



Стратегия SAP по работе с данными

SAP HANA Data Management Suite

Денис Шувалов, SAP
Денис Савкин, SAP

14 июня 2018 г.





2007

ExxonMobil

General Electric

Microsoft

Citigroup

AT&T

Bank of America

Toyota Motor

Gazprom

Petro China

Royal Dutch Shell

2017

Apple

Alphabet

Microsoft

Amazon.com

Berkshire Hathaway

Facebook

ExxonMobil

Johnson & Johnson

JPMorgan Chase

Tencent Holdings

Данные растут вместе с ожиданиями клиентов

Цифровая вселенная



5 EB новых данных появляются каждые 2 дня¹
4 PB новых данных загружают в Facebook²

Интеллектуальные решения



1 GB данных / **sec** обрабатывает машина без водителя Google³
200 GB данных / **день** собирают с ветряка⁴

Скорость



8 NVIDIA V100 GPUs
960 TeraFLOPS⁵

Google TPU в 15-30x раз быстрее
30-80x TOPS/Ватт⁶

Ожидания пользователей



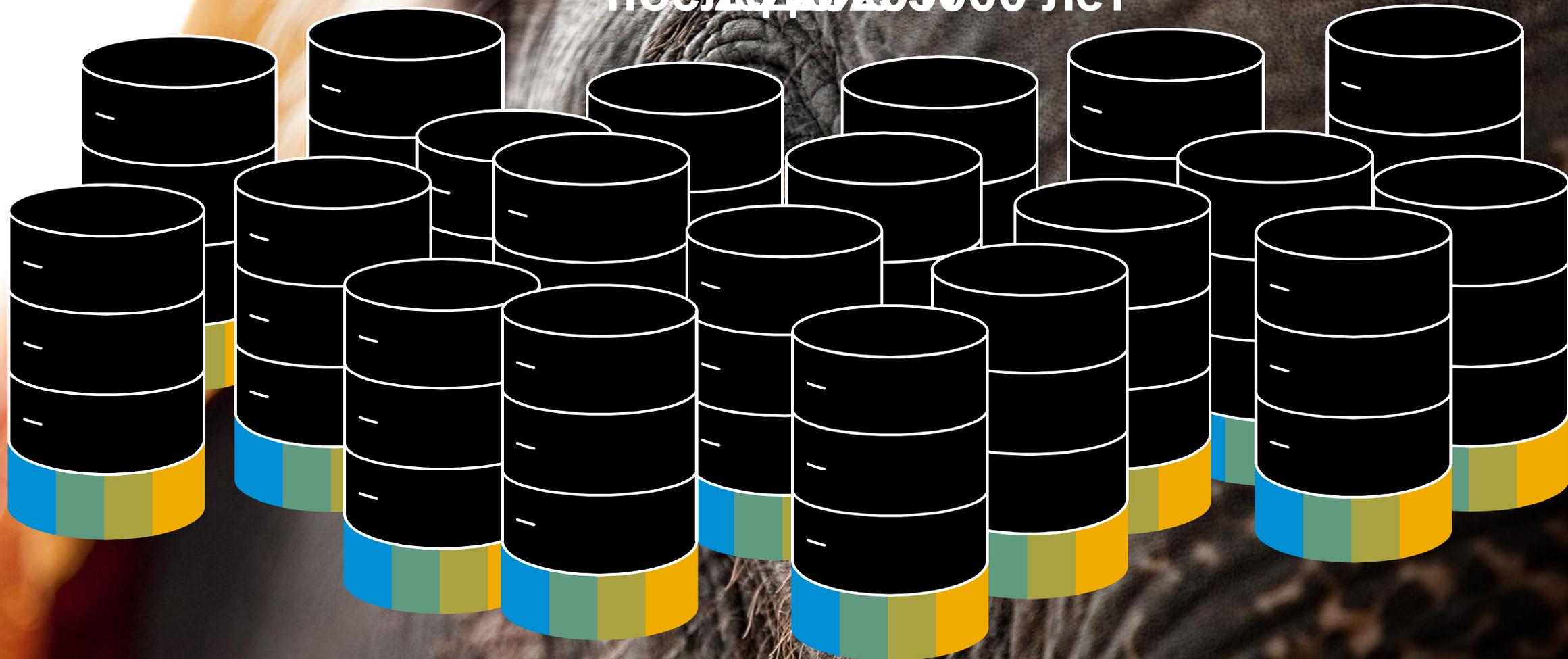
- § Как мы видим
- § Как говорим
- § Как покупаем
- § Как мы играем
- § Как спрашиваем

