5 уроков, полученных при маскировании баз данных

Антон Ноздрин, эксперт по продукту Сфера.Обезличивание данных

Кто мы?

НОТА (ИТ-холдинг Т1) — один из крупнейших ИТ-холдингов в РФ.
Производитель и интегратор ПО **Сфера.Обезличивание данных**

- Промышленное использование ПО обезличивания, в т.ч. один из крупнейших банков РФ (120 баз данных, 600 приложений)
- Десятки выполненных пилотов
- Десятки ответов на Функциональные Требования и интервью потенциальных клиентов

Маскирование (обезличивание)* данных для использования в тестировании ПО

Значительно снизить время вывода бизнес-ПО в промышленную эксплуатацию при одновременном сокращении

- Риска утечки конфиденциальных данных
- Риска зависимости от конкретных исполнителей
- Затрат на создание и обновление сред разработки и тестирования

* в рамках данного доклада

Специализированное ПО vs. Скрипты в БД

Почему скрипты в БД это не очень здорово

- Экономика процесса в пользу специализированного ПО
- Общее время на получение первой копии обезличенных данных у скриптов существенно выше
- Требуется разработчик на конкретном диалекте SQL Script очень высокого уровня
- Скрипты несопоставимо сильнее нагружают СУБД
- Безопасная промышленная разработка для скриптов — очевидное требование
- Нет кросс-платформенной целостности

Специализированное ПО vs. Скрипты в БД

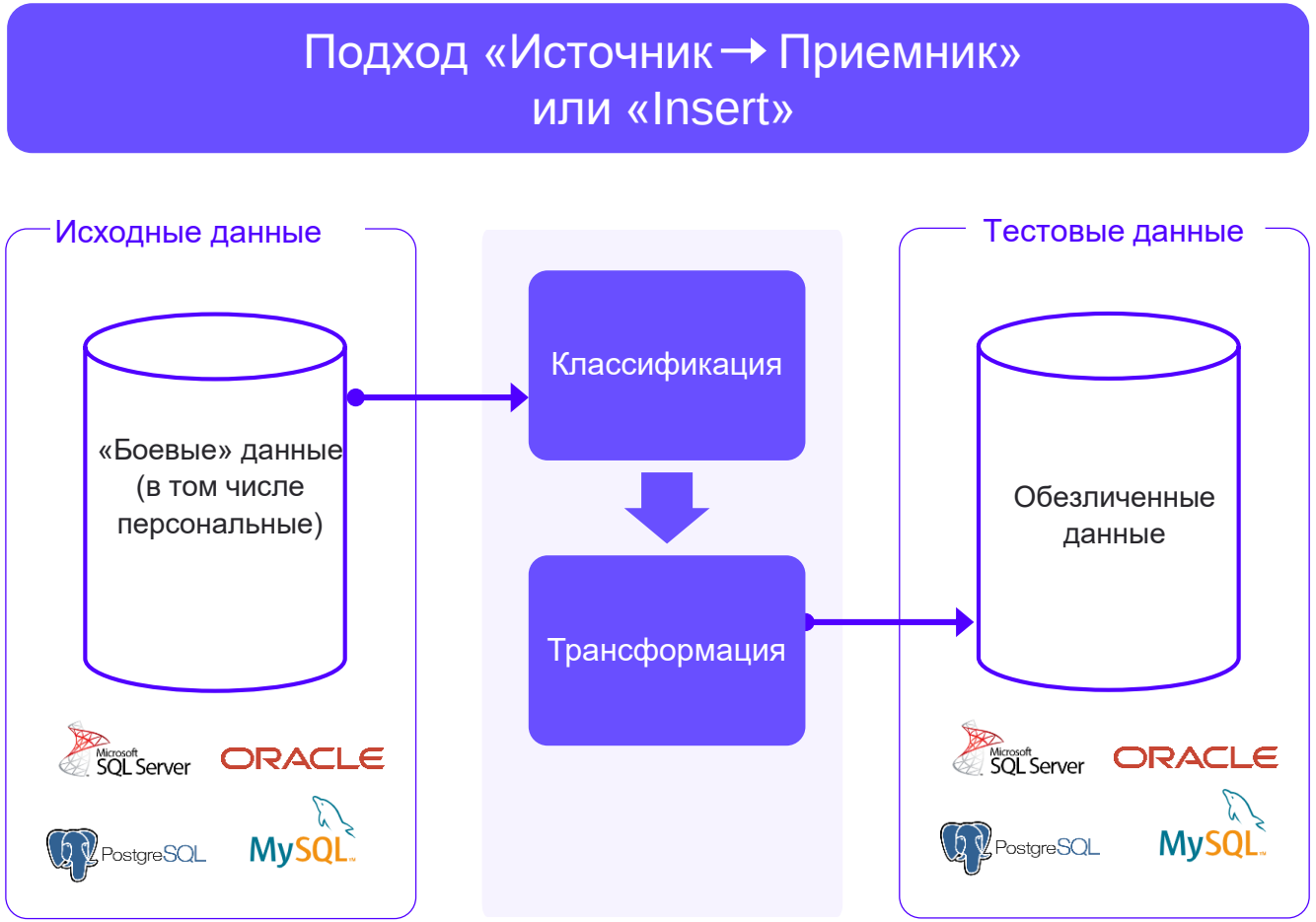
Почему скрипты в БД это не очень здорово

