

Построение бизнес-кейса для инициатив качества данных

Часть 1. Начало пути. Экспресс-оценка и паттерны

DIS Group Kazakhstan

2023

200+

КЛИЕНТОВ

КОМПАНИЙ-ЛИДЕРОВ КАЗАХСТАНА,
РОССИИ И СТРАН СНГ В РАЗЛИЧНЫХ
ОТРАСЛЯХ



18

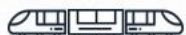
ЛЕТ

250+

ПРОЕКТОВ

УПРАВЛЕНИЕ ДАННЫМИ,
УПРАВЛЕНИЕ ЗНАНИЯМИ,
РЕШЕНИЯ ДЛЯ ПОДРАЗДЕЛЕНИЙ
КОМПЛАЕНСА

★ МОСКВА



АЛМАТЫ АСТАНА



★ АСТРАХАНЬ

★ САНКТ- ПЕТЕРБУРГ



★ ИРКУТСК

МАСТЕР-ДИСТРИБЬЮТОР В КАЗАХСТАНЕ И СНГ

ЧТО ОБЫЧНО ВИДИМ, КОГДА ХОТИМ ОБОСНОВАТЬ ИНВЕСТИЦИИ В КАЧЕСТВО ДАННЫХ?

ОБЩИЕ ПОКАЗАТЕЛИ ОБ ОГРОМНЫХ ПРОБЛЕМАХ СВЯЗАННЫХ С КАЧЕСТВОМ ДАННЫХ

*Ежегодно низкое качество
данных обходится
организациям в среднем
в 12,9 млн. долл.*

Gartner

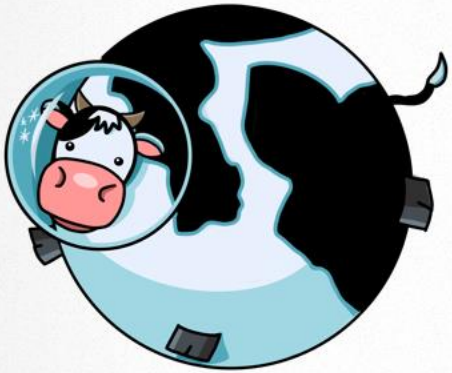
*Только у 3% компаний
данные соответствуют
основным стандартам
качества**