Современный ландшафт корпоративных данных разделен



Взрывной рост объемов данных

последние 50 лет



Очень важно посмотреть на данные «в целом», а не на их часть



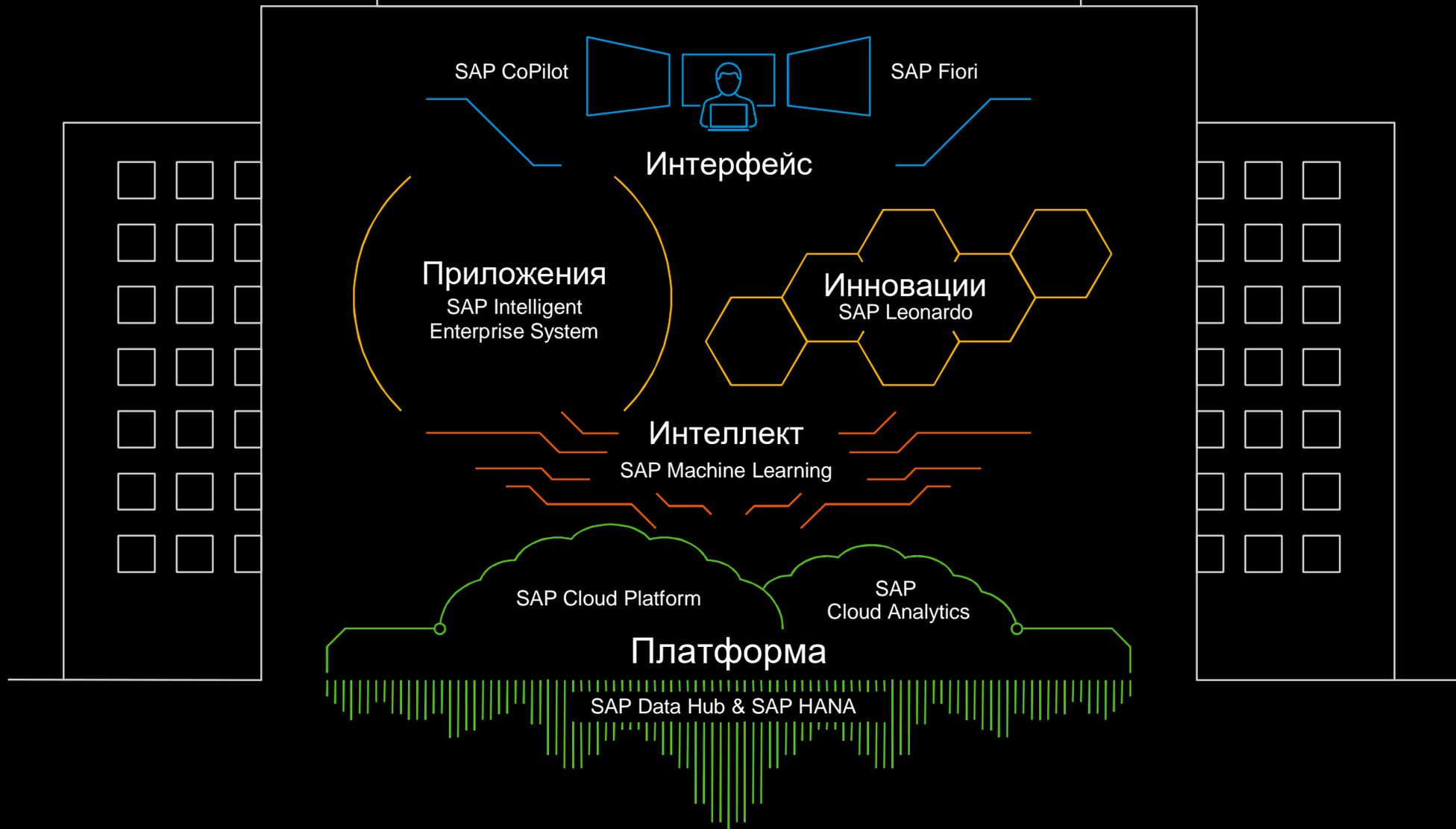
Добавим море



Добавим траву

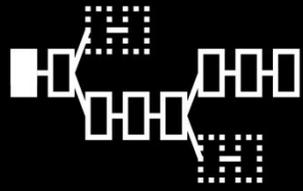


Интеллектуальная компания



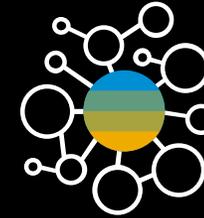
От управления предприятием по принципам системы ERP...

МАШИННОЕ ОБУЧЕНИЕ



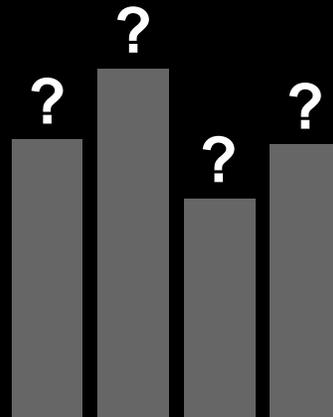
Платёжная дисциплина,
условия из договора

ИНТЕРНЕТ ВЕЩЕЙ



Информация
с оборудования и датчиков
(трансп. средства, контейнеры)

ГРАФИК ПЛАТЕЖЕЙ

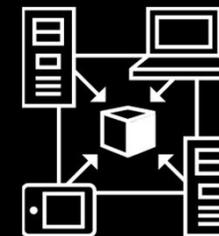


БОЛЬШИЕ ДАННЫЕ



Информация из интернета,
биржи, тренды, слухи

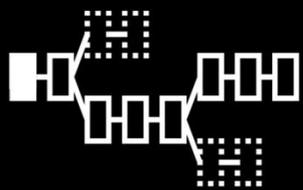
БЛОКЧЕЙН



Отслеживание перемещений
и статусов (логистика)

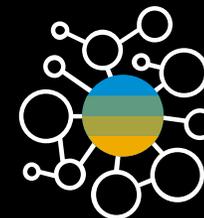
...к созданию системы **Интеллектуального Предприятия**

