers are playing a role in making decisions where the consequences of errors are unacceptable, one ends up with a somewhat circular problem: How to check these checkers?

The “Watching the Watchers” LDRD-ER project, led by CASC researcher Matthew Sottile, was initiated to study this problem: What techniques can be used to build an assurance case for the tools used to build assurance cases? This research involves considering this problem from a few perspectives: What assurance level is required? What tradeoffs come with different technical approaches to solving this problem? What are the pragmatic considerations of putting these techniques into practice?

At the start of the project, the researchers looked broadly at a few use cases where assurance tools play an important role. These included structured prompting tools for reliable usage of large language models (LLMs) [1]; model-checking techniques for reasoning about concurrency control correctness of C++ programs; and checking different stages in the AM workflow. They focused the bulk of the project on the AM problem area because they identified several points where assurance is necessary but is poorly covered in existing tools and techniques.

The AM workflow has multiple points where the researchers can assess if a design meets the application requirements. Given a design to print, a tool-path planner (or slicer) generates a program in a low-level command language (“G-code”) that is executed by the AM system. This command program is executed by the printer to fabricate the physical object that is then assessed against its requirements via some combination of in-situ measurements during printing and ex-situ measurements afterwards. These last two phases were the primary focus: How to build G-code simulators that correctly model the execution of the tool path, and how to approach the analysis of data that were collected during and after fabrication?

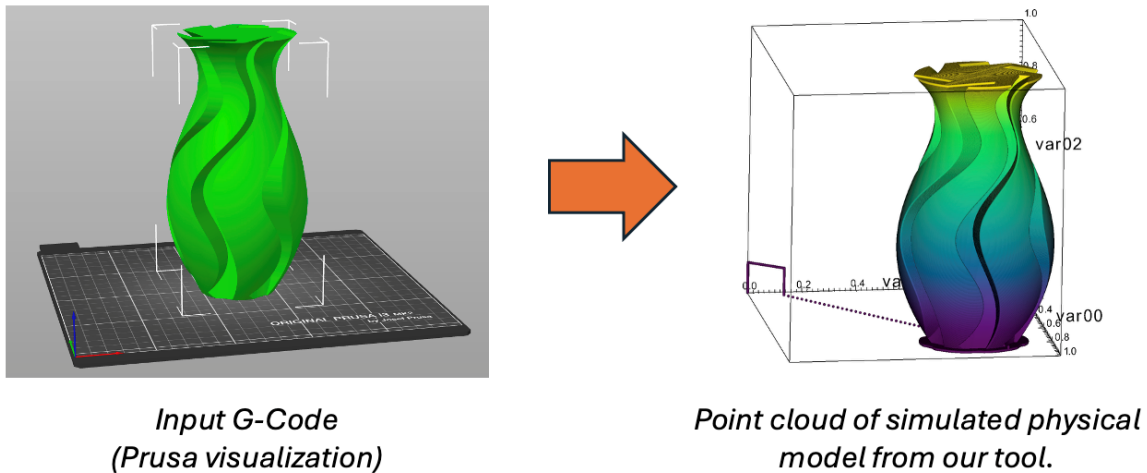


**Figure 5:** Structure of verified G-code simulator. AST = abstract syntax tree; VC = verified core.

To address the G-code simulation problem, the researchers took an approach inspired by formal verification of interpreters and compilers [2,3]. The design of the simulator is shown in Figure 5. They formally specified the semantics of each command in the G-code language in the Rocq proof assistant. Each command was then implemented as a function that modifies a state model representing the logical and physical state of the printer. This allowed them to write proofs in the proof assistant to ensure that these implementations adhered to the semantics defined by the G-code language. These proofs took approximately 6 months to complete and spanned approximately 15,000 lines of code with a large amount of the complexity arising from the use of floating point in the simulator logic.

They were able to obtain a correct-by-construction implementation in the Ocaml programming language that could be compiled and executed to generate a kinematic action sequence describing the action of the machine over time. An example is shown in Figure 6. This approach provided a very high confidence in the validity of the path trace and allowed them to analyze and test it for various correctness properties—from the simple

(“does the nozzle stay within a bounding box”) to the complex (“does the tool path intersect with the printed material”).



**Figure 6:** Reconstruction of a sample print from the verified simulator.

The second focus has been on the analysis of this path trace and the data obtained via sensors during and after printing. The physical printing process has inherent error and uncertainty from sources, such as motor precision limits and vibrational noise. The research applied specialized probabilistic programming languages (specifically, the Stan language) for constructing graphical models representing the sources of error that can be combined with simulated path traces to augment them with uncertainties.