**Harvard
Business
Review**

*Обычная организация
затрачивает 80% времени
на поиск и подготовку
данных*

IDC

МНОГИЕ ВОСПРИНИМАЮТ ОБЩУЮ СТАТИСТИКУ ПО ЗАТРАТАМ НЕПРАКТИЧНОЙ КТО-ТО ВСЕ ВРЕМЯ ИЩЕТ СЕКРЕТНЫЙ ИНГРЕДИЕНТ



SPHERICAL COW



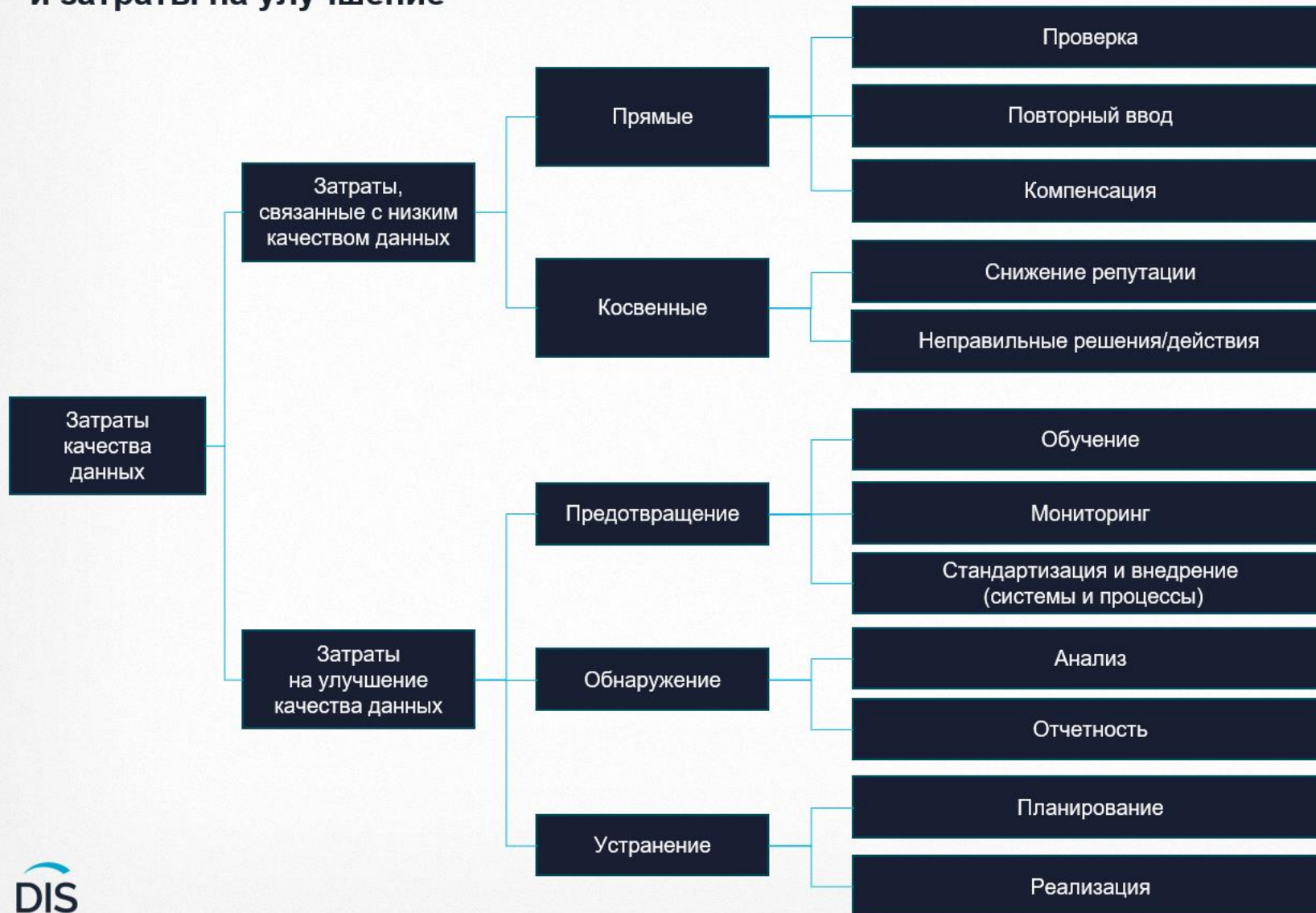
ПОСТОЯННЫЙ ПОИСК
СЕКРЕТНОГО ИНГРЕДИЕНТА



СФЕРИЧЕСКИЙ КОНЬ

ПОСТРОЕНИЕ ДЕРЕВА ЗАТРАТ КАЧЕСТВА ДАННЫХ

Затраты на качество данных состоят из двух основных типов: затраты, связанные с низким качеством данных и затраты на улучшение



- **Дебаты о достоверности:** Всегда будут обсуждения или разногласия относительно точности оценки затрат на данные низкого качества
- **Видимые затраты:** Некоторые затраты, связанные с данными низкого качества, являются очевидными и могут быть прямо наблюдаемыми организацией
- **Операционные расходы:** Это постоянные расходы на поддержание работы продукта, бизнеса или системы
- **Капитальные расходы:** Это затраты, которые компания несет на приобретение, обслуживание или улучшение своих основных средств
- **Косвенные (невидимые) затраты:** Это затраты, которые не сразу становятся очевидными или наблюдаемыми. Они становятся явными только после возникновения проблемы из-за данных низкого качества
- **Оценка косвенных затрат:** Поскольку косвенные затраты не наблюдаются непосредственно и могут варьироваться в зависимости от последствий низкого качества данных, их часто оценивают, а не базируют на точных данных

ПРАВИЛО 1:10:100 (ИЛИ ПРАВИЛО 10) ОЦЕНОЧНАЯ СТОИМОСТЬ РЕШЕНИЯ ПРОБЛЕМЫ С ДАННЫМИ



Инвестировать в предотвращение проблем заранее (предпроверка) – дешевле. Платить за исправление и устранение последствий позднее (постпроверка) – дороже