МАШИННОЕ ОБУЧЕНИЕ



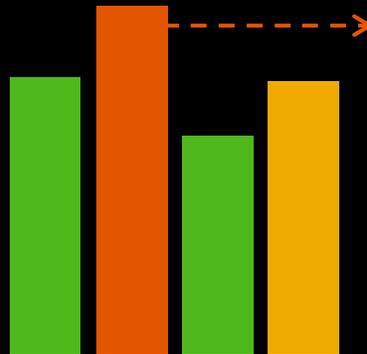
Платёжная дисциплина,
условия из договора

ИНТЕРНЕТ ВЕЩЕЙ



Информация
с оборудования и датчиков
(трансп. средства, контейнеры)

ГРАФИК ПЛАТЕЖЕЙ



БОЛЬШИЕ ДАННЫЕ



Информация из интернета,
биржи, тренды, слухи

БЛОКЧЕЙН



Отслеживание перемещений
и статусов (логистика)

Фокус 1: SAP HANA Data Management Suite

Единая платформа для создания интеллектуального предприятия

ПЛАТФОРМА SAP HANA

РАЗРАБОТКА ПРИЛОЖЕНИЙ



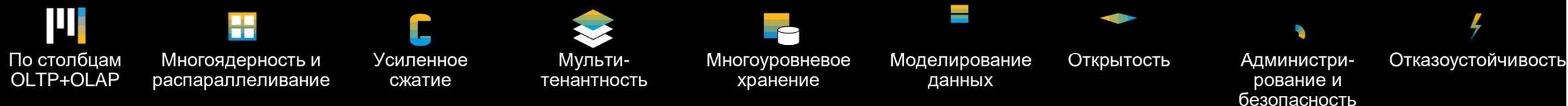
РАСШИРЕННАЯ АНАЛИТИЧЕСКАЯ ОБРАБОТКА



ИНТЕГРАЦИЯ И КАЧЕСТВО ДАННЫХ

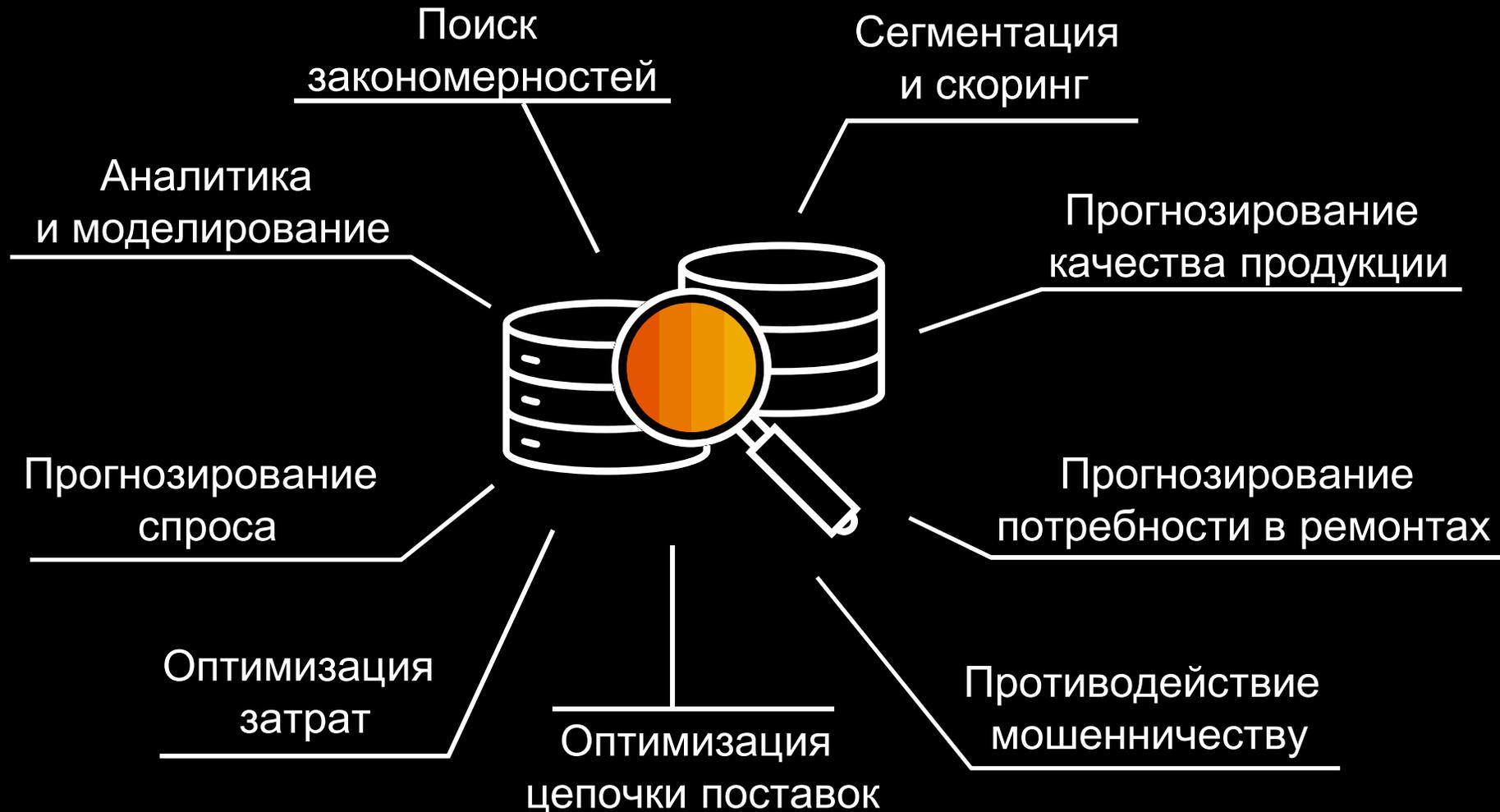


УПРАВЛЕНИЕ БАЗОЙ ДАННЫХ



SAP HANA Data Management Suite

Что можно сделать с данными?



