

DATA GOVERNANCE

ГЕРОЙ ЭПОХИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

■ Практика Data Governance. Выводы. Мысли.



Канат Наурызбаев

Data Governance Expert

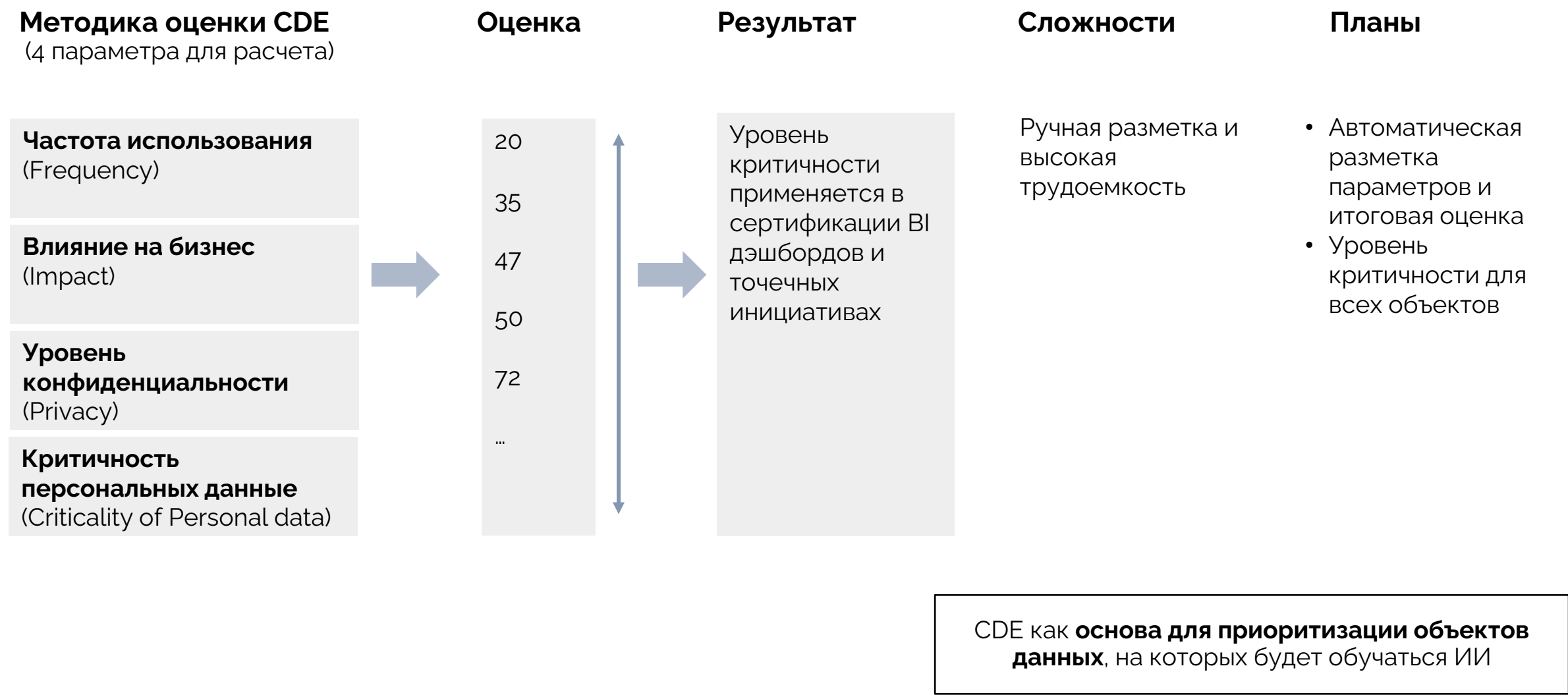
Tele2/Altel, 2025



ALTEL 5G

TELE2 5G

Критические элементы данных (CDE) как основа приоритизации важнейших данных



Бизнес-гlossарий - это не просто набор определений, а база смысла данных

Бизнес-гlossарий – это бизнес определение терминов, расчеты метрик, взаимосвязь с другими элементами, правила использования, источники и т.п.

