Digital Discovery

ROYAL SOCIETY OF CHEMISTRY

EDITORIAL

View Article Online
View Journal

Cite this: DOI: 10.1039/d5dd90057c

Introduction to the "Accelerate Conference 2023–2024" themed collection

Janine George, [ab] Claudiane Ouellet-Plamondon [c] and Kristofer Reyes[d]

DOI: 10.1039/d5dd90057c

rsc.li/digitaldiscovery

The field of experimental science is undergoing a profound transformation. Across chemistry, materials science, and manufacturing, machine learning and automation are redefining how we design, execute, and interpret experiments. Traditional trial-and-error workflows give way to adaptive, closed-loop systems that combine predictive modelling, optimisation algorithms, and robotic execution. This shift reflects more than speed—it marks a change in how we conceptualise experimentation as a continuous learning, computation-integrated process.

This themed collection gathers leading contributions to that transformation. The papers span innovations in algorithms, decision-making, and integrated self-driving laboratories—from efficient experimental design and probabilistic programming to orchestration frameworks coordinating sensing, actuation, and learning. Collectively, they illustrate new principles for accelerating and scaling discovery.

The Accelerate Conference serves as the foundation for this collection. Bringing together researchers at the intersection of AI, robotics, materials discovery, and sustainable design, the forum promotes not only technical progress but also reproducibility, interoperability, and ethical practice. Each paper in this collection contributes to a modular, transparent, and scalable infrastructure for automated discovery. Originating from the 2023 and 2024 conferences, the collection addresses researchers, developers, and policy-makers who view autonomy as the evolution of scientific practice rather than a novelty.

We hope this collection acts as both a snapshot of current capabilities and a call to future work. These technologies are tools that enable a new form of scientific inquiry—guided by human insight and computational foresight—built one automated experiment at a time.

Automated experiments

Hung *et al.* (https://doi.org/10.1039/D4DD00059E) survey automated laboratories across three dimensions: accelerated materials discovery, the role of human researchers, and levels of automation. They emphasise balancing data collection with decision frequency in active learning, highlight the need for trust between humans and machines, and discuss the role of autonomous research in bridging the critical gap between theory and experiment.

Lo *et al.* (https://doi.org/10.1039/D3DD00223C) introduce the "frugal twin" concept offering a low-cost physical analogue to digital twins, enabling rapid, low-risk prototyping of physical experiments for education and research.

Doloi *et al.* (https://doi.org/10.1039/D4DD00411F) review low-cost 3D printing for laboratory automation, showing how open source, modular liquid handlers and imaging systems—built with microcontrollers such as Arduino and Raspberry Pi—expand access to automated experimentation globally.

Ismail *et al.* (https://doi.org/10.1039/D4DD00381K) develop an automated system for preparing and drop-casting electrodes, producing 48 samples per day with high reproducibility and realistic mass loadings for battery research.

Hernández-del-Valle *et al.* (https://doi.org/10.1039/D4DD00198B) present an open-source hardware platform automating polymer-nanocomposite preparation using two modular 3D-printed devices costing under €400. Their release of full design files and firmware exemplifies democratised experimental automation.

Bayesian optimisation

Myung *et al.* (https://doi.org/10.1039/D4DD00281D) introduce AM-ARES, a syringe-extrusion 3D printer implementing multi-objective Bayesian optimisation to balance speed and accuracy while ensuring safe exploration of parameter space.

Deneault *et al.* (https://doi.org/10.1039/D4DD00320A) embed human

[a]Federal Institute for Materials Research and Testing (BAM), Germany. E-mail: janine.george@bam.de
[b]Friedrich Schiller University Jena, Germany
[c]École de technologie supérieure, Université du Québec, Canada. E-mail: Claudiane.Ouellet-Plamondon@etsmtl.ca
[d]University at Buffalo, USA. E-mail: kreyes3@buffalo.edu

preference into the optimisation loop *via* a conveyor-belt 3D printer that learns from user rankings to match aesthetic goals—demonstrating that preference-based learning can rival sensor-based approaches in efficiency.

Fitzner *et al.* (https://doi.org/10.1039/D5DD00050E) present BayBE, a Python package bringing Bayesian optimisation to scarce-data experimental settings through chemical encodings, transfer learning, multi-objective scalarisation, and automatic stopping rules.

Tamura *et al.* (https://doi.org/10.1039/D5DD00017C) propose a flexible batch Bayesian optimisation framework addressing unequal experimental throughput, validated on robotic synthesis of sulfonated 9-fluorenone derivatives relevant to redox-flow batteries.