- Экономика процесса в пользу специализированного ПО
- Общее время на получение первой копии обезличенных данных у скриптов существенно выше
- Требуется разработчик на конкретном диалекте SQL Script очень высокого уровня
- Скрипты несопоставимо сильнее нагружают СУБД
- Безопасная промышленная разработка для скриптов — очевидное требование
- Нет кросс-платформенной целостности

Но это все мелочи

Основное: скрипты — это потенциальный конфликт и потеря контроля для ИБ (да и для ИТ тоже)

Специализированное ПО: подходы к процессу



Урок 1

Качественная классификация* чувствительных данных важнее скорости непосредственной трансформации

- Каждое пропущенное поле — риск для ИБ
- Бизнес-процесс всегда можно скорректировать, утечка необратима
- Точное (100%) определение всех полей удорожает процесс, надо понимать баланс
- Ложноположительные срабатывания лучше ложноотрицательных

\+ Возможен сценарий использования только классификации (проверки пакета данных на наличие чувствительных данных)

* Также используются термины Профилирование, Поиск ПД, Идентификация ПД

Урок 2

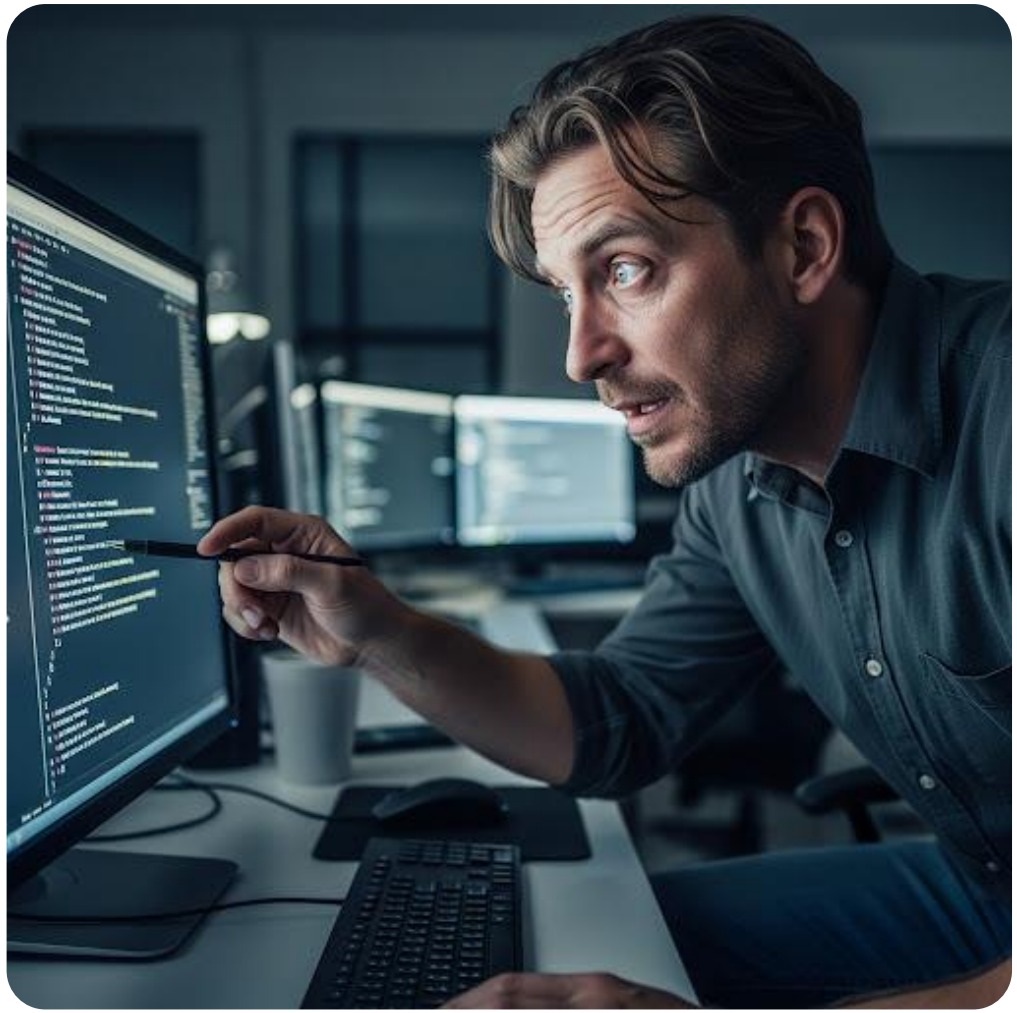
Не бывает единого подхода к организации процесса обезличивания

- Выделенный корпоративный сервис **«Обезличивание по заявкам»** или **встраивание в CI/CD** — это все абсолютно корректные подходы
- Возможно одновременное использование разных подходов, но нужно соблюдать баланс
