

# Digital Discovery

rsc.li/digitaldiscovery

The Royal Society of Chemistry is the world's leading chemistry community. Through our high impact journals and publications we connect the world with the chemical sciences and invest the profits back into the chemistry community.

## IN THIS ISSUE

ISSN 2635-098X CODEN DDIIAI 3(6) 1073–1248 (2024)



**Cover**  
See Michael Nosonovsky, Ekaterina V. Skorb *et al.*, pp. 1101–1107. Image reproduced by permission of the authors from *Digital Discovery*, 2024, 3, 1101.



**Inside cover**  
See Francisco J. Martin-Martinez *et al.*, pp. 1108–1122. Image reproduced by permission of Francisco J. Martin-Martinez from *Digital Discovery*, 2024, 3, 1108.

## EDITORIAL

1081

### Perspective on AI for accelerated materials design at the AI4Mat-2023 workshop at NeurIPS 2023

Santiago Miret,\* N. M. Anoop Krishnan, Benjamin Sanchez-Lengeling, Marta Skreta, Vineeth Venugopal and Jennifer N. Wei

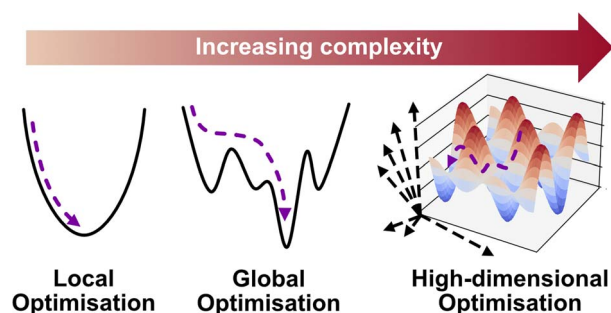


## PERSPECTIVE

1086

### Race to the bottom: Bayesian optimisation for chemical problems

Yifan Wu, Aron Walsh\* and Alex M. Ganose\*



# EES Catalysis

GOLD  
OPEN  
ACCESS

Exceptional research on energy  
and environmental catalysis

Open to everyone. Impactful for all

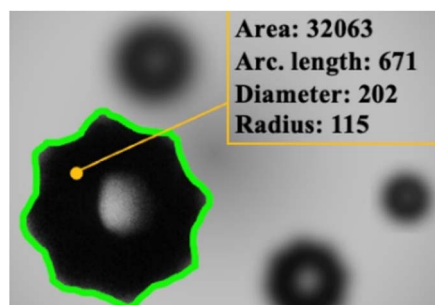
[rsc.li/EESCatalysis](https://rsc.li/EESCatalysis)

Fundamental questions  
Elemental answers

1101

## Automatic image processing of cavitation bubbles to analyze the properties of petroleum products

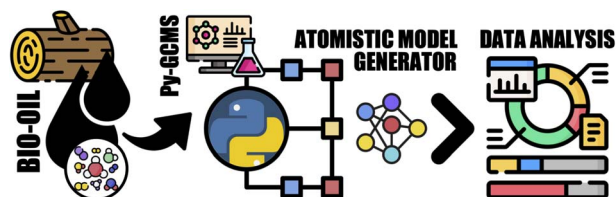
Timur Aliev, Ilya Korolev, Olga Burdulenko, Ekaterina Alchinova, Anton Subbota, Mikhail Yasnov, Michael Nosonovsky\* and Ekaterina V. Skorb\*



1108

## Data-driven representative models to accelerate scaled-up atomistic simulations of bitumen and biobased complex fluids

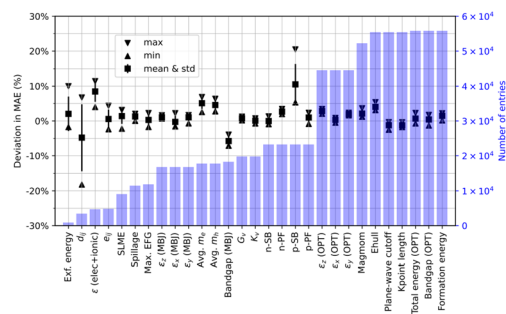
Daniel York, Isaac Vidal-Daza, Cristina Segura, Jose Norambuena-Contreras and Francisco J. Martin-Martinez\*



1123

## A reproducibility study of atomistic line graph neural networks for materials property prediction

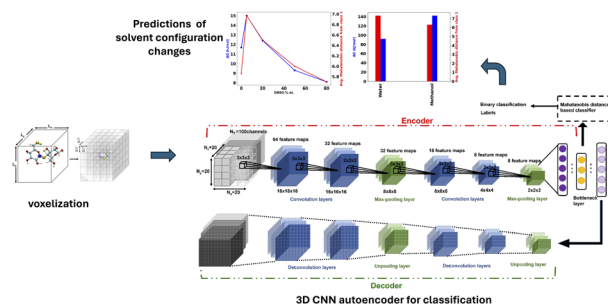
Kangming Li,\* Brian DeCost, Kamal Choudhary and Jason Hattrick-Simpers



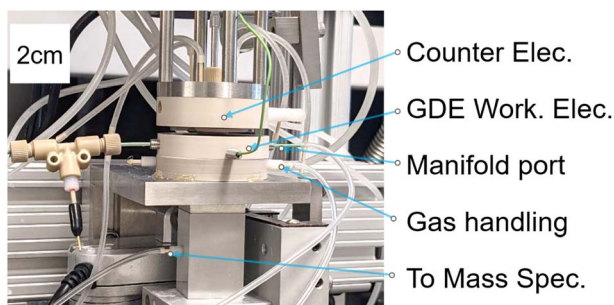
1130

## A 3d convolutional neural network autoencoder for predicting solvent configuration changes in condensed phase biomass reactions

