

ЦИФРОВЫЕ ТЕХНОЛОГИИ В ХИМИИ И МАТЕРИАЛОВЕДЕНИИ

© 2025 г. В.П. Анаников^{а,*}

^аИнститут органической химии им. Н.Д. Зелинского РАН, Москва, Россия

*E-mail: val@ioc.ac.ru

Поступила в редакцию 19.02.2025 г.

После доработки 10.03.2025 г.

Принята к публикации 10.04.2025 г.

В статье на фоне глобального химического ландшафта рассмотрены проблемы и перспективы цифровизации химических процессов и технологий. Особое внимание уделено таким вопросам, как низкая скорость обработки данных, проблема спящих и потерянных данных, трансфер знаний в технологии микро- и малотоннажной химии, цифровой цикл производства химической реакции. Таким образом, речь идёт о применении специализированных цифровых алгоритмов, необходимых для решения конкретных задач в химии, ускорении анализа данных, автоматизации химических исследований, создании новых подходов к проектированию материалов для катализа, разработке цифровых инструментов для науки и образовательных программ. Чрезвычайно важно актуальное практическое приложение — использование искусственного интеллекта для масштабирования малотоннажных процессов — и трансфер технологий в промышленные решения.

Ключевые слова: химия, материаловедение, цифровизация химических процессов и технологий, обработка данных, спящие и потерянные данные, цифровой цикл проектирования химических реакторов, микро- и малотоннажная химия.

DOI: 10.31857/S0869587325040079, **EDN:** EFKQWG

Современный глобальный химический ландшафт оперирует соединениями, полученными благодаря химической промышленности. В первую очередь обращают на себя внимание продукты нефтепереработки. Полученные компоненты и мономеры повсеместно используются в производстве полимеров, пластиков, высокомолекулярных соединений, которые являются неотъемлемой частью современной жизни. Лакокрасочные изделия, цветные пигменты для финального доведения продуктов, строительные материалы — все эти вещества и материалы получаются в результате химического

производства. Металлургия является важнейшей частью современной химии. Что касается тонкого органического синтеза, то ключевым его направлением является получение лекарственных препаратов. Эти и другие примеры подчёркивают важность химической промышленности.

Если говорить об огромном многообразии химической продукции, то практически всё, сделанное руками человека, содержит в себе компоненты, полученные на химическом заводе. Это обеспечивается благодаря трём масштабам современного химического синтеза:

- крупнотоннажный и среднетоннажный — с объёмом производства на уровне 150 000 т в год; ассортимент продукции невелик — примерно 500 наименований во всём мире, но это самые массовые продукты химической промышленности [1];
- микротоннажное и малотоннажное производство — с объёмом производства на уровне 10 000 т в год (или менее) [1]; порядка 100 000 соединений задействовано в этом диапазоне в мире;
- исследовательский и инновационный диапазон очень широкий — от миллиграммов до тонн вещества; химическое пространство в настоящее вре-



АНАНИКОВ Валентин Павлович — академик РАН, заведующий лабораторией металлокомплексных и наноразмерных катализаторов ИОХ РАН.

мя включает в себя до 100 000 000 соединений [1]; именно из этого диапазона выбирают соединения для практической реализации.

Как эти соединения получают? Основные технологии — это каталитические процессы. Около 80% химических технологий требуют использования катализаторов. По оценкам, примерно 35% мирового ВВП обеспечивает катализ. Нобелевский лауреат Б. Лист сравнивает катализ с наиболее значимыми достижениями человечества за всю его историю.

Ключевая современная проблема в разработке новых каталитических процессов — низкая скорость обработки данных. Это можно проиллюстрировать на конкретном примере Pd-катализируемой реакции кросс-сочетания, активно задействованной в тонком органическом синтезе для производства лекарств и агрохимии. Чтобы синтезировать 1 грамм вещества с молекулярной массой 290 путём гетерогенного катализа, понадобится около 10^{16} атомов катализатора (из расчёта один грамм — это 3.5 миллимоль или 10^{21} молекул продукта; количество катализатора составляет 0.001 мол%). Одна наночастица катализатора — это около 1000 атомов палладия, что означает применение 10^{13} наночастиц катализатора для получения одного грамма продукта. Если воспользоваться самым современным подходом к анализу наноразмерных объектов в химических технологиях (со скоростью анализа и обработки одна наночастица в минуту), понадобится 38 000 000 лет для изучения одной реакции, которая приводит к получению одного грамма вещества.

Возьмём другие процессы, гомогенные каталитические реакции, которые можно изучать с помощью спектрального мониторинга, например, с помощью масс-спектрометрии. Регистрация одного спектра занимает около 2 секунд, и за стандартное время мониторинга реакции (8 часов) регистрируется 14 400 спектров. Каждый спектр реакционной массы каталитической реакции может содержать до 1000 сигналов. При применении самого современного метода анализа и идентификации структур один сигнал можно интерпретировать в 5 минут (приблизительная оценка). Таким образом, при ручной обработке всего массива данных понадобится 137 лет для анализа одной реакции. Такого соотношения между реальной скоростью накопления данных и их анализом. То есть, чтобы вручную проанализировать данные, собранные прибором за 8 часов, требуется 137 лет работы человека.