sappi

Разрыв полотна бумаги на скорости 1000 м/мин

A large roll of glowing orange-red metal is being processed in a dark industrial setting. The metal is bright and textured, contrasting sharply with the dark background. The scene is illuminated by a strong, warm light, likely from the metal itself or nearby machinery, creating a dramatic, high-contrast environment. The background shows some industrial structures and equipment, but they are mostly in shadow.

Северсталь

SAP HANA, SAP MaxAttention

Увеличение скорости создания отчетов в **100 раз**
250% сокращение объема данных за счет
компрессии

Фокус 2: SAP Data Hub



SAP Data Hub

Унификация данных

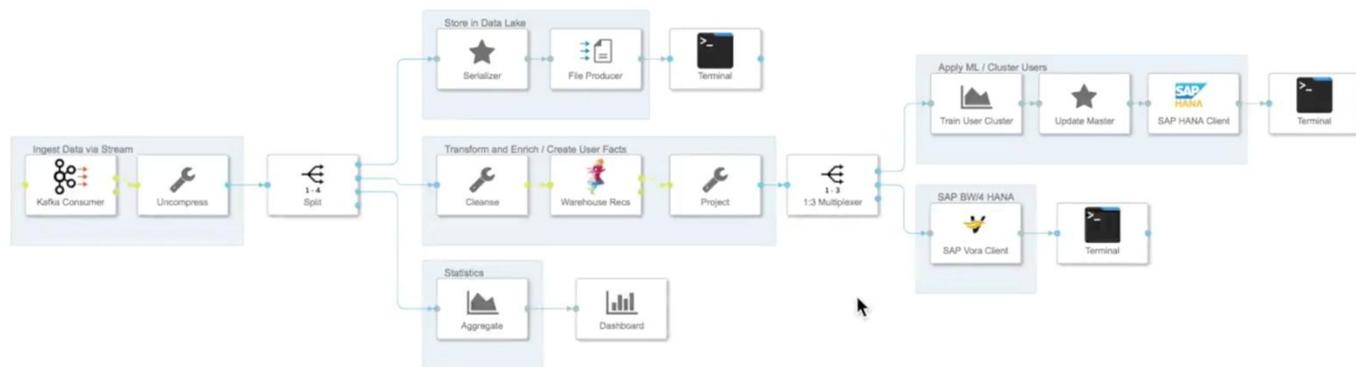
SAP Data Hub



Написание кода или **Data Hub**

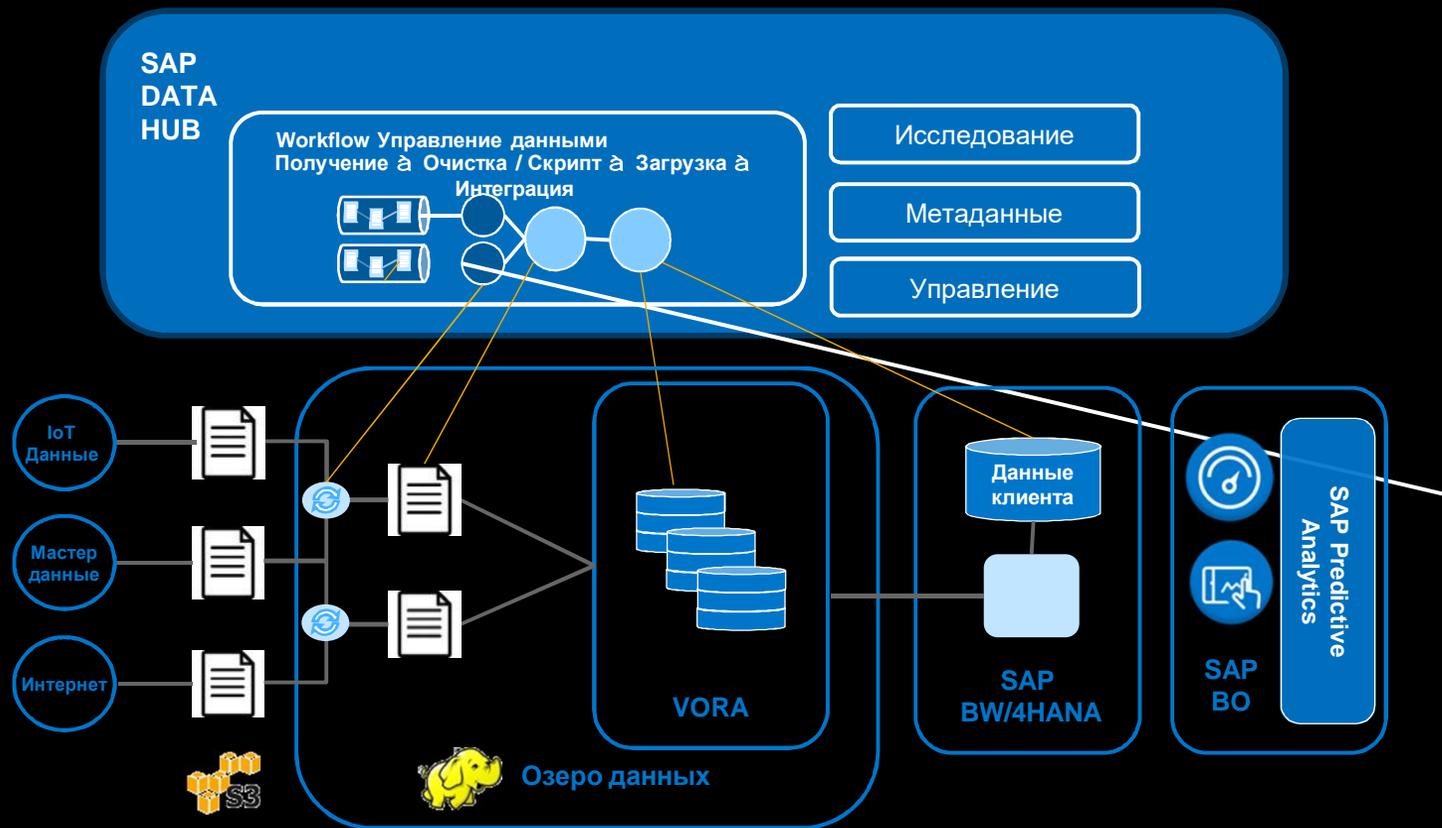
```
1 from __future__ import division
2 from collections import Counter, defaultdict
3 from functools import partial
4 from linear_algebra import shape, get_row, get_column, make_matrix, \
5     vector_mean, vector_sum, dot, magnitude, vector_subtract, scalar_multiply
6 from statistics import correlation, standard_deviation, mean
