

АЛЕКСАНДР ЮРЬЕВИЧ ЧЕСАЛОВ

Atlansys EUS MedPharma

ЛЕКАРСТВО НАЧИНАЕТСЯ С ГИПОТЕЗЫ

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ, ДАННЫЕ И ЭКСПЕРИМЕНТ

Александр Чесалов

Atlansys EUS MedPharma: Лекарство начинается с гипотезы. Искусственный интеллект, данные и эксперимент

«Издательские решения»

Чесалов А. Ю.

Atlansys EUS MedPharma: Лекарство начинается с гипотезы. Искусственный интеллект, данные и эксперимент / А. Ю. Чесалов — «Издательские решения»,

Как превратить научную гипотезу в перспективное лекарственное соединение? В данной книге дается ответ на вопрос, как объединить искусственный интеллект, данные и лабораторный эксперимент в единую исследовательскую систему. На основе опыта разработки Atlansys EUS MedPharma автор проводит читателя от поиска мишеней и дизайна молекул до проверки результатов, защиты интеллектуальной собственности и внедрения платформы. Книга написана для исследователей, разработчиков и руководителей проектов.

© Чесалов А. Ю.
© Издательские решения

Содержание

От Автора 7
Благодарность 11
Об авторе 12
Введение 14
Часть I. Научные основания и методология цифровой трансформации 22
Глава 1. Фармацевтическая разработка как объект цифровой трансформации 22
1.1 Этапы поиска и разработки лекарств и критерии перехода от активного соединения к лидеру и кандидату 23
1.2 Фрагментация данных и решений в исследовательских организациях 25
1.3 Возможности и ограничения искусственного интеллекта в фармацевтических исследованиях 27
1.4 Разрыв между скоростью вычислительного поиска и пропускной способностью экспериментальной проверки 29
1.5 Сравнение классов цифровых платформ по функциям и уровню подтверждения результатов 32
1.6 Постановка научной проблемы и исследовательских вопросов монографии 34
Выводы 36
Глава 2. Методология сквозного управления разработкой новых соединений 37
2.1 Целевой профиль будущего лекарственного продукта и требования к соединению-кандидату 37
2.2 Цикл «дизайн — синтез — тестирование — анализ» и уровни доказательности исследовательских результатов 39
2.3 Совместное использование данных, физических моделей, машинного обучения и экспертных решений 41
2.4 Многокритериальное принятие решений при ограниченных ресурсах и неопределенности 43
2.5 Критерии продолжения и прекращения проекта и правила пересмотра исследовательской гипотезы 45
2.6 Принципы измерения научной, технологической и организационной результативности 47
Выводы 50
Глава 2 Методология сквозного управления разработкой новых соединений 52
2.1 Целевой профиль будущего лекарственного продукта и требования к соединению-кандидату 52
2.2 Цикл «дизайн — синтез — тестирование — анализ» и уровни доказательности исследовательских результатов 54
2.3 Совместное использование данных, физических моделей, машинного обучения и экспертных решений 57
2.4 Многокритериальное принятие решений при ограниченных ресурсах и неопределенности 59

2.5 Критерии продолжения и прекращения проекта и правила пересмотра исследовательской гипотезы 61
2.6 Принципы измерения научной, технологической и организационной результативности 63
Выводы 66
Конец ознакомительного фрагмента. 67

Atlansys EUS MedPharma: Лекарство начинается с гипотезы Искусственный интеллект, данные и эксперимент

Александр Юрьевич Чесалов

Дизайнер обложки Александр Юрьевич Чесалов
Иллюстратор Александр Юрьевич Чесалов
Иллюстрации Нейросеть Stable Diffusion WebUI Forge

© Александр Юрьевич Чесалов, 2026
© Александр Юрьевич Чесалов, дизайн обложки, 2026
© Александр Юрьевич Чесалов, иллюстрации, 2026

ISBN 978-5-0071-5578-6
Создано в интеллектуальной издательской системе Ridero

От Автора

К этой книге я пришел, работая над платформой Atlansys EUS MedPharma.

Все началось в 2021 году. В компании «Программные системы Атлансис» стартовали работы по созданию отечественной цифровой платформы Atlansys EUS MedPharma для фармацевтической отрасли. В 2023 году первый вариант платформы был зарегистрирован в Роспатенте (Свидетельство №2023619721).

Целью разработки Atlansys EUS MedPharma — стало создание передовой платформы искусственного интеллекта для достижения на ее основе результатов мирового уровня в сфере разработки инновационных фармацевтических препаратов и медицинских технологий.

Название «EUS — Electronic Universal System» было выбрано не случайно. Оно является отсылкой к 80 и 90-м годам двадцатого века и данью уважения к инженерам и программистам того времени.

В 2025 году нами были получены первые практические результаты: созданы новые алгоритмы машинного обучения (Свидетельство о государственной регистрации программ для ЭВМ №2025665605 и №2025687503) и разработаны новые соединения — кандидаты (Патент на изобретение №2860793, №2861491 и №2865864). Мы подали соответствующие заявки на регистрацию патентов.

На сегодняшний день, к проекту развития платформы Atlansys EUS MedPharma присоединились ведущие ученые и специалисты из более чем десяти ведущих институтов и университетов Российской Федерации.

Проектирование ее архитектуры поставило передо мной вопросы, которые выходят за пределы выбора алгоритмов и программных средств. Меня интересовало, как донести смысл данных при передаче между специалистами и связать вычислительные расчеты с реальным веществом и результатом его испытания. Отдельного внимания требовало понимание принятия решения о продолжении исследования. Рассмотрение этих вопросов определило содержание настоящей монографии.

Эта монография основана на глубоком переосмыслении опыта разработки платформы Atlansys EUS MedPharma. Работа над ней стала поводом заново рассмотреть исходные предпосылки, согласовать понятия и определить, какие положения уже получили методологическое подтверждение, а какие требуют дальнейшей проверки. Я стремился превратить опыт проектирования конкретной системы в доступный для обсуждения и применения подход к организации фармацевтических исследований и применения новых иноформенных технологий для них.

Эта книга посвящена тому, как интегрировать данные, знания, вычислительные модели, лабораторные процедуры и решения специалистов в последовательный процесс ранней разработки новых соединений. Ее назначение — дать читателю общее основание для постановки исследовательских задач, проектирования подобных платформ и оценки получаемых результатов. В работе технологические возможности рассматриваются вместе с условиями, при которых ими можно обоснованно пользоваться.

Я адресую монографию исследователям в области фармацевтики, медицинской химии, биологии и фармакологии, специалистам по биоинформатике, разработчикам алгоритмов и моделей искусственного интеллекта и архитекторам информационных систем. Она также предназначена руководителям научных подразделений, лабораторий и технологических проектов, которым необходимо согласовывать научные цели, ресурсы и ответственность участников в подобного рода проектах.

Отдельную аудиторию составляют преподаватели, аспиранты и студенты, выбирающие исследования на пересечении естественных наук и информационных технологий. Для них книга может служить ориентиром в междисциплинарной области: показывать, как медицинский вопрос становится вычислительной задачей, каким образом результат модели поступает в эксперимент и почему следующий вывод зависит от условий получения данных.

Мне хотелось, чтобы специалист из одной области мог понять основания работы коллег из другой. Для этого в книге последовательно раскрываются связи между исследовательскими действиями. Формальные модели и архитектурные решения сопровождаются объяснением их назначения. Такое изложение позволяет читать монографию и как целостное исследование, и использовать как основу для обсуждения конкретного проекта.

Недавно, 21 октября 2026 года, в журнале ВАК категории К1 «Информационные технологии» вышла моя научная статья «Архитектура платформы Atlansys EUS MedPharma, как инструмент цифровой трансформации процессов фармацевтической разработки новых соединений на основе технологий искусственного интеллекта»: том 32, №10, с. 542—550, DOI: 10.17587/it.32.542—550 [1]. В ней представлена исходная архитектурная постановка, получившая развитие в этой книге.

Формат статьи позволил обозначить проблему фрагментации фармацевтической разработки и представить структуру платформы. Монография дает возможность подробнее исследовать связи между ее частями: от качества исходных сведений и формирования гипотез до получения материала, проверки свойств и оценки результативности процесса. При таком переходе архитектурная схема становится основанием для обсуждения методов, ограничений и способов проверки решений.

Авторский опыт занимает в исследовании определенное место. Он служит источником вопросов, проектных требований и предметных примеров. Его научное обобщение предполагает ясное описание условий, сопоставление с опубликованными работами и возможность независимого рассмотрения предложенных решений. Именно поэтому в книге отдельно определяются статусы вычислительного прогноза, экспериментального наблюдения и принятого решения, а также границы выводов об эффективности рассматриваемого решения.

Для меня переосмысление опыта означает готовность уточнять первоначальные представления. По мере развития проекта меняется понимание того, какие связи между компонентами существенны, где возникают потери контекста и какой результат необходимо считать основанием следующего шага. В книге эта изменчивость представлена как естественная часть исследовательской работы, для которой нужны прослеживаемая история и понятные правила пересмотра исследований.

Научный вклад монографии состоит в согласовании методологии сквозного управления ранней разработкой с предложенной мной в [1] архитектурой исследовательской платформы. В единой формате рассматриваются постановка гипотезы, подготовка данных, применение моделей, экспериментальная проверка и оценка последующего решения. Такой подход позволяет определить, что именно должен обеспечивать каждый компонент и каким образом проверять полезность их совместной работы между собой.

Для фармацевтической отрасли существенна связь между цифровым описанием и исследуемым материальным объектом. В монографии она раскрывается через идентичность соединения, форму вещества, образец, условия измерения и требования к результату. За счет этого моделирование мишеней, генеративный дизайн, виртуальный скрининг, прогнозирование фармакокинетики и безопасности, оценка синтетической осуществимости рассматриваются в общем исследовательском маршруте.

Вклад в развитие ИИ-технологий связан с условиями их содержательного применения. Для модели определяются назначение, допустимые входы, основания оценки и поря-

док использования результата. Для интегрированной системы — правила обмена, сохранения неопределенности и возврата новых наблюдений. Современная дискуссия о применении ИИ в разработке лекарств подчеркивает необходимость связывать оценку методов с их способностью улучшать исследовательские решения. Эта постановка соответствует общей направленности книги [3].

Особое значение имеет воспроизводимость. Она требует восстановить связь результата с данными, программной версией и фактическим исполнением. Работа [4] показывает, как описывать происхождение вычислительного процесса и связывать его входы, выходы и программные компоненты. В предлагаемой методологии эта линия продолжается до экспериментального материала и решения специалиста.

Предложенные положения объединяют научные и организационные условия применения ИИ. Совместные принципы FDA и EMA 2026 года связывают надлежащую практику с понятным контекстом использования, управлением данными, междисциплинарной экспертизой и проверкой, соразмерной риску. В книге эти вопросы рассматриваются через обязанности участников, правила допуска результатов и сопровождение изменений [5].

Содержание научного вклада определяется разработанными методологическими конструкциями, архитектурными принципами и программой их проверки. Величина эффекта конкретного внедрения устанавливается в сравнительных исследованиях. Такое соотношение помогает сохранять точность научных утверждений и развивать платформу по мере накопления новых данных.

Считаю очень важным отметить: редактирование и корректура авторского текста, а также работа над содержанием данной монографии выполнены при помощи ИИ-ассистента научно-образовательной платформы «Королев ИИ» МГТУ им. Н. Э. Баумана. Прошу читателя обратить внимание именно на слово «авторского»: исходный замысел, исследовательская проблематика и авторский материал составляют основу этой работы. Использование интеллектуального помощника является частью открыто описываемого процесса подготовки рукописи.

Для меня это важно, как для руководителя проекта «Королев ИИ», разработчика платформы и пользователя решения, которое мы создаем. Я хочу понимать, насколько наш продукт эффективно справляется с поставленными задачами и чем он может быть полезен широкой аудитории университета. Работа с собственным текстом дает возможность увидеть, где ассистент помогает автору, какие задачи выполняет лучше и какие решения нуждаются в дополнительном внимании.

В отдельных местах после редактирования, возможно, изложение стало более формальным или, если так можно сказать, «синтетическим». Это тоже часть опыта применения ИИ, которую я считаю полезным обозначить. При этом ассистент не просто сгенерировал текст, а существенно переработал исходный авторский материал. Лучше всего у него получилось обновить содержание работы и структуру уже написанных глав, подсказать, где сделать последовательность изложения более согласованной и выявить повторения.

Я не стесняюсь говорить об этом открыто: считаю такой результат достижением нашей команды. Для нас важно создавать продукт, полезность которого видна в конкретной работе преподавателя, исследователя и студента. Взаимодействие с ассистентом помогает точнее понимать требования к этой работе и совершенствовать решение на основании собственного пользовательского опыта.

Ответственность за научное содержание, точность существенных утверждений и окончательный выбор формулировок я сохраняю за собой. Мне важно, чтобы редакционная помощь поддерживала ясность и последовательность авторского исследования. Открытое описание роли ИИ позволяет читателю понимать, как готовилась книга и какое место интеллектуальный помощник занимал в этом процессе.

Монография включает двадцать глав, объединенных в пять частей. Последовательное чтение ведет от научных оснований цифровой трансформации к архитектуре, методам поиска соединений, экспериментальной проверке и условиям внедрения. Введение раскрывает исследовательскую постановку, а заключение связывает полученные результаты с целью работы и направлениями дальнейших исследований.

При обращении к отдельной предметной теме полезно сохранять общий вопрос книги: какое решение предстоит принять и какие основания для него необходимы? Этот вопрос помогает соотнести технический метод с данными, лабораторной задачей и доступными ресурсами. Я надеюсь, что предложенный подход будет полезен читателю при создании собственных исследовательских систем и в профессиональном обсуждении того, каким должно быть обоснованное применение ИИ в фармацевтической отрасли.

Благодарность

Выражаю глубокую благодарность всем тем людям, которые вдохновили меня своим примером, поддержали в мое работе и помогли мне поверить в свои силы, чтобы провести работу над накопленным за последние годы материалом и приступить к работе над проектом создания центра компетенций в области ИТ и искусственного интеллекта для разработки инновационных фармацевтических препаратов и медицинских технологий на основе платформы Atlansys EUS MedPharma.

Особую благодарность выражаю выдающемуся Советскому и Российскому ученому, доктору химических наук, профессору, заведующей отделом химии и технологии Всесоюзного научного центра по безопасности биологически активных веществ (АО «ВНЦ БАВ»), которая проработала более шестидесяти лет в области создания лекарственных препаратов Скачиловой Софии Яковлевне, а также колегам из АО «ВНЦ БАВ».

Об авторе



Александр Юрьевич Чесалов родился 10 февраля 1977 года в городе Тверь, Российская Федерация.

Экономист по образованию, со специализацией «Информационные системы в экономике». Окончил с отличием Тверской Государственный Технический Университет. Защитил докторскую диссертацию на тему «Методология определения операционных характеристик и рациональной структуры региональных распределенных сервисных сетей передачи, обработки и хранения данных».

Автор более 35 книг, 3 учебников по искусственному интеллекту, 4 монографий, 6 патентов на изобретения и 25 свидетельств о регистрации программ для ЭВМ. Всего более 100 различных публикаций.

Среди них наиболее известны следующие: «Как создать центр искусственного интеллекта за 100 дней»; «Генеративный искусственный интеллект #Forge&flux»; «Невероятный искусственный интеллект Easy Diffusion 3.0»; «Интеллектуальные системы на службе заводов: как предсказать поломку и сэкономить миллионы»; «Цифровая среда Института омбудсмена: концепция, технологии, практика»; «Глоссариум по искусственному интеллекту: 2500 терминов»;

«Методология построения распределенных сетей передачи, обработки и хранения данных» в двух томах, и многие другие.

Член Экспертного совета при Комитете Государственной Думы по науке и высшему образованию по вопросам развития информационных технологий в сфере образования и науки, а также член Российской ассоциации искусственного интеллекта (РАИИ). Эксперт по стратегии программы «Приоритет 2030», руководитель и разработчик проекта «Королев ИИ» МГТУ им. Н. Э. Баумана.

Сертификаты в области ИТ: IBM Professional certificate foundations of AI; IBM Professional certificate Essential Technologies for Business; Rutgers the State University of New Jersey: New Technologies for Business Leaders; University of London; Deeplearning.ai; Microsoft Azure; BSI ISO/IEC 27001; IBM DB2 и другие.

Научные интересы: искусственный интеллект, автоматизированные системы управления предприятиями, автоматизированные системы прогнозируемого и предписывающего обслуживания, системный анализ, моделирование сложных систем, корпоративные информационные системы, системы управления информационной и промышленной безопасностью.

Блог автора: chesalov.com

Введение

Лекарственный препарат начинается с исследовательского вопроса, но становится возможным благодаря последовательности проверенных решений. Между предположением о механизме заболевания и выбором соединения для дальнейшей разработки необходимо связать биологические основания, химическую структуру, свойства полученного вещества и результаты его испытаний. Каждое новое наблюдение может изменить первоначальный замысел. Поэтому качество фармацевтической разработки определяется тем, насколько обоснованно исследовательская группа переходит от одного состояния знания к другому.

Искусственный интеллект расширяет возможности такого перехода. Он помогает находить связи в биомедицинских данных, предсказывать структуры и свойства молекул, формировать варианты соединений и выбирать направления экспериментального поиска. Вместе с этим возникает вопрос, которому посвящена настоящая монография: каким образом объединить вычислительные методы, лабораторные процедуры и решения специалистов в управляемый исследовательский процесс, результат которого можно проверить и воспроизвести?

Основой работы служит концепция Atlansys EUS MedPharma, в которой поставлена задача преодоления информационной и методологической фрагментарности разработки новых соединений посредством интеграции специализированных технологий искусственного интеллекта. Исходная статья связывает эту задачу с модульной архитектурой платформы, управлением данными, поддержкой научного анализа и организацией экспериментальной проверки. В монографии данная концепция развивается в методологию сквозного управления ранней фармацевтической разработкой [1].

Современный исследователь располагает несколькими способами описать один и тот же объект. Соединение может быть представлено химической структурой, набором дескрипторов, пространственной конфигурацией и профилем клеточного ответа. Для белковой мишени доступны последовательности, структурные модели, функциональные аннотации и сведения о связи с заболеванием. Совместное использование этих представлений полезно лишь при сохранении условий, в которых получены исходные данные. Несопоставимые измерения, неустановленная идентичность образца или потеря происхождения результата способны сделать убедительный расчет непригодным для последующего решения.

Актуальность цифровой трансформации определяется необходимостью повысить результативность исследовательского процесса при ограниченных ресурсах. Вычислительный отбор способен быстро сформировать множество предложений, тогда как получение вещества, аналитическое подтверждение и биологическое испытание требуют собственного времени и доступных мощностей. Поэтому ускорение отдельного расчета приобретает практический смысл через его влияние на выбор объектов, экспериментальную очередь и время получения пригодного результата. Предсказательная валидность исследовательских моделей, то есть связь получаемых оценок с последующим значимым исходом, рассматривается в современной литературе как существенное основание продуктивности разработки лекарств [6].

Особенно важна эта постановка для проектов, в которых перспективность соединения зависит одновременно от активности, селективности, экспозиции, безопасности и возможности получения материала. Улучшение одного свойства может сопровождаться ухудшением другого. Для нейропротекторного направления значение имеют доступность вещества в целевой ткани и условия реализации эффекта; для противоопухолевого направления — биологический контекст, избирательность действия и пригодность экспериментальной модели. Эти вопросы задают требования к данным и программе проверки уже на ранних этапах работы.

Критическая оценка состояния области также меняет критерии успеха цифровых решений. В опубликованной в 2026 году работе A. Bender и соавторов предлагается оценивать методы ИИ через их способность улучшать исследовательские решения и учитывать условия будущего применения. Авторы обращают внимание на ограниченность доказательств клинически значимого вклада многих вычислительных подходов и на необходимость более точной постановки исследуемых задач. Для настоящей монографии это служит основанием связывать качество алгоритма с его ролью в конкретном процессе разработки [3].