Machine learning

Duangdangchote *et al.* (https://doi.org/10.1039/D4DD00140K) assess the transferability of graph-based machine-learned force fields across lithium compounds, identifying architectures that preserve accuracy across compositions.

Rakotonirina *et al.* (https://doi.org/10.1039/D4DD00228H) propose a Hammett-inspired linear-free-energy model to improve data efficiency for ligand tuning in Suzuki–Miyaura catalysis, identifying 145 promising catalysts, including cost-saving Ni-based ones.

Amariamir *et al.* (https://doi.org/10.1039/D4DD00394B) present a co-training procedure of two different machine learning models for synthesizability prediction of oxide crystals to leverage positive and unlabelled data, enhancing recall without needing explicit negatives.

Tang *et al.* (https://doi.org/10.1039/D5DD00079C) introduce a central-environment feature representation improving locality and data efficiency in graph-based deep learning, validated on substitution energies in Nb–Si systems.

Collinard-Granero *et al.* (https://doi.org/10.1039/D5DD00158G) show a deep-learning framework for bubble analysis in vanadium redox-flow batteries that uses multi-class segmentation of synchrotron X-ray tomography to quantify bubble dynamics and electrode performance, offering insights for electrode optimisation.

High-throughput experiments

Soh *et al.* (https://doi.org/10.1039/D4DD00368C) develop a proxy viscometer using automated aspiration–dispense cycles and machine-learning regression to estimate viscosity in under two minutes per sample.

Siemenn *et al.* (https://doi.org/10.1039/D4DD00249K) present Archerfish, a retrofitted 3D printer achieving continuous combinatorial printing of 250 unique gradients per minute at a low cost, shifting the bottleneck in additive-manufacturing discovery from fabrication to experiment planning.

Materials-discovery frameworks and literature analysis

Khajeh *et al.* (https://doi.org/10.1039/D4DD00293H) introduce a computational polymer-design framework integrating generative models, molecular-dynamics evaluation, and feedback loops, identifying 14 polymer units exceeding PEO ionic conductivity.

Gallagher and Webb (https://doi.org/10.1039/D4DD00298A) compare 31 material-classification sources and 100 learning algorithms in an algorithmic benchmarking study, showing that neural networks and random forests are the most data-efficient.

Ansari and Moosavi (https://doi.org/10.1039/D4DD00252K) present a chemist-AI agent powered by GPT-4 that extracts structured datasets from the literature using a Chain-of-Verification tool to avoid hallucinations, performing zero-shot learning on multiple data-extraction tasks.

Zhang *et al.* (https://doi.org/10.1039/D4DD00248B) convert literature protocols into machine-readable scripts for automated reactors, demonstrating end-to-end prompt-to-experiment translation in three case studies.

Dreger *et al.* (https://doi.org/10.1039/D4DD00362D) employ a semi-automated pipeline that uses large-language models and rule-based feedback to generate validated knowledge graphs from tables, forming part of a larger semantic-search platform.

Bai *et al.* (https://doi.org/10.1039/D5DD00069F) present a Python framework linking class hierarchies with RDF knowledge graphs, enabling applications such as simplifying automated geometry assembly in molecular design and synthesis-protocol extraction.

Finally, Ganose *et al.* (https://doi.org/10.1039/D5DD00019J) propose a new framework for automated and reproducible materials-property calculations, atomate2, which unifies machine-learned interatomic potentials and first-principle workflows, and interfaces with multiple electronic-structure codes for modular, scalable materials simulations.

Conclusions

The contributions to this collection demonstrate how autonomous experimentation is moving from isolated demonstrations toward a comprehensive framework for accelerating discovery. They highlight technical sophistication and full-system integration—addressing not only *what* to automate but *how* to do so responsibly, reproducibly, and at scale.

Recurring themes include maximising information gain under real-world constraints, integrating uncertainty and cost into decision-making, and designing modular, interpretable architectures that foster reuse and transparency.

The Accelerate Conference continues to shape this vision, providing a forum for discussing both innovation and accountability in automated discovery. Collectively, these works advance science that is faster, more robust, and more open—building a future where human insight and computational intelligence advance hand in hand.

Use of AI declaration

A longer draft version of the editorial was written. The editorial was then shortened by ChatGPT and reviewed by the authors and further refined for consistency. Antidote software was used to check the English.

Acknowledgements

Janine George author photograph from BAM, used with permission.