Работающий подход

- Фокус на разрыве в логике и понимании данных у технарей
- Создание шаблонов и площадки для дискуссии, чтобы вовлечь технарей в предметную область и выровнять понимание с бизнесом (vice versa)
- Сопровождение и методологическая поддержка до момента, пока стороны не выстроили устойчивое взаимодействие
- Формализация договоренности и внесение в единый бизнес-гlossарий
- От фасилитации к самостоятельности. Закрепление культуры и практики

Ценность для людей

Определение терминов



Расчеты метрик

Ценность для ИИ

Централизованная база знаний



Связь бизнес определений с кодом

Бизнес-гlossарий формирует единое понимание терминов и расчетов - основу, на которой ИИ-агенты могут корректно интерпретировать выдавать ответы

Data Stewardship – вызов корпоративной культуре

Работающий подход

- Определяем, что именно делает куратор и зачем
- Вместе с ним погружаемся в бизнес-контекст – важно, чтобы он сам увидел ценность описания данных
- Stewardship by example – не объясняем, а показываем
- Даем инструменты: шаблоны, площадку и консультацию, чтобы работать с данными было просто

Сложности классического кураторство данных

- Конфликт с основными обязанностями сотрудников («вторая работа»)
- Data Stewardship воспринимается как нагрузка без прямой выгоды

Благодаря роли кураторов, четко определенным метрикам качества и прозрачным потокам данных, ИИ-системы получают стабильный и интерпретируемый источник информации, а бизнес – уверенность, что результаты алгоритмов основаны на корректных и осмысленных данных.



DDD (Domain-Driven Design)

DDD — это рабочий подход к **проектированию информационных систем**, при котором бизнес-домен (предметная область) становится основой архитектуры.

КЛЮЧЕВЫЕ ПРИНЦИПЫ DDD:

- Фокус на **предметной области (domain)** — то, как бизнес реально работает.
- **Совместный язык** (Ubiquitous Language) между бизнесом и IT.
- Разделение системы на домены и **bounded contexts** (границы смыслов и ответственности).
- Модели данных строятся **из языка и процессов бизнеса**, а не наоборот.

Domain-Driven DESIGN

Tackling Complexity in the Heart of Software



Eric Evans
Foreword by Martin Fowler

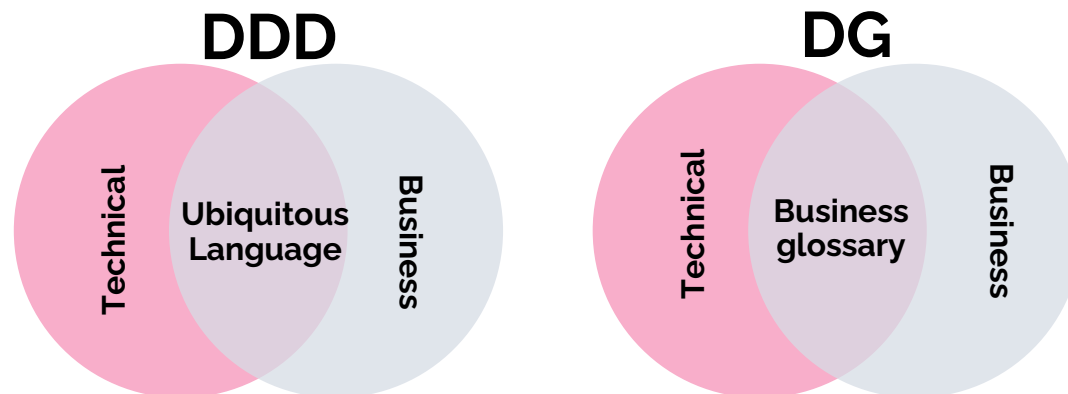
Bounded context - управление границами ответственности и интеграциями

Главные принципы DDD:

- бизнес-домен первичен
- каждая команда отвечает за свой "bounded context" (ограниченный контекст)
- коммуникация между контекстами строго определена