7 from probability import inverse_normal_cdf
8 from gradient_descent import maximize_batch
9 import math, random, csv
10 import matplotlib.pyplot as plt
11 import dateutil.parser
12
13 def bucketize(point, bucket_size):
14     """floor the point to the next lower multiple of bucket_size"""
15     return bucket_size * math.floor(point / bucket_size)
16
17 def make_histogram(points, bucket_size):
18     """buckets the points and counts how many in each bucket"""
19     return Counter(bucketize(point, bucket_size) for point in points)
20
21 def plot_histogram(points, bucket_size, title=""):
22     histogram = make_histogram(points, bucket_size)
23     plt.bar(histogram.keys(), histogram.values(), width=bucket_size)
24     plt.title(title)
25     plt.show()
26
27 def compare_two_distributions():
28
29     random.seed(0)
30
31     uniform = [random.randrange(-100,101) for _ in range(200)]
32     normal = [57 * inverse_normal_cdf(random.random())
```

Написание кода или **Data Hub**



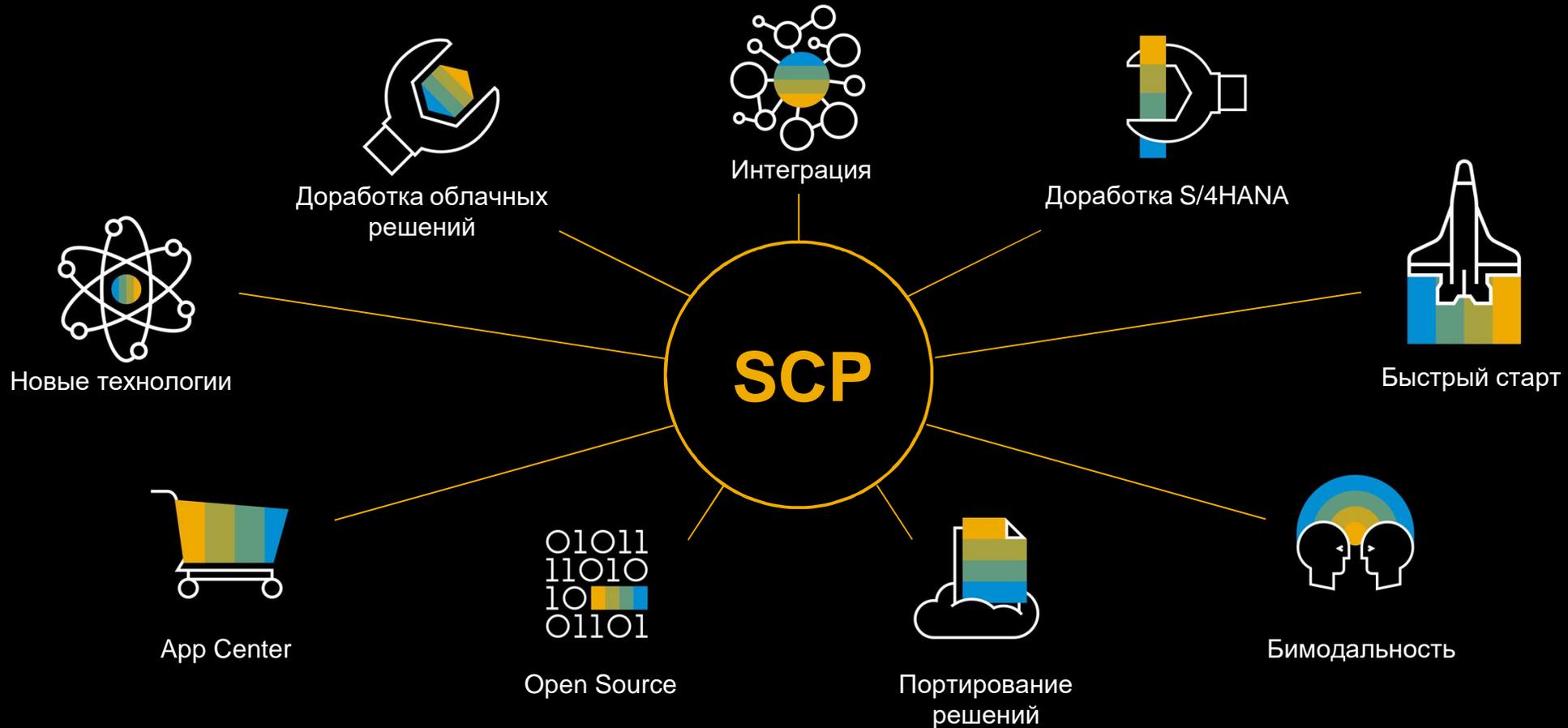
SAP Data Hub как интеграционный уровень для комплексного управления данными

