EDITORIAL

Check for updates

Introduction to Digital Catalysis

Evgeny A. Pidko [iD] *[a] and Núria López [iD] *[b]

Catalysis Science & Technology, Evgeny Pidko and Núria López would like to acknowledge Weixue Li for their contributions to the Digital Catalysis themed collection as a Guest Editor.

[a] *Inorganic Systems Engineering group, Department of Chemical Engineering, Delft University of Technology, van der Maasweg 9, 2629 HZ Delft, The Netherlands. E-mail: e.a.pidko@tudelft.nl*

[b] *Institute of Chemical Research of Catalonia (ICIQ-CERCA), The Barcelona Institute of Science and Technology (BIST), Av. Països Catalans 16, 43007 Tarragona, Spain. E-mail: nlopez@iciq.es*

We are witnessing a digital revolution marching across every aspect of our lives and technological fields, and catalysis is no exception. As workflows become increasingly digital and data-driven, catalysis research and development are now being transformed and accelerated. This themed collection on digital catalysis reflects how these innovations are reshaping the field, integrating advanced computational tools, machine learning, and FAIR data principles into catalyst science and technology. It highlights how digital tools can accelerate discovery, address unresolved challenges, and open new avenues for sustainable chemical innovation.

Over decades, catalysis research has generated a vast body of empirical data and mechanistic knowledge. Now, with the rise of digital methods and artificial intelligence, this accumulated knowledge can potentially be integrated directly into advanced computational and statistical models to provide a more realistic and nuanced picture of catalytic phenomena. By blending established physical principles with emerging data-driven approaches, we enhance our ability to navigate chemical

Evgeny A. Pidko

Evgeny A. Pidko is Full Professor and head of the Inorganic Systems Engineering group at Delft University of Technology. He was trained in physical chemistry at the Higher Chemical College of the Russian Academy of Sciences and obtained his PhD in computational catalysis with Prof. Rutger A. van Santen at Eindhoven University of Technology (2008). After serving as Assistant Professor at TU Eindhoven, he joined TU Delft in 2017, becoming Full Professor in 2020. He has also held visiting professorships at ETH Zürich (2012) and Hokkaido University (2023). His research focuses on the development of catalytic systems for sustainable chemistry, combining computational modeling, kinetic and mechanistic studies, and automation. He has authored over 260 papers and 6 patents. He is the recipient of an ERC Consolidator Grant (2016) and an NWO VICI Grant (2025), and he leads the national FAIR4ChemNL initiative on open and data-driven research.

Núria López

Núria López graduated in Chemistry and obtained her PhD in Theoretical Chemistry from the University of Barcelona (1999). After postdoctoral research with Prof. Jens K. Nørskov in Denmark, she started her independent career at the Institute of Chemical Research of Catalonia (ICIQ) in 2005. Her research group focuses on theoretical studies of heterogeneous and photoelectrocatalysis, with a particular emphasis on bridging fundamental theory and practical applications. Prof. López has authored more than 300 scientific publications, including several highly cited works, and contributed to the development of computational tools and data infrastructures such as ioChem-BD. She received an ERC Starting Grant (2010) and an ERC Proof-of-Concept Grant (2015), and was awarded the Prize for Excellence by the Spanish Royal Society of Chemistry in 2015. She has collaborated extensively with European industry and EU consortia and chaired the PRACE Scientific Steering Committee.

This journal is © The Royal Society of Chemistry 2025

Catal. Sci. Technol., 2025, **15**, 6925–6927 | **6925**

complexity and to move closer toward catalysts by design.

The comprehensive review by Lapkin and co-workers (https://doi.org/10.1039/D3CY01160G) presents a future vision on how integration of advanced computational methods and AI/ML techniques can enable predictive design and autonomous discovery of heterogeneous catalysts. In a complementary work, Parveen and Slater (https://doi.org/10.1039/D4CY01525H) stress the broader importance of digital frameworks and FAIR principles for enabling sustainable chemical production and exploring wider chemical spaces.

Probably the most common and widely practiced digital tool in contemporary catalysis research is computational modeling. Modern quantum chemical methods have reached the level of accuracy, accessibility and cost to provide indispensable support in interpreting complex spectroscopic data and building detailed mechanistic models to explain kinetic trends and guide the design of new catalysts *via* descriptors.

