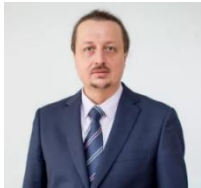


# БОЛЬШИЕ ДАННЫЕ И ИХ ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ В КЛЮЧЕВЫХ ОТРАСЛЯХ: КАК ПРЕВРАТИТЬ ДАННЫЕ В УПРАВЛЯЕМЫЕ БИЗНЕС-ИНСАЙТЫ

## Управление архитектурой данных

Первый шаг для цифровой  
трансформации



Коптелов Андрей Константинович  
Бизнес-тренер  
[koptelov.info](http://koptelov.info)



# Представляюсь

## **Коптелов Андрей Константинович**

- Консультант в области управления бизнес-процессами и бизнес-архитектурой
- Более 25 лет специализации в области процессного управления
- Более 60 завершенных проектов в области описания и оптимизации процессов
- Более 200 выступлений на конференциях и других мероприятиях
- Более 200 публикаций в российских и зарубежных изданиях
- Преподавание в бизнес-школах: ВШЭ, ИБДА РАНХИГС, Школа Бизнеса Синергия
- Преподавание в учебных центрах: IBS, IT-Expert, OTUS и др.
- Со-основатель ABPMP Russia (сообщество профессионалов BPM)
- CEO Школы бизнес-анализа Андрея Коптелова
- Исполнительный директор в банковской организации

[a@koptelov.info](mailto:a@koptelov.info)

# Про что рассказ – про цифровизацию обучения

Администрирование	Управление финансами	Управление контентом	Маркетинг	Продажи	Обучение	Управление ИТ
Бизнес-планирование (MS Excel)	Ценообразование и скидки (MS Excel)	Определение учебного контента для разработки	Маркетинг в группах (Facebook)	Управление воронкой продаж (MS Excel)	Управление учебными курсами	Управление хостингом доменными именами (Beget.com)
Управление задачами (Trello)	Финансовый учет (MS Excel)	Создание учебного контента (ZOOM)	Управление рассылками (RUSENDER)	Взаимодействие с клиентами (MS Excel)	Управление слушателями (MOODLE)	Управление сайтом (koptelov.online)
Управление бум. документами (СДЕК, Почта России)	Налоговый учет (Тинькофф Банк)	Создание текстового контента (MS Word)	Интернет продвижение (Yandex Direct)	Взаимодействие с учебными центрами	Проверка домашних заданий (MOODLE)	Управление сайтом (koptelov.info)
Управление эл. документами (ДИАДОК)	Контроль платежей (MS Excel)	Монтаж учебного контента (VideoPad)	Проведение маркетинговых вебинаров (ZOOM)	Подготовка Коммерческих предложений (MS Word)	Контроль успеваемости слушателей (MOODLE)	Управление сайтом (analyst.su)
Управление архивом документов (Шкаф в офисе)	Эквайринг (Модуль Банк)	Размещение учебного контента (MOODLE)	Управление отзывами слушателей (MS Excel)	Управление договорами и актами (MS Word)	Оценка выпускных работ (MOODLE)	Управление сайтом (edu.analyst.su)
Управление предложениями по улучшению (MS Excel)	Управление платежами (Банковские сервисы)	Управление архивом учебного контента (жесткий диск)	Управление отзывами компаний (Yandex Disk)	Поиск тендеров (Yandex.ru)	Проведение вебинаров (ZOOM)	Управление резервным копированием
Консультационные услуги (ZOOM)	Финансовый анализ деятельности (MS Excel)	Создание тестов (MOODLE)	Управление маркетинговыми материалами (Yandex Disk)	Управление жалобами	Подготовка сертификатов (PDF)	Управление обновлениями (MOODLE, WordPress)

Что в целях .... Data-driven decision making



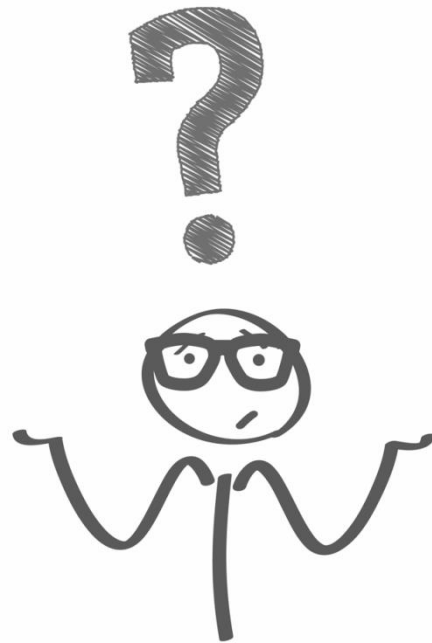
# ВЫЗОВЫ

## Низкое качество данных

- Нет методологии и процессов мониторинга качества данных
- Нет инструментов для автоматического контроля качества
- Нет SLA по качеству данных между владельцами и потребителями
- Нет процессов по решению инцидентов качества данных
- ...

