

Пособие: Старший руководитель проектов (Data Science & Human Research)

Введение

Роль старшего руководителя проектов в области Data Science и исследований с участием людей объединяет навыки классического **Project Management** с глубоким пониманием специфики научно-исследовательской работы. Проекты, связанные с анализом данных и экспериментами с участием людей, становятся стратегически важными для организаций, и их успешная реализация требует особого подхода. В этом пособии рассмотрены ключевые особенности управления такими проектами, лучшие методологии, реальные кейсы и примеры, а также взаимодействие проектных и продуктовых менеджеров.

Роль руководителя проекта

Иллюстрация: проджект-менеджер управляет исполнением проекта как дирижёр, координируя работу всей команды (разработчики, дата-сайентисты, тимлиды и др.)^[1]. Главная задача руководителя проекта – обеспечить достижение целей проекта в заданные сроки и бюджет. Он отвечает за планирование, координацию и контроль всех этапов выполнения. Проект-менеджер не пишет код и не разрабатывает модели, но знает **кто** и **когда** должен выполнять каждую задачу. Он следит, чтобы команда двигалась по плану от идеи до результата, и фокусируется на соблюдении сроков и координации работы команды^[2]. Фактически ему ставят требуемый результат, а он организует его достижение в срок и в полном объёме^[3]. Руководитель проекта обеспечивает связь между всеми участниками, быстро решает возникающие проблемы и корректирует план, если что-то идёт не так^[4]. Благодаря его контролю бизнес получает готовое решение вовремя и без лишних затрат.

Основные обязанности проектного менеджера охватывают весь жизненный цикл проекта: определение четких целей совместно с заинтересованными сторонами, детальное планирование задач и ресурсов, управление командой и распределение задач, мониторинг прогресса и бюджета, коммуникация с заказчиками и командой, управление рисками и изменениями, мотивация участников и разрешение конфликтных ситуаций^{[5][6]}. Эти задачи требуют не только знаний методологий и техники, но и превосходных организаторских способностей, лидерства и умения принимать решения в условиях неопределенности. В проектах Data Science руководителю важно понимать особенности процесса разработки моделей и анализа данных, чтобы говорить с командой на одном языке и адекватно оценивать состояние проекта.

Специфика проектов Data Science

Проекты в сфере Data Science имеют ряд отличий от классических ИТ-проектов. Во-первых, работа дата-сайентистов носит исследовательский характер, результат не всегда предсказуем, а задачи трудно заранее точно оценить. **Стандартные подходы**

разработки ПО напрямую не подходят для Data Science – создание моделей отличается от написания типового софта, и неправильный подход может привести к проблемам[7]. Успешное управление такими проектами предполагает понимание того, что требуется **собственная методология и жизненный цикл** проекта Data Science[8]. Многие компании ошибочно пытаются применять к анализу данных те же процессы, что и к разработке приложений, однако модели – это другой тип продукта, требующий адаптации процесса управления.

Одно из ключевых отличий – **неопределённость и итеративность**. Data Science-проекты зачастую развиваются нелинейно: исследования могут потребовать возврата к более ранним этапам, пересмотра гипотез или сбора дополнительных данных. Это нормальная ситуация: практически все проекты проходят через итерации, и менеджеру важно изначально заложить цикл обратной связи[9]. Например, модель, находясь на этапе валидации, может обнаружить проблему в данных, и тогда команде приходится вернуться к этапу подготовки данных – это не сбой, а естественный процесс улучшения решения. Опыт показывает, что успешные команды **принимают итерации как часть процесса**, но при этом стараются не допустить бесконтрольного затягивания сроков из-за повторных циклов[9].

Еще одна особенность – **сложность планирования и оценок**. Дата-сайентисты редко могут дать точный прогноз по времени: их задачи включают анализ и ресерч, результаты которого неизвестны заранее. В одном из кейсов отмечалось отсутствие точных оценок задач – специалисты по данным не могли назвать количество часов, необходимых на решение, поскольку задачи исследовательские и непредсказуемые[10]. Руководителю проекта приходится мириться с определенной долей неопределенности и искать способы сделать планирование гибким. Например, можно договориться о лимитах и декомпозировать большие задачи на мелкие подзадачи с оценкой не более 20 часов каждая – такой подход сделал сроки более управляемыми даже при неточных оценках[11].