- Основной критерий сбалансированного решения — ИБ должно быть удобно контролировать процесс при максимально возможной степени свободы для ИТ

Урок 3

Пользовательские (кастомные) алгоритмы маскирования — необходимое зло

- **Каждый пользовательский алгоритм** — это время, затраченное сотрудником ИБ на его проверку
- **Сбалансированный подход** — это библиотека утвержденных пользовательских алгоритмов, используемая для всех БД в контуре обезлички
- **Гибкая система полномочий на управление кастомными алгоритмами** — это способ обеспечить контроль ИБ



Урок 4

Не все интуитивно привлекательные функции имеют практический смысл

- Чужой опыт использования – вот лучший маркетинг
- Референс к существующему клиенту – возможность копнуть поглубже и примерить решение на себя
- Пилотный проект – в нише маскирования данных это обязательный шаг
- Чудес не бывает и за все надо платить

Урок 5

ПО для маскирования данных — это только «станок в цехе», нужен «забор, колючая проволока, отдел режима и охраны и обеспечение порядка ввоза и вывоза материальных ценностей»

- Небрежный подход к защите самой системы обезличивания — путь к успешной атаке
- ИБ-регламент по защите ландшафта в целом* — обязательная часть проекта внедрения Маскирования
- Опытная команда и лучшие практики (референс-визиты) — экономия времени и снижение рисков