Anjana Puliyaanda, Arul Mozhi Devan Padmanathan, Samir H. Mushrif and Vinay Prasad\*



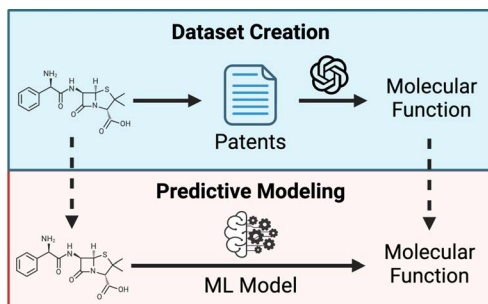
1144



### Accelerated screening of gas diffusion electrodes for carbon dioxide reduction

Ryan J. R. Jones, Yungchieh Lai, Dan Guevarra, Kevin Kan, Joel A. Haber and John M. Gregoire\*

1150



### Mining patents with large language models elucidates the chemical function landscape

Clayton W. Kosonocky, Claus O. Wilke, Edward M. Marcotte and Andrew D. Ellington\*

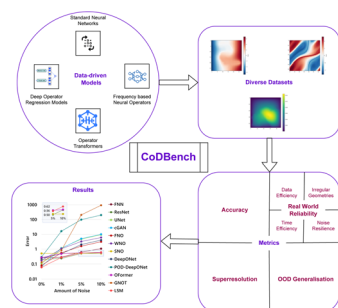
1160



### iSIM: instant similarity

Kenneth López-Pérez, Taewon D. Kim and Ramón Alain Miranda-Quintana\*

1172



### CoDBench: a critical evaluation of data-driven models for continuous dynamical systems

Priyanshu Burark, Karn Tiwari, Meer Mehran Rashid, Prathosh A. P.\* and N. M. Anoop Krishnan\*

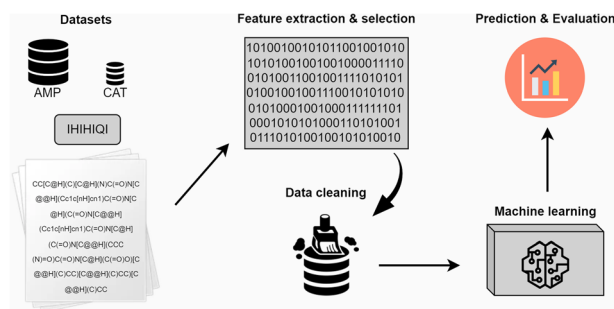




1182

## Efficiently solving the curse of feature-space dimensionality for improved peptide classification

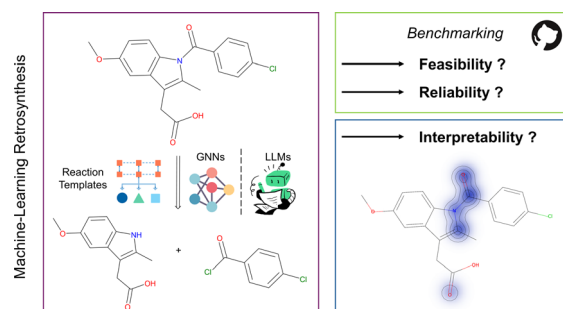
Mario Negovetić, Erik Otović, Daniela Kalafatović\* and Goran Mauša\*



1194

## Investigating the reliability and interpretability of machine learning frameworks for chemical retrosynthesis

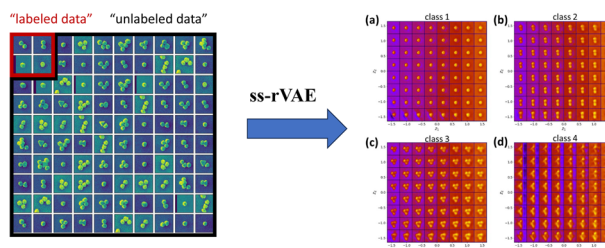
Friedrich Hastedt,\* Rowan M. Bailey, Klaus Hellgardt, Sophia N. Yaliraki, Ehecatt Antonio del Rio Chanona\* and Dongda Zhang\*



1213

## Semi-supervised learning of images with strong rotational disorder: assembling nanoparticle libraries

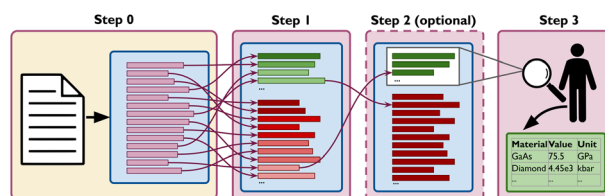
Maxim A. Ziatdinov,\* Muammer Yusuf Yaman, Yongtao Liu, David Ginger and Sergei V. Kalinin



1221

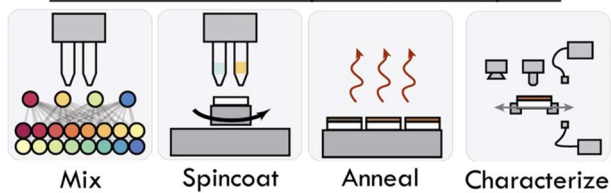
## Flexible, model-agnostic method for materials data extraction from text using general purpose language models

Maciej P. Polak,\* Shrey Modi, Anna Latosinska, Jinming Zhang, Ching-Wen Wang, Shaonan Wang, Ayan Deep Hazra and Dane Morgan\*



1236

## Perovskite Automated Spin Coat Assembly Line



### PASCAL: the perovskite automated spin coat assembly line accelerates composition screening in triple-halide perovskite alloys

Deniz N. Cakan, Rishi E. Kumar, Eric Oberholtz, Moses Kodur, Jack R. Palmer, Apoorva Gupta, Ken Kaushal, Hendrik M. Vossler and David P. Fenning\*