Распределение времени и ресурсов в DS-проектах тоже существенно отличается. Известно, что основная часть усилий уходит на подготовку данных: очистку, сбор, исследование данных. Этот этап может занимать **до 70–90% от всего времени проекта**[12]. Качество данных напрямую влияет на успех модели, поэтому менеджеру нужно закладывать достаточные резервы времени на data engineering и data cleaning. Неправильная оценка этого этапа – типичная причина срывов: если не учесть длительность подготовки данных, проект неизбежно выбьется из графика. Кроме того, после внедрения модели работа не заканчивается – **Data Science-проект никогда не бывает окончательно завершён**. Модель в продуктиве требует мониторинга качества, может появляться необходимость дообучения на новых данных или доработки под изменившиеся бизнес-требования[13]. Таким образом, цикл проекта Data Science часто повторяется: после этапа **Deployment** начинается сбор новых данных и улучшение решения, что превращает проект в итеративный процесс постоянного развития, а не разовый акт создания продукта.

Наконец, **команда DS-проекта** обычно междисциплинарная: в неё входят специалисты по анализу данных, ML-инженеры, разработчики программного обеспечения, аналитики

и т.д. Руководитель проекта играет роль связующего звена между разными ролями. Ему приходится разбираться в терминологии и подходах разных областей – от бизнес-анализа до машинного обучения – чтобы эффективно координировать всех. В реальной практике проджект-менеджеры отмечают, что **дата-сайентисты – не совсем обычные разработчики**, и привычные процессы разработки на них напрямую не натянешь[14]. Например, менеджеру, ранее работавшему только с software engineers, пришлось учиться понимать рабочие подходы DS-специалистов, иначе он не мог качественно контролировать результаты их работы[15]. Часто помогает вовлечение технического лида или наставника, который «переводит» сложные аспекты машинного обучения на понятный язык для менеджера[15]. Таким образом, успех DS-проекта во многом зависит от способности руководителя погрузиться в предметную область и говорить с командой data scientists на одном языке.

Проекты с исследованиями на людях (Human Research)

Если проект включает **исследования с участием людей** (например, эксперименты с пользователями, клинические испытания, опросы или сбор данных у добровольцев), к управлению добавляются особые требования. Руководитель проекта в этом случае несет ответственность не только за техническую часть, но и за соблюдение этических и юридических норм проведения исследований. Необходимо обеспечить защиту данных участников, конфиденциальность и получить все требуемые разрешения перед началом работ.

Практически это выражается в дополнительных задачах менеджера. Он должен планировать и вести исследования подобно отдельным проектам внутри основного проекта. **Подготовка и запуск исследования** с людьми включает разработку протокола, бюджетирование, взаимодействие с регулирующими органами. Руководитель проекта организует получение одобрения этических комитетов (IRB) и других регуляторов, готовит необходимые контракты и согласования, продумывает логистику (набор участников, расписание сессий, сбор данных) и контролирует качество получаемых данных[16]. Во многих случаях требуется нанимать или управлять командой координаторов исследований, которые занимаются практической работой с участниками (набор, проведение сессий, сбор анкет и т.д.) – менеджер проекта руководит их работой и масштабирует эти возможности по мере роста числа исследований[17]. Также он налаживает сотрудничество с внешними партнерами: научно-исследовательскими институтами, контрактными организациями (CRO) или другими компаниями, которые помогают провести исследование на высоком уровне[17].

Важная часть такой работы – **координация науки и инжиниринга**. Руководитель проекта служит мостом между исследователями (учеными, врачами, аналитиками) и инженерами (разработчиками продукта, дата-инженерами). Совместно с научными специалистами и командой разработки он создает инструменты для сбора данных, протоколы экспериментов, формы анкет и обеспечивает надежное документирование всех этапов исследования[18]. При этом менеджер заботится о том, чтобы ускорять проведение исследований без потери научной строгости: внедряет **стандартизированные процессы и шаблоны**, которые позволяют команде работать

быстрее, но сохранять высокое качество данных[19]. Например, можно автоматизировать часть рутинных операций по работе с участниками (электронные информированные согласия, напоминания о визитах), что уменьшит вероятность ошибок и ускорит сбор информации.