Для исследовательских организаций практическая актуальность такого подхода состоит в возможности согласовать работу химиков, биологов, специалистов по данным и разработчиков программного обеспечения. Общая цифровая среда должна сохранять основания передачи задания, происхождение результата и полномочия участников. Тогда становится возможным установить, какое изменение метода или организации работ повлияло на итог исследования и какие условия необходимо воспроизвести на другой площадке.

Научная основа монографии сформирована с учетом публикаций и методических документов последних пяти лет по состоянию на 25 октября 2026 года. Основное внимание уделяется работам 2022—2026 годов, в которых описаны методы, условия проверки и границы интерпретации результатов. Обзор имеет проблемно ориентированный характер: источники сопоставляются по их вкладу в построение и оценку сквозного исследовательского процесса.

Первое направление связано с расширением возможностей структурного моделирования. AlphaFold 3 представляет подход к совместному предсказанию структуры комплексов, включающих белки, нуклеиновые кислоты, малые молекулы и другие компоненты. Такие методы увеличивают доступность структурных гипотез для молекулярного дизайна. Исследовательская ценность конкретной модели при этом определяется ее соответствием объекту, состоянию биологической системы и вопросу последующей проверки [7].

Второе направление охватывает генеративный дизайн и предварительную оценку свойств. В REINVENT 4 объединены средства построения и оптимизации молекулярных структур для нескольких постановок задачи, включая изменение заместителей и химического каркаса. ADMET-AI демонстрирует возможность масштабной оценки ряда фармакокинетических и токсикологических характеристик химических библиотек. Для интегрированной платформы эти работы показывают, как формировать вычислительные предложения и сопровождать их оценками; решение о дальнейшем исследовании требует учета применимости каждой модели и доступности экспериментального подтверждения [8, 9].

Третье направление связано с организацией взаимодействия вычислений и лаборатории. В работе D. A. Boiko и соавторов система Coscientist объединяет языковую модель, поиск информации, программные инструменты и экспериментальное оборудование для выполнения исследовательских химических задач. Этот пример раскрывает возможности программной координации действий и использования результатов опытов. Перенос такого подхода в фармацевтический проект требует установить допустимые операции, критерии качества материала и измерений, а также порядок научного руководства [10].

Четвертое направление относится к независимой проверке вычислительного поиска. В первом исследовании CACHE разные команды выбирали предполагаемые лиганды одной мишени, после чего соединения проходили общую экспериментальную оценку. Данный подход важен возможностью сопоставить способы отбора на согласованном основании. Для методологии Atlansys EUS MedPharma он поддерживает переход от сравнения расчетных рейтингов к проверке реально выбранных исследовательских объектов [11].

Развитие области включает и клинические исследования соединений, при создании которых применялся ИИ. В 2025 году опубликовано рандомизированное исследование фазы 2a рентосертиба, ингибитора TNIK, при идиопатическом легочном фиброзе. Оно представляет

клинический результат определенной программы разработки. Для оценки цифровых платформ существенно сохранять предмет этого результата: испытание препарата характеризует соответствующее вмешательство и само по себе не определяет величину вклада отдельного алгоритма или интеграционной архитектуры в затраты и сроки поиска [12].

Одновременно совершенствуются способы проверки надежности моделей. Исследование D. Graber и соавторов, дополненное в 2026 году, показывает существенное влияние смещений и пересечений данных на оценку обобщения при прогнозировании аффинности. В DataSAIL предложено разделение данных с учетом сходства биологических объектов; дополнение 2026 года расширяет сравнение с разделениями, опубликованными вместе с исходными наборами. Эти работы обосновывают необходимость согласовать контрольную выборку с предполагаемым сценарием применения модели [13, 14].

На уровне научной инфраструктуры развиваются средства описания происхождения результатов. Workflow Run RO-Crate позволяет связывать входы, выходы, программные компоненты и сведения о выполнении вычислительного процесса. Для фармацевтической платформы такой подход создает основу воспроизводимой регистрации расчетов; его необходимо дополнить прослеживаемостью материального образца, протокола измерения и экспертного решения [4].

Методические документы последних лет поддерживают оценку моделей в определенном контексте. Второе издание OECD (Q) SAR Assessment Framework 2024 года разграничивает рассмотрение модели, отдельного предсказания и результата, основанного на нескольких прогнозах. Документ EMA 2024 года рассматривает применение ИИ на протяжении жизненного цикла лекарственного продукта с учетом риска и назначения метода. Совместные принципы FDA и EMA 2026 года связывают надлежащую практику с управлением данными, междисциплинарной экспертизой, оценкой характеристик системы и сопровождением ее жизненного цикла [5, 15, 16].

Сопоставление этих направлений позволяет сформулировать исходную позицию монографии. Современная литература дает основания для применения отдельных вычислительных методов и содержит примеры организации замкнутых исследовательских процессов. Для разработки Atlansys EUS MedPharma необходимо определить, как согласовать такие возможности в общей архитектуре, сохранить смысл передаваемых результатов и установить условия проверки добавочного эффекта интеграции.

Научная проблема работы заключается в разработке согласованной методологии связывания исследовательских гипотез, разнородных данных, вычислительных моделей, материальных образцов и решений специалистов в воспроизводимом процессе поиска и оптимизации новых соединений. Методология должна учитывать неопределенность прогнозов, ограничения экспериментальной базы и изменение знания по мере выполнения исследований.

Основное противоречие проявляется между расширением пространства вычислительных возможностей и ограниченной способностью исследовательской организации подтверждать предлагаемые варианты. Увеличение числа сгенерированных структур может повышать нагрузку на лабораторию. Более точная модель одного свойства может мало влиять на выбор кандидата, если решающим ограничением оказывается другое свойство. Дополнительная интеграция способна сократить потери при передаче заданий, но потребовать значительных расходов на сопровождение. Следовательно, эффект платформы необходимо устанавливать для определенного исследовательского маршрута и согласованного состава ресурсов.

В предлагаемой постановке фрагментарность имеет содержательный и организационный характер. Один и тот же показатель может получить различную интерпретацию в вычислительной и экспериментальной группах; одна химическая структура может соответствовать нескольким материальным формам; результат без сведений о протоколе может быть ошибочно

использован как универсальное свойство вещества. Преодоление таких разрывов требует правил идентификации объектов, передачи неопределенности, допуска данных и документирования оснований следующего действия.

Центральный исследовательский вопрос состоит в том, при каких условиях согласованная работа этих правил и компонентов улучшает результат ранней разработки по сравнению с доступными альтернативами. Вопрос охватывает архитектуру взаимодействия, качество исходных оснований и способ оценки изменений. Его решение требует общей единицы анализа, которая сохраняет связь между научным замыслом и фактически выполненной работой.

Объектом исследования являются процессы раннего поиска, проектирования, оптимизации и экспериментального подтверждения новых соединений и многокомпонентных систем, предназначенных для последующей разработки лекарственных средств. Эти процессы рассматриваются во взаимодействии вычислительной среды, экспериментальной лаборатории и специалистов, принимающих исследовательские решения.

Предмет исследования составляют методы и архитектурные принципы сквозного управления указанными процессами: согласование целевых требований, верификация и интеграция данных, выбор и совместное применение моделей, учет неопределенности, планирование проверок, прослеживаемость результатов и оценка научной, экономической и организационной результативности платформы.

Единицей сквозного управления в монографии выступает проверяемая исследовательская гипотеза, связанная с идентифицированными объектами, программой получения данных и возможным решением. Для ее сопровождения используется паспорт гипотезы. По мере движения по исследовательскому маршруту к нему присоединяются сведения о структуре, материале, измерениях, версиях моделей и основаниях выбора. При сравнительной оценке платформы единица статистического анализа определяется дизайном исследования; для оценки целого процесса ею может быть независимая кампания или проект.

Цель исследования состоит в разработке и методологическом обосновании принципов построения и применения платформы Atlansys EUS MedPharma, обеспечивающей прослеживаемую интеграцию данных, вычислительных моделей и экспериментальных процедур для управляемой ранней разработки новых соединений с проверяемой оценкой ее результативности.

Для достижения этой цели решаются следующие взаимосвязанные задачи.

1. Определить содержание ранней фармацевтической разработки как объекта цифровой трансформации, разграничить состояния исследуемого объекта и установить связь каждого перехода с необходимыми основаниями решения.

2. Сформировать методологию сквозного управления, объединяющую целевой профиль разработки, паспорт гипотезы, учет неопределенности и критерии продолжения, уточнения или прекращения исследования.

3. Согласовать функциональную и логическую архитектуру платформы, распределить ответственность между системами, подсистемами, исследовательскими модулями и общими сервисами, определить контракты обмена данными и результатами.

4. Разработать правила подготовки и допуска биомедицинских и химических данных, интеграции знаний и формирования исследовательских гипотез с сохранением происхождения, контекста и ограничений исходных сведений.

5. Определить порядок совместного использования структурного моделирования, молекулярных представлений, генеративного дизайна, виртуального скрининга, прогнозирования фармакокинетики и безопасности, а также оценки синтетической осуществимости.

6. Сформировать требования к замыканию вычислительного и экспериментального цикла и программу проверки моделей, исследовательских решений и добавочного эффекта интегрированной платформы.

7. Обосновать организацию управления данными и инфраструктурой, научной отчетности, информационной безопасности и экспертной ответственности на протяжении исследовательского маршрута.

8. Разработать согласованные основания патентных исследований, экономической оценки и поэтапного внедрения платформы с учетом подтвержденных результатов, затрат, рисков и условий переноса на новые проекты.

Последовательность задач отражает развитие аргументации книги. Требования к исследовательскому решению определяют содержание архитектуры; архитектура задает правила взаимодействия методов и лаборатории; программа проверки устанавливает, какие выводы можно сделать о результативности полученной организации работ.

Рабочая гипотеза состоит в том, что при сопоставимых моделях, исходных данных, критериях подтверждения и ресурсных ограничениях прослеживаемая интеграция вычислительных и экспериментальных этапов повышает выход подтвержденных пригодных объектов ранней разработки по сравнению с раздельным применением тех же инструментов. Предполагаемый механизм улучшения связан с уменьшением потерь контекста при передаче заданий, более обоснованным распределением экспериментальной очереди и своевременным возвратом результатов в процесс принятия решений.

Выход подтвержденных объектов определяется для заранее установленной категории: активное соединение, соединение-лидер или иной согласованный результат выбранного этапа. Их содержание и критерии приемки фиксируются до сравнения. Дополнительно оцениваются календарное время до соответствующего результата, полные затраты в согласованных границах и воспроизводимость ключевых наблюдений. Квалифицированные решения об остановке бесперспективного направления сохраняют самостоятельный учет и собственное экономическое обоснование.

Для проверки гипотезы в монографии предусмотрено сравнение трех исследовательских режимов: принятого в организации процесса; применения выбранных ИИ-инструментов с существующим порядком взаимодействия; использования сопоставимых инструментов в интегрированном процессе Atlansys EUS MedPharma. Сравнение второго и третьего режимов позволяет оценивать добавочный эффект интеграции. При изменении моделей, данных или доступных ресурсов эти различия должны быть учтены в интерпретации результата.

Гипотеза допускает опровержение. Если при согласованном бюджете интегрированный режим не увеличивает выход подтвержденных объектов либо ухудшает выполнение критических требований к их пригодности, оснований считать ее подтвержденной для данного контекста нет. Ускорение отдельного расчета, рост числа сгенерированных молекул или полнота цифровой документации характеризуют отдельные стороны работы, но требуют самостоятельной связи с основным исследовательским исходом.

Методологическую основу работы составляет системный подход к фармацевтической разработке как к взаимодействию научных объектов, данных, технических средств и решений специалистов. Используются функциональный и процессный анализ, архитектурное моделирование, онтологическое описание предметной области и анализ происхождения результатов. Эти методы применяются для установления границ ответственности и согласования переходов между этапами исследования.

Научные публикации сопоставляются по назначению метода, составу данных, устройству контрольной проверки и уровню подтверждаемого вывода. Отдельно рассматриваются описа-

ния программных средств, вычислительные исследования, лабораторные работы и клинические испытания. Такое сопоставление позволяет определить, какое положение применимо к проектированию платформы и какие дополнительные основания необходимы для его переноса в конкретный проект.

В вычислительном контуре рассматривается совместное использование химоинформатики и биоинформатики, физических моделей, графовых нейронных сетей, генеративных методов, машинного обучения и вероятностной оценки неопределенности. Выбор средства подчиняется исследовательской задаче и доступным данным. Методы многокритериальной оптимизации применяются для сохранения существенных компромиссов между свойствами; активное обучение и байесовское планирование — для определения информативной следующей проверки при ограниченном экспериментальном бюджете. Их алгоритмическое содержание раскрывается в соответствующих главах.

Программа проверки включает ретроспективное воспроизведение, оценку на независимых данных и проспективные исследования с фиксацией прогнозов до получения новых измерений. Разделение данных должно соответствовать предполагаемому переносу: во времени, между химическими сериями, мишенями, лабораторными партиями или организациями. Для оценки вклада отдельных компонентов рассматриваются исследования с их исключением при сопоставимом порядке настройки моделей.

При сравнении процессов необходимы заранее заданный основной исход, правила учета неудач и незавершенных работ, обоснование объема наблюдений и статистический анализ на уровне независимых единиц. Общий протокол экспериментальной проверки уменьшает различия в критериях подтверждения. При невозможности случайного распределения используются сопоставимые кампании с анализом существенных исходных различий; оставшиеся ограничения причинной интерпретации раскрываются явно.

Экономическая оценка строится на связи исследовательских событий с полными ресурсными затратами, календарными сроками и подтвержденными исходами. Сценарный анализ и анализ чувствительности используются для проверки зависимости выводов от предпосылок. Организационная оценка рассматривает освоение процесса, распределение ответственности и возможность его самостоятельного воспроизведения подготовленной командой. Формулы и условные расчеты в книге поясняют методику; фактический эффект конкретного внедрения устанавливается по его собственным данным.

Основная область применения предлагаемой методологии охватывает ранний исследовательский этап: формирование и уточнение гипотезы, выбор мишени, поиск и оптимизацию соединений, получение и проверку материала, оценку свойств и принятие решения о дальнейшей программе. Переход к разработке лекарственной формы, технологии производства и последующим исследованиям рассматривается через требования к передаваемым данным и результатам. Полный объем доклинической и клинической разработки определяется отдельными программами.

Рассматриваемые структурные, фармакокинетические и токсикологические прогнозы имеют собственные области применимости. Вывод об аффинности не устанавливает требуемую экспозицию в ткани; сведения о проникновении не определяют терапевтическую эффективность; изменение клеточного показателя сохраняет связь с условиями соответствующего опыта. Для многокомпонентных систем отдельно устанавливаются состав и организация материала, а утверждение о синергизме проверяется относительно выбранной модели совместного действия.

Авторские нейропротекторные и противоопухолевые разработки задают предметные ориентиры и материал для описания требований к исследовательскому досье. В новой структуре книги сведения о них используются в тематических главах и при анализе интеллектуаль-

ной собственности. Публикации, регистрации программ и патентные документы характеризуют соответствующие результаты и правовые объекты; объем научного вывода о свойствах вещества определяется доступными методами, первичными наблюдениями и условиями проверки [1].

Применимость архитектуры зависит от качества и правомерной доступности данных, возможностей лаборатории, компетенций участников и условий эксплуатации вычислительных средств. Перенос на новую мишень, химический класс, биологическую модель или организацию требует оценки изменившихся предпосылок. Степень автоматизации устанавливается для конкретных действий; ответственность за научную интерпретацию и решение сохраняется в системе управления исследованием.

Документы OECD, EMA и FDA используются в пределах их предмета и статуса. Методические ориентиры оценки ИИ помогают определить контекст применения, требования к данным и объем проверки. Обязательность конкретных положений для организации и исследовательского процесса устанавливается по применимому праву и сфере деятельности; соответствующие вопросы развиваются в главе 17 [5, 15, 16].

Научный результат монографии имеет методологический и архитектурный характер. Его содержание раскрывается через взаимосвязанные положения, для которых определены исследовательский смысл, форма представления и способ дальнейшей проверки.

Первый результат — методология сквозного управления ранней разработкой новых соединений. Она объединяет целевые требования, паспорт гипотезы, программу получения данных и правила исследовательского решения. Благодаря этому изменение гипотезы, модели или экспериментального плана может быть связано с конкретным основанием и с последствиями для следующих этапов.

Второй результат — согласованная референсная архитектура Atlansys EUS MedPharma. В ней раскрыта связь четырех базовых систем, восемнадцати подсистем и сорока двух базовых исследовательских модулей системы исследований и разработок с логическими слоями, общими сервисами и контрактами обмена. Архитектура определяет функциональные обязанности компонентов и позволяет описывать конкретные конфигурации реализации без утраты общего содержания исследовательского маршрута.

Третий результат — система правил подготовки, интеграции и использования данных и знаний. Она связывает химическую и биологическую идентичность, материальный образец, содержание измерения, происхождение данных и условия их допуска. Паспорта наборов данных и исследовательских объектов обеспечивают возможность восстановить основания вычислительного вывода и определить последствия обнаруженной ошибки.

Четвертый результат — порядок согласования вычислительного дизайна с экспериментальной проверкой. Он предусматривает передачу неопределенности между методами, учет синтетической осуществимости и свойств материала, планирование следующего опыта и различение отрицательного наблюдения, отсутствующего измерения и технической неудачи. Этот порядок сохраняет пригодность обратной связи для пересмотра гипотез и обновления моделей.

Пятый результат — программа многоуровневой валидации. Она разграничивает проверку модели, обоснованности исследовательского решения и результативности целого платформенного процесса; включает независимые данные, проспективную проверку и сравнительный дизайн оценки вклада интеграции. Для каждого уровня устанавливаются собственные критерии, единицы анализа и границы допустимого вывода.

Шестой результат — согласованные основания экономической и организационной оценки, управления интеллектуальной собственностью и поэтапного внедрения. Они связывают затраты и сроки с подтвержденными исследовательскими исходами, определяют требо-

вания к передаче результатов и позволяют рассматривать расширение платформы как последовательность проверяемых переходов.

Специфика предлагаемого подхода заключается в согласовании этих положений вокруг одной исследовательской гипотезы и ее проверок. В монографии сформированы методологические конструкции и требования к реализации, а также программа установления их результативности. Количественная величина эффекта интеграции и переносимость конкретного внедрения остаются предметом сравнительных исследований; условные примеры расчетов используются для объяснения способов оценки.

Практическое значение работы состоит в создании общего основания для проектирования, пилотирования и сопровождения фармацевтической исследовательской платформы. Для научного руководителя оно определяет связь цели с программой проверок; для лаборатории — требования к заданию, материалу и результату; для разработчика — обязанности компонентов и правила обмена; для организации — основания оценки затрат, готовности и последующего расширения. Такая согласованность помогает обсуждать цифровую трансформацию через конкретные исследовательские действия и проверяемые результаты.

Книга включает двадцать глав, объединенных в пять частей. Ее последовательность выстроена от постановки исследовательской проблемы и методологии к архитектуре, вычислительным методам, экспериментальной проверке и условиям внедрения.

В первой части, главы 1 и 2, фармацевтическая разработка рассматривается как объект цифровой трансформации. Формируется методология сквозного управления: от целевых требований и исследовательских статусов к правилам работы с неопределенностью и оценке результата.

Во второй части, главы 3—6, эти положения раскрываются через архитектуру платформы, подготовку данных, интеграцию знаний и формирование проверяемых гипотез. Определяются объекты обмена между компонентами и основания их использования.

Третья часть, главы 7—12, посвящена мишеням и белковым структурам, молекулярным представлениям, генеративному дизайну, виртуальному скринингу, фармакокинетике, безопасности и синтетической осуществимости. Показано, как совместно применять эти методы, сохраняя ограничения каждого из них.