## Высокий T2M при разработке для аналитики

- Поиск данных
- ...



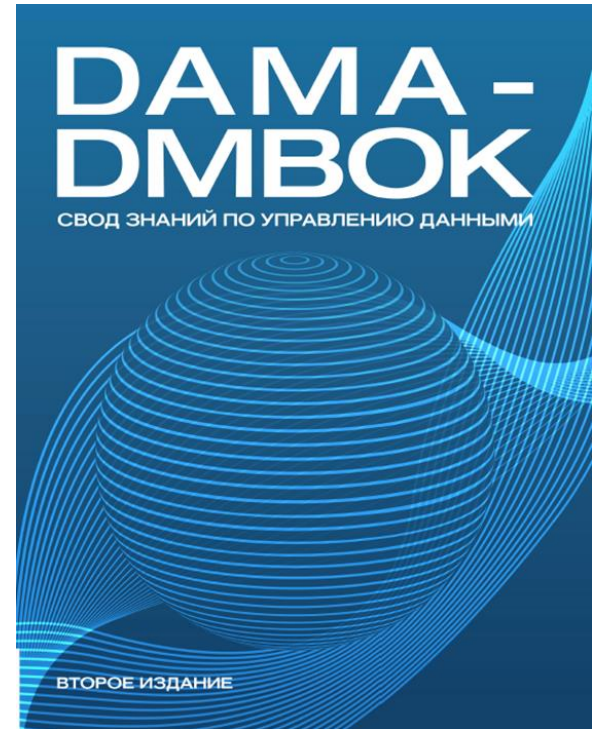


# Что такое архитектура данных

DAMA-DMBOK: архитектура данных «включает

- **методы, правила и способы**, позволяющие
- **описать** текущее состояние данных,
- **сформировать требования** к данным,
- **обеспечить интеграцию** данных и
- **контролировать использование** данных
- **как активов** в соответствии со
- **стратегией управления данными»**

Данные являются **стратегическим активом**, которым необходимо управлять на уровне организации



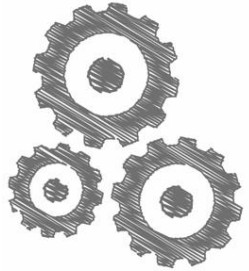
# Архитектура данных

Архитектура данных отвечает на вопросы: **какие данные есть в организации, как они организованы, как ими управляют и как их распространяют**

Если не управлять данными, то организация может столкнуться с проблемами:

- неуправляемого роста объемов данных и затрат на их хранение
- отсутствия информации о том, какие данные у организации есть, где они располагаются, кем и как используются
- низкой эффективности обработки данных

Значительно возрастают правовые риски, риски утечки данных, риски санкций и штрафов