Особое внимание следует уделять **этике и безопасности участников**. Руководитель проекта обязан строго соблюдать протоколы, гарантирующие, что ни один человек не подвергнется необоснованному риску в ходе эксперимента. Перед началом исследований необходимо получить информированное согласие от каждого участника, разъяснить цели и процедуру. На протяжении проекта менеджер отслеживает, чтобы данные участников хранились безопасно, обезличивались и использовались только в рамках оговоренных целей. В случае клинических или психологических исследований он также планирует меры на случай непредвиденных обстоятельств (например, если участнику станет плохо). Таким образом, управление проектом с человеческими испытуемыми сочетает в себе черты проектного менеджмента и ответственности исследователя-координатора по стандартам **Good Clinical Practice (GCP)** и другим нормам. В конечном счёте успешный менеджер обеспечивает, чтобы исследования были проведены эффективно, этично и привели к ценным результатам для проекта.

Методологии управления проектами Data Science

Учитывая перечисленные особенности, для Data Science-проектов разрабатываются специальные методологические подходы. На практике команды часто комбинируют элементы Agile (гибких методологий) с этапами, характерными для научного исследования. Существуют признанные стандарты, помогающие структурировать работу с данными. **Самой популярной методологией** выполнения проектов анализа данных является стандарт **CRISP-DM** (Cross-Industry Standard Process for Data Mining)[20]. Этот подход был разработан еще в конце 1990-х и зарекомендовал себя как независимый от предметной области универсальный процесс решения задач Data Science[21]. Согласно определению, **CRISP-DM описывает общие процессы и подходы к аналитике данных**, применимые в дата-майнинг проектах независимо от конкретной индустрии или типа задачи[21]. Опросы показывают, что **CRISP-DM регулярно занимает первое место** по популярности среди методологий анализа данных, опережая другие варианты (SEMMA, KDD и др.)[20].

График: популярность методологий анализа данных по опросу KDnuggets. CRISP-DM доминирует среди подходов к Data Science-проектам (43% респондентов в 2014 г.), опережая собственные методы компаний и альтернативы вроде SEMMA и KDD[20]. Популярность CRISP-DM обусловлена его полнотой и детальной проработкой шагов; для сравнения, KDD считается более общей методикой, а SEMMA – скорее последовательностью действий в инструменте SAS Enterprise Miner[22].

Методология **CRISP-DM** разбивает выполнение проекта на **шесть этапов** с чётко определенными целями на каждом этапе. Эти основные фазы выполняются последовательно, но с возможностью возврата при необходимости. Шаги CRISP-DM следующие[23]:

- **Бизнес-анализ (Business Understanding)** – формулировка целей проекта с точки зрения бизнеса, определение вопросов, на которые нужно ответить, и критериев успеха.
- **Анализ данных (Data Understanding)** – сбор исходных данных, их изучение, выявление проблем с качеством, формирование первоначальных гипотез.
- **Подготовка данных (Data Preparation)** – очистка и обработка данных, создание необходимого набора для моделирования (интеграция разнородных источников, отбор признаков и пр.).
- **Моделирование (Modeling)** – выбор и применение моделей машинного обучения/аналитики, настройка алгоритмов на подготовленных данных.
- **Оценка результатов (Evaluation)** – оценка качества модели и результата с точки зрения изначальной бизнес-задачи, проверка, удовлетворяет ли решение критериям успеха, и принятие решения о готовности к внедрению.
- **Внедрение (Deployment)** – развертывание полученного решения в реальной среде, передача его пользователям или заказчику, организация использования модели и сбор обратной связи.

Рисунок: циклическая модель CRISP-DM. Шесть фаз жизненного цикла проекта аналитики данных – от понимания бизнес-задачи до деплоя модели – образуют итеративный процесс[23]. Методология гибкая: допускаются возвраты к предыдущим этапам или пропуски некоторых шагов, если они не актуальны для проекта[24]. Например, после оценки модели может выясниться, что нужны дополнительные данные – тогда команда возвращается к этапу Data Understanding и обновляет выборку. Или для простого проекта можно пропустить отдельные стадии, сконцентрировавшись на ключевых шагах. CRISP-DM подчеркивает, что такой **негосударственный** (нежесткий) подход обеспечивает адаптируемость под особенности каждого проекта.