Четвертая часть, главы 13—17, связывает вычислительный маршрут с лабораторией, валидацией, инфраструктурой, научной отчетностью и информационной безопасностью. Устанавливаются условия воспроизводимости исследовательского цикла и доверия к его результатам.

Пятая часть, главы 18—20, объединяет патентные исследования и управление интеллектуальной собственностью, оценку экономической и организационной результативности, внедрение и развитие платформы. Здесь научные положения связываются с условиями их практического использования.

Дальнейшее изложение начинается с анализа фармацевтической разработки как процесса, в котором данные, методы и организация работы совместно определяют качество исследовательского решения.

Часть I. Научные основания и методология цифровой трансформации

Глава 1. Фармацевтическая разработка как объект цифровой трансформации

Лекарство начинается с предположения о том, какое изменение в биологической системе способно помочь человеку. Между этим предположением и обоснованным выбором соединения для дальнейшей разработки лежит последовательность исследований. Каждый следующий опыт уточняет первоначальный замысел, иногда существенно меняя его. Цифровая трансформация приобретает научный и практический смысл тогда, когда помогает организовать эту последовательность: сохранить основания гипотезы, выбрать информативный эксперимент и связать полученный результат с последующим решением.

В исходной концепции Atlansys EUS MedPharma поставлена задача преодоления информационной и методологической фрагментарности разработки новых соединений посредством интеграции специализированных методов искусственного интеллекта в общую платформенную архитектуру. В ней выделены системы исследований и разработок, управления данными и инфраструктурой, визуализации и отчетности, информационной безопасности. Такая постановка задает направление исследования, в котором предметом преобразования становится согласованная работа вычислительных и экспериментальных процессов [1].

Для развития этой концепции необходимо определить, какие именно процессы платформа должна изменять и по каким результатам следует судить об изменениях. Современная литература показывает, насколько существенна связь между характеристиками исследовательской модели и ее пригодностью для последующего решения. Предсказательная валидность определяется отношением результата модели к тому свойству или исходу, ради которого выполняется исследование. Для разработки лекарств особенно сложен перенос от доступного лабораторного показателя к ожидаемой пользе и безопасности у человека [6].

В этой главе фармацевтическая разработка рассматривается как объект управления, включающий научные гипотезы, материальные образцы, данные, модели и действия специалистов. Основное внимание уделяется ранним этапам: от выбора биологического направления до решения о дальнейшем развитии соединения-кандидата. Клинические исследования и последующая регистрация определяют требования к ранним решениям и показывают границы выводов, которые допустимо делать по вычислительным и доклиническим результатам.

Главный тезис главы состоит в следующем: результативность цифровой платформы определяется качеством и воспроизводимостью исследовательских решений при заданных ресурсах. Количество рассчитанных молекул, число подключенных моделей и скорость выполнения отдельного алгоритма характеризуют технические возможности системы. Их влияние на получение экспериментально подтвержденного результата требует самостоятельной проверки. Такой подход согласуется с обсуждаемым в современной литературе переходом от оценки вычислительных моделей к оценке их вклада в принятие решений [3].

1.1 Этапы поиска и разработки лекарств и критерии перехода от активного соединения к лидеру и кандидату

В широком контексте настоящей монографии термин «фармацевтическая разработка» охватывает поиск, проектирование, оптимизацию и экспериментальную оценку новых соединений и многокомпонентных систем, предназначенных для дальнейшего создания лекарственных средств. При таком употреблении необходимо сохранять различия между ранним поиском биологически активных объектов, выбором кандидата, разработкой лекарственной формы и последующими исследованиями. У каждой стадии свой объект работы, собственный состав доказательств и отдельное решение о продолжении проекта.

Отправной точкой служит медицинская потребность. Исследователь определяет заболевание, предполагаемую группу пациентов и ожидаемое действие будущего продукта. Уже на этой стадии возникают ограничения: возможный путь введения, длительность воздействия, приемлемая частота приема, требования к проникновению в ткани. Пока эти параметры предварительны, однако именно они придают смысл молекулярным расчетам. Высокое сродство к белку представляет различную ценность в проектах, где требуются действие в центральной нервной системе, местная активность или ограниченное системное воздействие.

Между медицинской потребностью и химической структурой располагается биологическая гипотеза. При поиске на основе мишени необходимо обосновать, почему воздействие на определенный белок, рецептор или иной биологический объект может изменить течение заболевания. При фенотипическом поиске исходным ориентиром выступает наблюдаемый ответ клеток, ткани или организма. Эти стратегии могут дополнять друг друга. В опубликованной работе по ингибитору TNIK вычислительная идентификация направления исследования была продолжена экспериментальным изучением антифибротического действия и ранними клиническими исследованиями [17].

Слово «мишень» поэтому обозначает биологический объект с определенным обоснованием его роли. Предсказание пространственной структуры помогает сформулировать предположение о возможном взаимодействии с молекулой. Оно оставляет открытым вопрос о том, приведет ли такое взаимодействие к нужному изменению заболевания. Для цифровой платформы эти сведения должны существовать раздельно: гипотеза о роли мишени, структурная модель и экспериментальные данные о воздействии на нее.

Активное соединение, или первичная находка (hit), в предлагаемой терминологии представляет собой объект, для которого в заданной экспериментальной системе обнаружен воспроизводимый сигнал, отвечающий заранее установленным условиям отбора. Вычислительно отобранная структура получает статус виртуальной находки до получения соответствующих измерений. Такое рабочее разграничение позволяет избежать ситуации, когда одно и то же слово обозначает высокий расчетный рейтинг, однократное наблюдение и подтвержденную активность.

Подтверждение первичной находки требует проверки идентичности и качества образца, повторного измерения и анализа зависимости эффекта от концентрации. Важно установить, что наблюдаемый сигнал связан с исследуемым объектом и выбранным биологическим процессом. Например, снижение жизнеспособности опухолевых клеток может отражать разные виды воздействия, а изменение показаний оптического теста иногда связано с физико-химическими свойствами вещества. Поэтому условия повторной и независимой проверки следует определять одновременно с первичным скринингом.

Соединение-лидер (lead) рассматривается как отправная точка для направленной оптимизации. К этому моменту у исследовательской группы должны появиться основания для продолжения химической серии: воспроизводимая активность, начальные сведения о селек-

тивности и свойствах, приемлемая возможность получения вещества. Лидер может уступать будущему кандидату по растворимости, метаболической устойчивости или удобству синтеза. Существенно то, что выявленные ограничения допускают проверяемую программу улучшения.

Переход к соединению-кандидату (drug candidate) означает решение о дальнейшем развитии определенного объекта на основании совокупности данных. Состав этой совокупности зависит от медицинской задачи и принятой программы исследований. Для раннего решения важны согласованность фармакологической активности с экспозицией, сведения о безопасности, аналитическая определенность объекта и возможность воспроизводимого получения. Номинация кандидата открывает следующую программу работ; ее нельзя отождествлять с завершением всех необходимых доклинических исследований.

Фармакокинетические свойства связывают молекулярную активность с условиями ее реализации в организме. Вещество должно достичь места действия в достаточной концентрации и сохраняться там необходимое время. Под сокращением ADMET понимаются всасывание, распределение, метаболизм, выведение и токсичность. Современные вычислительные системы позволяют предварительно оценивать ряд этих характеристик, однако каждая оценка относится к конкретной измеряемой величине и набору исходных данных [9].

Для проектов с предполагаемым нейропротекторным действием эта логика требует особенно точного описания доставки и экспозиции в целевой ткани. Для противоопухолевых исследований необходимо сопоставлять действие на выбранные модели с ответом соответствующих контрольных систем. Это примеры того, как медицинская задача меняет критерии раннего отбора. Установленный эффект в одной клеточной линии или одной модели поражения следует сохранять вместе с условиями его получения, чтобы последующее обобщение оставалось проверяемым.

Если объект представляет собой многокомпонентную систему, маршрут разработки дополняется самостоятельными вопросами о составе и свойствах материала. Смесь компонентов, физически охарактеризованный супрамолекулярный комплекс и лекарственная форма требуют разных процедур описания. Предположение о полезном сочетании нуждается в сравнении с отдельными компонентами и соответствующими контролями. В рамках монографии идентичность объекта и взаимодействие компонентов рассматриваются как отдельные исследовательские задачи, предшествующие сильным выводам о преимуществе комплекса.

Связь этапов имеет итеративный характер. Результаты проверки могут потребовать изменения структуры, пересмотра гипотезы или прекращения направления. Цикл «дизайн — синтез — тестирование — анализ» (design—make—test—analyse, DMTA) описывает эту повторяемость. Его результатом становится обновленное состояние знания о соединении и обоснованное следующее действие. Принципиальная схема такого цикла представлена на рисунке 1.1.



Рисунок 1.1. Исследовательский цикл ранней разработки соединений и основания перехода между его стадиями.

Цифровое представление этапов должно отражать изменение доказательного статуса объекта. Полученная структура, расчет свойств, синтезированный образец и результат испытания имеют собственные идентификаторы и версии. Изменение статуса допускается после проверки соответствующих оснований. Благодаря этому исследователь может восстановить, почему конкретное соединение было выбрано, какое измерение изменило его оценку и кто утвердил дальнейший план.

1.2 Фрагментация данных и решений в исследовательских организациях

Фармацевтический проект объединяет результаты, возникающие в разных исследовательских контекстах. Химическая структура описывает один аспект объекта, профиль экспрессии генов — другой, изображение клеточной культуры — третий. Значение каждого наблюдения зависит от протокола, состояния модели, времени измерения и обработки сигнала. При интеграции приходится сохранять эти различия и одновременно устанавливать связи, необходимые для совместного анализа.

Открытые научные ресурсы хорошо показывают масштаб задачи. В публикации об обновлении PubChem за 2025 год описана база, объединяющая сведения более чем из тысячи источников, включая 119 млн соединений. Эти числа относятся к состоянию, представленному в указанной публикации. Они характеризуют размер информационного ресурса, но сами по себе не определяют объем сопоставимых данных для конкретной мишени, экспериментальной системы или проектной задачи [18].

ChEMBL предоставляет структурированные сведения о биологической активности и связывает их с химическими объектами, мишенями и испытаниями. В описании базы подчеркиваются ручная курация и разнообразие типов данных. Для платформенной разработки это важный ориентир: значение активности следует интерпретировать в связи с происхождением записи и характеристиками теста. Простое объединение таблиц по названию соединения не обеспечивает научной сопоставимости результатов [19].

Первый вид фрагментации связан с идентичностью объектов. Одна структура может фигурировать под внутренним кодом, систематическим названием, регистрационным номером и строковым представлением. Дополнительные различия возникают между исходным веществом и его солью, стереоизомерами, таутомерными формами, образцами разных серий. Если при объединении записей эти различия исчезают, система способна приписать одному объекту свойства нескольких. Если связи между эквивалентными записями не установлены, часть полезных сведений останется недоступной для анализа.

Второй вид относится к значениям и условиям измерений. Концентрация, единицы, время экспозиции и тип наблюдаемого эффекта определяют содержание записи. Значения IC50, EC50 и константы связывания отвечают на разные вопросы. Для их совместного использования требуется явное описание отношения между показателями. Даже одинаковое обозначение конечной точки не исключает различий между клеточными системами, составом среды или способом регистрации ответа.

Третий вид фрагментации возникает на границе расчетов и материальных образцов. Модель получает молекулярное представление и возвращает прогноз. Лаборатория работает с определенной партией вещества, имеющей установленную чистоту, историю хранения и аналитические характеристики. Если связать эти уровни только названием файла, расхождение расчетного и экспериментального результата будет трудно объяснить. Поэтому цифровая прослеживаемость должна соединять химический объект, физический образец, эксперимент и обработанный результат.

Четвертый вид связан с решениями. В итоговом отчете обычно виден выбранный вариант, тогда как рассмотренные альтернативы и причины отказа могут оставаться в переписке, индивидуальных записях или памяти участников. Для последующего обучения такой массив оказывается неполным. Невозможно достоверно установить, какие данные были доступны в момент выбора и какие ограничения повлияли на него. В предлагаемой методологии решение становится самостоятельной записью с датой, основанием, ответственным специалистом и перечнем допустимых следующих действий.

Характерный пример возникает, когда биологический тест и расчетная модель дают разные оценки одного соединения. Без общей истории проекта группа может повторно запустить прежний расчет, сопоставить результат с неподходящей серией опытов и потратить время на обсуждение несопоставимых значений. При наличии связанной записи становится доступна проверка конкретных причин: версии структуры, состава образца, параметров модели, условий опыта. Такой порядок помогает превратить расхождение в исследовательский вопрос.

Отрицательные результаты нуждаются в столь же точном описании. Отсутствие эффекта при ограниченной растворимости, значение ниже предела обнаружения и воспроизводимое отсутствие активности в подходящих условиях имеют разный смысл. При обучении модели эти ситуации нельзя автоматически кодировать одинаковой меткой. В монографии предлагается сохранять причину присвоения отрицательного статуса и границы наблюдения. Тогда система сможет учитывать, какое предположение действительно проверено, а какое осталось неопределенным.

Особое значение имеет разделение проектной истории по времени. Прогноз, сделанный после ознакомления с результатом опыта, не воспроизводит исходную ситуацию выбора. Для проверки платформы необходимо фиксировать состав данных и версию модели до проведения оцениваемого эксперимента. Работа по устранению смещений в прогнозировании аффинности показывает, насколько сходство обучающих и проверочных структур способно влиять на оценку обобщения. Авторы предложили очистку разделения данных на основе структурных критериев [13].

В организационном отношении фрагментация проявляется в несовпадении локальных показателей работы. Вычислительная группа может стремиться увеличить число оцененных

структур, химическая лаборатория — выпуск синтезированных веществ, биологи — число завершенных тестов. Для проекта требуется согласовать эти действия с вопросом о пригодности кандидата. Единицей совместного планирования целесообразно сделать исследовательский пакет: определенный объект, проверяемое предположение, необходимые данные, запланированный опыт и решение, которое его результат позволит принять.

Общее пространство данных при этом не требует физического переноса всех сведений в одно хранилище. Часть материалов может оставаться у лабораторного партнера или в защищенном контуре организации. На уровне методологии существенно наличие согласованных идентификаторов, правил доступа, описания происхождения и воспроизводимого способа получить разрешенные данные. Архитектурные варианты реализации такого пространства рассматриваются в последующих главах.

Следовательно, цифровая трансформация предполагает работу одновременно с содержанием данных и с порядком взаимодействия специалистов. Исправление форматов делает записи технически совместимыми. Согласование определений, протоколов и оснований решений делает их пригодными для исследования. Именно эту задачу должны поддерживать интерфейсы между подсистемами Atlansys EUS MedPharma.

1.3 Возможности и ограничения искусственного интеллекта в фармацевтических исследованиях

Разные методы искусственного интеллекта изменяют разные участки исследовательского процесса. Одни строят представления молекул, другие прогнозируют измеряемые свойства, третьи предлагают структуры или помогают работать с научными текстами. Оценка пользы начинается с определения решения, в котором результат модели будет использован. Для одного проекта достаточно ранжирования доступных соединений, для другого требуется выбор нового эксперимента, а для третьего — обоснование переноса на ранее не исследованную химическую серию.

Одно из наиболее заметных направлений — моделирование биомолекулярных структур. В работе по AlphaFold 3 представлена диффузионная архитектура для совместного предсказания комплексов, включающих белки, нуклеиновые кислоты, малые молекулы, ионы и модифицированные остатки. Авторы показали улучшение ряда структурных прогнозов на использованных наборах проверки. Такие возможности расширяют доступность структурных гипотез при проектировании соединений [7].

Структурная гипотеза приобретает прикладную ценность при проверке пригодности для конкретной задачи. В проспективной работе Lyu и соавторов виртуальный скрининг по моделям AlphaFold2 был продолжен экспериментальным тестированием новых молекул для двух рецепторов. Результаты показали, что предсказанные структуры могут служить продуктивной основой для поиска лигандов в исследованных условиях. Вывод относится к проверенным мишеням и процедуре отбора; его перенос требует самостоятельного обоснования [20].

Для проектирования платформы это означает необходимость хранить происхождение структурной модели и критерии ее допуска в расчет. Уверенность в геометрии отдельных участков, возможные конформационные изменения и состояние места связывания могут влиять на дальнейшие выводы. Показатель структурной уверенности не следует автоматически преобразовывать в вероятность фармакологического эффекта. Между геометрией комплекса и действием вещества располагаются дополнительные физические и биологические процессы.

Графовые нейронные сети позволяют учитывать отношения между атомами и связями, а также объединять молекулярную информацию с признаками биологических систем. В XGDP молекулярные графы использованы совместно с данными экспрессии генов опухолевых клеточных линий для прогнозирования лекарственного ответа. Авторы также применили методы

атрибуции признаков и представили области молекул и гены, существенные для прогнозов модели [21].

Такая интерпретация помогает сформировать проверяемые предположения о причинах различий между соединениями или клеточными системами. Однако вклад признака в расчет и причинная роль биологического механизма требуют разных доказательств. В монографии карта значимости рассматривается как основание для последующего эксперимента. Утверждение о механизме действия должно опираться на соответствующие биохимические, генетические или иные исследования, выбранные для конкретной гипотезы.

Практическую возможность поиска новых химических направлений демонстрирует работа Wong и соавторов. Объяснимое глубокое обучение было использовано для отбора структурных фрагментов и соединений с антибактериальными свойствами, после чего авторы провели экспериментальную оценку активности, токсичности и действия в моделях инфекции. Значимость примера состоит в прослеживаемом переходе от вычислительного отбора к проверке выбранного химического класса [22].

Генеративный дизайн расширяет набор структур, доступных для рассмотрения. Исследователь получает возможность искать варианты, удовлетворяющие нескольким требованиям, и исследовать области химического пространства, слабо представленные в исходной библиотеке. Вместе с этим возникает необходимость проверять химическую определенность, доступность исходных веществ и пригодность маршрута синтеза. Предложенная моделью структура должна пройти тот же путь аналитического и биологического подтверждения, что и соединение, выбранное другим способом.

Оптимизация генератора зависит от качества его оценочных моделей. Если расчетная функция систематически завышает преимущество определенной группы структур, генератор способен усиливать это смещение. Поэтому в предлагаемом подходе вычислительный отбор дополняется независимыми ограничениями и проверкой разнообразия выбранных объектов. Часть экспериментального бюджета целесообразно направлять на изучение областей, где прогноз особенно неопределенен, если полученные данные способны изменить решение по проекту.

Виртуальный докинг и оценка взаимодействий также требуют нескольких критериев качества. Исследование PoseBusters показало, что близость рассчитанной позы к известной структуре по геометрической метрике может сочетаться с физически некорректными характеристиками. В работе проверялись геометрия, стерические столкновения и другие ограничения, а результаты зависели от выбранного набора тестирования. Это обосновывает отдельную проверку физической правдоподобности получаемых поз [23].

Даже корректная поза оставляет открытыми вопросы о силе взаимодействия, скорости связывания, селективности и действии в клетке. В российской работе о глубоком обучении в моделировании белок-лигандных взаимодействий показаны возможности разных архитектур и зависимость результатов от свойств исследуемых объектов. Для платформы важно сохранять назначение каждого вычислительного вывода и условия, при которых его можно использовать [24].

Отдельная группа задач связана с прогнозированием ADMET. Такие модели помогают заранее выявлять потенциальные ограничения и распределять объекты по последующим проверкам. Однако модель проницаемости, классификатор риска определенного токсического эффекта и прогноз клиренса имеют разные выходы, единицы и области применимости. Их объединение в одну итоговую оценку требует объяснения того, как ошибки каждого прогноза повлияют на выбор соединения. Методика такого объединения должна оставаться доступной исследовательской группе.