- **Затраты на предупреждение (профилактика):** Инвестируем в качественные процессы и обучение сотрудников, чтобы избежать ошибок с первой попытки. Это эквивалентно **1 единице работы**
- **Затраты на исправление:** Это стоимость брака и переделок. Исправление стоит **в 10 раз дороже, чем профилактика, или 10 единиц работы**
- **Затраты из-за неудач:** Если проблему не решить, последующие убытки могут быть **в 10 раз больше, чем затраты на исправление, или 100 единиц работы**. Более того, они могут привести к неудовлетворенности клиентов и ущербу для репутации организации

ЭКСПРЕСС-ОЦЕНКА (1-3)

ОБЩИЕ ПОКАЗАТЕЛИ ОБ УБЫТКАХ СВЯЗАННЫХ С КАЧЕСТВОМ ДАННЫХ

1 ВЫРУЧКА: Стоимость некачественных данных для большинства компаний составляет от **15% до 25%** от выручки

Формула	Расчет (консервативный)	Описание
Выручка x 15-25%	100 млрд. x 15% = 15 млрд. тенге	В 2022 году, оценочная стоимость потерь связанных с проблемами качества данных составила 15 млрд. тенге

2 ПОТЕРЯ ВРЕМЕНИ: До **50%** времени работников умственного труда (УТ) (тех, кто использует данные и информацию) тратится впустую из-за некачественных данных

Формула	Расчет (консервативный)	Описание
Количество работников УТ x Количество рабочего времени в месяц x 50%	Орг. ед. "А": 15 чел. x 160 час. x 50% = 1 200 часов в месяц	В 2022 году согласно оценке, орг. единицей "А" (департаментом/командой) было потеряно 1 200 часов в связи с проблемами поиска и некачественными данными

ЭКСПРЕСС-ОЦЕНКА (2-3)

ОБЩИЕ ПОКАЗАТЕЛИ ОБ УБЫТКАХ СВЯЗАННЫХ С КАЧЕСТВОМ ДАННЫХ

3 ОШИБКИ ПРИ ЗАВЕДЕНИИ ДАННЫХ: В среднем 47% вновь созданных записей данных содержат как минимум одну критическую ошибку

Формула	Расчет (консервативный)	Описание
Количество новых записей в месяц x 47%	$30\ 000 \times 47\% = 14\ 100$ записей	Из 30 000 записей, создаваемых каждый месяц, примерно 14 100 содержат по крайней мере одну критическую ошибку, которая негативно влияет на наш бизнес

4 ДОВЕРИЕ К ДАННЫМ: 16% менеджеров полностью доверяют данным, которые они используют для принятия важных решений.

Формула	Расчет (консервативный)	Описание
Количество менеджеров x 16%	Организация. “Б”: 150 чел. x 16% = 24 чел.	В 2022 году согласно оценке в нашей организации, из 150 менеджеров лишь 24 полностью доверяют данным, которые они используют для принятия решений

МОДЕЛЬ БИЗНЕС-КЕЙСА (ПАТТЕРН 1) ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ РАБОТЫ С ДАННЫМИ



1 ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ РАБОТЫ С ДАННЫМИ: В среднем сотрудник Управления данными (УД) компании А затрачивает 50% времени на рутинные задачи и устранение ошибок связанных с данными

N	Компонент	Мин	Сред	Макс	Тип
1	Количество сотрудников УД, ч		25		Ввод
2	Средний годовой ФОТ на сотрудника, тенге		12 000 000		Ввод
3	Средний процент времени, затрачиваемого на рутину и устранение ошибок		50%		Ввод
4	Общий уровень FTE		12,5		1*3
5	Ср. время оптимизации	40%	50%	60%	Ввод
6	Высвобождение FTE на выполнение других работ	5	6,25	7,5	4*5*
7	Выгода, тенге	60 000 000	75 000 000	90 000 000	6*2

Informatica оптимизирует работу сотрудников по данным, автоматизирует рутинные задачи. Это позволяет высвободить время для более важных задач и ускоряет реализацию проектов, связанных с данными. **Годовой эффект согласно среднему сценарию составляет: 75 000 000 тенге**