Отличительные черты CRISP-DM – **ориентация на бизнес-цели и тщательное документирование**. С самого начала большое внимание уделяется пониманию того, **зачем** этот проект нужен компании и какую ценность принесёт решение. Авторы методологии отмечают, что это помогает руководству воспринимать проект анализа данных не как эксперименты «в песочнице», а как полноценный элемент бизнес-процессов[25]. Второй важный момент – подробная документация на каждом шаге: после выполнения каждой фазы фиксируются полученные результаты, выводы, принятые решения. Это делает процесс прозрачным для менеджмента и позволяет дата-сайентистам влиять на решения, опираясь на зафиксированные факты[25]. В итоге проект становится более управляемым, а его ценность понятна всем участникам.

Помимо CRISP-DM, существуют и другие методологии (KDD, SEMMA), но они, как уже отмечалось, менее распространены. Многие команды разрабатывают **собственные гибридные процессы**, особенно если компания имеет наработки в области Agile. **Agile-подходы** – Scrum, Kanban – также применимы в проектах Data Science, хотя и требуют адаптации. Например, можно вести бэклог задач и итерации спринтов, но спринт-планирование будет учитывать исследовательские спайки и эксперименты. Руководителю проекта полезно знать классические стандарты управления (PMBOK) и современные гибкие методологии, чтобы комбинировать их инструменты под нужды

команды[26]. На практике подход часто бывает таким: **первый спринт** посвящается бизнес-анализу и подготовке данных, **второй** – экспериментам с моделями, **третий** – внедрению и тестированию, затем цикл повторяется. Главное – сохранять ориентир на ценность для бизнеса и регулярно демонстрировать промежуточные результаты стейкхолдерам.

Независимо от методологии, есть общие **принципы успеха DS-проекта**:

- **Problem First:** всегда начинайте с четкого определения бизнес-проблемы, которую нужно решить, и ожидаемой ценности. Не стоит брать данные и «искать в них что-нибудь интересное» без ясной цели – проект должен быть привязан к конкретной задаче бизнеса[27].
- **Итеративное развитие:** планируйте проект циклами и регулярно пересматривайте гипотезы. Метрики успеха и даже сами цели могут уточняться по ходу работы – это нормально. Важно, чтобы заказчики понимали эту особенность Data Science. Хорошей практикой считается заранее договориться о допущениях на случай неудачи модели и включать пункт о возможных доработках (iterative refinement) в план проекта.
- **Оценка выгод:** еще на старте полезно прикинуть *порядок величины* потенциального эффекта от проекта. Даже если точный ROI посчитать сложно, попытка оценить, например, «сколько компания сэкономит, если повысим точность прогноза на 5%» поможет расставить приоритеты[28][29]. Такой приблизительный расчет пользы и издержек дает ориентир и мотивирует команду и бизнес.
- **Управление рисками и качеством данных:** идентифицируйте типичные риски Data Science-проекта – нехватка или плохое качество данных, отсутствие поддержки со стороны заказчика, слишком высокий барьер для внедрения модели в бизнес-процесс – и подготовьте план их снижения. Например, заранее проверьте доступность необходимых источников данных и вовлеченность бизнес-экспертов[30][31]. Следите за качеством данных на этапе подготовки – лучше потратить время на тщательную проверку, чем получить модель, не несущую ценности из-за грязных исходных данных.
- **Прозрачность и коммуникация:** регулярно информируйте всех стейкхолдеров о прогрессе, находках и трудностях. Демонстрируйте промежуточные результаты (прототип модели, выводы разведочного анализа) на понятном языке. Это удерживает доверие и помогает скорректировать цели, если выясняются новые обстоятельства. Некоторые компании проводят обучающие сессии для бизнес-команд, объясняя, почему в DS много итераций, и тем самым вырабатывают общее понимание процесса[32][33].

Взаимодействие проектного и продуктового менеджера

В крупных инициативах, особенно по созданию новых **data-driven продуктов**, в работе участвуют как **Project Manager**, так и **Product Manager**. Несмотря на схожесть названий, эти роли различны и дополняют друг друга. **Продуктовый менеджер** (Product Manager) отвечает за стратегическое развитие продукта, формирование видения и ценности для

пользователей, определение того, *что* должно быть создано. **Проектный менеджер** (Project Manager) отвечает за операционное выполнение конкретного проекта, фокусируется на том, *как* и *когда* это будет сделано[34]. Говоря образно, продакт – это **стратег**, смотрящий на долгосрочное развитие продукта, а проджект – **тактик**, обеспечивающий реализацию поставленных задач вовремя и в нужном объеме[34].