Развитие альтернативных экспериментальных методов расширяет возможности проверки вычислительных предположений. В работе В. Н. Перфиловой рассмотрены клеточные

системы, тканевые чипы, органоиды и вычислительные подходы в контексте программы Tox21. Их использование открывает направления повышения информативности оценки безопасности и сокращения применения животных. При этом автор подчеркивает необходимость дальнейшей валидации и внедрения новых методов [25].

Большие языковые модели полезны для поиска и организации сведений, подготовки программного кода и взаимодействия с вычислительными инструментами. Работа Coscientist продемонстрировала систему, объединяющую языковую модель с поиском, выполнением кода и автоматизацией химических экспериментов. В исследовании были выполнены задачи планирования и экспериментальной оптимизации химических реакций. Эти результаты относятся к представленным химическим задачам и составу подключенных инструментов [10].

При работе с научными текстами необходимо учитывать разницу между связным ответом и проверенным знанием. Для предлагаемой платформенной методологии значимы возможность открыть исходный документ, проверить контекст утверждения и установить, каким способом оно вошло в исследовательскую гипотезу. Сведения, полученные языковой моделью, целесообразно связывать с источниками и сохранять отдельно от экспериментальных результатов. Ответственность за утверждение программы исследования остается у уполномоченного специалиста.

Качество моделей следует оценивать с учетом структуры задачи. Общая доля правильных ответов может скрывать неудовлетворительную работу с редким, но значимым классом. В отечественном обзоре методов оценки классификации систематизированы показатели, основанные на матрице ошибок, и показана необходимость выбирать критерии в соответствии с содержанием задачи. Для ранней разработки особенно существенна цена пропуска перспективного объекта и цена ошибочного направления вещества в дорогостоящую проверку [26].

Наряду с этими ограничениями существуют примеры перехода ИИ-поддержанной разработки к клинической проверке. После публикации доклинических и ранних клинических результатов ингибитора TNIK в Nature Medicine были представлены данные многоцентрового рандомизированного двойного слепого плацебо-контролируемого исследования рентосертиба фазы 2a при идиопатическом легочном фиброзе. Это отдельный клинический результат определенного проекта; он не позволяет установить универсальное сокращение сроков разработки или причинный вклад каждого примененного инструмента [12].

Для Atlansys EUS MedPharma из рассмотренных исследований следует принцип распределения задач по уровню подтверждения. ИИ может предлагать, рассчитывать, ранжировать и помогать планировать. Материальный образец, измерение и решение специалиста создают следующие звенья исследования. Функция платформы состоит в том, чтобы сохранить связи между этими звеньями и сделать переходы между ними управляемыми.

1.4 Разрыв между скоростью вычислительного поиска и пропускной способностью экспериментальной проверки

Вычислительный поиск способен сформировать больше предложений, чем лаборатория может проверить в ближайший период. Возникает задача управления потоком: какие структуры передать химикам, какие образцы испытать первыми и какая проверка даст наиболее полезную информацию. При неограниченном увеличении числа предложений очередь может расти, тогда как число завершенных исследований будет определяться доступными экспериментальными ресурсами.

Ограничение имеет несколько возможных источников. Для одной химической серии наиболее продолжителен синтез и очистка, для другой — подготовка биологической системы или получение данных о распределении вещества. Дополнительное время уходит на закупки, передачу образцов, согласование протоколов и экспертное рассмотрение. Поэтому анализ произ-

водительности начинается с восстановления фактического маршрута исследования, включая периоды ожидания и повторные операции.

Введем простое представление длительности одного завершенного исследовательского цикла. Обозначим время вычислительных операций через Tc, материально-экспериментальных операций через Te, ожидания через Tq, анализа и согласования решения через Td. Если на рассматриваемом маршруте эти интервалы не перекрываются, общая длительность равна их сумме, представленной в выражении (1.1).

$$T_{цикл} = T_c + T_e + T_q + T_d \quad (1.1)$$

При параллельной работе необходимо анализировать критический путь, поскольку механическое сложение длительности всех задач завысит срок проекта. Тем не менее предложенное разложение полезно для первичной диагностики. Оно показывает, какие интервалы действительно изменились после внедрения вычислительного метода, и предотвращает перенос ускорения одной операции на весь процесс.

Рассмотрим условный расчет, который поясняет этот эффект и не описывает результаты Atlansys EUS MedPharma. Пусть последовательный цикл длится 20 рабочих дней: 2 дня занимают расчеты, 10 дней — получение и экспериментальная проверка материала, 6 дней — ожидание, 2 дня — анализ и решение. Ускорение расчетов в десять раз уменьшит их длительность до 0,2 дня. Общий срок составит 18,2 дня, то есть сократится на 9%. Дополнительное уменьшение ожидания с 6 до 3 дней сократит цикл до 15,2 дня, или на 24% относительно исходного срока. Эти значения следуют только из принятых условий примера.

Таким образом, потенциальный эффект интеграции возникает на границах работ. Передача корректного задания в лабораторию, доступность исходных аналитических данных и своевременное рассмотрение результата могут влиять на срок исследования наряду с вычислительной скоростью. Для установления такого эффекта требуется измерять соответствующие интервалы до и после изменения процесса. Наблюдение за средним временем выполнения модели для этого недостаточно.

Связь числа проверенных объектов с выходом подтвержденных результатов также можно описать простой учетной моделью. Для фиксированной когорты испытанных соединений обозначим число завершенных проверок через Ntest, а долю объектов, прошедших заранее заданные критерии подтверждения, через q. Тогда число подтвержденных объектов определяется выражением (1.2). При использовании выражения как прогноза q становится оценкой с соответствующей неопределенностью.

$$N_{подтв} = N_{test} \times q \quad (1.2)$$

Этот показатель требует точного знаменателя. В него следует включать объекты с завершенной процедурой проверки, отдельно фиксируя не полученные вещества, сорванные опыты и незавершенные исследования. Исключение неудачных случаев способно создать видимость более высокой результативности. Если ресурсы расходуются на несколько типов тестов, в дальнейшем понадобится учитывать их разную длительность, стоимость и информативность.

На уровне всей системы производительность ограничивается доступными мощностями синтеза, аналитики и биологических испытаний с учетом выбранного маршрута. Снижение ложноположительных направлений может освободить часть этих мощностей. Однако слишком жесткий отбор способен исключить перспективные химические направления. Поэтому решение о составе экспериментальной очереди должно учитывать одновременно прогнозируемую полезность, неопределенность и возможность получить новые знания.

В предлагаемой методологии выделяются два основания для назначения следующего опыта. Первое связано с подтверждением наиболее перспективного объекта. Второе — с уменьшением неопределенности, которая препятствует выбору между альтернативами. Иногда проверка соединения со средним расчетным рейтингом оказывается полезнее повторного изучения лидера, если она позволяет отвергнуть неверную гипотезу о целой химической серии. Конкретные правила такого выбора будут раскрыты в главе о взаимодействии моделей и лаборатории.

Не всякое сокращение числа экспериментов означает улучшение разработки. Исследование может завершиться быстрее из-за уменьшения полноты проверки или изменения критериев успеха. Поэтому показатель времени должен сопровождаться описанием достигнутого результата. Аналогично низкая стоимость отдельного расчета мало говорит о затратах проекта, если предложенные структуры трудно получить или их результаты плохо воспроизводятся.

Для оценки платформы в монографии предлагается совокупность взаимосвязанных показателей: календарное время до определенного подтвержденного результата, полные затраты в согласованных границах, выход пригодных объектов и воспроизводимость ключевых наблюдений. Для каждого показателя заранее задаются начало отсчета, окончание, единица анализа и правила учета незавершенных работ. Состав такой оценки представлен на рисунке 1.2.



Рисунок 1.2. Связь качества вычислительной модели с результативностью исследовательского процесса.

Качество выбора следующего эксперимента является связующим звеном между вычислительной точностью и производительностью проекта. Если прогноз не меняет ни одного решения либо его изменения не улучшают результат проверок, высокий показатель на тестовом наборе остается локальным достижением. Напротив, умеренное улучшение модели может быть значимым, когда оно надежно поддерживает решение на наиболее затратном участке.

Наконец, завершение исследования с обоснованным отрицательным выводом также представляет ценность. Для этого должны быть определены проверенная гипотеза, достаточность выполненных опытов и причина прекращения. Смешивать такие решения с получением нового активного соединения в одном числовом показателе нельзя. Их организационный

эффект следует оценивать отдельно, например через предотвращенные повторные работы и высвобожденные ресурсы.

1.5 Сравнение классов цифровых платформ по функциям и уровню подтверждения результатов

Для выбора платформенного подхода полезно различать классы цифровых решений по их основной функции. Такое сравнение позволяет определить, какое звено исследовательского процесса поддерживается и какие вопросы остаются за его пределами. Перечень классов ниже представляет аналитическую классификацию настоящей монографии. Приведенные примеры относятся к опубликованным работам и не являются рейтингом актуальных коммерческих продуктов.

Первый класс образуют информационные ресурсы и средства управления исследовательскими данными. PubChem и ChEMBL обеспечивают доступ к химической и биологической информации, а лабораторные информационные системы и электронные журналы помогают фиксировать локальные наблюдения. Для этого класса существенны полнота описания, качество связей и возможность повторно использовать сведения. Сам факт накопления большого массива данных оставляет открытым вопрос о пользе конкретного решения по соединению.

Второй класс включает специализированные вычислительные системы: моделирование структур, молекулярное ранжирование, прогноз отдельных свойств. Их можно сопоставлять на наборах данных с определенными условиями и метриками. Однако результат такого сопоставления относится к вычислительной задаче. Для оценки влияния на проект необходимо установить, какие соединения были бы выбраны с использованием модели и без нее, а затем проверить эти объекты по сопоставимой процедуре.

Третий класс связан с генеративным проектированием и многокритериальной оптимизацией. Его выходом являются предлагаемые структуры и план их дальнейшего уточнения. Для оценки важны химическое разнообразие, соответствие требованиям, возможность получения и результаты последующей проверки. Только часть этих характеристик допускает надежную оценку до синтеза. Поэтому число созданных структур характеризует объем предложений, а экспериментальный выход — продуктивность их отбора.

Четвертый класс объединяет средства автоматизированного экспериментирования и управления лабораторным циклом. Они связывают планирование с выполнением операций и обработкой результатов. Публикация Coscientist показывает один из вариантов такого взаимодействия для химических задач. Для фармацевтического проекта дополнительного подтверждения потребует пригодность всей процедуры к выбранным биологическим системам, условиям измерения и требованиям безопасности.

Пятый класс составляют интегрированные платформы, связывающие несколько исследовательских функций с общим управлением данными и решениями. Их предполагаемое преимущество возникает благодаря согласованию этапов. Проверять это преимущество следует на уровне целого рабочего процесса. Высокое качество отдельных компонентов является необходимым основанием для их применения, но общая результативность зависит и от интерфейсов, очередей, полноты обратной связи и участия специалистов.

Таблица 1.1. Классы цифровых решений и предмет проверки их результативности.

Класс и пример	Непосредственный результат	Что требует дополнительной проверки
Информационные ресурсы PubChem и ChEMBL	Структурированные сведения об объектах и испытаниях [18, 19].	Пригодность выбранных записей для решения по проекту
Специализированные модели AlphaFold 3 и ADMET-AI	Структурный прогноз или оценка свойства [7, 9].	Качество выбора соединений при независимой проверке
Генеративное проектирование	Предложенные структуры и порядок оптимизации	Получаемость и экспериментальный выход выбранной серии
Автоматизированное экспериментирование Coscientist	Выполненные химические операции [10].	Перенос процедуры на заданный фармацевтический процесс
Интегрированная платформа Atlansys EUS MedPharma	Связанный маршрут данных, расчетов и экспериментов	Добавочный эффект интеграции при сопоставимых ресурсах

В исходных материалах Atlansys EUS MedPharma исследовательская система включает семь подсистем: верификации и анализа данных, предсказания структуры белков, дизайна и оптимизации молекул, виртуального анализа и тестирования, анализа закономерностей, генерации идей, ускорения и оптимизации разработки. В приложении они раскрыты через функциональные модули. В монографии эта структура рассматривается как проектная основа интеграции исследовательского цикла [27].

Для первой главы существенно функциональное назначение этих компонентов. Проверка архитектуры должна показать, сохраняется ли контекст при переходе между ними и возвращаются ли новые измерения в последующий анализ. Подробный состав модулей и согласование конфигураций относятся к архитектурной главе. Количество компонентов само по себе не устанавливает зрелость платформы или степень завершенности ее экспериментальной проверки.

Наряду с классами решений необходимо различать уровни подтверждения результата. Ретроспективная оценка показывает работу на ранее собранных данных. Независимая проверка выявляет устойчивость при изменении источника или условий. Проспективное исследование оценивает предсказания, зафиксированные до получения новых измерений. Сравнение рабочих процессов отвечает на вопрос о добавочном эффекте технологии при сопоставимом задании. Клиническое исследование дает сведения о конкретном вмешательстве у человека. Эти процедуры образуют разные основания для выводов.

Результаты могут сочетаться: платформа способна иметь убедительно проверенную модель одного свойства и одновременно находиться на ранней стадии интеграции с лабораторией. Поэтому для каждого утверждения следует указывать объект подтверждения. Формулировка «прошла валидацию» требует пояснения: валидированы данные, модель, программная процедура, определенный рабочий процесс или фармакологическое действие вещества. Такая определенность позволяет сопоставлять решения, находящиеся на разных стадиях развития.

Этот подход соответствует направлению международных методических документов. В принятых FDA и EMA в январе 2026 года десяти принципах надлежащей практики применения ИИ выделены контекст использования, соразмерная риску оценка, управление данными, показатели работы модели, управление жизненным циклом и понятная информация для пользователей. Принципы задают общие ориентиры применения ИИ на протяжении жизненного цикла лекарственного продукта [5].

Документ EMA 2024 года также рассматривает применение ИИ на разных стадиях жизненного цикла лекарств. Для проектирования цифровой среды это означает необходимость связывать требования к модели с той ролью, которую она выполняет. Автоматизированный поиск литературы, приоритизация лабораторного опыта и формирование данных для значимого решения требуют разных объемов подтверждения и контроля [16].

В проекте руководства FDA, опубликованном в январе 2025 года, предложен подход к оценке достоверности модели с учетом вопроса исследования, контекста использования и риска. На странице FDA документ обозначен как проект, содержащий необязательные рекомендации. Его статус следует учитывать при использовании методических положений. В настоящей монографии эти положения служат ориентиром для построения проверяемой аргументации [28].

Сопоставление классов платформ, таким образом, приводит к двум практическим требованиям. Во-первых, описание функции должно сопровождаться указанием проверяемого результата. Во-вторых, сведения о качестве компонента необходимо связывать с условиями его применения. Эти требования позволяют перейти от перечня возможностей Atlansys EUS MedPharma к постановке исследовательской программы, раскрываемой далее.

1.6 Постановка научной проблемы и исследовательских вопросов монографии

Рассмотренные процессы определяют научную проблему настоящей монографии: каким образом согласовать неоднородные данные, вычислительные модели, экспериментальные ресурсы и решения специалистов так, чтобы интегрированная платформа давала измеримый и воспроизводимый результат ранней разработки. Проблема включает технические интерфейсы и научное содержание переходов между этапами. Особенно существенны сохранение контекста наблюдений, учет неопределенности и проверка вклада интеграции.

Объектом исследования являются процессы поиска, проектирования, оптимизации и экспериментального подтверждения новых соединений и многокомпонентных систем на ранних этапах создания лекарственных средств. Предмет исследования составляют методы и архитектурные решения интеграции данных, моделей искусственного интеллекта и экспериментальных процедур в платформе Atlansys EUS MedPharma.

Цель исследования состоит в разработке и обосновании методологии построения и применения платформы, обеспечивающей прослеживаемость решений и измеримую результативность исследовательского цикла. В качестве проверяемого результата рассматривается получение объекта, удовлетворяющего заранее установленным критериям соответствующей стадии, либо обоснованное решение о прекращении или изменении направления. Эти виды результата учитываются раздельно.

Рабочая гипотеза заключается в том, что согласование качества данных, неопределенности вычислительных прогнозов и доступных экспериментальных ресурсов улучшает результативность ранней разработки по сравнению с раздельным применением тех же инструментов. Предполагаемый механизм такого улучшения состоит в уменьшении потерь при передаче заданий, снижении числа непригодных направлений в экспериментальную очередь и более своевременном использовании новых измерений. Наличие и величина эффекта должны быть установлены исследованием.

Первый исследовательский вопрос связан с качеством исходных оснований. Какие сведения о химическом объекте, образце и протоколе необходимы, чтобы результаты разных источников можно было совместно использовать для определенного решения? Ответ должен включать правила идентификации, требования к метаданным, порядок работы с противоречиями и условия допуска записей в обучение и проверку моделей.

Второй вопрос относится к управлению вычислительными прогнозами. Как оценивать пригодность модели за пределами данных, на которых она создавалась, и передавать неопределенность между этапами анализа? Здесь требуется связать качество прогнозирования с условиями последующего применения. Вопрос о том, какая модель лучше на определенном наборе, дополняется вопросом о том, в каких ситуациях исследователь вправе использовать ее результат.

Третий вопрос касается экспериментальной очереди. Как выбирать следующую проверку при ограниченных мощностях синтеза, аналитики и биологических испытаний? Ответ должен учитывать ценность ожидаемого результата, стоимость и длительность опыта, а также возможность уменьшить неопределенность. Такое планирование должно сохранять основания выбора и допускать экспертный пересмотр.

Четвертый вопрос связан с организацией обратной связи. Какие изменения данных, модели или исследовательской гипотезы должны следовать из положительного, отрицательного либо неопределенного результата? Для ответа необходимо определить, когда новый результат действительно опровергает прежнее предположение, а когда требует дополнительного контроля качества измерения или образца. Автоматизация должна поддерживать различение этих ситуаций.

Пятый вопрос относится к доказательству добавочного эффекта архитектуры. Насколько интегрированная работа компонентов меняет время, полные затраты и выход подтвержденных объектов при сопоставимых условиях? Для его изучения предлагается сравнивать три варианта: действующий исследовательский процесс, использование выбранных ИИ-инструментов раздельно и их применение в составе интегрированной платформы. Такое сравнение позволяет раздельно оценить эффект вычислительных методов и эффект организации их взаимодействия.

Сравнение должно иметь общую медицинскую задачу, согласованные требования к соединениям и сопоставимый ресурсный бюджет. До начала оценки фиксируются исходные данные, версии инструментов, правила выбора опытов и критерии подтверждения. При невозможности прямого распределения задач между вариантами потребуется обосновать другой дизайн и описать оставшиеся различия. Проспективная проверка особенно ценна тем, что сохраняет реальную ситуацию выбора до получения экспериментального ответа.

Шестой вопрос связан с переносимостью результатов. Какие части методологии сохраняют пользу при изменении химического класса, мишени, типа биологической модели или организации? Для этого требуется различать общие механизмы интеграции и ограничения конкретного кейса. Результат нейропротекторного проекта не следует автоматически переносить на противоопухолевую задачу, а успешное применение одной модели — на все компоненты платформы.

Граница исследования проходит по решению о дальнейшем развитии кандидата и передаче его в последующие программы. Отдельные материалы могут включать данные более поздних стадий, однако их использование будет соотноситься с фактическим объемом исследования. Доклиническая программа занимает самостоятельное место в цикле разработки лекарственных средств, что подчеркивается в профильной российской публикации М. Н. Макаровой [29].

Научный результат монографии должен включать методологию связывания данных и решений, архитектурное описание ее реализации и воспроизводимый способ проверки. При этом статус программного компонента, научная обоснованность модели и фармакологические свойства вещества будут описываться отдельно. Документы об интеллектуальной собственности характеризуют соответствующие результаты и права; сведения о действии исследуемого объекта раскрываются через методы, наблюдения и границы переноса выводов.