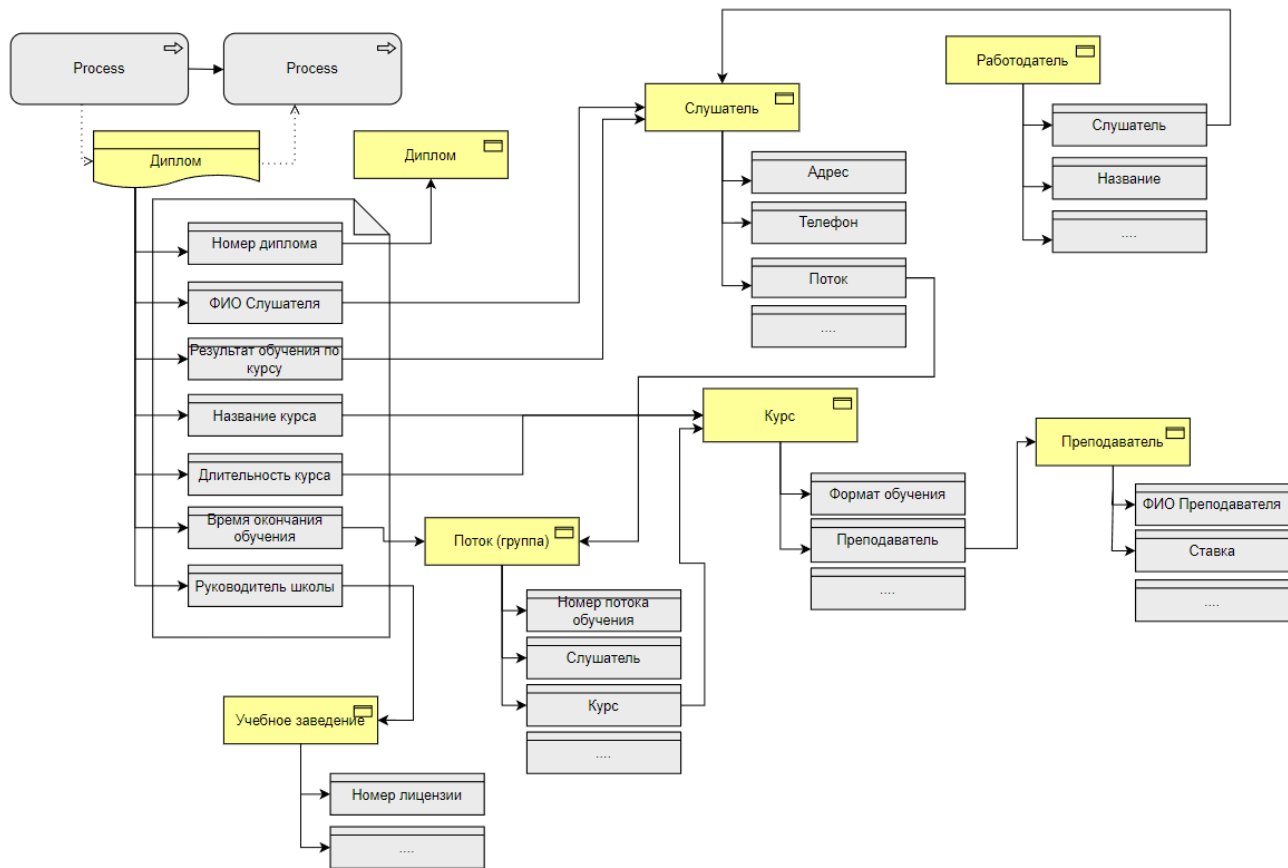
## Проблема – договориться о терминологии

- ...
  - Курс
  - Дисциплина
  - Программа
  - Урок
  - Тренинг
  - Предмет
  - ...
- ....
  - Информационная система
  - Автоматизированная система
  - Приложение
  - Программное обеспечение
  - Программный продукт
  - ИТ-сервис
  - ...

