A Multi-Modal Deep Learning Platform for Cross-Domain Property Prediction in Chemical Process Design

Mikhail Tsitsvero [1] **Atsuyuki Nakao** [1]

[1] *CrowdChem, Inc., Tokyo, Japan*
Correspondence to: Mikhail Tsitsvero m.tsitsvero@crowdchem.net, Atsuyuki Nakao a.nakao@crowdchem.net.

1. Introduction

Experimental validation of chemical processes remains the primary bottleneck in materials discovery. Each trial is costly and time-consuming, limiting how quickly researchers can explore the vast design space of materials, formulations, and processing conditions. Yet decades of patents and scientific literature contain detailed experimental records spanning text, tables, molecular structures, and numerical measurements. These heterogeneous data sources offer tremendous potential for data-driven discovery if they can be unified into a machine-readable format suitable for modern deep learning.

We present a multi-modal deep learning platform that addresses this challenge by learning universal representations of chemical processes and predicting target properties across diverse application domains. Our approach builds on the directed-tree representation for chemical processes introduced in [1], which unifies molecular structures (encoded as SMILES), text descriptions, and numerical quantities into a single hierarchical graph structure. This work extends the framework with a flexible platform architecture that supports customer-specific fine-tuning and integration with experimental workflows.

2. Approach

2.1 Universal Process Representation

Our framework models chemical processes as directed trees where nodes represent materials, processing steps, conditions, and measured properties, while edges encode their relationships. Each node carries multi-modal attributes: molecular structures represented as SMILES, text embedded using large language models, and numerical values transformed into dense vector representations. This representation naturally captures the hierarchical flow of real experiments, from raw materials through processing steps to final products and their properties.

The key insight is that chemical processes across domains share common structural patterns: material combinations, processing operations, and condition-property relationships. By encoding these patterns explicitly in graph form, our representation enables models to learn transferable features that generalize beyond any single application domain.

2.2 Multi-Modal Learning Architecture

The platform employs a property-conditioned attention mechanism that dynamically focuses on the most relevant materials, conditions, and processing steps for the target property being predicted. Unlike traditional fixed-fingerprint approaches that lose structural information, our architecture preserves the causal dependencies between process steps and enables the model to distinguish between identical operations performed at different stages of a workflow.

The model is trained on hundreds of thousands of processes curated by domain experts from thousands of diverse documents spanning polymers, electronics, energy materials, and industrial formulations. This human-curated, large-scale pretraining corpus produces universal process embeddings that transfer effectively to specialized prediction tasks with minimal additional data.

2.3 Platform for Customer Data Integration

A central contribution is the development of a complete platform that allows organizations to leverage pretrained models on their proprietary data. The platform provides a graphical interface for curating experimental data into the universal process representation, multiple fine-tuning strategies optimized for different data regimes, uncertainty quantification for guiding experimental design, and integration capabilities with laboratory automation and self-driving experiment platforms.

3. Results

We validated the approach through fine-tuning experiments on domain-specific datasets. Even with limited domain-specific samples, the pretrained model achieves strong predictive performance [1], demonstrating effective transfer from the universal representation. Analysis of learned latent spaces shows that the model captures meaningful chemical similarity: experiments with related materials and conditions cluster together, while train/validation/test samples remain well-mixed within clusters, indicating genuine generalization rather than memorization.

The platform has been deployed for industrial applications including polymer property prediction, battery materials optimization, and coating formulation design. Users can upload curated process data, fine-tune models to their specific prediction tasks, and obtain uncertainty-calibrated predictions. Future work will extend the representation to 3D structural and spectral data, and integrate with autonomous experimentation platforms for closed-loop discovery.

Acknowledgments

This paper is based on results obtained from project JPNP23019, subsidized by the New Energy and Industrial Technology Development Organization (NEDO), Japan.

References

[1] Mikhail Tsitsvero, Atsuyuki Nakao, and Hisaki Ikebata. Accelerating materials discovery: Learning a universal representation of chemical processes for cross-domain property prediction. *arXiv preprint arXiv:2512.05979*, 2025.