Структура дальнейшего изложения следует из сформулированных вопросов. Во второй главе определяется методология сквозного управления, после чего раскрываются архитектура и пространство данных и знаний. Затем рассматриваются методы проектирования и оценки соединений. Следующий блок связывает модели с экспериментом, валидацией, эксплуатацией и защитой информации. Авторские исследования позволяют проверить применимость подхода на конкретных задачах, а экономическая и организационная оценка завершает анализ условий внедрения.

Такое построение сохраняет общее основание для всех глав: исследовательская гипотеза должна быть связана с доступными данными, вычислительным выводом, процедурой проверки и решением. При изменении любого звена необходимо понимать последствия для остальных. Atlansys EUS MedPharma рассматривается как среда, в которой эти связи могут быть представлены явно и проверены на практике.

Выводы

Фармацевтическая разработка представляет собой последовательность решений с изменяющимся составом доказательств. На ранних этапах исследователь переходит от биомедицинской гипотезы к активному соединению, лидеру и кандидату. Эти статусы требуют собственных критериев и сохраняют связь с конкретными объектами, образцами и условиями наблюдения. Итеративность процесса делает необходимым возврат новых результатов к исходным предположениям.

Главное препятствие интеграции связано с потерей контекста при передаче данных и заданий. Установление идентичности объектов, сохранение протоколов, регистрация отрицательных результатов и фиксация оснований выбора создают условия для воспроизводимого анализа. Общая цифровая среда должна поддерживать эти действия независимо от того, в каких системах физически размещены исходные материалы.

Исследования последних лет демонстрируют полезность ИИ для структурного моделирования, отбора соединений и организации отдельных экспериментальных задач. Допустимый вывод зависит от того, что именно проверено: расчетный показатель, новое лабораторное наблюдение, рабочий процесс или клиническое вмешательство. Для платформы принципиально сохранять эти различия и связывать каждый вывод с подтверждающими материалами.

Скорость вычислений влияет на общую длительность разработки в соответствии с местом расчетов в исследовательском маршруте. При ограниченной лабораторной мощности существенны качество отбора, ожидание и полнота обратной связи. Поэтому оценка цифровой трансформации должна включать время и полные затраты на определенный подтвержденный результат, выход пригодных объектов и воспроизводимость. Критерии результата устанавливаются до сравнения процессов.

Для Atlansys EUS MedPharma сформулирована рабочая гипотеза о пользе согласованной интеграции данных, моделей и лабораторных ресурсов. Ее проверка предполагает сопоставление действующего процесса, раздельного применения тех же ИИ-инструментов и интегрированной платформы. Такая постановка позволяет установить добавочный эффект архитектуры и определить условия его переноса.

Полученная карта процессов и ограничений задает требования к методологии следующей главы. В центре дальнейшего рассмотрения находятся целевой профиль разрабатываемого объекта, правила работы с неопределенностью и порядок выбора очередного исследования. Их согласование превращает функциональную архитектуру платформы в проверяемый способ организации фармацевтической разработки.

Глава 2. Методология сквозного управления разработкой новых соединений

Соединения могут занимать соседние строки в таблице расчетов и требовать совершенно разных исследовательских решений. Для одного остается проверить воспроизводимость активности, для другого — выяснить, удастся ли получить вещество нужной чистоты. В третьем случае главным становится вопрос о достижении необходимой концентрации в целевой ткани. Общий рейтинг полезен лишь тогда, когда исследователь понимает, какие основания стоят за каждой оценкой и какое действие должно последовать.

В первой главе определены процессы и ограничения ранней фармацевтической разработки. Теперь необходимо связать их в методологию, которая устанавливает порядок перехода от медицинской потребности к исследовательскому заданию, от задания — к эксперименту, от результата — к следующему решению. В исходной концепции Atlansys EUS MedPharma эта задача поставлена через сквозную интеграцию данных, специализированных методов ИИ и функций управления исследованиями. Настоящая глава развивает ее на уровне правил организации исследовательского процесса [1].

Под сквозным управлением будем понимать согласованное изменение требований, исследовательских гипотез, вычислительных моделей и экспериментальных действий на основании прослеживаемых данных. Единицей управления выступает проверяемая гипотеза об определенном объекте разработки. Для нее задаются условия подтверждения, возможные альтернативы и решение, которое можно принять после получения результата. Благодаря этому план исследования сохраняет связь с конечной целью даже при изменении химической серии или состава применяемых алгоритмов.

Современные обзоры рассматривают совместную работу вычислительной и экспериментальной лабораторий как необходимую составляющую развития ИИ в создании лекарств. В работе Fu и Chen интеграция биологических знаний и алгоритмических методов связана с организацией обмена данными между этими средами. Для методологии платформы это означает необходимость определить содержание такого обмена и ответственность за использование его результатов [30].

Предлагаемый подход включает шесть взаимосвязанных элементов: целевые требования, паспорт исследовательской гипотезы, согласование источников знания, выбор эксперимента при ограниченных ресурсах, правила пересмотра проекта и программу оценки результативности. Эти элементы образуют общую рамку последующих глав. Их конкретная программная реализация может изменяться, тогда как основания исследовательских решений должны оставаться проверяемыми.

2.1 Целевой профиль будущего лекарственного продукта и требования к соединению-кандидату

Разработка начинается с описания результата, который будет иметь значение для будущего применения. Целевой профиль лекарственного продукта (target product profile, TPP) связывает предполагаемое назначение, группу пациентов, желаемое действие и ограничения использования. В образовательном материале NIH SEED выделены минимальные и предпочтительные характеристики продукта, а сам профиль рассматривается как обновляемый инструмент планирования. Этот материал объясняет подход к работе с TPP и не является руководством регулирующего органа [31].

Для раннего поиска профиль продукта необходимо преобразовать в требования к исследуемому объекту. В настоящей методологии целевой профиль соединения-кандидата (target candidate profile, TCP) описывает свойства и подтверждающие данные, которые нужны для решения о его дальнейшем развитии. TPP сохраняет ориентацию на медицинскую задачу, а TCP переводит ее в проверяемые условия выбора. Профиль требований к качеству будущего продукта решает отдельную задачу; его нельзя считать полным заменителем этих двух документов.

Например, предполагаемая необходимость воздействия в центральной нервной системе формирует вопрос о доставке и экспозиции в соответствующей ткани. Требование длительного действия меняет значение метаболической устойчивости и времени удержания эффекта. Предпочтительный путь введения влияет на требования к растворимости и возможности создания лекарственной формы. В обзоре Serrano и соавторов задачи поиска соединений и доставки рассматриваются в общем контексте разработки, что подчеркивает необходимость согласования молекулярных и фармацевтических характеристик [32].

При этом формулировка требования должна включать способ его проверки. Указание «хорошая проницаемость» не определяет измеряемую величину, экспериментальную систему и допустимый результат. Для управления требуется установить, какой показатель будет использован, в каких условиях его получат и какое решение он поддерживает. На раннем этапе часть значений может оставаться неизвестной. Такие пробелы следует обозначать явно и связывать с планом получения данных.

Предлагается различать обязательное условие, желательный уровень и исследовательское предположение. Обязательное условие определяет допустимость дальнейшего выбора в данном контексте. Желательный уровень позволяет сравнить уже допустимые варианты. Исследовательское предположение фиксирует ожидаемую связь, которую еще предстоит проверить. Это разделение препятствует превращению расчетного ожидания в фактически выполненное требование.

Для каждого требования необходимы идентификатор, содержательная формулировка, измеряемый показатель, единица, условия наблюдения и критерий интерпретации. Дополнительно указываются источник обоснования, доступные сведения и ответственный за решение. Если один критерий зависит от другого, эту связь фиксируют в профиле. Вопрос о запасе безопасности, например, связан с предполагаемой эффективной экспозицией; раздельные благоприятные оценки без такого сопоставления оставляют исследовательский вопрос открытым.

Требования полезно разделять и по моменту принятия решения. На раннем отборе допустима предварительная расчетная оценка части свойств. Перед дорогостоящим исследованием требуется более определенное состояние сведений о веществе и его действии. Подход не предполагает одинакового набора испытаний для всех проектов. Состав проверки должен соответствовать биологической гипотезе и последствиям возможной ошибки.

В онкологическом направлении особенно существенен контекст биологического ответа. Обзор Ocana и соавторов связывает интеграцию ИИ с выбором мишеней и ранней разработкой, одновременно отмечая ограничения переноса результатов неподходящих доклинических моделей. Поэтому в профиле кандидата нужно фиксировать, какую сторону заболевания воспроизводит выбранная модель и какие свойства за пределами этой модели пока не установлены [33].

Ниже приведена методологическая матрица перевода требований. Она показывает логику постановки задач и не задает универсальных числовых нормативов. Конкретные пороги устанавливаются для проекта на основании медицинского замысла, характеристик исследуемого объекта и обоснованной программы проверки.

Таблица 2.1. Перевод целевых требований в исследовательские решения.

Требование профиля	Проверяемый вопрос	Основание решения
Медицинская задача	Какое изменение состояния должно иметь значение?	Релевантная модель и определенная конечная точка
Фармакологическое действие	Воспроизводится ли нужный эффект?	Измерение с контролями и установленными критериями
Экспозиция	Достигает ли вещество целевой ткани?	Связь экспозиции с условиями реализации эффекта
Безопасность и селективность	В каких условиях проявляется нежелательное действие?	Сопоставимые системы и план проверки рисков
Идентичность объекта	Что именно получено и испытано?	Состав, чистота, серия и стабильность материала
Получение и применение	Доступен ли воспроизводимый путь получения?	Маршрут, ресурсы и ограничения последующей разработки

У каждого требования должна существовать история изменений. Если после неудачного результата группа меняет критерий, первоначальная оценка сохраняется вместе с причиной пересмотра. Иначе улучшение показателей может оказаться следствием изменения цели. Версия целевого профиля входит в состав любого существенного решения по соединению, серии и проекту.

Пересмотр профиля допустим при появлении новых сведений о заболевании, механизме действия, доступности технологии или условиях будущего применения. Такое изменение требует согласования последствий для текущей программы. Если меняется предполагаемый путь введения, часть прежних требований сохраняет значение, а часть должна быть уточнена. Платформа должна позволять установить, какие данные остались пригодными и какие исследования необходимо выполнить заново.

Таким образом, профиль служит основанием для согласования работы специалистов. Он помогает связать различающиеся по форме задания биоинформатика, химика и фармаколога с общей целью. Дальнейшее управление строится вокруг вопроса о том, какие данные нужны, чтобы обоснованно изменить состояние выполнения этих требований.

2.2 Цикл «дизайн — синтез — тестирование — анализ» и уровни доказательности исследовательских результатов

В первой главе цикл «дизайн — синтез — тестирование — анализ» был рассмотрен как форма повторяемости исследования. В методологии управления каждому его прохождению необходимо придать определенный вопрос и условие завершения. Итерация начинается с зафиксированного состояния знания и завершается решением, основания которого можно восстановить. Между этими точками располагаются конкретные задания, результаты выполнения и проверки качества.

Для этого вводится паспорт исследовательской гипотезы. Он содержит формулировку предполагаемой связи, объект проверки, исходные основания и альтернативные объяснения. Следующая часть паспорта определяет эксперимент, необходимые контроли, измеряемый результат и ожидаемые варианты интерпретации. Завершающая часть задает возможные решения и условия пересмотра. Паспорт представляет собой предлагаемый в монографии способ

организации исследования, который может быть реализован разными программными средствами.

Примером служит предположение о том, что изменение определенной части структуры улучшит устойчивость соединения, сохранив его активность. Для проверки важно заранее установить, с каким объектом сравнивается новый вариант и в каких условиях измеряются оба свойства. Возможны разные исходы: оба требования выполнены, одно улучшилось ценой другого, результаты противоречивы либо образец не отвечает требованиям качества. Эти исходы должны вести к различающимся действиям.

Особое значение имеет предсказательная валидность выбранного измерения. Scannell и соавторы связывают продуктивность разработки с тем, насколько исследовательские инструменты помогают обнаруживать объекты, полезные на последующих стадиях. В предлагаемом паспорте поэтому фиксируется отношение измеряемого показателя к целевому требованию. Если опыт характеризует промежуточный признак, границы его интерпретации сохраняются вместе с результатом [6].

Для вычислительного задания необходимо сохранить версию структуры, состояние мишени, набор исходных данных, параметры и версию модели. Для получения вещества — сведения о маршруте, партии, очистке и аналитическом подтверждении. Для биологического опыта — характеристики системы, протокол, контрольные условия и исходные измерения. Объем записи определяется возможностью воспроизвести существенные основания вывода и установить причину расхождения.

Предлагается различать четыре состояния отдельного результата. «Прогноз» обозначает вычислительный вывод с определенной областью применимости. «Наблюдение» фиксирует полученное измерение после контроля его качества. «Подтверждение» означает выполнение заранее установленной процедуры проверки конкретного утверждения. «Решение» определяет дальнейшее действие на основании доступной совокупности сведений. Эти состояния описывают роль информации в процессе и не образуют универсальной шкалы клинической доказательности.

Ретроспективная оценка модели, повторное измерение и независимый эксперимент решают разные задачи. Повтор в тех же условиях помогает оценить устойчивость наблюдения. Исследование иным методом может проверить альтернативное объяснение. Проверка в другой лаборатории характеризует переносимость процедуры. Переход к более сложной биологической системе расширяет исследуемый контекст, но не устраняет автоматически все ограничения предшествующих данных.

Для состава многокомпонентного объекта формируется собственная ветвь подтверждения. Данные о компонентах, воспроизводимости материала и его стабильности связываются с сериями биологических опытов. Гипотеза о взаимодействии компонентов требует программы сравнений, выбранной модели совместного действия и учета неопределенности. Предположение о фармакологическом синергизме и предполагаемый эффект интеграции программных модулей относятся к разным объектам исследования и должны иметь разные доказательные основания.

Существенен и порядок работы с несостоявшимся экспериментом. Неудачный синтез характеризует возможность получить объект выбранным путем. Нарушение контроля качества ставит вопрос о пригодности измерения. Отсутствие эффекта при корректно выполненной проверке относится к биологической гипотезе. В информационной системе эти исходы должны получать разные статусы, чтобы последующее обучение не смешивало причины неуспеха.

Дискуссионный документ FDA по использованию ИИ и машинного обучения, пересмотренный в феврале 2025 года, выделяет управление с участием человека, качество данных и вопросы разработки и оценки моделей. Сам документ прямо обозначен как материал для

обсуждения, а не руководство или политика FDA. Для предлагаемой методологии он представляет методический ориентир в описании оснований и ответственности [34].

Состояние исследования в момент принятия решения можно представить выражением (2.1). Здесь R обозначает действующие требования, H — набор проверяемых гипотез, D — допущенные к использованию данные, M — версии моделей и методов, B — доступный бюджет, L — экспериментальные мощности и ограничения. Такое представление является авторской схемой описания процесса.

$$St = (Rt, Ht, Dt, Mt, Bt, Lt) \quad (2.1)$$

По завершении итерации обновляется конкретная часть этого состояния. Новый опыт может изменить D и оценку H, переобучение — M, расход материалов — B и L. Изменение R требует отдельного основания. Благодаря этому пересмотр гипотезы, исправление модели и изменение цели остаются различимыми событиями.

Решение по итерации принимается после сопоставления результата с условиями, определенными до его получения. Если условия выполнены частично, группа устанавливает, какой неопределенности достаточно для следующего шага, а какую необходимо уменьшить. Контур принятия решений представлен на рисунке 2.1. Он связывает целевой профиль с программой опытов и сохраняет возможность обоснованного возврата к предыдущему состоянию.



Рисунок 2.1. Контур сквозного управления исследовательской гипотезой.

2.3 Совместное использование данных, физических моделей, машинного обучения и экспертных решений

Сквозная методология предполагает явное распределение функций между источниками знания. Измерения описывают поведение объекта в заданных условиях. Физические модели выражают принятые закономерности и допущения. Машинное обучение выявляет зависимо-

сти в доступных данных. Экспертное решение связывает полученные выводы с исследовательской задачей и ответственностью за следующий шаг.

Такое распределение необходимо, поскольку разные методы могут иметь общий источник ошибки. Несколько моделей, обученных на пересекающихся данных, не образуют автоматически независимые подтверждения. Сходные входные структуры и процедуры подготовки способны приводить к согласованным, но ошибочным прогнозам. При объединении оценок требуется анализировать происхождение данных и зависимости между способами расчета.

Исследование Graber и соавторов по прогнозированию аффинности показывает, что структурное пересечение обучающих и проверочных данных способно существенно влиять на оценку обобщения. Предложенное авторами очищенное разделение данных иллюстрирует необходимость контролировать подобные связи. Для сквозного процесса важно учитывать происхождение проверки до включения показателя качества в решение [13].

Вычислительный результат следует передавать вместе с его назначением. Оценка геометрии комплекса, прогноз аффинности и вероятность активности в определенной системе имеют разные смыслы. Физически обоснованный расчет может уточнить предположение о взаимодействии, но сохраняет зависимость от выбранного состояния объекта и допущений. Объединение результатов допустимо после согласования измеряемых величин и условий их применения.

В случае ADMET подобное согласование особенно важно для связи прогноза с программой последующих опытов. ADMET-AI представляет пример вычислительной платформы для оценки ряда свойств больших химических библиотек. Для методологии управления существенны конкретный прогнозируемый показатель и его область применения. Решение о дальнейшем исследовании должно учитывать, какое измерение подтвердит или изменит первоначальную оценку [9].

Предлагается различать неопределенность, связанную с вариабельностью наблюдений, и неопределенность, возникающую из-за недостатка знания модели. Их практическое разделение зависит от выбранного метода и данных. Yang и Li предложили подход к оценке и атрибуции неопределенности молекулярных прогнозов, а также процедуру последующей калибровки. Работа показывает возможность анализировать надежность прогноза на более детальном уровне [35].

Высокая уверенность модели требует проверки соответствия наблюдаемым ошибкам. Если интервалы или вероятности используются для отбора, необходимо установить их фактическое поведение на данных, близких к планируемому применению. При изменении химического класса или экспериментальной системы прежняя калибровка может оказаться недостаточной. В таком случае допускается ограничить использование модели и назначить дополнительную проверку.

Экспертное участие организуется через явные полномочия и основания вмешательства. Специалист может отклонить вычислительную рекомендацию из-за доступности синтеза, недостаточной аналитической определенности или несоответствия биологической гипотезе. Причина отклонения записывается вместе с использованными сведениями. Эта запись помогает последующему анализу различить полезное профессиональное суждение и систематическое предпочтение, не подтвержденное данными.

Языковые модели могут помогать формировать исследовательские задания, извлекать сведения из документов и вызывать специализированные инструменты. В работе Boiko и соавторов система Coscientist объединяла языковую модель с поиском, выполнением кода и автоматизацией химических опытов. Авторы продемонстрировали выполнение конкретных задач экспериментального планирования и оптимизации реакций. Это пример организации взаимодействия инструментов в определенном химическом контексте [10].

При использовании подобных средств в Atlansys EUS MedPharma требуется сохранять различие между предложением плана и допуском к его выполнению. В монографии предлага-

ется связывать каждое существенное извлеченное утверждение с проверяемым источником, а машинно сформированное задание — с утвержденной версией протокола. Наблюдения, полученные при выполнении опыта, регистрируются отдельно от текста, который объясняет их возможный смысл.

Значение экспериментальной обратной связи показано и в работе Koscher и соавторов, где замкнутая автоматизированная платформа соединяла прогнозирование свойств с синтезом и измерениями. Демонстрационными объектами выступали молекулярные красители. Результат подтверждает работоспособность соответствующего подхода для исследованной химической задачи; перенос на фармакологические свойства требует собственной проверки [36].

Согласование вычислительных и экспериментальных методов поэтому должно начинаться с постановки вопроса. Если расхождение вызвано недостаточным знанием о конформации, полезен один набор действий. Если не определен состав образца, нужен другой. Если наблюдаемая активность не связана с предполагаемым механизмом, требуется биологическая проверка альтернатив. Платформа должна поддерживать выбор метода по причине неопределенности, а не только по его доступности.