Проджект-менеджер получает от продакт-менеджера требования и конечную цель (результат), после чего организует работу команды для ее достижения к определенному сроку[3]. Его зона ответственности – план, сроки, ресурсы, риски, процесс. Он следит, чтобы все необходимые шаги проекта были выполнены, и команда пришла к намеченной цели в срок[2]. В то время как продакт-менеджер концентрируется на **ценности и качестве продукта** для пользователя и бизнеса. У продуктового менеджера часто нет жестких дедлайнов – его цель сделать продукт востребованным, улучшать его итерационно. Он формулирует задачи и приоритеты для развития продукта, основываясь на исследованиях рынка и потребностей аудитории[35]. В конце концов именно продакт отвечает на вопрос, какую пользу принес продукт пользователям и компании[35], а проджект – как эффективно и своевременно выпустить запланированный функционал.

На практике эти специалисты работают **рука об руку**. Продукт-менеджер определяет **что** должно быть реализовано (фичи, модели, эксперименты), а проектный менеджер планирует **как** это реализовать и приводит команду к результату. Для одного продукта может выполняться несколько проектов (например, последовательность версий или отдельных модулей), поэтому продакт и проджект постоянно взаимодействуют[36]. Продуктовый менеджер задает приоритеты и принимает ключевые продуктовые решения, а проектный менеджер предоставляет ему информацию о прогрессе, ограничениях и рисках реализации. Проектный менеджер также координирует несколько команд или потоков работ, чтобы все части продукта сошлись к релизу.

Рассмотрим пример: компания решает внедрить на своем сервисе систему рекомендаций на основе машинного обучения. **Product Manager** изучает пользовательские потребности, конкурентов и формулирует, какую метрику необходимо улучшить (скажем, вовлеченность или конверсию продаж) с помощью новой модели. Он же определяет, что успешность решения будет оцениваться, например, ростом времени, проведенного на сайте пользователями, на 15%. **Project Manager**, получив эту цель, организует проект по разработке и интеграции рекомендательной системы: набирает команду data science, планирует этапы (сбор данных, разработка модели, A/B-тест, внедрение), устанавливает сроки и координирует задачи между дата-сайентистами, инженерами бэкенда и командами фронтенда для вывода рекомендаций на сайте. В ходе работы продуктовый менеджер участвует в принятии решений – какие именно элементы рекомендаций важны пользователю, какие сценарии тестировать – а проектный менеджер следит, чтобы техническая реализация этих решений шла по графику и укладывалась в ресурсы.

В реальных командах границы ролей могут размываться, но **цели у них разные**. Продакт-менеджер ориентирован на конечный **продукт и пользователя**, проджект-менеджер – на **проект и процесс**. Например, в кейсе Lamoda, посвященном внедрению

ML-модели ценообразования, над одним направлением работали одновременно один проджект-менеджер и два продакт-менеджера[37]. При расширении команды до 40+ человек потребовалось присутствие нескольких менеджеров продукта, каждый из которых отвечал за свой участок развития продукта, тогда как проджект-менеджер единолично координировал все параллельные потоки разработки и исследований[37]. Такой раздел труда позволил масштабировать проект: продуктовые менеджеры фокусировались на функциональности и эффективности ML-решения для бизнеса, а руководитель проекта выстраивал процессы в увеличившейся команде, чтобы выполнение шло синхронно и с сохранением общего контекста.

Взаимодействие продакта и проджекта строится на постоянном обмене информацией. Проектный менеджер регулярно предоставляет продукту данные о статусе проекта, выявленных рисках, ресурсных ограничениях – чтобы продакт мог скорректировать ожидания или требования. Со стороны продукта менеджер обеспечивает проект приоритетами и видением конечного результата, участвует в приемке результатов каждого этапа. Такая связка помогает, с одной стороны, держать проект в русле бизнес-целей, а с другой – реальность выполнения согласовать с бизнес-планами. В итоге и продукт достигает нужных метрик, и проект укладывается в рамки.