# Бизнес-гlossарий - задачи

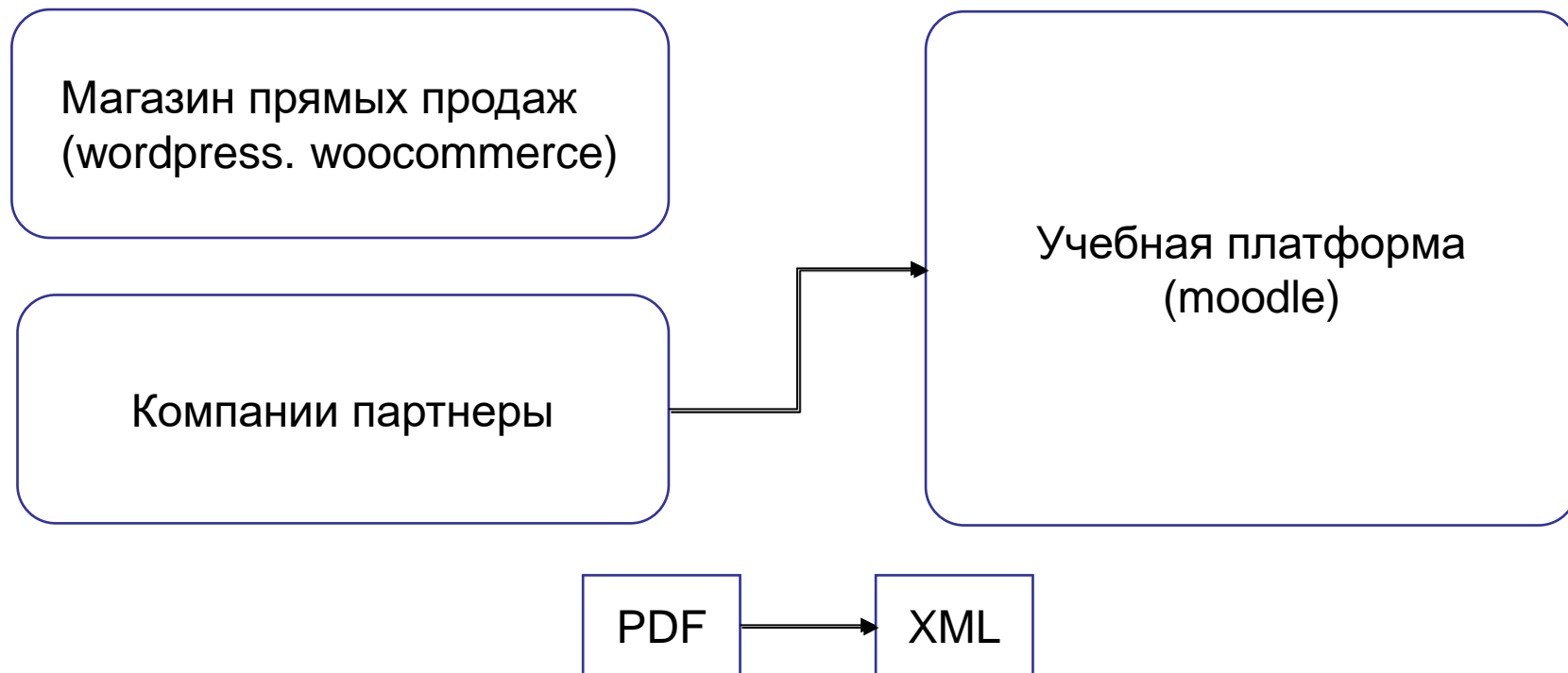
- обеспечение единого общего понимания основных понятий и терминов
- снижение риска неправильного использования данных из-за неверного понимания связанных с бизнесом понятий и концепций
- повышение терминологической согласованности между подразделениями, отвечающими за технологические ресурсы (с собственными соглашениями об именовании), и подразделениями, отвечающими за организацию и ведение бизнеса
- максимизацию возможностей поиска и обеспечения доступа к документированным знаниям организации

# Процессы, информационные потоки и объекты



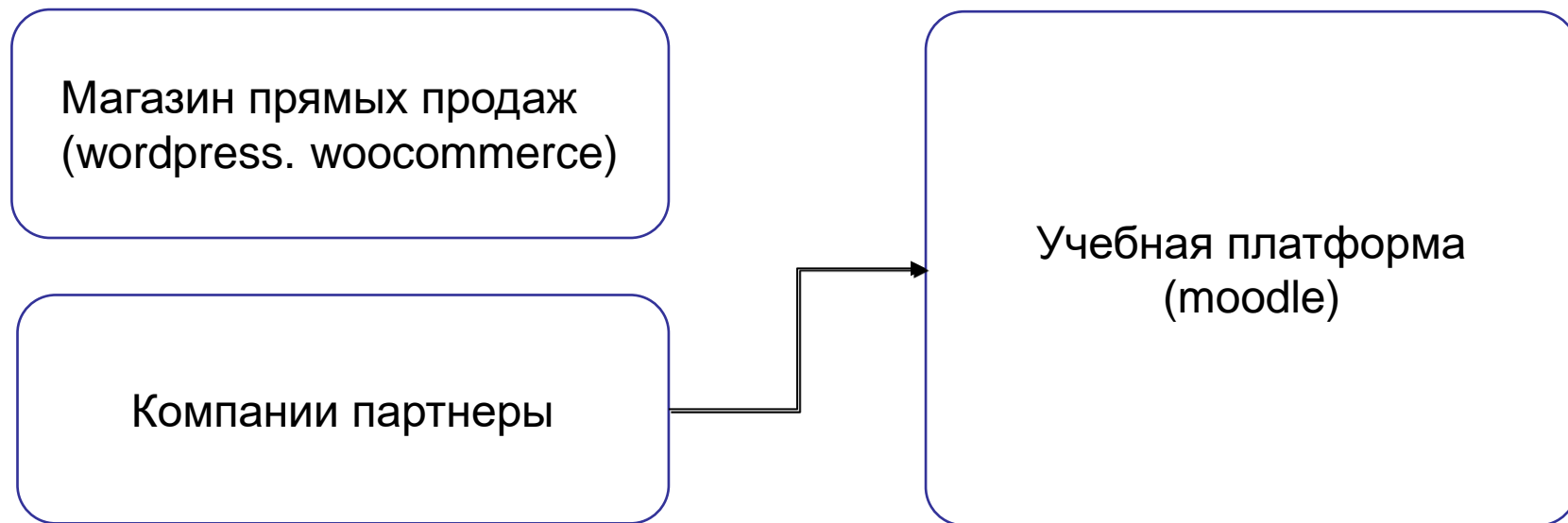
# Для начала – обмен в машиночитаемом формате