Currently, the researchers are transitioning some of the tools and techniques developed for this project into practice. They have identified programmatic use cases where this work can aid in assuring real AM systems at LLNL and other DOE sites. They have also learned some important lessons about practical applications of formal verification methods, especially balancing the complexity of verification against the benefits [4].

[1] T. Vanderbruggen and M. Sottile. “The challenges of software assurance and supply chain risk management.” In *High Confidence Software and Systems Conference*, 2024.

[2] M. Tekriwal and M. Sottile. “Mechanized semantics for correctness of the RS274 additive manufacturing command language.” In *NASA Formal Methods*, 2025. [doi.org/10.1007/978-3-031-93706-4\\_20](https://doi.org/10.1007/978-3-031-93706-4_20).

[3] M. Sottile and M. Tekriwal. “Design and implementation of a verified interpreter for additive manufacturing programs.” In *FUNARCH 2024: Proceedings of the 2nd ACM SIGPLAN International Workshop on Functional Software Architecture*, 2024. [doi.org/10.1145/3677998.367822](https://doi.org/10.1145/3677998.367822).

[4] M. Sottile, M. Tekriwal, and J. Sarracino. “Towards richer challenge problems for scientific computing correctness.” In *International Workshop on Verification of Scientific Software*, 2025. [doi.org/10.48550/arXiv.2510.13423](https://doi.org/10.48550/arXiv.2510.13423).

---

## Machine Learning & Applications | Using AI to Turbocharge Chemistry

Contact: [Tal Ben Nun](#)

Advances in AI are poised to transform how scientists explore, understand, and engineer chemical systems. CASC researchers, including Tal Ben Nun, are harnessing novel neural architectures to push the boundaries of molecular discovery.

LLMs, the driving force behind many writing assistants, coding tools, and search engines, belong to a broader class of neural networks known as *foundation models*. These powerful networks learn representations from massive data volumes and adapt to a wide range of downstream tasks. A key research question is: How to go beyond text and coding tasks to model molecules, reactions, and material properties? By bridging the gap between language models and chemical science, they can accelerate innovation across material design and sustainable chemistry.

At the heart of computational chemistry lies the ability to accurately represent and analyze molecules. The Simplified Molecular Input Line Entry System (SMILES) is the most widely used textual encoding notation. For example, the SMILES string for caffeine is CN1C=NC2=C1C(=O)N(C(=O)N2C)C. While SMILES enables rapid data handling and integration with text-based foundation models, its linear nature imposes limitations:

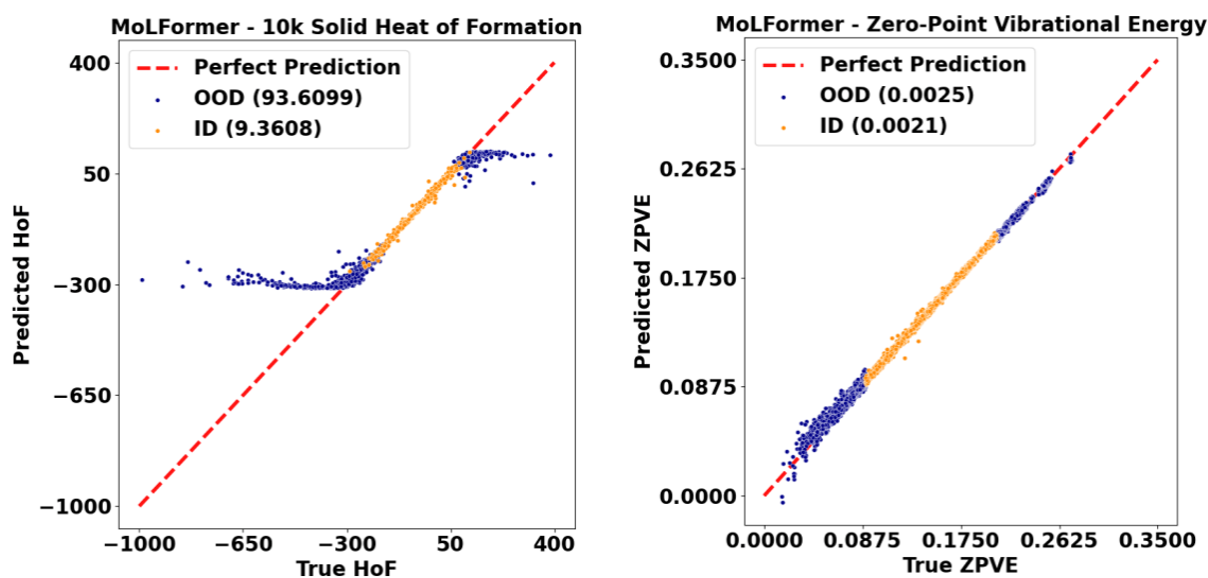
- It is inherently sequential, losing some of the spatial and graph-like relationships in real molecules.
- Models trained on SMILES struggle in low-data regimes typical of highly specialized chemical systems.
- Out-of-distribution generalization remains poor, even when closed-form solutions exist; current models rarely extrapolate to novel chemical spaces.