Этот пример иллюстрирует, какие большие объёмы данных (и низкие скорости их ручной обработки) задействованы в химических исследованиях, в химическом производстве, при создании новых систем. Коллектив из 10 исследователей в год обрабатывает от $10^{-8}\%$ до 0.1% полученных для каталитической системы данных. Усреднённые временные затраты на разработку новых катализаторов составляют 5 лет, причём найденные каталитические си-

стемы, как правило, являются локальными минимумами в пространстве оптимизации химических реакций. Дальнейшая дооптимизация может длиться несколько лет и требовать существенных усилий.

Новые перспективы открылись в химических исследованиях с появлением цифровых методов. В Отделении химии и наук о материалах РАН (ОХНМ РАН) проведено тестирование этих методов, чтобы убедиться, насколько они применимы и где от них может быть практическая польза в разработке новых химических технологий. Проведено исследование в области полностью определённого катализа [2]. В этом случае все данные, которые получены при характеристике катализаторов, обрабатываются цифровыми алгоритмами. Была отработана система нейронных сетей по распознаванию изображения, построению карты катализатора, проведению анализа всех наночастиц и воссозданию структуры каталитической системы. Удалось создать матрицы оцифрованных в автоматическом режиме наночастиц катализатора и полную трёхмерную модель микрочастиц катализатора. Это первый пример работы такого рода в мировой практике [2].

Одна из сложнейших задач, от которой зависит активность и стабильность катализаторов, — изучение процесса деградации. В этой связи долгое время нерешённой проблемой оставался алгоритм позиционирования: нужно находить на огромном пространстве всей поверхности катализатора (от сотен квадратных метров до квадратных километров) одни и те же атомы. С помощью цифровых технологий такой алгоритм был отработан — это пошаговое последовательное позиционирование [3]. Показано, что с помощью цифровых технологий можно определять одно и то же место в пространстве с точностью до одной наночастицы. Ещё раз стоит обратить внимание на тот факт, что на площади, измеряемой квадратными километрами, удаётся обнаружить отдельные атомы с атомарной точностью позиционирования. Практическое применение такого метода в химических проектах позволит провести исследования по повышению стабильности катализаторов. Рассмотрение одних и тех же каталитических центров до и после реакции показало, что катализатор может терять свою активность в результате образования мостиков между наночастицами, сглаживания неровностей и потери одноатомных каталитических центров [3].

Целенаправленный проект в области цифровой химии был начат в ОХНМ РАН несколько лет назад. Мы хотели на собственном опыте проверить, насколько цифровые алгоритмы эффективны и представляют практическую пользу. В проекте полностью определённого катализа ручная обработка данных занимает около четырёх лет на одну реакцию, в то время как для разработки химической технологии нужно провести и проанализировать как минимум несколько сотен реакций — можно посчи-

тать временные затраты, и они будут огромны (годы работы). Анализ с помощью цифровых технологий требует намного меньше времени. Полноразмерный анализ массива данных и построение функциональных карт мониторинга химических процессов с помощью алгоритмов машинного обучения занимают несколько часов.

Следует отметить, что речь идёт о специализированных алгоритмах именно для химии. Это не популярные продукты на базе искусственного интеллекта, а специально разработанные и обученные алгоритмы и программы, основанные на глубоком машинном обучении.

Серьёзной проблемой является низкий уровень анализа данных, проблема спящих или потерянных данных. Скорость накопления информации в химии высокая — 1 терабайт в год для активных проектов. Анализ данных, полученных в центрах коллективного пользования, показал, что нередко обрабатывается и публикуется примерно 10% данных, а 90% остаются неиспользованными. По диссертациям аналогичная картина — 20% данных обработано и опубликовано, порядка 80% данных не подвергается подробному анализу. На фоне очень высокой скорости накопления информации, скорость её ручной обработки остаётся низкой. Эта ситуация характерна для современного состояния химии ряда других областей, даже появились специальные термины: спящие данные (sleeping data), потерянные данные (lost data).

Следует отметить важную задачу в области трансфера знаний в микро- и малотоннажной химии. Для данной области характерна существенная сложность — это малая численность исследователей, в то время как в области изучения граммовых и крупнотоннажных процессов исследователей намного больше [1]. Это касается и объёма информации, минимальный её объём сохраняется при разработке процессов микротоннажной и малотоннажной химии.

Алгоритмы цифрового преобразования позволяют осуществлять трансфер знаний: по имеющимся реакциям в малых и больших диапазонах предсказывать оптимальные условия для процессов в среднем диапазоне. Для этого нужна консолидация данных, нужны цифровые системы обработки информации, и такие проекты будут активно реализовываться в ближайшем будущем. Также важно развивать методы отработки рисков, передачи, обработки и построения системы искусственного интеллекта в анализе данных в химии.