В Data Governance схожие задачи:

- разделить компанию на домены данных (Customer, Product, Network, Finance и др.)
- назначить владельцев и кураторов
- **установить правила взаимодействия данных между доменами**
- обеспечить понятные термины и смысл через глоссарий и каталог



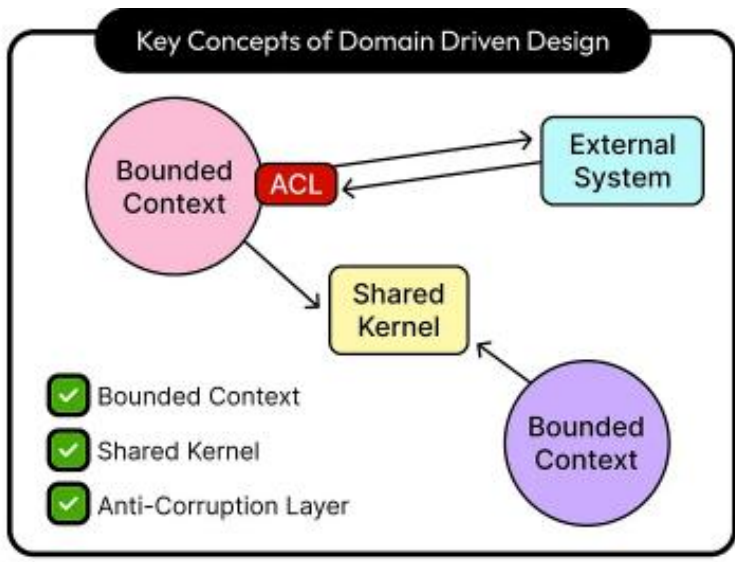
В Data Governance bounded context – это Data Domain

В Data Governance Ubiquitous Language – это Business glossary

Типы связи между контекстами – аналог Data Contracts и политик качества

Архитектура взаимодействия доменов через данные, и как правильно моделировать отношения между ними, чтобы данные были устойчивыми, предсказуемыми и управляемыми.

Тип связи в DDD	Описание	Пример в DG	Риски
Shared Kernel	Оба контекста используют одну и ту же модель данных (или её часть)	Общая таблица/модель используется в двух доменах	высокий (общая ответственность)
Conformist	Потребитель данных принимает модель поставщика как есть	Потребитель следует правилам поставщика	низкий
Anticorruption Layer (ACL)	Потребитель не доверяет модели поставщика и создает свой слой адаптации	Потребитель адаптирует под свои нужды	управляемый



Закон Конвея и сквозные бизнес-процессы

Многие крупные организации приходят к функциональным анклавам.

Закон Конвея – «Организации проектируют системы, которые копируют структуру коммуникаций в этой организации»

Другими словами, если структура вашей организации – слабо взаимодействующие функциональные подразделения, то скорее всего, вы для каждого подразделения купите или разработаете собственные системы, и которые будут друг с другом взаимодействовать так же плохо



Кейс: управления данными в конкретном под-домене (Bounded context)

Подход

1. Обследование и выявление болей и ограничений
2. Устранение текущих проблем с данными
3. Разработка стандарта DQ, инструкции
4. Определение и обучение кураторов данных
5. Создание базы знаний под-домена в Confluence
6. Описание бизнес-гlossария, выявление CDE
7. Разработка правил качества данных
8. Реализация мониторинга загрузок данных (тип ACL)

Результаты

1. Устранения проблем в типах данных, в справочниках **на 80%**
2. Повышение знаний предметной области **на 30%**
3. Сокращение риска потери экспертизы **на 70%**
4. Повышение уровня качества данных **на 10%**
5. Сокращение «костылей» **на 30%**
6. Сокращение технического долга **на 50%**

Применение процессов DG на практике

Data Governance создает фундамент для эффективного внедрения ИИ: **она обеспечивает достоверность, понятность и управляемость данных**, без которых модели ИИ **теряют точность и доверие**

Спасибо за внимание!

■ С радостью готов ответить на вопросы