Пример архитектуры клиента



- Машинные данные загружаются в AWS S3
- Данные (файлы) сохраняются в «озере данных» и передаются в хранилище
- Очищенные данные передаются в Vora и могут обрабатываться в оперативной памяти
- Данные теперь могут быть использованы в обычном SQL запросе
- По финальным данным строится отчетность в SAP BO
- Расширенная аналитика используется для построения прогнозных моделей (получение данных из Vora)

Фокус 3: SAP Cloud Platform

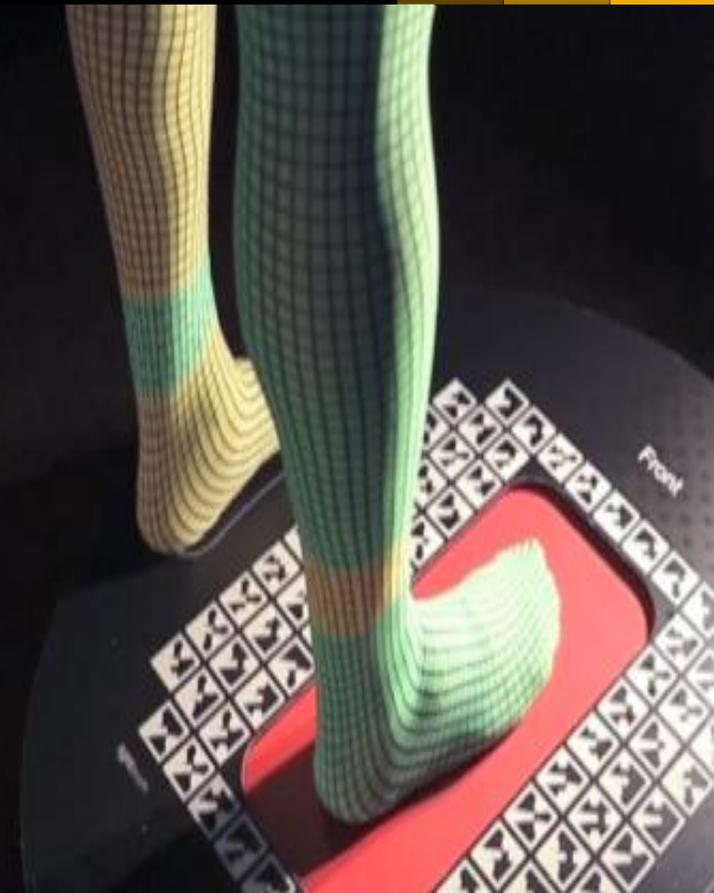


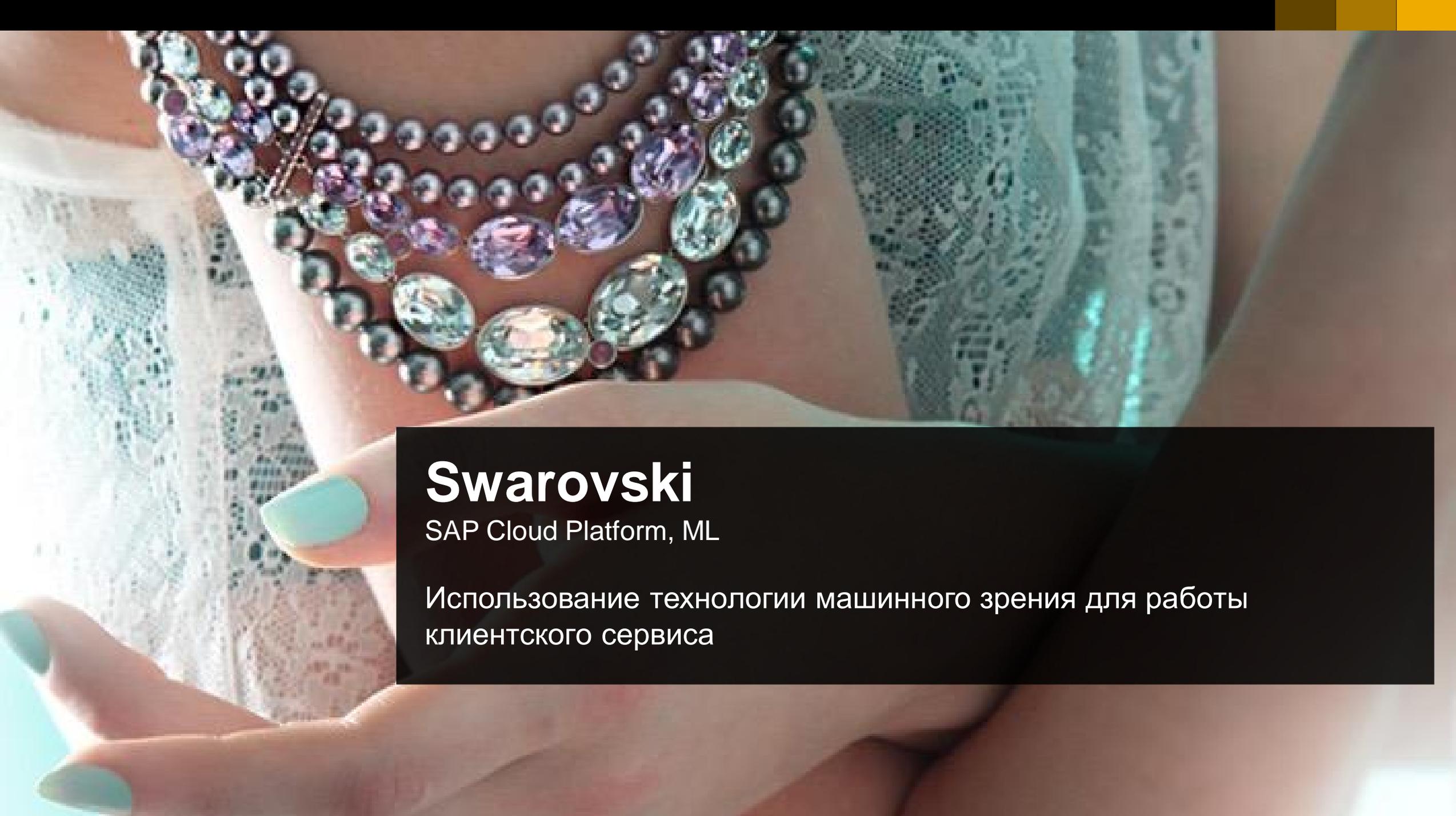


Risch

SAP Cloud Platform

**Использование концепции «цифрового двойника»
при производстве обуви**



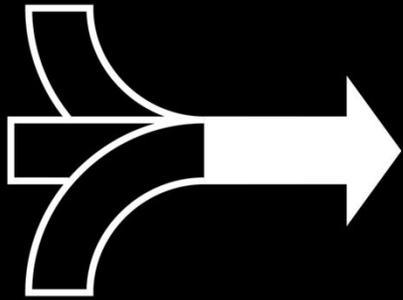
A close-up photograph of a hand holding a Swarovski necklace. The necklace features a string of pearls with several large, faceted crystals interspersed. The hand holding the necklace has light blue nail polish. The background is a white lace fabric. A black text box is overlaid on the bottom right of the image.