Общий контекст применения связывает эту организацию с современными принципами доверия к ИИ. В совместных принципах FDA и EMA 2026 года названы междисциплинарная экспертиза, управление данными, оценка работы модели с учетом риска и сопровождение на протяжении жизненного цикла. В методологии монографии эти ориентиры отражаются в определении роли модели, правил ее использования и состава подтверждающих материалов [5].

Предлагаемая система ответственности допускает автоматическое выполнение рутинных операций в установленных границах. За пределами этих границ формируется запрос на экспертное рассмотрение с конкретным предметом неопределенности. Такая организация делает взаимодействие человека и алгоритмов воспроизводимой частью процесса, которую можно исследовать и улучшать.

2.4 Многокритериальное принятие решений при ограниченных ресурсах и неопределенности

На ранних стадиях редко существует одно свойство, улучшение которого гарантирует полезность следующего шага. Активность связана с селективностью, экспозиция — с рисками нежелательного воздействия, доступность синтеза — с длительностью и стоимостью проверки. Методология должна показывать, какие ограничения являются обязательными и какие уступки допустимы между вариантами, уже признанными пригодными для рассмотрения.

Предлагается начинать отбор с проверки обязательных условий. Если объект не определен аналитически либо не выполнено установленное условие пригодности биологического опыта, его высокий расчетный рейтинг не устраняет соответствующий пробел. После такой проверки сравниваются допустимые альтернативы. Для них могут использоваться несколько показателей и заранее согласованные предпочтения исследовательской группы.

Полезным понятием становится недоминируемое решение. Оно означает, что в рассматриваемом наборе нет другого допустимого варианта, который не хуже по всем выбранным показателям и лучше хотя бы по одному. Совокупность таких вариантов образует множество Парето. Оно помогает сделать компромиссы видимыми, однако окончательный выбор требует учета цели проекта, неопределенности и доступного эксперимента.

Сведение всех характеристик к единому баллу допустимо только при определенном содержательном обосновании. Вес показателя должен соответствовать значению изменения этого свойства для решения. Способ нормирования сохраняется вместе с расчетом. Особо важно исключить ситуацию, когда улучшение второстепенных характеристик компенсирует

нарушение обязательного условия. Для подобных ограничений предусматривается самостоятельная проверка допустимости.

Ресурсная задача включает выбор объекта и глубины его изучения. Дешевый расчет, первичный тест и развернутая серия измерений дают разную информацию и требуют разных ресурсов. В исследовании McDonald и соавторов байесовская оптимизация была объединена с такими уровнями проверки при поиске ингибиторов гистондеацетилаз. Платформа использовала докинг, измерения ингибирования в одной точке и кривые зависимости ответа от концентрации. Высший из этих уровней характеризовал лабораторную оценку активности, а не клиническую доказанность [37].

Практический выбор определяется тем, какое незнание препятствует следующему решению. Для одного соединения достаточно уточнить вероятность успешного синтеза. Для другого требуется измерить зависимость эффекта от концентрации. Иногда необходимо проверить контрольный объект, чтобы установить пригодность целой серии опытов. Поэтому единицей планирования становится пара «объект — исследование», дополненная условиями выполнения и возможными исходами.

Активное обучение направляет получение данных на улучшение знания модели, тогда как байесовская оптимизация обычно ориентирована на поиск вариантов с наиболее полезными значениями заданной цели. Эти задачи могут сочетаться, но их критерии следует различать. Сокращение неопределенности о малоинтересном участке химического пространства может улучшить модель и при этом мало повлиять на текущий выбор кандидата.

В BATCHIE предложено последовательное планирование серий опытов с комбинациями препаратов на основе байесовской модели и оценки ожидаемого уменьшения неопределенности. После получения новой серии данных модель обновляется, затем выбирается следующая серия. Опубликованная работа включает проверку подхода в ретроспективных и проспективных исследованиях комбинаций. Этот пример показывает возможность выбирать информативную серию опытов с учетом ограничений дизайна [38].

Для управления конкретным проектом предлагается оценивать ожидаемую ценность изменения решения после опыта. В выражении (2.2) EVSI обозначает ожидаемую ценность выборочной информации: прирост ожидаемой полезности наилучшего последующего действия после получения результата. C обозначает стоимость исследования, а L — потери, связанные с его ожиданием. Все величины должны быть приведены к одной шкале полезности или стоимости. Формула представляет правило постановки задачи, а не готовый оценочный алгоритм платформы.

$$V(e \mid D) = EVSI(e \mid D) — C(e) — L(e) \quad (2.2)$$

Положительная оценка означает, что при принятых предпосылках ожидаемая польза информации превышает связанные с опытом затраты и потери времени. Расчет требует модели возможных исходов и их влияния на выбор. Если достоверно определить эти величины нельзя, предпочтительнее использовать документированное качественное сравнение сценариев. Числовое выражение экспертной интуиции без проверки калибровки создало бы необоснованную определенность.

Рассмотрим условную ситуацию без привязки к результатам Atlansys EUS MedPharma. Соединение А дает более высокий прогноз активности, однако его экспозиция в целевой ткани неизвестна. Соединение Б имеет несколько меньшую измеренную активность и более определенные сведения о доставке. Третье измерение активности А может мало изменить выбор, тогда как исследование его экспозиции способно определить судьбу всей серии. Выбор опыта в этом случае зависит от того, какой исход действительно изменит решение.

При планировании серий опытов учитываются общие затраты на подготовку, доступность приборов и зависимость результатов друг от друга. Несколько близких соединений могут оказаться дешевле в совместной проверке, но дать меньше новой информации, чем более разнообразный набор. Уже выполняющиеся исследования также влияют на очередность: повторное назначение опыта до получения ранее запланированного ответа требует отдельного обоснования.

Политика отбора должна сохранять возможность исследовать новые области. Если система направляет ресурсы только на хорошо изученные серии, она может повысить локальный выход положительных результатов и сузить поиск. В монографии предлагается заранее определять долю либо условия назначения исследовательских опытов, ориентированных на проверку новых предположений. Конкретное распределение ресурсов устанавливается в протоколе проекта.

Выбор сохраняет устойчивость, когда он мало меняется при разумном изменении предпочтений, стоимости или оценки неопределенности. Поэтому при существенных компромиссах полезен анализ чувствительности. Если небольшое изменение исходного допущения полностью меняет порядок приоритетов, это становится основанием для дополнительной проверки или экспертного рассмотрения. Такой результат характеризует надежность решения и не сводится к качеству одной модели.

2.5 Критерии продолжения и прекращения проекта и правила пересмотра исследовательской гипотезы

Контрольная точка завершает определенный объем проверки и задает следующий допустимый шаг. Ее содержание устанавливается до получения оцениваемых результатов. Для ранней разработки необходимо определить проверяемое требование, минимальный состав данных, допустимую неопределенность и возможные действия. Контрольная точка может относиться к соединению, химической серии, биологической гипотезе или проекту в целом.

Решение о продолжении принимается, если данные достаточны для следующего этапа и сохраняется содержательно обоснованная перспектива выполнения цели. Такое решение может сопровождаться условиями: получить дополнительное аналитическое подтверждение, проверить иной метод регистрации эффекта или уточнить ограничивающее свойство. Условия должны иметь срок, ответственного и способ проверки выполнения.

Состояние приостановки используется, когда препятствие связано с отсутствием необходимых данных или ресурсов. Пересмотр гипотезы означает изменение исследовательского объяснения и программы его проверки. Прекращение относится к объекту или направлению, для которого выполнены заранее установленные основания отказа. Эти решения различаются по смыслу и последствиям, поэтому в предлагаемом подходе имеют отдельные статусы.

Таблица 2.2. Решения по результатам контрольной точки ранней разработки.

Решение	Основание	Следующее действие
Продолжить	Данных достаточно для следующего этапа	Утвердить объем работ и условия перехода
Приостановить	Не хватает определенных данных или ресурсов	Задать условие и срок возобновления
Пересмотреть гипотезу	Результат требует иного объяснения	Сохранить исходный вывод и утвердить новую проверку
Прекратить	Выполнен критерий отказа от объекта или направления	Зафиксировать научную либо ресурсную причину

Основания прекращения могут быть научными и ресурсными. К научным относится достаточная проверка, которая не поддержала критическую для проекта связь. К ресурсным — отсутствие выполнимого маршрута получения данных в согласованных пределах. Исчерпание бюджета не является доказательством отсутствия биологического эффекта. Отчет о завершении должен сохранять это различие.

До назначения опыта полезно сформулировать, какое наблюдение ослабит исходную гипотезу. Если любое возможное измерение можно объяснить в ее пользу, гипотеза плохо подходит для управления исследованием. Альтернативные объяснения следует включать в паспорт одновременно с основным. Тогда отрицательный или противоречивый результат сможет изменить программу, а не только увеличить число дополнительных расчетов.

При расхождении прогноза и измерения сначала проверяются идентичность объекта, пригодность образца и качество экспериментальной процедуры. Затем рассматриваются применимость вычислительной модели и возможные биологические объяснения. Такой порядок помогает не перекладывать автоматически ошибку на лабораторию либо алгоритм. Причина расхождения может относиться к нескольким звеньям и требовать отдельных проверок.

Изменение механистического объяснения требует данных о причинной связи. Работа Lee и соавторов посвящена методическим и регуляторным вопросам причинного ИИ в разработке лекарств, включая контроль смешения при анализе наблюдательных данных. Этот исследовательский контекст показывает значение явных допущений причинного анализа. В применении к ранней разработке в монографии предлагается раздельно фиксировать прогноз ответа и утверждение о механизме, подтверждаемое соответствующими экспериментами [39].

Для нейропротекторных и противоопухолевых исследований это правило имеет конкретные следствия. Отсутствие ожидаемой экспозиции может потребовать изменения доставки, даже если активность в исходной модели воспроизводится. Потеря селективности может изменить приоритет химической серии. Несоответствие состава многокомпонентного объекта предполагаемой структуре требует пересмотра материального объекта исследования. В каждом случае уточняется, какое требование нарушено и какая гипотеза остается проверяемой.

Условия завершения доклинической программы следует отделять от ранней номинации кандидата. Полнота исследований определяется задачей следующей стадии и соответствующей программой. Российская публикация М. Н. Макаровой рассматривает доклинические исследования как самостоятельную часть цикла разработки лекарственных средств. Для методологии управления это служит основанием связывать раннее решение с явно определенным объемом последующих работ [29].

Пересмотр модели проводится по установленному порядку. Вновь полученные данные могут служить основанием для переобучения, но результаты первоначальной проверки сохраняются относительно исходной версии. После изменения алгоритма выполняется новая оценка

его пригодности. Иначе система рискует объединить в одном показателе прогноз, сделанный до опыта, и объяснение, сформированное после знакомства с ответом.

Проект руководства FDA по использованию ИИ для поддержки регуляторных решений, опубликованный в январе 2025 года, связывает оценку достоверности модели с контекстом применения и риском. На официальной странице он обозначен как проект с необязательными рекомендациями. В рамках монографии этот подход используется как методический ориентир для определения объема проверки при изменении роли или версии модели [28].

Приостановленный проект может быть возобновлен при появлении заранее определенных оснований: доступного метода синтеза, нового способа доставки или данных, меняющих биологическую гипотезу. Возобновление оформляется новой версией плана с указанием того, какие ограничения сняты. Такая история помогает использовать ранее полученные отрицательные результаты и предотвращает повторение уже проверенных предположений без нового основания.

Профессиональная ответственность требует сохранять возможность прекращения направления. Наличие патента, программного свидетельства, публикации или значительных уже понесенных расходов не определяет будущую фармакологическую пригодность объекта. В предлагаемой методологии эти сведения входят в разные части досье. Решение о следующем исследовании опирается на его ожидаемую ценность и действующие требования.

2.6 Принципы измерения научной, технологической и организационной результативности

Методология становится проверяемой, когда заранее определено, какой результат считается существенным и с чем он сравнивается. Научная результативность характеризует полученное знание и обоснованность вывода. Технологическая — способность воспроизводимо выполнять исследовательские операции и получать пригодные результаты. Организационная — согласованность работы, затраты ресурсов и условия принятия решений. Эти измерения связаны, но требуют собственных показателей.

Для научной оценки необходимо установить, какие гипотезы получили поддержку, какие отвергнуты и какие остались неопределенными. Число рассмотренных молекул само по себе мало сообщает о таком результате. Полезны сведения о подтвержденных объектах, устойчивости эффектов и полноте проверки альтернативных объяснений. Отдельно фиксируются ситуации, когда полученная информация позволила обоснованно прекратить бесперспективную работу.

Технологическая оценка включает воспроизводимость расчетов и измерений, полноту цифровой истории, длительность согласованного исследовательского маршрута и долю работ, прошедших контроль качества. Для каждого показателя определяются начало отсчета, единица анализа и правила учета повторов. Сокращение длительности операции следует интерпретировать вместе с изменением общего маршрута и достигнутого объема подтверждения.

Организационные показатели относятся к передаче заданий, ожиданию, доступности специалистов и причинам ручного вмешательства. Уменьшение времени согласования может представлять полезный эффект, если качество решения сохраняется. Снижение числа вмешательств само по себе неоднозначно: оно может означать как понятную автоматизацию, так и недостаточный экспертный контроль. Поэтому количественный показатель должен сопровождаться содержанием соответствующих событий.

Современная критическая оценка ИИ в разработке лекарств обращает внимание на необходимость связывать вычислительные достижения с улучшением решений. Bender и соавторы рассматривают разрыв между проверкой модели и ее прикладной ценностью. Для Atlansys EUS

MedPharma из этого следует требование включать в программу оценки конечную исследовательскую задачу и доступные альтернативы, а не ограничиваться сравнением алгоритмов [3].

Ранее предложенное сравнение трех процессов получает здесь операционное содержание. Вариант А описывает действующий процесс организации. В варианте Б используются выбранные ИИ-инструменты с существующим порядком взаимодействия специалистов. Вариант В объединяет те же классы инструментов и согласованные правила работы в интегрированной платформе. Протокол должен точно описать, какие элементы изменяются в каждом варианте и какие условия остаются сопоставимыми.

Всем вариантам задаются общий профиль требований, сопоставимый исходный доступ к данным и согласованный ресурсный бюджет. Критерии обновления моделей и выбора опытов описываются до начала оценки. По мере выполнения исследования наборы полученных данных могут различаться, поскольку политика отбора влияет на выполняемые эксперименты. Это различие учитывается как часть сравниваемого процесса и сохраняется в истории его версий.

Сопоставление А и Б характеризует изменение, связанное с применением выбранных инструментов в определенных условиях. Сопоставление Б и В направлено на оценку добавочного эффекта интеграции. Различие А и В отражает совокупное изменение. Причинная интерпретация требует обоснованного дизайна и контроля иных различий между вариантами. Схема разделения эффектов приведена на рисунке 2.2.



Рисунок 2.2. Сравнение процессов для оценки вклада инструментов и их интеграции.

По возможности задания или сопоставимые исследовательские единицы распределяются между вариантами до получения оцениваемых ответов. Экспериментальные серии балансируются по существенным условиям, а образцы могут кодироваться для уменьшения влияния ожиданий исполнителя. Если распределение выполняется на уровне проектов, статистическая оценка должна учитывать зависимость результатов внутри проекта. Большое число расчетов одной серии не заменяет множество независимых проверок.

Контрольная процедура подтверждения должна быть общей для сравниваемых вариантов. Иначе один процесс может оказаться быстрее из-за менее полного исследования. До начала работы устанавливаются правила разрешения пограничных и противоречивых результатов, а

также учета не полученных образцов и незавершенных опытов. Изменения протокола сохраняются отдельно, чтобы их последствия можно было проверить.

Результаты поперечного исследования Druedahl и соавторов описывают сообщавшееся применение ИИ в разработке препаратов по данным Pharmaprojects. Такой дизайн помогает характеризовать распространение подходов и типы задач. Он не воспроизводит прямое сравнение одинаковых проектов с ИИ и без него. Поэтому сведения об использовании технологии следует отделять от доказательства ее причинного влияния на срок или выход разработки [40].

Для затрат задается согласованная граница учета. Она включает подготовку данных, вычисления, труд специалистов, синтез, аналитику, биологические испытания и относимую к проекту эксплуатацию платформы. Учитываются неуспешные попытки и повторные работы. Затраты прошлых периодов, общая инфраструктура и стоимость партнерских услуг отражаются по заранее описанному правилу, одинаковому для сопоставляемых вариантов.

В выражении (2.3) Сполная обозначает затраты в установленных границах, а Nподтв — число объектов одной согласованной категории, подтвердивших требования в общей процедуре. Показатель не смешивает первичные активные соединения и кандидаты с разным объемом исследования. При нулевом выходе значение отношения не определяется; в отчете указываются понесенные затраты и отсутствие подтвержденного результата.

$$C_{объект} = C_{полная} / N_{подтв} \quad (2.3)$$

Временные показатели требуют учета незавершенных работ. Если к моменту анализа часть проектов продолжает исследование, их нельзя автоматически исключать или считать завершенными неуспешно. Протокол должен задавать срок наблюдения и способ описания таких случаев. Медианное время до события, доля завершенных проверок в установленный период и ресурсные затраты отвечают на разные вопросы и представляются раздельно.

Таблица 2.3. Основные показатели проверки сквозного управления.

Показатель	Что измеряется	Правило сопоставления
Подтвержденный выход	Число объектов, прошедших установленную проверку	Единая категория результата и общие критерии
Время до результата	Календарное время до согласованного события	Общее начало отсчета и учет незавершенных работ
Полные затраты	Ресурсы в установленных границах учета	Включение неуспехов, повторов и подготовки
Стоимость объекта	Полные затраты на подтвержденный объект	При нулевом выходе отношение не определяется
Качество прогноза	Ошибка и калибровка в заданном контексте	Независимые данные и зафиксированная версия
Воспроизводимость	Устойчивость результата в повторной проверке	Предварительно заданная процедура и ее полнота
Прослеживаемость	Полнота связей данных и оснований решения	Единый перечень обязательных сведений

Для оценки различий приводятся неопределенность и устойчивость результата, а не только одно среднее значение. Проверяется влияние сложности химического пространства,

доступности исходных веществ, состава лабораторных мощностей и опыта исполнителей. При небольшом числе проектов целесообразно ограничить выводы наблюдаемыми условиями и дополнить количественную оценку разбором причин изменений.

Экономический эффект пилота также необходимо отличать от прогноза его масштабирования. Обзоры, в том числе работа Wu и соавторов, объединяют широкий круг направлений применения ИИ — от скрининга до клинических исследований. Такие обзоры помогают определить возможные области использования. Для расчета эффекта определенного проекта требуются его собственные данные, единый базовый процесс и согласованная методика учета [41].

Воспроизводимый отчет должен позволять восстановить требования, состав исследовательских единиц, доступные на момент выбора данные, версии инструментов и выполненные проверки. Его выводы связываются с конкретной конфигурацией процесса. При изменении модели, лаборатории или целевой задачи устанавливается, какая часть подтверждения сохраняет силу и какую нужно получить заново.

Предлагаемая методология тем самым задает требования к архитектуре Atlansys EUS MedPharma. Платформа должна связывать версии целевого профиля и гипотез с данными, вычислительными заданиями, материальными образцами и контрольными решениями. Общие службы управления данными, отчетности и безопасности обеспечивают сохранность этих связей. В следующей главе данная рамка будет раскрыта через распределение функций между системами, подсистемами и сервисами.

Выводы

Сквозное управление ранней разработкой основано на согласовании исследовательских действий с целевыми требованиями. Профиль будущего продукта определяет медицинский замысел, а требования к кандидату переводят его в проверяемые условия выбора. Для каждого условия задаются показатель, контекст наблюдения, критерий интерпретации и история изменений.