We have seen the development of multiscale models that merge molecular simulations, kinetic modeling, and quantum chemistry, allowing us to tackle complexity across scales. Several contributions illustrate the power of such methods. Tong *et al.* (https://doi.org/10.1039/D3CY01590D) and Dunn *et al.* (https://doi.org/10.1039/D4CY00506F) employ molecular dynamics simulations to resolve how zeolite morphology and molecular interactions shape transport phenomena that ultimately govern catalytic turnover. Thomas *et al.* (https://doi.org/10.1039/D4CY00284A) combine DFT and molecular dynamics to capture the speciation of manganese catalysts under oxidative conditions, providing atomistic insight into dynamic coordination environments that control stability and reactivity under realistic operation. Ureel *et al.* (https://doi.org/10.1039/D4CY00973H) develop a predictive group additive model for β-scission kinetics in zeolites, introducing a pore-confinement descriptor that connects local structure with macroscopic cracking rates. Chen *et al.* (https://doi.org/10.1039/D4CY00586D) further demonstrate the strength of multiscale modeling by showing the role of TiO_2 polymorphs in dictating Ni cluster morphology and reactivity in CO_2 hydrogenation. He *et al.* (https://doi.org/10.1039/D4CY01076K) integrate DFT and microkinetic modeling to establish design principles and identify Ni_3Fe alloys as selective quinoline hydrogenation catalysts. At the electrochemical interface, Iida *et al.* (https://doi.org/10.1039/D5CY00369E) combine DFT and a statistical mechanical theory of liquids (3D-RISM) to explain the disappearance of double-layer effects, offering fundamental understanding of electrode–electrolyte interactions. Together, these studies demonstrate that multiscale modeling now routinely describes diffusion, adsorption, and condition-dependent reactivity with a resolution inaccessible to experiment.

Despite the great success of quantum chemistry models, critical challenges remain. Capturing catalyst dynamics, competing pathways, and condition-dependent equilibria is a formidable task. The increasing size of datasets and configuration spaces demands new strategies that combine physical fidelity with scalable efficiency. Miyazaki *et al.* (https://doi.org/10.1039/D4CY00685B) provide a systematic assessment of exchange–correlation functionals by comparing predicted vibrational frequencies with experiment. Such studies establish clear reference points and allow researchers to quantify the uncertainty of popular methodologies. Hühn *et al.* (https://doi.org/10.1039/D4CY01152J) combine ^{31}P NMR, *ab initio* molecular dynamics, and machine learning to characterize phosphate speciation on alumina. Their results highlight how disorder and dynamic effects challenge standard models, but also how hybrid approaches can bring simulations in line with measurable observables.

Abdelmaqsoud *et al.* (https://doi.org/10.1039/D4CY00615A) extend this discussion to machine-learning interatomic potentials, demonstrating that inconsistencies due to surface reconstruction in large DFT datasets result in biased models and propose that total-energy references provide more robust training data. Rey *et al.* (https://doi.org/10.1039/D4CY00548A) introduce a hybrid ML-thermodynamic perturbation theory framework that achieves near *ab initio* accuracy in free-energy barriers at a fraction of the cost, making predictive kinetics feasible for complex zeolite reactions. Ting *et al.* (https://doi.org/10.1039/D4CY01000K) illustrate the role of unsupervised learning in revealing surface patterns in nanoparticle simulations, offering a path toward systematic identification of complex catalytic motifs serving as active sites.

As the community generates ever-larger datasets from both experiments and simulations, the need for data-driven methodologies is becoming critical for analyzing this multifaceted data, identifying patterns, and guiding the development of catalysts and catalytic processes. These digital tools not only help us navigate much wider chemical spaces and mechanistic landscapes but close the gap between *operando* catalysts and our models. Yet, with this data-rich paradigm comes the challenge of ensuring data integrity and adopting FAIR (Findable, Accessible, Interoperable, Reusable) principles. Several contributions in this collection discuss how the community is starting to address these barriers. Trunschke *et al.* (https://doi.org/10.1039/D4CY00693C) outline a framework for digital and automatic acquisition, storage, and linking of catalysis data and metadata. They present machine-readable SOPs and automation to capture experimental workflows and their associated data in a form that supports reproducibility and direct integration with machine learning. Behr *et al.* (https://doi.org/10.1039/D4CY00369A) introduce automated knowledge graphs that structure

6926 | *Catal. Sci. Technol.*, 2025, **15**, 6925–6927

This journal is © The Royal Society of Chemistry 2025

information extracted from catalysis literature, making hidden connections explicit and providing a foundation for autonomous discovery. Li *et al.* (https://doi.org/10.1039/D4CY01159G) illustrate how ML and text mining can be applied at scale to extract synthesis and performance data from literature on SCR catalysts, directly enabling performance prediction and synthesis optimization.

The final set of contributions demonstrate how machine learning and physics-based methods can be combined into hybrid workflows. These approaches draw their strength directly from the advances highlighted in the preceding sections: accurate and benchmarked electronic structure methods, mechanistic and multiscale models that define descriptors of catalytic function, and structured datasets that ensure reproducibility and reusability. Without reliable data and validated reference methods, machine learning remains a black box; without acceleration from data-driven models, high-level simulations remain too costly to drive discovery. Guo and Harvey (https://doi.org/10.1039/D3CY01625K) provide a clear example by coupling *ab initio* calculations with microkinetic modeling and data fitting to experiments, achieving predictive accuracy for catalytic rates.

Saha *et al.* (https://doi.org/10.1039/D4CY00763H) employ machine-learning potentials to analyze atomic arrangements in zeolites, generating statistically meaningful insights into topology and synthesis–property relations. Kuddusi *et al.* (https://doi.org/10.1039/D4CY00873A) take this a step further by combining ML with active learning and automated experimentation, exemplifying how scientist-in-the-loop strategies can accelerate design cycles for CO_2 hydrogenation catalysts.

The contributions in this collection show that hybrid data-driven strategies are not merely accelerators but enablers of the new catalysis science.