Кейс: управление Data Science-проектом (пример Lamoda Tech)

Рассмотрим реальный пример из практики компании Lamoda Tech, где руководитель проекта внедрял процесс разработки ML-модели в быстро растущей команде[1]. Проект касался создания внутреннего продукта – модели машинного обучения для оптимизации цен. **Исходные условия:** небольшая кросс-функциональная команда (6 человек), включающая 3 дата-сайентиста, тимлида, одного Go-разработчика и продуктового менеджера[38]. Сроки стояли сжатые, и перед проджект-менеджером стояла задача наладить процессы с нуля. При этом это был его первый опыт работы с Data Science-командой, и вскоре выяснилось, что методы, привычные для разработки ПО, не работают напрямую – специфика работы DS-специалистов потребовала новых подходов[14].

В начальной фазе менеджер столкнулся с несколькими **ключевыми проблемами**, которые приходилось решать на ходу:

- **Проблема:** отсутствие точных оценок по задачам – дата-сайентисты не могли спрогнозировать время, поскольку их задачи носили исследовательский характер и были непредсказуемы[10]. **Решение:** договориться о лимите времени на одну задачу и максимально декомпозировать крупные задачи. Команда приняла правило: каждая подзадача не должна оцениваться более чем в 20 часов. Вместо тщетных попыток получить точные числа решили смириться с неопределённостью и управлять ею – дробили работу на мелкие части и ограничивали время на анализ каждой частью. Это сделало прогнозирование более **управляемым** и позволило контролировать прогресс по коротким этапам[11].
- **Проблема:** слишком длительные общие обсуждения – ежедневные командные созвоны (стендапы) превращались в часовые дискуссии по техническим

вопросам[39]. **Решение:** развести форматы встреч по целям. Выяснилось, что разработчики пытались решать сложные вопросы прямо на утренних стендапах, отчего те затягивались. Менеджер установил чёткую цель стендапа – синхронизация статусов – и вынес все обсуждения в отдельные **брейншторм-встречи**. Брейнштормы стали проводить 2 раза в неделю, заранее готовя список вопросов по текущим задачам[40]. Это позволило сократить ежедневные собрания до необходимых 15 минут, а сложные технические проблемы решать адресно на брейнштормах. Со временем количество вопросов снизилось, команда стала многие проблемы решать самостоятельно без вовлечения всех участников, и регулярные брейншторм-сессии вскоре вообще перестали требоваться[41].

- **Проблема:** руководителю проекта трудно контролировать результаты работы, не обладая знаниями в области ML – он не понимал деталей, поскольку впервые работал с Machine Learning[42]. **Решение:** погружение в предметную область при помощи команды. Менеджер стал активно изучать основы ML и попросил тимлида выступать в роли наставника, переводящего сложные термины на понятный язык. Без такого погружения было невозможно качественно управлять проектом[15], поэтому часть своего времени РМ тратил на самообразование и общение с дата-сайентистами об их работе. В результате через некоторое время он уже лучше понимал прогресс по моделям и мог отслеживать качество результатов, не упуская важных деталей.
- **Проблема:** единственный backend-разработчик (Go) чувствовал себя изолированным – в команде не было других специалистов его профиля, с кем он мог бы посоветоваться, оценить задачи или провести code review[43]. **Решение:** интеграция разработчика в другое подразделение на часть времени. Менеджер договорился, что Go-разработчик будет участвовать во встречах и процессах команды разработки другого стрима, где были коллеги с тем же стеком[44]. Там он получал профессиональное общение, помощь в обзоре кода и не чувствовал себя одиноко. Одновременно на основном проекте для синхронизации его подключали к общим созвонам с DS-командой дважды в неделю. Этот шаг помог сохранить мотивацию разработчика и качество его работы, несмотря на то что в core-команде он был единственным.

Спустя некоторое время описанные меры дали результат: команда успешно создала **MVP модели**, провела A/B-тестирование и перешла к полноценной разработке решения. Эффективность процессов возросла, и руководство расширило инвестиции в проект. **Команда начала быстро расти** – из 6 человек превратилась в группу из **40 сотрудников** примерно за два года[45]. Новый расширенный состав включал три подкоманды: команду дата-сайентистов, две команды разработчиков (каждая со своим тимлидом), а также ряд привлеченных специалистов (аналитики, инженеры больших данных, дизайнеры и др.) по необходимости[37]. Появились также дополнительные **продуктовые менеджеры** – двое, вместе с одним проектным менеджером, стали работать над этим направлением[46].