Партнеры присылают данные в «странных» для меня форматах



# Дальше - ответственность за качество данных

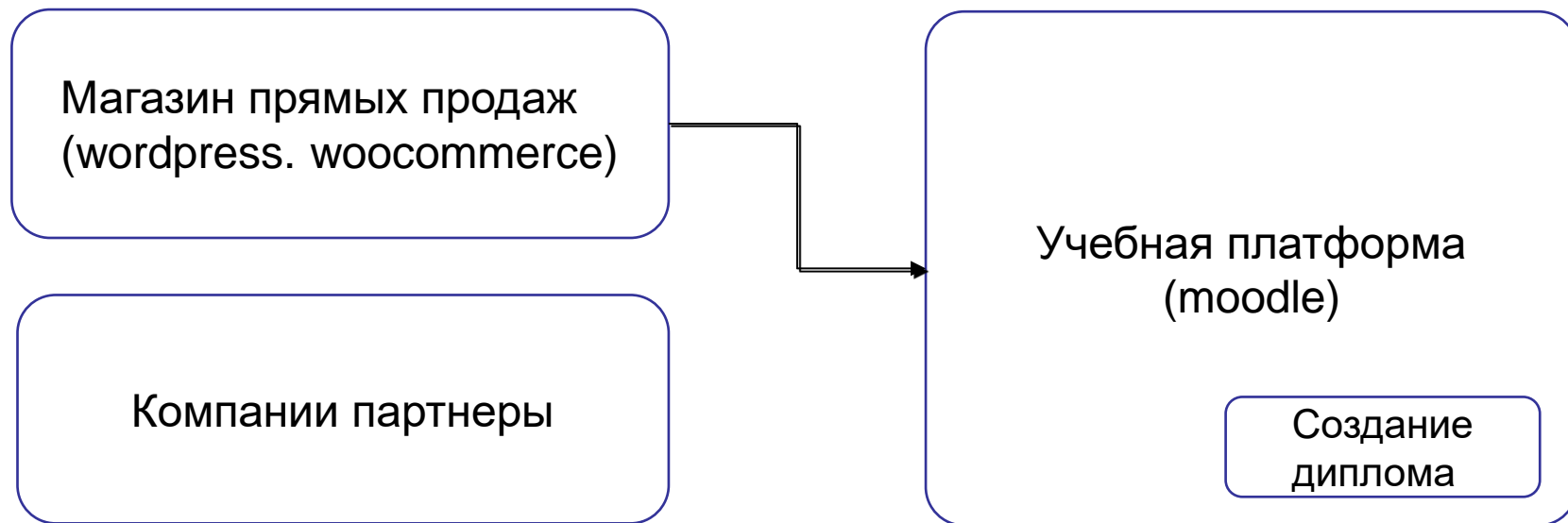
Партнеры присылают некачественные данные (ошибки)



Атрибут: эл. почта (контроль качества)

Атрибут: Фамилия Имя Отчество (контроль формата, качества)

# Проблемы внутри: автоматическое создание диплома



Атрибут: Имя

Атрибут: Фамилия

# Контроль ввода данных

**Основная цель контроля** — предотвратить ввод некорректных данных пользователем:

- **контроль формата вводимых данных** – указывается допустимый тип данных: целое, дата, время, текст определенной длины
- **контроль диапазона значений вводимых данных** – значение между заданными границами, вне заданных границ, равно, неравно, больше, меньше, больше или равно, меньше или равно значению
- **контроль принадлежности вводимого значения фиксированному списку** – выбор из справочника, поле со списком для выбора