Паспорт гипотезы связывает исходные основания с альтернативными объяснениями, программой опытов и возможными решениями. Завершение итерации означает определенное изменение состояния знания и плана действий. Вычислительный прогноз, измерение, подтверждение и решение сохраняются как различающиеся информационные объекты, что делает переходы между ними прослеживаемыми.

Данные, физические модели, машинное обучение и экспертное участие выполняют взаимодополняющие функции. Их согласование требует учета зависимостей между источниками, области применимости прогнозов и надежности оцененной неопределенности. Экспертное вмешательство становится частью воспроизводимого процесса, когда его предмет, основание и последствия зафиксированы.

Многокритериальный выбор начинается с обязательных ограничений и продолжается анализом допустимых компромиссов. Планирование охватывает одновременно соединение и способ его проверки. Информативность опыта оценивается по тому, насколько его результат способен улучшить последующее решение при доступных ресурсах. При недостаточной определенности количественных оценок применяется документированное сравнение сценариев.

Продолжение, приостановка, пересмотр гипотезы и прекращение имеют собственные основания. Различение научных, технических и ресурсных причин позволяет сохранять полезное знание даже при завершении направления. Изменение цели или модели требует новой версии плана и не отменяет результатов первоначальной проверки.

Проверка методологии включает научную, технологическую и организационную результативность. Для установления добавочного эффекта интеграции сравниваются процессы с общими требованиями и сопоставимыми условиями. Время, затраты, выход подтвержденных

объектов и воспроизводимость оцениваются по заранее определенным правилам с учетом незавершенных и неуспешных работ.

Главный результат главы — единая рамка, связывающая требования к кандидату, выбор следующего исследования и проверку итогового эффекта. Она определяет, какие информационные связи и функции управления должна реализовать архитектура Atlansys EUS MedPharma. Дальнейшее изложение переходит от этих методологических требований к устройству платформы.

Глава 2 Методология сквозного управления разработкой новых соединений

Соединения могут занимать соседние строки в таблице расчетов и требовать совершенно разных исследовательских решений. Для одного остается проверить воспроизводимость активности, для другого — выяснить, удастся ли получить вещество нужной чистоты. В третьем случае главным становится вопрос о достижении необходимой концентрации в целевой ткани. Общий рейтинг полезен лишь тогда, когда исследователь понимает, какие основания стоят за каждой оценкой и какое действие должно последовать.

В первой главе определены процессы и ограничения ранней фармацевтической разработки. Теперь необходимо связать их в методологию, которая устанавливает порядок перехода от медицинской потребности к исследовательскому заданию, от задания — к эксперименту, от результата — к следующему решению. В исходной концепции Atlansys EUS MedPharma эта задача поставлена через сквозную интеграцию данных, специализированных методов ИИ и функций управления исследованиями. Настоящая глава развивает ее на уровне правил организации исследовательского процесса [1].

Под сквозным управлением будем понимать согласованное изменение требований, исследовательских гипотез, вычислительных моделей и экспериментальных действий на основании прослеживаемых данных. Единицей управления выступает проверяемая гипотеза об определенном объекте разработки. Для нее задаются условия подтверждения, возможные альтернативы и решение, которое можно принять после получения результата. Благодаря этому план исследования сохраняет связь с конечной целью даже при изменении химической серии или состава применяемых алгоритмов.

Современные обзоры рассматривают совместную работу вычислительной и экспериментальной лабораторий как необходимую составляющую развития ИИ в создании лекарств. В работе Fu и Chen интеграция биологических знаний и алгоритмических методов связана с организацией обмена данными между этими средами. Для методологии платформы это означает необходимость определить содержание такого обмена и ответственность за использование его результатов [30].

Предлагаемый подход включает шесть взаимосвязанных элементов: целевые требования, паспорт исследовательской гипотезы, согласование источников знания, выбор эксперимента при ограниченных ресурсах, правила пересмотра проекта и программу оценки результативности. Эти элементы образуют общую рамку последующих глав. Их конкретная программная реализация может изменяться, тогда как основания исследовательских решений должны оставаться проверяемыми.

2.1 Целевой профиль будущего лекарственного продукта и требования к соединению-кандидату

Разработка начинается с описания результата, который будет иметь значение для будущего применения. Целевой профиль лекарственного продукта (target product profile, TPP) связывает предполагаемое назначение, группу пациентов, желаемое действие и ограничения использования. В образовательном материале NIH SEED выделены минимальные и предпочтительные характеристики продукта, а сам профиль рассматривается как обновляемый инструмент планирования. Этот материал объясняет подход к работе с TPP и не является руководством регулирующего органа [31].

Для раннего поиска профиль продукта необходимо преобразовать в требования к исследуемому объекту. В настоящей методологии целевой профиль соединения-кандидата (target candidate profile, TCP) описывает свойства и подтверждающие данные, которые нужны для решения о его дальнейшем развитии. TPP сохраняет ориентацию на медицинскую задачу, а TCP переводит ее в проверяемые условия выбора. Профиль требований к качеству будущего продукта решает отдельную задачу; его нельзя считать полным заменителем этих двух документов.

Например, предполагаемая необходимость воздействия в центральной нервной системе формирует вопрос о доставке и экспозиции в соответствующей ткани. Требование длительного действия меняет значение метаболической устойчивости и времени удержания эффекта. Предпочтительный путь введения влияет на требования к растворимости и возможности создания лекарственной формы. В обзоре Serrano и соавторов задачи поиска соединений и доставки рассматриваются в общем контексте разработки, что подчеркивает необходимость согласования молекулярных и фармацевтических характеристик [32].

При этом формулировка требования должна включать способ его проверки. Указание «хорошая проницаемость» не определяет измеряемую величину, экспериментальную систему и допустимый результат. Для управления требуется установить, какой показатель будет использован, в каких условиях его получат и какое решение он поддерживает. На раннем этапе часть значений может оставаться неизвестной. Такие пробелы следует обозначать явно и связывать с планом получения данных.

Предлагается различать обязательное условие, желательный уровень и исследовательское предположение. Обязательное условие определяет допустимость дальнейшего выбора в данном контексте. Желательный уровень позволяет сравнить уже допустимые варианты. Исследовательское предположение фиксирует ожидаемую связь, которую еще предстоит проверить. Это разделение препятствует превращению расчетного ожидания в фактически выполненное требование.

Для каждого требования необходимы идентификатор, содержательная формулировка, измеряемый показатель, единица, условия наблюдения и критерий интерпретации. Дополнительно указываются источник обоснования, доступные сведения и ответственный за решение. Если один критерий зависит от другого, эту связь фиксируют в профиле. Вопрос о запасе безопасности, например, связан с предполагаемой эффективной экспозицией; раздельные благоприятные оценки без такого сопоставления оставляют исследовательский вопрос открытым.

Требования полезно разделять и по моменту принятия решения. На раннем отборе допустима предварительная расчетная оценка части свойств. Перед дорогостоящим исследованием требуется более определенное состояние сведений о веществе и его действии. Подход не предполагает одинакового набора испытаний для всех проектов. Состав проверки должен соответствовать биологической гипотезе и последствиям возможной ошибки.

В онкологическом направлении особенно существенен контекст биологического ответа. Обзор Ocana и соавторов связывает интеграцию ИИ с выбором мишеней и ранней разработкой, одновременно отмечая ограничения переноса результатов неподходящих доклинических моделей. Поэтому в профиле кандидата нужно фиксировать, какую сторону заболевания воспроизводит выбранная модель и какие свойства за пределами этой модели пока не установлены [33].

Ниже приведена методологическая матрица перевода требований. Она показывает логику постановки задач и не задает универсальных числовых нормативов. Конкретные пороги устанавливаются для проекта на основании медицинского замысла, характеристик исследуемого объекта и обоснованной программы проверки.

Таблица 2.1. Перевод целевых требований в исследовательские решения.

Требование профиля	Проверяемый вопрос	Основание решения
Медицинская задача	Какое изменение состояния должно иметь значение?	Релевантная модель и определенная конечная точка
Фармакологическое действие	Воспроизводится ли нужный эффект?	Измерение с контролями и установленными критериями
Экспозиция	Достигает ли вещество целевой ткани?	Связь экспозиции с условиями реализации эффекта
Безопасность и селективность	В каких условиях проявляется нежелательное действие?	Сопоставимые системы и план проверки рисков
Идентичность объекта	Что именно получено и испытано?	Состав, чистота, серия и стабильность материала
Получение и применение	Доступен ли воспроизводимый путь получения?	Маршрут, ресурсы и ограничения последующей разработки

У каждого требования должна существовать история изменений. Если после неудачного результата группа меняет критерий, первоначальная оценка сохраняется вместе с причиной пересмотра. Иначе улучшение показателей может оказаться следствием изменения цели. Версия целевого профиля входит в состав любого существенного решения по соединению, серии и проекту.

Пересмотр профиля допустим при появлении новых сведений о заболевании, механизме действия, доступности технологии или условиях будущего применения. Такое изменение требует согласования последствий для текущей программы. Если меняется предполагаемый путь введения, часть прежних требований сохраняет значение, а часть должна быть уточнена. Платформа должна позволять установить, какие данные остались пригодными и какие исследования необходимо выполнить заново.

Таким образом, профиль служит основанием для согласования работы специалистов. Он помогает связать различающиеся по форме задания биоинформатика, химика и фармаколога с общей целью. Дальнейшее управление строится вокруг вопроса о том, какие данные нужны, чтобы обоснованно изменить состояние выполнения этих требований.

2.2 Цикл «дизайн — синтез — тестирование — анализ» и уровни доказательности исследовательских результатов

В первой главе цикл «дизайн — синтез — тестирование — анализ» был рассмотрен как форма повторяемости исследования. В методологии управления каждому его прохождению необходимо придать определенный вопрос и условие завершения. Итерация начинается с зафиксированного состояния знания и завершается решением, основания которого можно восстановить. Между этими точками располагаются конкретные задания, результаты выполнения и проверки качества.

Для этого вводится паспорт исследовательской гипотезы. Он содержит формулировку предполагаемой связи, объект проверки, исходные основания и альтернативные объяснения. Следующая часть паспорта определяет эксперимент, необходимые контроли, измеряемый результат и ожидаемые варианты интерпретации. Завершающая часть задает возможные решения и условия пересмотра. Паспорт представляет собой предлагаемый в монографии способ

организации исследования, который может быть реализован разными программными средствами.

Примером служит предположение о том, что изменение определенной части структуры улучшит устойчивость соединения, сохранив его активность. Для проверки важно заранее установить, с каким объектом сравнивается новый вариант и в каких условиях измеряются оба свойства. Возможны разные исходы: оба требования выполнены, одно улучшилось ценой другого, результаты противоречивы либо образец не отвечает требованиям качества. Эти исходы должны вести к различающимся действиям.

Особое значение имеет предсказательная валидность выбранного измерения. Scannell и соавторы связывают продуктивность разработки с тем, насколько исследовательские инструменты помогают обнаруживать объекты, полезные на последующих стадиях. В предлагаемом паспорте поэтому фиксируется отношение измеряемого показателя к целевому требованию. Если опыт характеризует промежуточный признак, границы его интерпретации сохраняются вместе с результатом [6].

Для вычислительного задания необходимо сохранить версию структуры, состояние мишени, набор исходных данных, параметры и версию модели. Для получения вещества — сведения о маршруте, партии, очистке и аналитическом подтверждении. Для биологического опыта — характеристики системы, протокол, контрольные условия и исходные измерения. Объем записи определяется возможностью воспроизвести существенные основания вывода и установить причину расхождения.

Предлагается различать четыре состояния отдельного результата. «Прогноз» обозначает вычислительный вывод с определенной областью применимости. «Наблюдение» фиксирует полученное измерение после контроля его качества. «Подтверждение» означает выполнение заранее установленной процедуры проверки конкретного утверждения. «Решение» определяет дальнейшее действие на основании доступной совокупности сведений. Эти состояния описывают роль информации в процессе и не образуют универсальной шкалы клинической доказательности.

Ретроспективная оценка модели, повторное измерение и независимый эксперимент решают разные задачи. Повтор в тех же условиях помогает оценить устойчивость наблюдения. Исследование иным методом может проверить альтернативное объяснение. Проверка в другой лаборатории характеризует переносимость процедуры. Переход к более сложной биологической системе расширяет исследуемый контекст, но не устраняет автоматически все ограничения предшествующих данных.

Для состава многокомпонентного объекта формируется собственная ветвь подтверждения. Данные о компонентах, воспроизводимости материала и его стабильности связываются с сериями биологических опытов. Гипотеза о взаимодействии компонентов требует программы сравнений, выбранной модели совместного действия и учета неопределенности. Предположение о фармакологическом синергизме и предполагаемый эффект интеграции программных модулей относятся к разным объектам исследования и должны иметь разные доказательные основания.

Существенен и порядок работы с несостоявшимся экспериментом. Неудачный синтез характеризует возможность получить объект выбранным путем. Нарушение контроля качества ставит вопрос о пригодности измерения. Отсутствие эффекта при корректно выполненной проверке относится к биологической гипотезе. В информационной системе эти исходы должны получать разные статусы, чтобы последующее обучение не смешивало причины неуспеха.

Дискуссионный документ FDA по использованию ИИ и машинного обучения, пересмотренный в феврале 2025 года, выделяет управление с участием человека, качество данных и вопросы разработки и оценки моделей. Сам документ прямо обозначен как материал для

обсуждения, а не руководство или политика FDA. Для предлагаемой методологии он представляет методический ориентир в описании оснований и ответственности [34].

Состояние исследования в момент принятия решения можно представить выражением (2.1). Здесь R обозначает действующие требования, H — набор проверяемых гипотез, D — допущенные к использованию данные, M — версии моделей и методов, B — доступный бюджет, L — экспериментальные мощности и ограничения. Такое представление является авторской схемой описания процесса.

$$S_t = (R_t, H_t, D_t, M_t, B_t, L_t) \quad (2.1)$$

По завершении итерации обновляется конкретная часть этого состояния. Новый опыт может изменить D и оценку H, переобучение — M, расход материалов — B и L. Изменение R требует отдельного основания. Благодаря этому пересмотр гипотезы, исправление модели и изменение цели остаются различимыми событиями.

Решение по итерации принимается после сопоставления результата с условиями, определенными до его получения. Если условия выполнены частично, группа устанавливает, какой неопределенности достаточно для следующего шага, а какую необходимо уменьшить. Контур принятия решений представлен на рисунке 2.1. Он связывает целевой профиль с программой опытов и сохраняет возможность обоснованного возврата к предыдущему состоянию.



Рисунок 2.1. Контур сквозного управления исследовательской гипотезой.

2.3 Совместное использование данных, физических моделей, машинного обучения и экспертных решений

Сквозная методология предполагает явное распределение функций между источниками знания. Измерения описывают поведение объекта в заданных условиях. Физические модели выражают принятые закономерности и допущения. Машинное обучение выявляет зависимости в доступных данных. Экспертное решение связывает полученные выводы с исследовательской задачей и ответственностью за следующий шаг.

Такое распределение необходимо, поскольку разные методы могут иметь общий источник ошибки. Несколько моделей, обученных на пересекающихся данных, не образуют автоматически независимые подтверждения. Сходные входные структуры и процедуры подготовки способны приводить к согласованным, но ошибочным прогнозам. При объединении оценок требуется анализировать происхождение данных и зависимости между способами расчета.

Исследование Graber и соавторов по прогнозированию аффинности показывает, что структурное пересечение обучающих и проверочных данных способно существенно влиять на оценку обобщения. Предложенное авторами очищенное разделение данных иллюстрирует необходимость контролировать подобные связи. Для сквозного процесса важно учитывать происхождение проверки до включения показателя качества в решение [13].

Вычислительный результат следует передавать вместе с его назначением. Оценка геометрии комплекса, прогноз аффинности и вероятность активности в определенной системе имеют разные смыслы. Физически обоснованный расчет может уточнить предположение о взаимодействии, но сохраняет зависимость от выбранного состояния объекта и допущений. Объединение результатов допустимо после согласования измеряемых величин и условий их применения.

В случае ADMET подобное согласование особенно важно для связи прогноза с программой последующих опытов. ADMET-AI представляет пример вычислительной платформы для оценки ряда свойств больших химических библиотек. Для методологии управления суще-

ственны конкретный прогнозируемый показатель и его область применения. Решение о дальнейшем исследовании должно учитывать, какое измерение подтвердит или изменит первоначальную оценку [9].

Предлагается различать неопределенность, связанную с вариабельностью наблюдений, и неопределенность, возникающую из-за недостатка знания модели. Их практическое разделение зависит от выбранного метода и данных. Yang и Li предложили подход к оценке и атрибуции неопределенности молекулярных прогнозов, а также процедуру последующей калибровки. Работа показывает возможность анализировать надежность прогноза на более детальном уровне [35].

Высокая уверенность модели требует проверки соответствия наблюдаемым ошибкам. Если интервалы или вероятности используются для отбора, необходимо установить их фактическое поведение на данных, близких к планируемому применению. При изменении химического класса или экспериментальной системы прежняя калибровка может оказаться недостаточной. В таком случае допускается ограничить использование модели и назначить дополнительную проверку.

Экспертное участие организуется через явные полномочия и основания вмешательства. Специалист может отклонить вычислительную рекомендацию из-за доступности синтеза, недостаточной аналитической определенности или несоответствия биологической гипотезе. Причина отклонения записывается вместе с использованными сведениями. Эта запись помогает последующему анализу различить полезное профессиональное суждение и систематическое предпочтение, не подтвержденное данными.

Языковые модели могут помогать формировать исследовательские задания, извлекать сведения из документов и вызывать специализированные инструменты. В работе Boiko и соавторов система Coscientist объединяла языковую модель с поиском, выполнением кода и автоматизацией химических опытов. Авторы продемонстрировали выполнение конкретных задач экспериментального планирования и оптимизации реакций. Это пример организации взаимодействия инструментов в определенном химическом контексте [10].

При использовании подобных средств в Atlansys EUS MedPharma требуется сохранять различие между предложением плана и допуском к его выполнению. В монографии предлагается связывать каждое существенное извлеченное утверждение с проверяемым источником, а машинно сформированное задание — с утвержденной версией протокола. Наблюдения, полученные при выполнении опыта, регистрируются отдельно от текста, который объясняет их возможный смысл.

Значение экспериментальной обратной связи показано и в работе Koscher и соавторов, где замкнутая автоматизированная платформа соединяла прогнозирование свойств с синтезом и измерениями. Демонстрационными объектами выступали молекулярные красители. Результат подтверждает работоспособность соответствующего подхода для исследованной химической задачи; перенос на фармакологические свойства требует собственной проверки [36].

Согласование вычислительных и экспериментальных методов поэтому должно начинаться с постановки вопроса. Если расхождение вызвано недостаточным знанием о конформации, полезен один набор действий. Если не определен состав образца, нужен другой. Если наблюдаемая активность не связана с предполагаемым механизмом, требуется биологическая проверка альтернатив. Платформа должна поддерживать выбор метода по причине неопределенности, а не только по его доступности.

Общий контекст применения связывает эту организацию с современными принципами доверия к ИИ. В совместных принципах FDA и EMA 2026 года названы междисциплинарная экспертиза, управление данными, оценка работы модели с учетом риска и сопровождение на протяжении жизненного цикла. В методологии монографии эти ориентиры отражаются в определении роли модели, правил ее использования и состава подтверждающих материалов [5].

Предлагаемая система ответственности допускает автоматическое выполнение рутинных операций в установленных границах. За пределами этих границ формируется запрос на экспертное рассмотрение с конкретным предметом неопределенности. Такая организация делает взаимодействие человека и алгоритмов воспроизводимой частью процесса, которую можно исследовать и улучшать.

2.4 Многокритериальное принятие решений при ограниченных ресурсах и неопределенности

На ранних стадиях редко существует одно свойство, улучшение которого гарантирует полезность следующего шага. Активность связана с селективностью, экспозиция — с рисками нежелательного воздействия, доступность синтеза — с длительностью и стоимостью проверки. Методология должна показывать, какие ограничения являются обязательными и какие уступки допустимы между вариантами, уже признанными пригодными для рассмотрения.