Chemical language models such as MolFormer and ether0 demonstrate promising capabilities on benchmark datasets, yet falter when experimental data are sparse. To learn from chemical data and propose untested compounds, one must introduce domain-specific inductive biases and rethink foundational neural architectures.

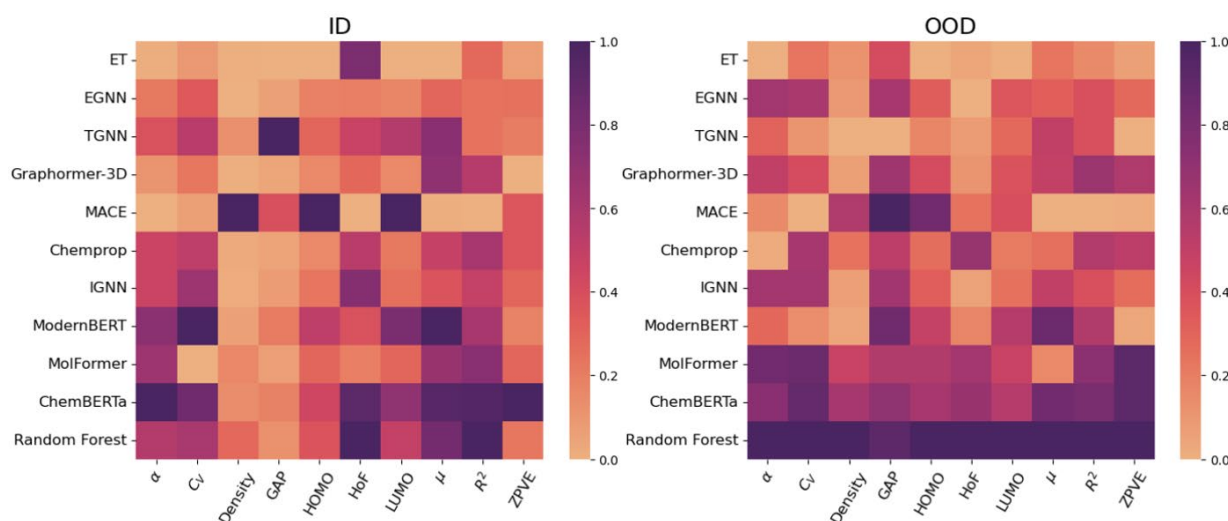
Through the LDRD-SI “FLASK” project, led by LC researcher Brian Van Essen, researchers are exploring alternative molecular encodings and network designs to improve generalization. As part of this effort, they have collaborated with Binghamton University to create the **Benchmarking Out-Of-distribution Molecules (BOOM)** dataset [1].

BOOM features (a) property-aware decompositions of molecular datasets that simulate realistic, out-of-distribution scenarios; and (b) new generalization metrics to quantify model performance beyond standard test splits. Figure 7 illustrates a typical test set parity plot from the BOOM benchmark, highlighting pronounced deviations when models predict properties of molecules outside the values found in their training distribution.

The researchers conducted a comprehensive study of neural architectures, including variants of ModernBERT, a modernized version of encoder-only transformers; Equivariant Transformers, which enforce rotational and translational symmetry in 3D molecular data; and MACE, designed to capture higher order interactions among atoms. Their findings, summarized in Figure 8, reveal that while these architectures yield gains over general transformer counterparts, none fully overcomes out-of-distribution generalization challenges. These results underscore the need for further innovation in incorporating stronger chemical priors into model design.