Такой быстрый рост принес новые вызовы. Процессы, которые отлично работали в коллективе до 10 человек, с трудом масштабировались на 40 человек[47]. Увеличилось число новичков, им требовался онбординг; ежедневные синхронизации опять стали затягиваться; общая картина стала распадаться на фрагменты по командам. Руководителю проекта пришлось перестраивать систему управления под новый масштаб. **Решением** стало разделение большой группы на несколько команд и введение новых структур коммуникации. Был нанят второй тимлид, и разработчиков разбили на две команды, распределив людей по навыкам и задачам так, чтобы каждая команда сфокусировалась на своём блоке проекта[48]. Дата-сайентисты остались отдельной командой аналитиков. Теперь уже **Project Manager** координировал работу трёх команд через тимлидов, а для поддержки продукта работали два **Product Manager**. Такой шаг значительно упростил управление людьми и позволил каждой группе сосредоточиться на своих задачах, ускорив разработку[48].

Рисунок: структура расширенной команды ML-проекта после масштабирования. В проекте Lamoda Tech, после роста до 40 человек, команда была реорганизована: выделены две команды разработки (Dev Team 1 и 2, каждая со своим тимлидом) и одна команда Data Science. Над всем направлением работали один проджект-менеджер (PM) и два продакт-менеджера, координирующие развитие продукта[37][48]. Такая структура позволила распределить зоны ответственности и сохранить управляемость проекта при его расширении.

Чтобы большие команды не теряли связь друг с другом и с общей целью, менеджер внедрил **новые процессы взаимодействия**. Были организованы регулярные *общие демо-сессии*, где каждая команда показывала результаты своей работы – это проходило раз в квартал и позволяло всем участникам понимать общий прогресс[49]. Также ввели контроль ключевых метрик эффективности команд (например, Cycle Time, показатель недооценки задач) с ежемесячным обсуждением – это помогло выявлять узкие места и стимулировать команды работать еще лучше[50]. Многие рутинные процедуры автоматизировали: настроили ботов в рабочих чатах, которые напоминали о логировании времени, о необходимости код-ревью для зависших задач и т.п.[50]. Для планирования сделали **сводный план-график (диаграмму Ганта)** по всему направлению и отдельные детальные планы для каждой команды с указанием нагрузки на каждого исполнителя[51]. Раз в квартал начали проводить общее *kick-off*-совещание всей большой команды, на котором обсуждали цели на предстоящий период и ключевые фокусы работы[51]. Кроме того, установили еженедельные *синки* (координационные встречи) между командами разработки и DS, где представители команд обменивались статусами задач и совместно решали проблемы интеграции[52]. Тимлиды разных команд тоже стали собираться на регулярной основе раз в неделю, чтобы синхронизироваться между собой. Все эти меры позволили поддерживать единый контекст в проекте и не терять скорость развития даже при работе большого коллектива. В итоге проект успешно вышел на промышленную эксплуатацию: модель была внедрена, бизнес-показатели улучшились, а выстроенные процессы продолжили служить основой для новых Data Science-проектов компании.

Заключение

Старший руководитель проектов в сфере Data Science & Human Research – это специалист, находящийся на стыке управления, технологий и науки. Ему необходимо обеспечивать эффективную работу команды в условиях высокой неопределенности, совмещать строгие сроки с исследовательской гибкостью и гарантировать, что результаты проекта приносят реальную пользу бизнесу. В таких проектах успех во многом зависит от умения менеджера адаптировать классические методики под новые реалии: принять итеративность процесса, уделять особое внимание качеству данных и методологии анализа, выстраивать коммуникации между разнородными специалистами.