Предлагается начинать отбор с проверки обязательных условий. Если объект не определен аналитически либо не выполнено установленное условие пригодности биологического опыта, его высокий расчетный рейтинг не устраняет соответствующий пробел. После такой проверки сравниваются допустимые альтернативы. Для них могут использоваться несколько показателей и заранее согласованные предпочтения исследовательской группы.

Полезным понятием становится недоминируемое решение. Оно означает, что в рассматриваемом наборе нет другого допустимого варианта, который не хуже по всем выбранным показателям и лучше хотя бы по одному. Совокупность таких вариантов образует множество Парето. Оно помогает сделать компромиссы видимыми, однако окончательный выбор требует учета цели проекта, неопределенности и доступного эксперимента.

Сведение всех характеристик к единому баллу допустимо только при определенном содержательном обосновании. Вес показателя должен соответствовать значению изменения этого свойства для решения. Способ нормирования сохраняется вместе с расчетом. Особо важно исключить ситуацию, когда улучшение второстепенных характеристик компенсирует нарушение обязательного условия. Для подобных ограничений предусматривается самостоятельная проверка допустимости.

Ресурсная задача включает выбор объекта и глубины его изучения. Дешевый расчет, первичный тест и развернутая серия измерений дают разную информацию и требуют разных ресурсов. В исследовании McDonald и соавторов байесовская оптимизация была объединена с такими уровнями проверки при поиске ингибиторов гистондеацетилаз. Платформа использовала докинг, измерения ингибирования в одной точке и кривые зависимости ответа от концентрации. Высший из этих уровней характеризовал лабораторную оценку активности, а не клиническую доказанность [37].

Практический выбор определяется тем, какое незнание препятствует следующему решению. Для одного соединения достаточно уточнить вероятность успешного синтеза. Для другого требуется измерить зависимость эффекта от концентрации. Иногда необходимо проверить контрольный объект, чтобы установить пригодность целой серии опытов. Поэтому единицей планирования становится пара «объект — исследование», дополненная условиями выполнения и возможными исходами.

Активное обучение направляет получение данных на улучшение знания модели, тогда как байесовская оптимизация обычно ориентирована на поиск вариантов с наиболее полезными значениями заданной цели. Эти задачи могут сочетаться, но их критерии следует различать. Сокращение неопределенности о малоинтересном участке химического пространства может улучшить модель и при этом мало повлиять на текущий выбор кандидата.

В BATCHIE предложено последовательное планирование серий опытов с комбинациями препаратов на основе байесовской модели и оценки ожидаемого уменьшения неопределенности. После получения новой серии данных модель обновляется, затем выбирается следующая серия. Опубликованная работа включает проверку подхода в ретроспективных и проспективных исследованиях комбинаций. Этот пример показывает возможность выбирать информативную серию опытов с учетом ограничений дизайна [38].

Для управления конкретным проектом предлагается оценивать ожидаемую ценность изменения решения после опыта. В выражении (2.2) EVSI обозначает ожидаемую ценность выборочной информации: прирост ожидаемой полезности наилучшего последующего действия после получения результата. C обозначает стоимость исследования, а L — потери, связанные с его ожиданием. Все величины должны быть приведены к одной шкале полезности или стоимости. Формула представляет правило постановки задачи, а не готовый оценочный алгоритм платформы.

$$V(e \vee D) = EVSI(e \vee D) - C(e) - L(e) \quad (2.2)$$

Положительная оценка означает, что при принятых предпосылках ожидаемая польза информации превышает связанные с опытом затраты и потери времени. Расчет требует модели возможных исходов и их влияния на выбор. Если достоверно определить эти величины нельзя, предпочтительнее использовать документированное качественное сравнение сценариев. Числовое выражение экспертной интуиции без проверки калибровки создало бы необоснованную определенность.

Рассмотрим условную ситуацию без привязки к результатам Atlansys EUS MedPharma. Соединение А дает более высокий прогноз активности, однако его экспозиция в целевой ткани неизвестна. Соединение Б имеет несколько меньшую измеренную активность и более определенные сведения о доставке. Третье измерение активности А может мало изменить выбор, тогда как исследование его экспозиции способно определить судьбу всей серии. Выбор опыта в этом случае зависит от того, какой исход действительно изменит решение.

При планировании серий опытов учитываются общие затраты на подготовку, доступность приборов и зависимость результатов друг от друга. Несколько близких соединений могут оказаться дешевле в совместной проверке, но дать меньше новой информации, чем более разнообразный набор. Уже выполняющиеся исследования также влияют на очередность: повторное назначение опыта до получения ранее запланированного ответа требует отдельного обоснования.

Политика отбора должна сохранять возможность исследовать новые области. Если система направляет ресурсы только на хорошо изученные серии, она может повысить локальный выход положительных результатов и сузить поиск. В монографии предлагается заранее определять долю либо условия назначения исследовательских опытов, ориентированных на проверку новых предположений. Конкретное распределение ресурсов устанавливается в протоколе проекта.

Выбор сохраняет устойчивость, когда он мало меняется при разумном изменении предпочтений, стоимости или оценки неопределенности. Поэтому при существенных компромиссах полезен анализ чувствительности. Если небольшое изменение исходного допущения полностью меняет порядок приоритетов, это становится основанием для дополнительной проверки или экспертного рассмотрения. Такой результат характеризует надежность решения и не сводится к качеству одной модели.

2.5 Критерии продолжения и прекращения проекта и правила пересмотра исследовательской гипотезы

Контрольная точка завершает определенный объем проверки и задает следующий допустимый шаг. Ее содержание устанавливается до получения оцениваемых результатов. Для ранней разработки необходимо определить проверяемое требование, минимальный состав данных, допустимую неопределенность и возможные действия. Контрольная точка может относиться к соединению, химической серии, биологической гипотезе или проекту в целом.

Решение о продолжении принимается, если данные достаточны для следующего этапа и сохраняется содержательно обоснованная перспектива выполнения цели. Такое решение может сопровождаться условиями: получить дополнительное аналитическое подтверждение, проверить иной метод регистрации эффекта или уточнить ограничивающее свойство. Условия должны иметь срок, ответственного и способ проверки выполнения.

Состояние приостановки используется, когда препятствие связано с отсутствием необходимых данных или ресурсов. Пересмотр гипотезы означает изменение исследовательского объяснения и программы его проверки. Прекращение относится к объекту или направлению, для которого выполнены заранее установленные основания отказа. Эти решения различаются по смыслу и последствиям, поэтому в предлагаемом подходе имеют отдельные статусы.

Таблица 2.2. Решения по результатам контрольной точки ранней разработки.

Решение	Основание	Следующее действие
Продолжить	Данных достаточно для следующего этапа	Утвердить объем работ и условия перехода
Приостановить	Не хватает определенных данных или ресурсов	Задать условие и срок возобновления
Пересмотреть гипотезу	Результат требует иного объяснения	Сохранить исходный вывод и утвердить новую проверку
Прекратить	Выполнен критерий отказа от объекта или направления	Зафиксировать научную либо ресурсную причину

Основания прекращения могут быть научными и ресурсными. К научным относится достаточная проверка, которая не поддержала критическую для проекта связь. К ресурсным — отсутствие выполнимого маршрута получения данных в согласованных пределах. Исчерпание бюджета не является доказательством отсутствия биологического эффекта. Отчет о завершении должен сохранять это различие.

До назначения опыта полезно сформулировать, какое наблюдение ослабит исходную гипотезу. Если любое возможное измерение можно объяснить в ее пользу, гипотеза плохо подходит для управления исследованием. Альтернативные объяснения следует включать в пас-

порт одновременно с основным. Тогда отрицательный или противоречивый результат сможет изменить программу, а не только увеличить число дополнительных расчетов.

При расхождении прогноза и измерения сначала проверяются идентичность объекта, пригодность образца и качество экспериментальной процедуры. Затем рассматриваются применимость вычислительной модели и возможные биологические объяснения. Такой порядок помогает не перекладывать автоматически ошибку на лабораторию либо алгоритм. Причина расхождения может относиться к нескольким звеньям и требовать отдельных проверок.

Изменение механистического объяснения требует данных о причинной связи. Работа Lee и соавторов посвящена методическим и регуляторным вопросам причинного ИИ в разработке лекарств, включая контроль смешения при анализе наблюдательных данных. Этот исследовательский контекст показывает значение явных допущений причинного анализа. В применении к ранней разработке в монографии предлагается раздельно фиксировать прогноз ответа и утверждение о механизме, подтверждаемое соответствующими экспериментами [39].

Для нейропротекторных и противоопухолевых исследований это правило имеет конкретные следствия. Отсутствие ожидаемой экспозиции может потребовать изменения доставки, даже если активность в исходной модели воспроизводится. Потеря селективности может изменить приоритет химической серии. Несоответствие состава многокомпонентного объекта предполагаемой структуре требует пересмотра материального объекта исследования. В каждом случае уточняется, какое требование нарушено и какая гипотеза остается проверяемой.

Условия завершения доклинической программы следует отделять от ранней номинации кандидата. Полнота исследований определяется задачей следующей стадии и соответствующей программой. Российская публикация М. Н. Макаровой рассматривает доклинические исследования как самостоятельную часть цикла разработки лекарственных средств. Для методологии управления это служит основанием связывать раннее решение с явно определенным объемом последующих работ [29].

Пересмотр модели проводится по установленному порядку. Вновь полученные данные могут служить основанием для переобучения, но результаты первоначальной проверки сохраняются относительно исходной версии. После изменения алгоритма выполняется новая оценка его пригодности. Иначе система рискует объединить в одном показателе прогноз, сделанный до опыта, и объяснение, сформированное после знакомства с ответом.

Проект руководства FDA по использованию ИИ для поддержки регуляторных решений, опубликованный в январе 2025 года, связывает оценку достоверности модели с контекстом применения и риском. На официальной странице он обозначен как проект с необязательными рекомендациями. В рамках монографии этот подход используется как методический ориентир для определения объема проверки при изменении роли или версии модели [28].

Приостановленный проект может быть возобновлен при появлении заранее определенных оснований: доступного метода синтеза, нового способа доставки или данных, меняющих биологическую гипотезу. Возобновление оформляется новой версией плана с указанием того, какие ограничения сняты. Такая история помогает использовать ранее полученные отрицательные результаты и предотвращает повторение уже проверенных предположений без нового основания.

Профессиональная ответственность требует сохранять возможность прекращения направления. Наличие патента, программного свидетельства, публикации или значительных уже понесенных расходов не определяет будущую фармакологическую пригодность объекта. В предлагаемой методологии эти сведения входят в разные части досье. Решение о следующем исследовании опирается на его ожидаемую ценность и действующие требования.

2.6 Принципы измерения научной, технологической и организационной результативности

Методология становится проверяемой, когда заранее определено, какой результат считается существенным и с чем он сравнивается. Научная результативность характеризует полученное знание и обоснованность вывода. Технологическая — способность воспроизводимо выполнять исследовательские операции и получать пригодные результаты. Организационная — согласованность работы, затраты ресурсов и условия принятия решений. Эти измерения связаны, но требуют собственных показателей.

Для научной оценки необходимо установить, какие гипотезы получили поддержку, какие отвергнуты и какие остались неопределенными. Число рассмотренных молекул само по себе мало сообщает о таком результате. Полезны сведения о подтвержденных объектах, устойчивости эффектов и полноте проверки альтернативных объяснений. Отдельно фиксируются ситуации, когда полученная информация позволила обоснованно прекратить бесперспективную работу.

Технологическая оценка включает воспроизводимость расчетов и измерений, полноту цифровой истории, длительность согласованного исследовательского маршрута и долю работ, прошедших контроль качества. Для каждого показателя определяются начало отсчета, единица анализа и правила учета повторов. Сокращение длительности операции следует интерпретировать вместе с изменением общего маршрута и достигнутого объема подтверждения.

Организационные показатели относятся к передаче заданий, ожиданию, доступности специалистов и причинам ручного вмешательства. Уменьшение времени согласования может представлять полезный эффект, если качество решения сохраняется. Снижение числа вмешательств само по себе неоднозначно: оно может означать как понятную автоматизацию, так и недостаточный экспертный контроль. Поэтому количественный показатель должен сопровождаться содержанием соответствующих событий.

Современная критическая оценка ИИ в разработке лекарств обращает внимание на необходимость связывать вычислительные достижения с улучшением решений. Bender и соавторы рассматривают разрыв между проверкой модели и ее прикладной ценностью. Для Atlansys EUS MedPharma из этого следует требование включать в программу оценки конечную исследовательскую задачу и доступные альтернативы, а не ограничиваться сравнением алгоритмов [3].

Ранее предложенное сравнение трех процессов получает здесь операционное содержание. Вариант А описывает действующий процесс организации. В варианте Б используются выбранные ИИ-инструменты с существующим порядком взаимодействия специалистов. Вариант В объединяет те же классы инструментов и согласованные правила работы в интегрированной платформе. Протокол должен точно описать, какие элементы изменяются в каждом варианте и какие условия остаются сопоставимыми.

Всем вариантам задаются общий профиль требований, сопоставимый исходный доступ к данным и согласованный ресурсный бюджет. Критерии обновления моделей и выбора опытов описываются до начала оценки. По мере выполнения исследования наборы полученных данных могут различаться, поскольку политика отбора влияет на выполняемые эксперименты. Это различие учитывается как часть сравниваемого процесса и сохраняется в истории его версий.

Сопоставление А и Б характеризует изменение, связанное с применением выбранных инструментов в определенных условиях. Сопоставление Б и В направлено на оценку добавочного эффекта интеграции. Различие А и В отражает совокупное изменение. Причинная интерпретация требует обоснованного дизайна и контроля иных различий между вариантами. Схема разделения эффектов приведена на рисунке 2.2.



Рисунок 2.2. Сравнение процессов для оценки вклада инструментов и их интеграции.

По возможности задания или сопоставимые исследовательские единицы распределяются между вариантами до получения оцениваемых ответов. Экспериментальные серии балансируются по существенным условиям, а образцы могут кодироваться для уменьшения влияния ожиданий исполнителя. Если распределение выполняется на уровне проектов, статистическая оценка должна учитывать зависимость результатов внутри проекта. Большое число расчетов одной серии не заменяет множество независимых проверок.

Контрольная процедура подтверждения должна быть общей для сравниваемых вариантов. Иначе один процесс может оказаться быстрее из-за менее полного исследования. До начала работы устанавливаются правила разрешения пограничных и противоречивых результатов, а также учета не полученных образцов и незавершенных опытов. Изменения протокола сохраняются отдельно, чтобы их последствия можно было проверить.

Результаты поперечного исследования Druedahl и соавторов описывают сообщавшееся применение ИИ в разработке препаратов по данным Pharmaprojects. Такой дизайн помогает характеризовать распространение подходов и типы задач. Он не воспроизводит прямое сравнение одинаковых проектов с ИИ и без него. Поэтому сведения об использовании технологии следует отделять от доказательства ее причинного влияния на срок или выход разработки [40].

Для затрат задается согласованная граница учета. Она включает подготовку данных, вычисления, труд специалистов, синтез, аналитику, биологические испытания и относимую к проекту эксплуатацию платформы. Учитываются неуспешные попытки и повторные работы. Затраты прошлых периодов, общая инфраструктура и стоимость партнерских услуг отражаются по заранее описанному правилу, одинаковому для сопоставляемых вариантов.

В выражении (2.3) Сполная обозначает затраты в установленных границах, а Nподтв — число объектов одной согласованной категории, подтвердивших требования в общей процедуре. Показатель не смешивает первичные активные соединения и кандидаты с разным объемом исследования. При нулевом выходе значение отношения не определяется; в отчете указываются понесенные затраты и отсутствие подтвержденного результата.

$$C_{объект} = \frac{C_{полная}}{N_{подтв}} \quad (2.3)$$

Временные показатели требуют учета незавершенных работ. Если к моменту анализа часть проектов продолжает исследование, их нельзя автоматически исключать или считать завершенными неуспешно. Протокол должен задавать срок наблюдения и способ описания таких случаев. Медианное время до события, доля завершенных проверок в установленный период и ресурсные затраты отвечают на разные вопросы и представляются раздельно.

Таблица 2.3. Основные показатели проверки сквозного управления.

Показатель	Что измеряется	Правило сопоставления
Подтвержденный выход	Число объектов, прошедших установленную проверку	Единая категория результата и общие критерии
Время до результата	Календарное время до согласованного события	Общее начало отсчета и учет незавершенных работ
Полные затраты	Ресурсы в установленных границах учета	Включение неуспехов, повторов и подготовки
Стоимость объекта	Полные затраты на подтвержденный объект	При нулевом выходе отношение не определяется
Качество прогноза	Ошибка и калибровка в заданном контексте	Независимые данные и зафиксированная версия
Воспроизводимость	Устойчивость результата в повторной проверке	Предварительно заданная процедура и ее полнота
Прослеживаемость	Полнота связей данных и оснований решения	Единый перечень обязательных сведений

Для оценки различий приводятся неопределенность и устойчивость результата, а не только одно среднее значение. Проверяется влияние сложности химического пространства, доступности исходных веществ, состава лабораторных мощностей и опыта исполнителей. При небольшом числе проектов целесообразно ограничить выводы наблюдаемыми условиями и дополнить количественную оценку разбором причин изменений.

Экономический эффект пилота также необходимо отличать от прогноза его масштабирования. Обзоры, в том числе работа Wu и соавторов, объединяют широкий круг направлений применения ИИ — от скрининга до клинических исследований. Такие обзоры помогают опре-

делить возможные области использования. Для расчета эффекта определенного проекта требуются его собственные данные, единый базовый процесс и согласованная методика учета [41].

Воспроизводимый отчет должен позволять восстановить требования, состав исследовательских единиц, доступные на момент выбора данные, версии инструментов и выполненные проверки. Его выводы связываются с конкретной конфигурацией процесса. При изменении модели, лаборатории или целевой задачи устанавливается, какая часть подтверждения сохраняет силу и какую нужно получить заново.

Предлагаемая методология тем самым задает требования к архитектуре Atlansys EUS MedPharma. Платформа должна связывать версии целевого профиля и гипотез с данными, вычислительными заданиями, материальными образцами и контрольными решениями. Общие службы управления данными, отчетности и безопасности обеспечивают сохранность этих связей. В следующей главе данная рамка будет раскрыта через распределение функций между системами, подсистемами и сервисами.

Выводы

Сквозное управление ранней разработкой основано на согласовании исследовательских действий с целевыми требованиями. Профиль будущего продукта определяет медицинский замысел, а требования к кандидату переводят его в проверяемые условия выбора. Для каждого условия задаются показатель, контекст наблюдения, критерий интерпретации и история изменений.

Паспорт гипотезы связывает исходные основания с альтернативными объяснениями, программой опытов и возможными решениями. Завершение итерации означает определенное изменение состояния знания и плана действий. Вычислительный прогноз, измерение, подтверждение и решение сохраняются как различающиеся информационные объекты, что делает переходы между ними прослеживаемыми.

Данные, физические модели, машинное обучение и экспертное участие выполняют взаимодополняющие функции. Их согласование требует учета зависимостей между источниками, области применимости прогнозов и надежности оцененной неопределенности. Экспертное вмешательство становится частью воспроизводимого процесса, когда его предмет, основание и последствия зафиксированы.

Конец ознакомительного фрагмента.

Текст предоставлен ООО «Литрес».

Прочитайте эту книгу целиком, купив полную легальную версию на Литрес.

Безопасно оплатить книгу можно банковской картой Visa, MasterCard, Maestro, со счета мобильного телефона, с платежного терминала, в салоне МТС или Связной, через PayPal, WebMoney, Яндекс.Деньги, QIWI Кошелек, бонусными картами или другим удобным Вам способом.