МОДЕЛЬ БИЗНЕС-КЕЙСА (ПАТТЕРН 2) МЕТРИКИ DEVOPS, DATAOPS



2 СНИЖЕНИЕ ВРЕМЕНИ НА ОБНАРУЖЕНИЕ И УСТРАНЕНИЕ СБОЕВ: В среднем сотрудник компании А затрачивает **30%** времени на инциденты с данными, а среднее время простоя сервисов данных составляет **36 ч/мес**

N	Компонент	Сейчас	План	Тип
1	Время необходимое на обнаружение инцидента (сумма), ч	120		Ввод
2	Количество инцидентов, ед	10		Ввод
3	Среднее время обнаружения (Mean-Time-To-Detect), ч	12	2	1/2
4	Время на устранение проблемы (сумма), ч	240		Ввод
5	Количество инцидентов, ед	10		Ввод
6	Среднее время разрешения (Mean-Time-To-Resolve), ч	24	13	4/5
7	Средняя стоимость простоя*, тенге/ч	80 000	80 000	Ввод*
8	Стоимость простоя из-за проблем с данными в год, тенге	34 560 000	14 400 000	(3+6)*7*12м

Проект DQ сократит средний ежемесячный простой с 36 ч. до 15 ч. (оценка), что приводит к экономии **20 160 000 тенге в год***. (* - вычтите стоимость инвестиций в проект)



1. В данном примере (паттерне) для упрощения не учитываются потери для бизнеса, издержки упущенной выгоды и т.д.
 2. Это лишь составной показатель для оценки инвестиций в проект управления данными. Инциденты фазы эксплуатации действующего ИТ-ландшафта.
 3. * Взят инцидент: Команда администрирования, команда инженеров данных, тестировщики, разработчики, инженеры СИ, разработчики BI-отчетов, группа сопровождения, департамент бизнеса

МОДЕЛЬ БИЗНЕС-КЕЙСА (ПАТТЕРН 3) ПРОВЕРКИ КАЧЕСТВА ДАННЫХ: ИСПРАВЛЕНИЕ VS ПРЕДУПРЕЖДЕНИЕ

N	Компонент	УД	Департамент отчетности		Тип
			Постпроверка (А)	Предпроверка* (50%) (В)	
1	Количество сотрудников, ед.	20	10		Ввод
2	Средний годовой ФОТ на сотрудника, тенге	12 000 000	15 000 000		Ввод
3	Стоимость бизнес-функции в год, тенге	240 000 000	150 000 000		1*2
4	Количество элементов данных БФ (использование/владение), ед.	150	80		Ввод
5	Стоимость одного элемента данных (ЭД), тенге	1 600 000	1 875 000		3/4
6	Количество некорректных элементов данных на входе из СИ, ед.	20	20		Ввод
7	Затраты БФ на исправление 20 некорректных ЭД, тенге	32 000 000	37 500 000		5*6
8	Количество некорректных элементов данных на входе в БФ*, ед.		20	10	
9	Затраты БФ на исправление оставшихся некорректных ЭД, тенге	32 000 000	37 500 000	18 750 000	
10	Экономия затрат за счет предварительной проверки (предпроверка), тенге	-	-	18 750 000	A7-B7
11	Повышение операционной эффективности БФ, тенге		18 750 000		10

Проект DQ для Департамента отчетности повысит его операционную эффективность на 18 750 000 тенге в год