Ключевые рекомендации для руководителя DS-проектов:

- **Фокус на цель и ценность.** Всегда начинайте с четкого определения бизнес-целей проекта и критериев успеха. Держите эти цели в центре внимания команды, чтобы даже в длительном исследовании все понимали, к какому эффекту стремятся.
- **Гибкое планирование.** Планируйте работу итерациями, закладывая время на исследования и возможные изменения направления. Будьте готовы пересматривать план по мере появления новых данных или результатов экспериментов. Управляйте неопределенностью через короткие циклы и регулярные контрольные точки.
- **Коммуникация и прозрачность.** Постоянно информируйте стейкхолдеров о ходе проекта, важных находках и рисках. Объясняйте сложные технические моменты простым языком, демонстрируйте промежуточные результаты (прототипы, отчеты анализа) бизнес-заказчикам. Это формирует доверие и позволяет избежать разочарований из-за завышенных ожиданий.
- **Совместная работа с продуктовой командой.** Тесно сотрудничайте с продакт-менеджерами или заказчиками продукта. Понимайте их видение ценности, помогайте им понять ограничения и возможности технологий. Совместная приоритизация задач и регулярная синхронизация с точки зрения продукта гарантируют, что команда тратит усилия на наиболее важные направления.
- **Управление данными и качеством.** Уделяйте первостепенное внимание качеству данных: заранее проверяйте доступность нужных источников, при необходимости планируйте проекты по сбору/очистке данных. В ходе работы внедряйте метрики качества данных и модели, отслеживайте их. Также готовьте план действий на случай, если данных окажется недостаточно или модель не достигнет требуемых показателей – это поможет избежать ступора, если эксперимент не удастся с первого раза.
- **Этика и комплаенс.** В проектах с человеческими данными или участием людей обеспечьте неукоснительное соблюдение нормативов. Позаботьтесь о конфиденциальности, получите все разрешения, донесите до команды важность этических стандартов. Это не только убережет от юридических рисков, но и повысит доверие к результатам проекта.

- **Постоянное обучение.** Data Science – быстро развивающаяся область. Эффективный РМ постоянно учится: осваивает новые инструменты анализа, вникает в основы машинного обучения, следит за лучшими практиками индустрии. Это позволяет лучше понимать команду и принимать более обоснованные решения. Также важно обучать и команду менеджерским навыкам – например, тому же планированию или оценке эффектов – чтобы все говорили на одном языке.

В заключение, управление проектами на стыке данных и исследований – увлекательная и сложная задача. Каждый такой проект уникален, но опора на проверенные методологии (как CRISP-DM), умелая координация команды и ориентация на ценность для бизнеса помогут довести дело до успешного результата. Старший руководитель проектов, умеющий соединить миры разработки, аналитики и научного эксперимента, становится незаменимым драйвером инноваций в современной организации. Используя изложенные подходы, кейсы и рекомендации, вы сможете уверенно руководить Data Science и Human Research проектами, преодолевать характерные сложности и добиваться впечатляющих результатов.

Источники: Использованы материалы и кейсы из открытых публикаций (Habr, Skillbox, Domino Data Lab и др.), а также практические рекомендации специалистов отрасли для подтверждения изложенных принципов [14][8][10][20] и других.

[1] [10] [11] [14] [15] [37] [38] [39] [40] [41] [42] [43] [44] [45] [46] [47] [48] [49] [50] [51] [52]
Data Science + Разработка =... или Как наладить процессы в растущей кросс-функциональной команде / Хабр

<https://habr.com/ru/companies/lamoda/articles/915126/>

[2] [3] [35] Чем продакт отличается от проджекта: мнение эксперта / Skillbox Media

<https://skillbox.ru/media/management/chem-prodaktmenedzher-otlichается-ot-prodzhektmenedzhera/>

[4] [5] [6] [26] [34] [36] Product manager vs Project manager: кто это такие и в чем разница – GoIT Global

<https://goit.global/ua-ru/articles/product-manager-vs-project-manager-kto-to-takye-y-v-chem-raznytsa/>

[7] [8] [9] [27] [28] [29] [32] [33] Data science project management | Domino Data Lab

<https://domino.ai/resources/field-guide/managing-data-science-projects>

[12] [13] Структура Data Science-проекта с высоты птичьего полета / Хабр

<https://habr.com/ru/articles/468493/>

[16] [17] [18] [19] Senior Research Project Manager Job in San Diego, CA at Oura

<https://www.ziprecruiter.com/c/Oura/Job/Senior-Research-Project-Manager/-in-San-Diego,CA?jid=7e94715c5edf7efa&lvk=QljkbYAxTfjMFML4ENKWUQ.--NvsrAtJAB>

[20] [21] [22] [23] [24] [25] [30] [31] CRISP-DM: проверенная методология для Data Scientist-ов / Хабр

<https://habr.com/ru/companies/lanit/articles/328858/>