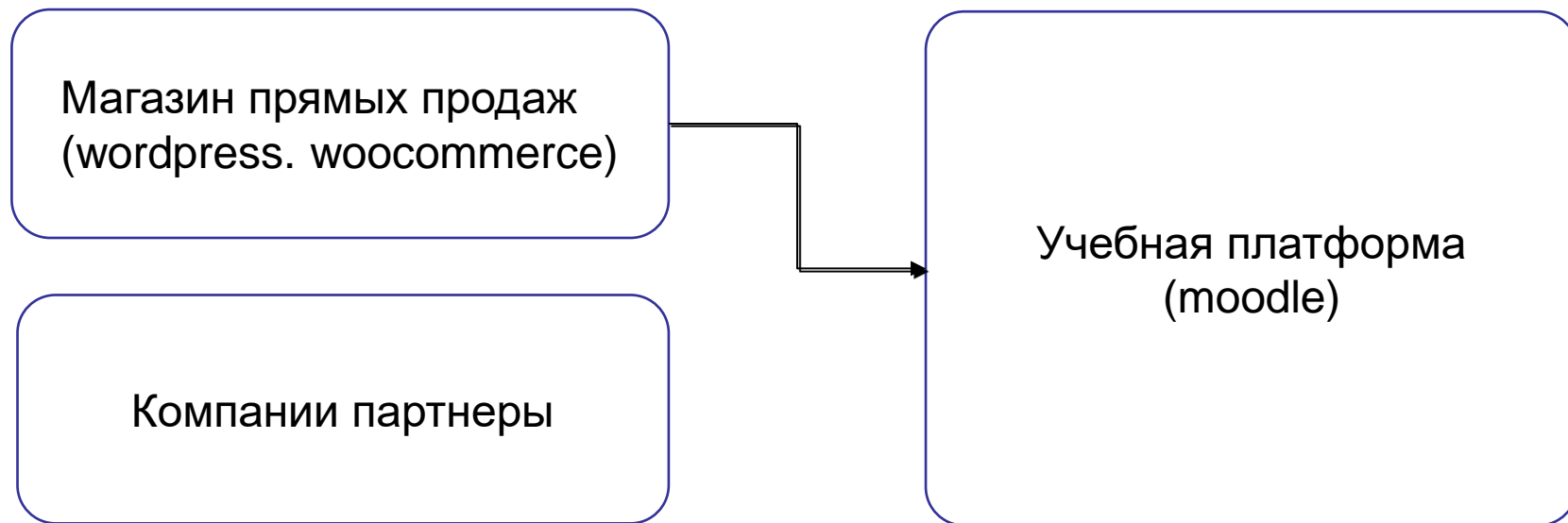
# Проблема ввода данных слушателя

- Какие данные вводятся при зачислении слушателя?
  - Как их проверить, что он не допустил ошибки?
  - Как в автоматическом режиме валидировать данные о слушателе?
  - Можем ли мы автоматически исправить ошибки в данных о слушателе?
- ФИО
  - Телефон личный
  - Почтовый адрес личный
  - Возраст
  - Пол
- ФИО -> Отсутствие любых символов, кроме русских букв и тире. Соответствие отчества и пола (по хорошему справочник имен и отчеств)
  - Дата рождения -> Возраст между 18 и 60
  - Телефон личный -> Количество цифр в номере (по хорошему нужен сервис подтверждения)
  - Почтовый адрес личный -> Присутствие @ в строке (по хорошему нужен сервис подтверждения)