ОБЩИЙ ИТОГ

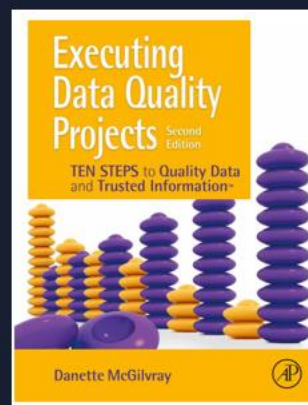
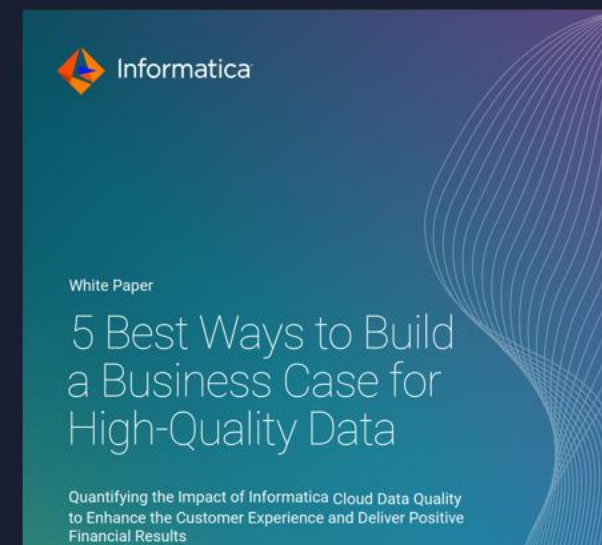
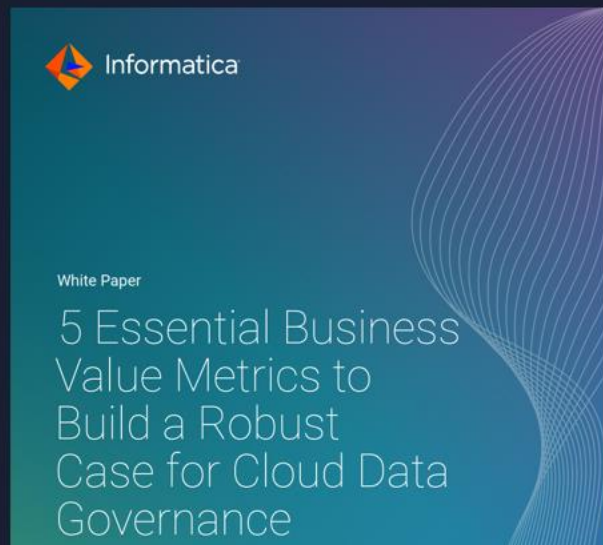
СБОРКА ИНТЕГРИРОВАННОГО БИЗНЕС-КЕЙСА ИЗ ПАТТЕРНОВ

Иницируемый проект управления качеством обеспечивает выгоды на всем жизненном цикле данных.
Суммарный экономический эффект составляет 113 910 000 тенге в год

Общий эффект/выгода инициативы управления качеством данных в год	Сумма, тенге
1 ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ РАБОТЫ С ДАННЫМИ	75 000 000
2 СНИЖЕНИЕ ВРЕМЕНИ НА ОБНАРУЖЕНИЕ И УСТРАНЕНИЕ СБОЕВ	20 160 000
3 ПРЕВЕНТИВНОЕ УПРАВЛЕНИЕ КАЧЕСТВОМ ДАННЫХ	18 750 000
ИТОГО	113 910 000

Допущения описаны в детализации каждого кейса. В общей оценке и представлен консервативный сценарий

Источники для вдохновения + опыт DIS GROUP



ВЫВОДЫ И РЕКОМЕНДАЦИИ

1

Признавайте субъективность и изменчивость оценок

Примите внутреннюю субъективность данных; приветствуйте дебаты и вызовы, убедившись, что они ведут к большей ясности и пониманию

2

Показывайте сценарии

Подчеркните результаты управления качеством данных в сравнении с его отсутствием или пренебрежением, позволяя стейкхолдерам видеть будущие риски и возможности

3

Идите в гемба и коммуницируйте

Общайтесь со стейкхолдерами, проводите опросы и оценивайте влияние данных на результаты бизнеса. Содействуйте культуре командного взаимодействия, смотря на проблему под разными углами

4

Стремитесь к простоте и используйте паттерны

Чем проще и понятней ваш бизнес-кейс, тем лучше. Используйте паттерны для создания собственных кейсов. Объединяйте их как блоки для интегрированного кейса

5

Делайте ставку на превентивное управление и последующий мониторинг

Проактивное управление данными снижает издержки. Отслеживайте результаты после внедрения и корректируйте стратегию и планы при необходимости

6

Думайте о людях, снабдите их инструментами

Люди - ключевой элемент успешного управления данными. Поддерживайте их с помощью современных инструментов для управления данными

**Закажите
демонстрацию решений по
управлению данными
на dis-group.world**



**СПАСИБО
ЗА ВНИМАНИЕ!**

Таймагамбетов Болат
<https://dis-group.world/>