Ещё один пример — это цифровой цикл в производстве химических реакторов. На практике общее время от начала проектирования до пуско-наладочных работ составляет от 20 до 50 месяцев, достаточно большой срок с учётом того, что за короткое время нужно вывести на рынок несколько сотен соединений. Это время можно существенно сократить с применением аддитивных технологий, потому что

именно для микротоннажного диапазона важная часть химического реактора целиком помещается в камеру 3D-принтера. Часть химического реактора можно напечатать, в том числе в случае сложных геометрических форм. Это было проверено экспериментально, реакторы сложной геометрической формы, которые трудно сделать иными методами, уверенно создаются с помощью аддитивных технологий. Более того, даже соединительные элементы, полученные в результате процесса 3D-печати, обладают достаточными характеристиками для тестирования реакторов.

При реализации таких проектов, можно провести прямой цифровой дизайн реактора под конкретную химическую реакцию. Это требует интегрированной базы знаний с высокими требованиями к безопасности информации на единой платформе, что остаётся сдерживающим фактором для развития подобных технологий, хотя их польза очевидна. Создание интегрированной базы знаний позволит миновать несколько времязатратных циклов и ускорить разработку химических технологий.

Нельзя не сказать о значимых рисках. Необходимо развивать собственные системы обработки химических данных, включая программы машинного обучения и большие языковые модели. Для повышения эффективности работы и защиты данных нужна отработка доверенных систем, стандартизация и регулирование.

Что касается предложений, то для эффективного развития цифровых технологий в химии и материаловедении необходимо разработать комплексную программу по цифровой химии, ориентированную на интеграцию соответствующих инструментов с фокусом на химические процессы. Следует учитывать, что химия оперирует большим количеством данных, которые имеют специфический формат и очень разнородны. Соответственно, нужны специфические алгоритмы их обработки. На стадии подготовки алгоритмов искусственного интеллекта следует предусмотреть системы и меры безопасности. Необходимые шаги — разработка цифровых платформ, создание гибкой цифровой экосистемы, повышение уровня безопасности, внедрение программных решений для автоматизации химических процессов и выполнение национальной научно-исследовательской программы в этой области. На широком круге задач было показано, что цифровые алгоритмы значительно ускоряют разработку химических процессов и помогают в выведении продуктов на рынок [4].

ЛИТЕРАТУРА

1. *Анаников В.П., Белецкая И.П., Максимов А.Л. и др.* Микротоннажная и малотоннажная химия // Химический эксперт. 2024. № 4 (12). С. 24–31. <http://zioc.ru/preprint.012024v1>

- Ananikov V.P., Beletskaya I.P., Maksimov A.L., et al.* Microtonnage and low-tonnage chemistry // Chemical expert. 2024, no. 4 (12), pp. 24–31. <http://zioc.ru/preprint.012024v1>
2. *Eremine D.B., Galushko A.S., Boiko D.A. et al.* Toward Totally Defined Nanocatalysis: Deep Learning Reveals the Extraordinary Activity of Single Pd/C Particles // J. Am. Chem. Soc. 2022, vol. 144, no. 13, pp. 6071–6079. <https://doi.org/10.1021/jacs.2c01283>
3. *Galushko A.S., Boiko D.A., Pentsak E.O. et al.* Time-Resolved Formation and Operation Maps of Pd Catalysts Suggest a Key Role of Single Atom Centers in Cross-Coupling // J. Am. Chem. Soc. 2023, vol. 145, no. 16, pp. 9092–9103. <https://doi.org/10.1021/jacs.3c00645>
4. *Ananikov V.P.* Top 20 influential AI-based technologies in chemistry // Art.Int.Chem. 2024, no. 2(2), 100075. <https://doi.org/10.1016/j.aichem.2024.100075>

DIGITAL TECHNOLOGIES IN CHEMISTRY AND MATERIALS SCIENCE

V.P. Ananikov^{a,*}

^a*Zelinsky Institute of Organic Chemistry of the Russian Academy of Sciences, Moscow, Russia*

^{*}*E-mail: val@ioc.ac.ru*

The article considers the problems and prospects of digitalization of chemical processes and technologies on the background of the global chemical landscape. Special attention is paid to such issues as low speed of data processing, the problem of lost data, knowledge transfer in micro- and low-tonnage chemistry technologies, and the digital cycle of chemical reaction production. Key areas are the application of specialized digital algorithms needed to solve specific problems in chemistry, accelerating data analysis, automating chemical research, creating new approaches to design materials for catalysis, and developing digital tools for science and educational programs. Important practical application is the use of artificial intelligence for scaling of chemical processes and technology transfer into industrial solutions.

Keywords: chemistry, materials science, digitalization of chemical processes and technologies, data processing, lost data, digital design cycle of chemical reactors, scaling processes in chemistry.