# Слушатели, купившие на распродаже – не учатся (80%)

Атрибут: Распродажа

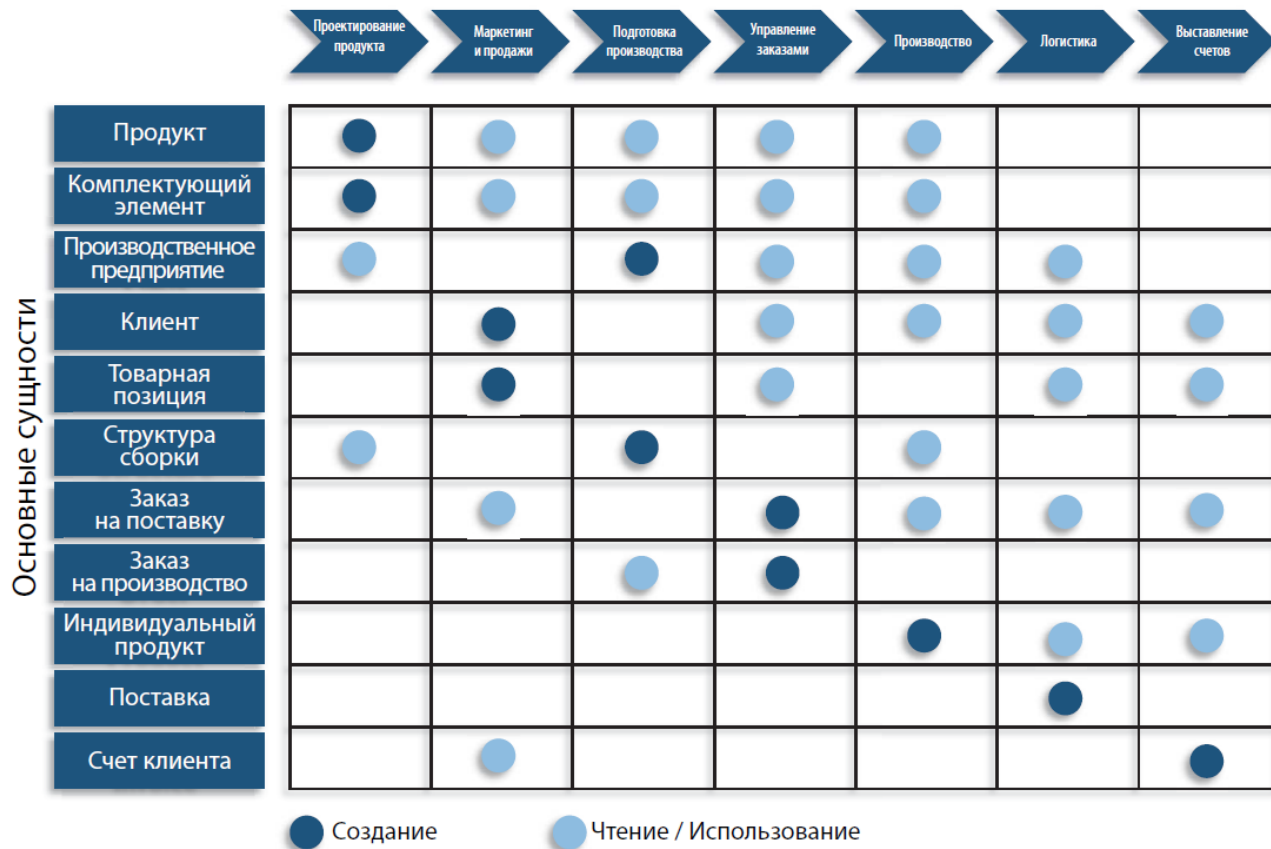


# Скромные выводы

- Ответственность за данные – это правильно
- Глоссарий нужен, но термины могут переопределяться локально
- Все что про качество данных – требования уходят в процессы (информационные потоки) или в ИТ (интеграционные потоки)
- Контроль ввода первичных данных – через проверки и идентификацию слушателя, либо человек, либо решение на уровне страны
- Вопросы управления качеством данных можно решить через Process Governance, не создавая Data Governance
- Нужна матрица – процессы/данные (до уровня атрибутов)
- Нужно учить владельцев процессов качеству данных