* В контексте задачи обезличивания данных

Сфера.Обезличивание данных

- Стандартизированный промышленный подход к процессу деперсонализации
- Процесс INSERT (Источник-Приемник) и UPDATE (Источник-Источник)
- Автоматический поиск КИ на основе ML и регулярных выражений
- Встроенные и пользовательские алгоритмы обезличивания (на Java)
- Кроссплатформенная ссылочная целостность БД
- Поддержка специфических типов данных (ИНН, ОКПО, СНИЛС, алгоритм Луна и т.д.)
- Гибкая система полномочий, ABAC+RBAC
- UI и API

Соответствует законодательству РФ: ФЗ №152 «О персональных данных», ФЗ №98 «О коммерческой тайне», ФЗ №395-1, ст.26 - «Банковская тайна».

Внесена в Реестр отечественного ПО, № 12277

Платформа производства «Сфера»

Инструменты для организации и автоматизации процессов производства на всех этапах жизненного цикла ИТ-продукта

Цели и задачи

Вывод на рынок технологически независимых инструментов для создания ИТ-продуктов

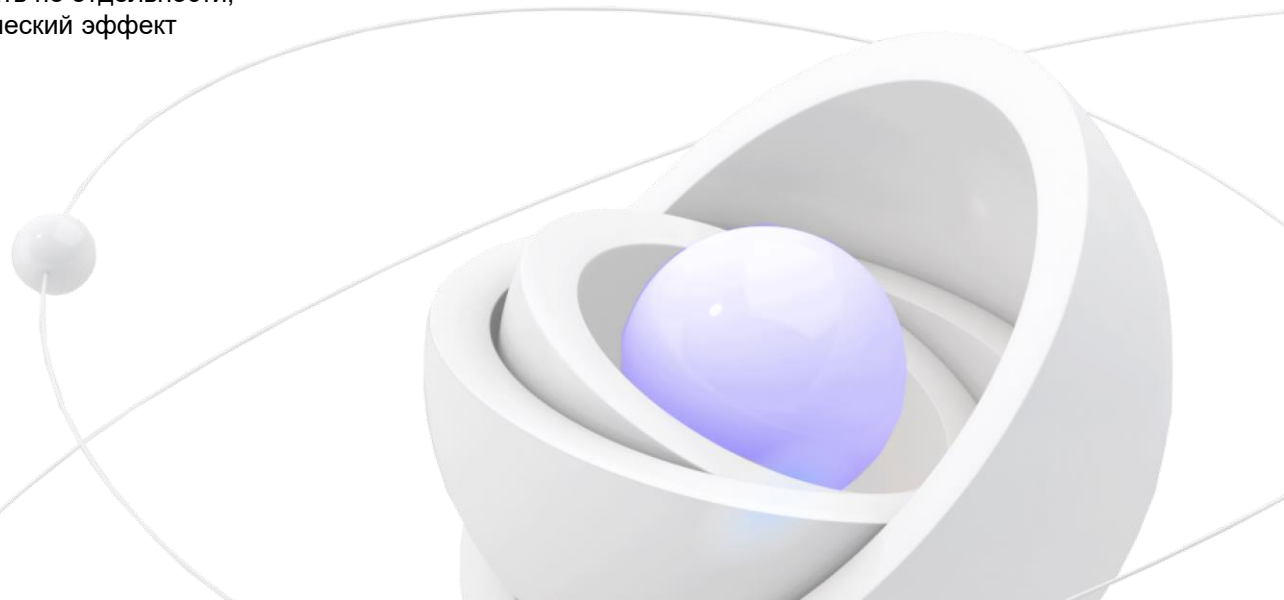
Модель дистрибуции

- **On-premise** для Enterprise и организаций с госучастием
- **SaaS*** для SMB

Инструменты могут работать по отдельности, но вместе дают синергетический эффект

Отечественные инструменты

Все продукты зарегистрированы в ЕРРП Минцифры



* Готовность SaaS — 2025

Вопросы?