

Стадия ЖЦ данных	Объект управления и описания	Описание метаданных	Как использовать
Первичный анализ	задача	описания требований в свободном формате	часто получение требований происходит по телефону или на встречах - но иногда можно в целом на базе протокола встречи написать тз фиксация ожиданий и феномен того что люди воспринимают все по разному и когда что-то фиксируешь письменно появляется автоматический стимул расписать детали
		заказчик и его контакты	а вот это кстати, далеко не всегда очевидный момент, и если в начале не осознать, кто твой конечный заказчик и чем же он все таки с точки зрения бизнеса занимается - можно сделать совсем не то ("испорченный телефон") эта информация пригодится для дальнейшего взаимодействие по sla tla проблемы с данными и уточнение тз
		функциональное подразделение или бизнес процесс пользователь	считаю эти параметры зачастую наиболее эффективными чем понятие предметной области, в будущем это поможет разложить данные на крупные папки создав некий верхний уровень группировки данных
		цель создания или бизнес драйвер	обоснование стоимости и приоритета работ, а так же end to end понимание аналитиком задачи
		условия и требования к обновлению и эксплуатации (tla sla)	capacity management, управление нагрузкой на железо, режим сопровождения, планирование сервисных работ end to end понимание аналитиком задачи понимание клиентом текущих технических ограничений(особенно актуально, если заказчик далек от ИТ)
		историчность и глубина данных	стоимость хранения данных и объем требуемого в будущем места
		критичные моменты в качестве данных	целостность и интероперабельность данных
Аналитика	задача	источники данных	планирование миграции систем, зависимости от ошибок и изменений на источниках, lineage
		использование ранее разработанных dataset	экономия мощностей и интероперабельность, непротиворечивость итоговой отчетности
		soulution архитектура и инструменты	управление производительностью, стандартизация подходов разработки и сопровождения, проведения анализа инфраструктурных ошибок и ошибок внешнего программного обеспечения
	задача -> data set	логическая модель и связи с бизнес терминами и сами термины	продвинутый уровень, эти метаданные увы не всегда типовое явление, но в переписках бизнес в основном оперирует именно методологией и ключевые термины бизнес сущностей просто необходимы в системе для нормального поиска
Прототипирова	data set	физическая модель	контроль изменений, генерация тестовых данных

ние		name convention	структуризация, ускорение процесса написания аналитических запросов, простота интерпретации показателей в таблицах, конечно, она сама по себе не появится - но зачастую имеет место быть как стандарт разработки
		инструменты загрузки	тут могут быть перечислены как etl средства, так и какие то кастомные инструменты, так и просто написанные классы библиотек кода так же это может быть ссылка на некий внутренний справочник - по которому можно почитать пояснения по инструментам, инструкции по эксплуатации и проч
Разработка	data set	идентификатор витрины	наверное, самый сложный вопрос, что такое витрина! пробовала ответить на него в предыдущей статье, но в рамках системы управления - это максимально удобный объект управления
		ссылка на версионный репозиторий хранения кода	тут все понятно, нет разработки - нет данных
		профилирование данных	это редкая история, но суть в том, чтобы понимать на уровне самих данных (а не структур наполнения) паттерны данных в таблицах, это поможет в создании тестовых данных и отслеживание изменений паттернов на источниках
		s2t. и соотв состав описаний и таблиц	ключевой документ описывающий логику преобразования данных , пригодится всегда и везде - чтобы рассказать что у в витрине, от чего она зависит, как рассчитывается показатель, чтобы запланировать миграцию данных, сгенерировать целостные тестовые данные часто в подобных документах так же указывается уровень чувствительности данных
		контакт аналитик	контакты для логики сбора данных и других деталей
		как получить доступ	это удобно как и внутри команды, когда приходят новые люди так и для пользователей. в зависимости от ит уровня зрелости вариантов заполнения поля может быть много от простой инструкции до ссылки на какого нибудь телеграм бота, который задаст вопросы - и выдаст права
	data set-> потоки	дерево или последовательность загрузок	понимание структуры расчета витрин и последовательности потоков загрузки - оно заложено в коде витрины, но визуализации помогает быстро ориентироваться - "где застряли". иногда дерево можно сформировать автоматически из логов завершения потоков
	поток (часто привязан к таблицам, а не к витрине в целом)	название	бизнес название, например, отображающее этапность загрузки, по которому команда может быстро ориентироваться на какой стадии расчета витрина
		актуальность обновления	понимание даты обновления данных говорит потенциальным пользователям, насколько витрина в целом актуальна и можно ли завязываться на эти расчеты, а так же можно понять фактическую частоту расчета . Это техническая информация - которая точно есть в системы - каким загрузчиком вы бы не пользовались, но простота сбора этой информации сильно зависит от количество различных инструментов загрузки и стандартизации подходов.

		статус	потоки могут быть "задизейблены" - обычно это технический статус из управляющего механизма. или, например, витрина может считаться не целиком - это тоже важная информация
Внедрение	релиз	бизнес результат	это сильно перекликается с пунктами выше про общее назначение данных, но в релизе требуется указывать именно инкремент изменение - например, добавили расчет такого-то показателя. кроме экономического обоснования, это может служить важным фактором в сопровождение релиза и витрины в целом
		ответственные команды	информация для сопровождения витрины при возникновение проблем
		среды на которых проводилась отладка - даты отладки	удобно для выпуска новых релизов
		тестовые (функциональные и бизнес) скрипты	обычно бывают крайне удобны для регресс тестирования и просто поиска ошибок, обычно источники также живой организм и эти скрипты точно пригодятся, как и референсные значения по их выполнению на этапе прохождения промышленных испытаний
Эксплуатация	data-set часто в интерпретации набора потоков	TLA и SLA	технические и бизнес условия по уровню эксплуатации сервиса - многие живут без них, но иногда люди не совсем могут выровнять ожидания к тому как потоки должны работать, и как быстро их должны чинить
		Владелец и ответственный за развитие	на стадии вывода потока этих людей не может не быть, другой вопрос - что они могут меняться и как отслеживать эту информацию это еще тот квест
		Потребители	перечень реальных потребителей, получивших доступ, как бизнес пользователи, так и автоматизированные (сюда можно наvertеть всяких рюшечек активность потребителей и тепловые карты, но пора мне остановится)