# Поток данных, представленный в виде матрицы

## Бизнес-процессы



# Паттерны оптимизации при автоматизации

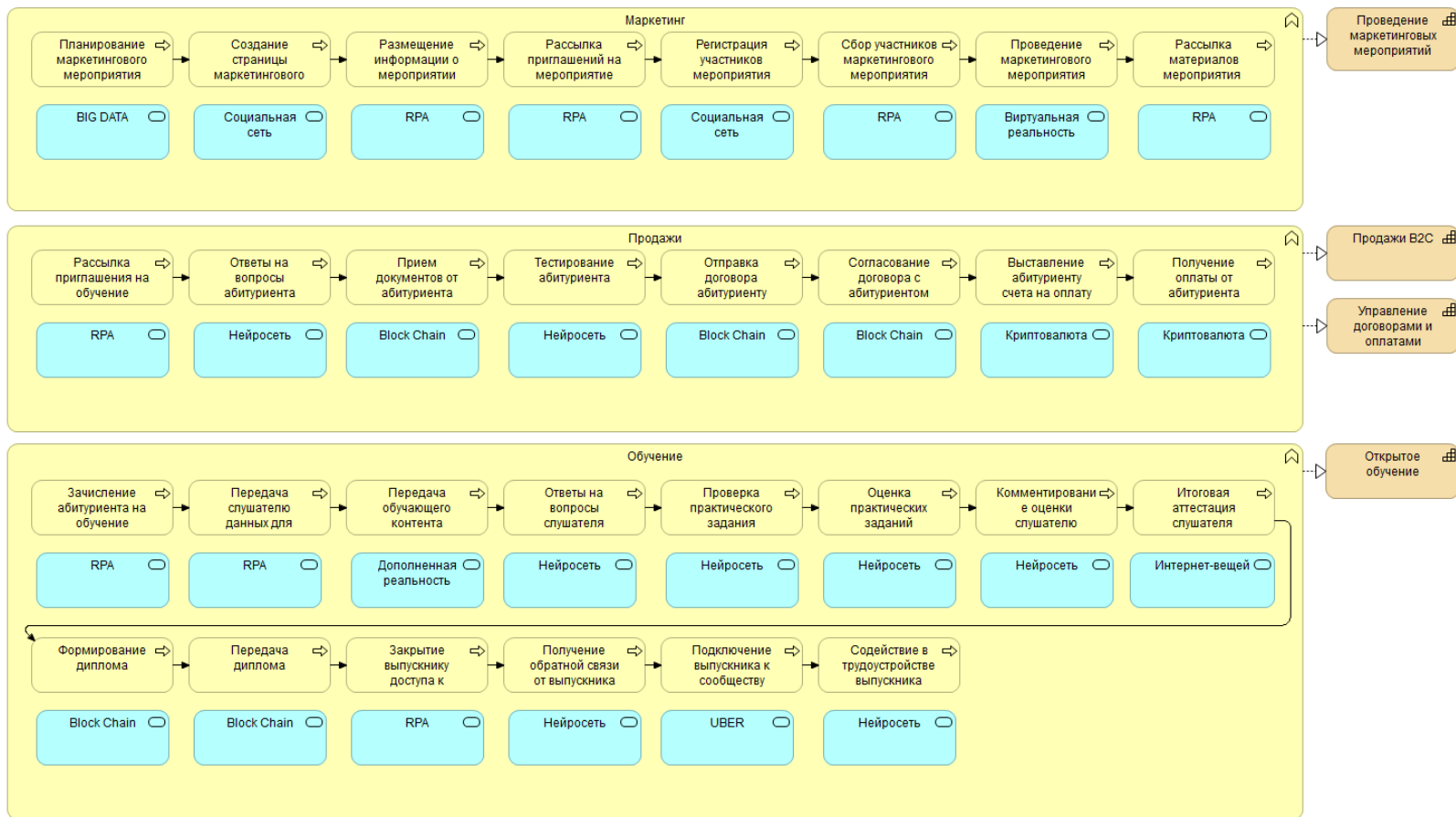
ПРИМЕР

- Искать естественные причинно-следственные связи
- Данные должны попадать в систему там, где они возникают
- Данные должны храниться в минимальном количестве мест хранения, для каждого из них должен быть определен статус данных (зачем и откуда они здесь)
- Избавляться от искусственных потоков данных, созданных специально для того чтобы управлять процессом и следить за статусом данных
- Создавать механизм управления процессом (Управлять процессом с помощью системы класса work-flow или регламента организующего процесс)
- Всем системам-потребителям данных брать данные в мастер-системе
- Избавляйтесь от документов, оставляйте только ту их составляющую, которая действительно необходима
- Избавляться от двойного ввода
- Издавайте и используйте правила, при которых действие или результат согласованы автоматически
- Делайте промежуточные результаты машиночитаемыми, и вообще автоматически обрабатываемыми

# Определение ценности данных

- Затраты на получение и хранение данных
- Затраты на восстановление данных в случае утери
- Потери организации из-за отсутствия нужных данных
- Затраты на минимизацию риска и потенциальные убытки, обусловленные рисками, связанными с данными
- Затраты на повышение качества данных
- Выгоды за счет обладания данными высокого качества
- Цена, которую конкуренты готовы заплатить за данные
- Стоимость данных в случае их продажи
- Ожидаемые доходы от инновационного использования данных

# Еще куча работы 😊



# Полезные ссылки

Группа Бизнес-  
аналитики.  
Вакансии. Резюме



Авторский канал на  
YouTube



Школа Бизнес-  
анализа Андрея  
Коптелова