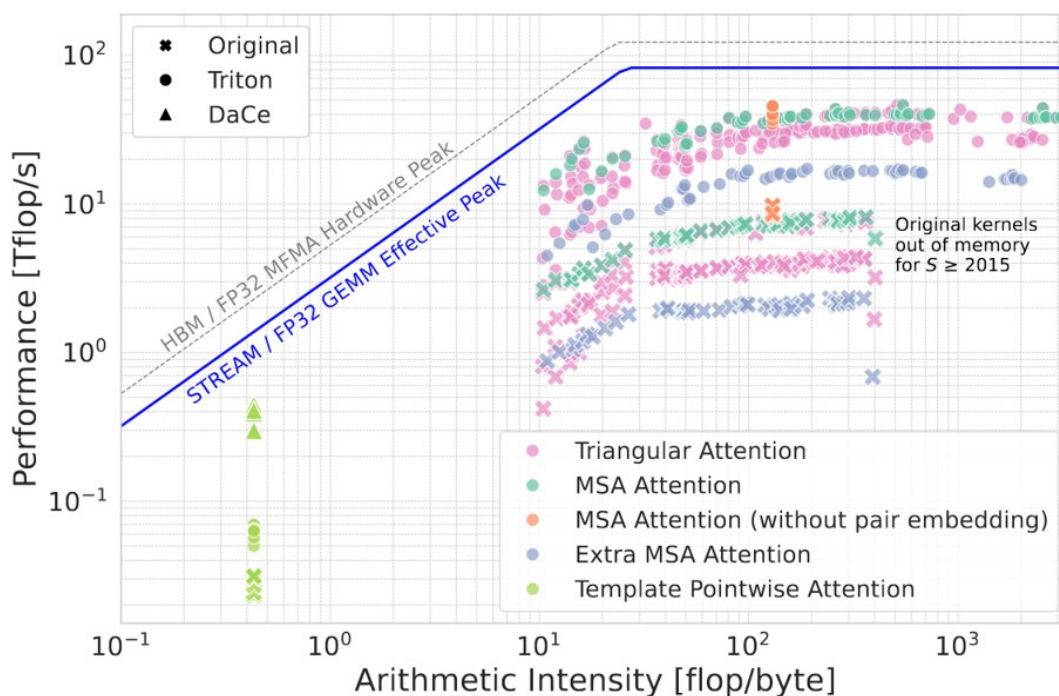
**Figure 7:** Parity plots of the BOOM test sets reveal how, for certain properties, chemical language models struggle with generalization beyond the regime they were trained on.



**Figure 8:** BOOM dataset normalized root mean square error scores across varying properties and neural architectures. Lower values are better. ID = in-distribution; OOD = out-of-distribution.

The FLASK project extends foundation model methodologies to biological macromolecules and large-scale data workflows. In collaboration with Columbia University and Advanced Micro Devices (AMD), researchers developed **ELMerFold**, a supercomputer-optimized pipeline that predicts high-quality, 3D protein structures at scale. ELMerFold has processed over 41 million proteins on El Capitan, currently the world's fastest supercomputer, in order to deliver open-source data to the global scientific community and democratize access to leading-edge structure prediction. To achieve this, many low-level optimizations were performed to the AI inference pipeline, from optimizing bespoke GPU kernels (Figure 9) using Triton and DaCe, to sophisticated scheduling of the work and I/O within the supercomputer.

AI assistants are already a force multiplier at LLNL, handling routine data analysis and hypothesis generation. The next frontier is empowering AI reasoning models to “think like a scientist”—*i.e.*, designing experiments and iterating autonomously through the scientific method. The ongoing research directions include AI-driven experimental planning and multimodal foundation models that jointly process text and simulation outputs.



**Figure 9:** Roofline plot with the specialized Evoformer attention kernel on the AMD MI300A processor. Variants of the kernel were reimplemented in Triton and DaCe to obtain speedups of up to an order of magnitude.

[1] E. Antoniuk, S. Zaman, T. Ben-Nun, P. Li, J. Diffenderfer, ..., and B. Van Essen. "BOOM: Benchmarking Out-Of-distribution Molecular property predictions of machine learning models." *The Thirty-ninth Annual Conference on Neural Information Processing Systems Datasets and Benchmarks Track (NeurIPS'25)*, 2025. [neurips.cc/virtual/2025/loc/san-diego/poster/121661](https://neurips.cc/virtual/2025/loc/san-diego/poster/121661).

## CASC Newsletter Sign-Up

Was this newsletter link passed along to you? Or did you happen to find it on social media? [Sign up](#) to be notified of future newsletters.

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. LLNL-MI-2018818. Edited by [Ming Jiang](#).