Swarovski

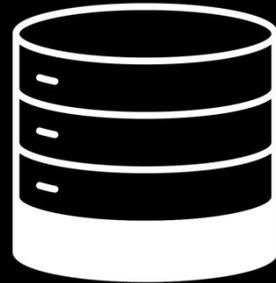
SAP Cloud Platform, ML

Использование технологии машинного зрения для работы клиентского сервиса

SAP Data Management Suite - технологии, которые работают вместе



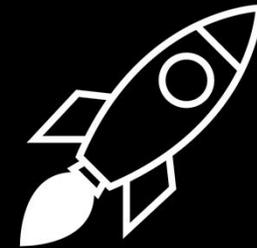
Собирать



Хранить



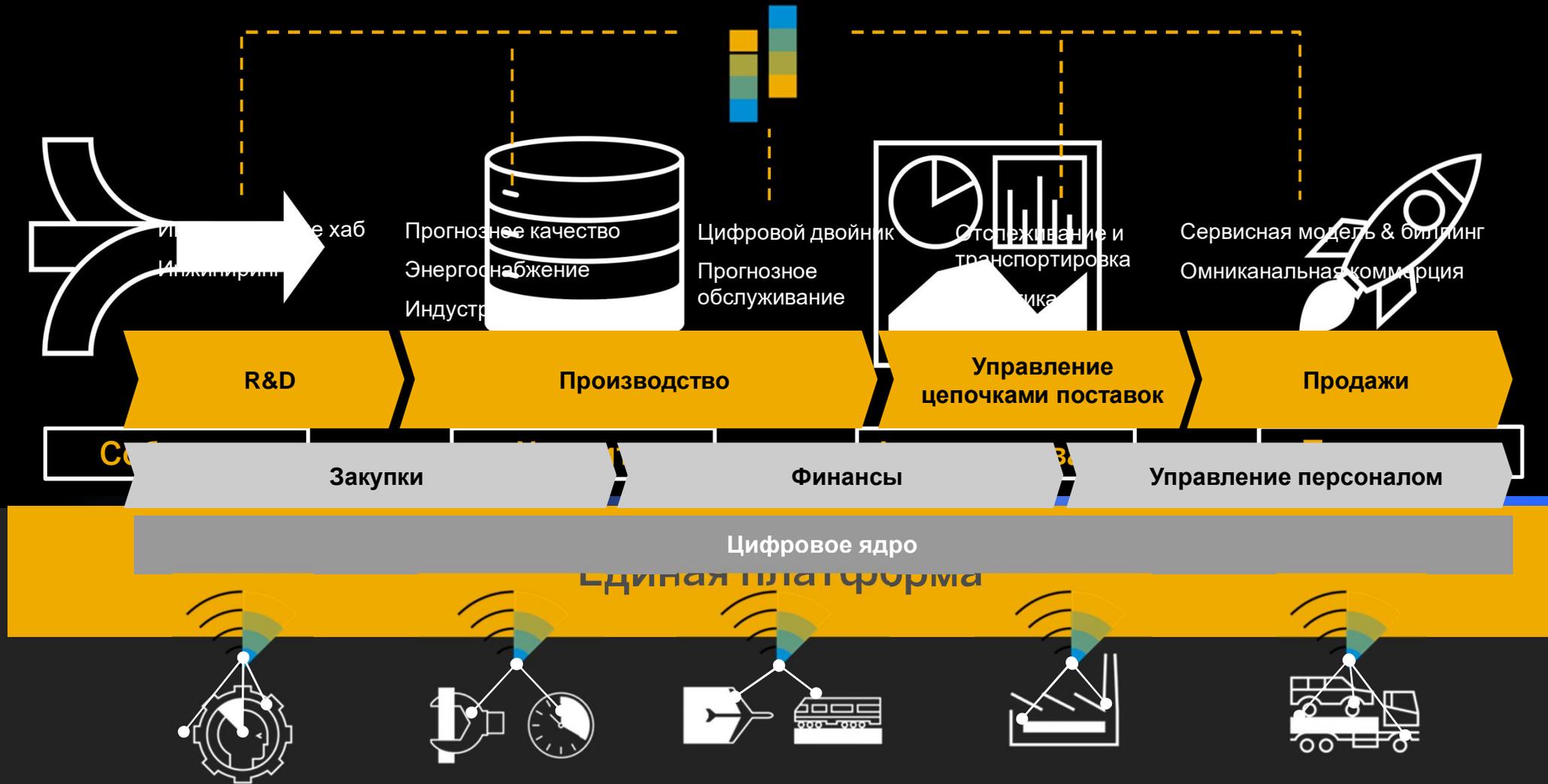
Анализировать



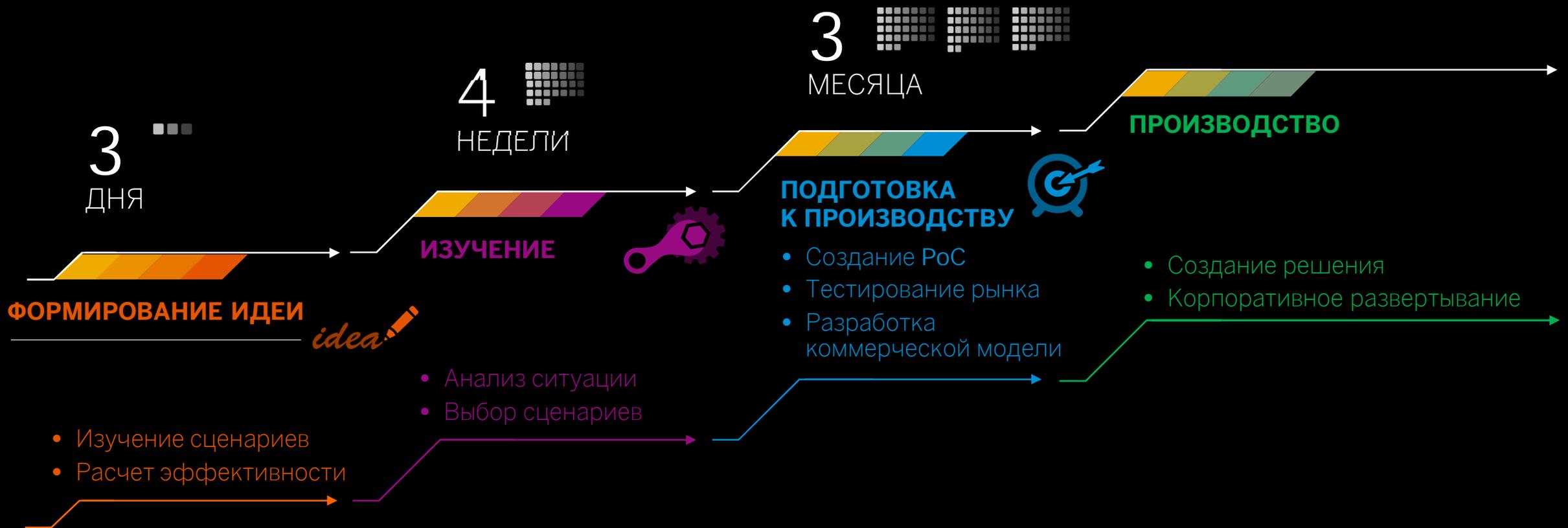
Применять

Единая платформа

Новые бизнес сценарии



Путь инноваций: от идеи к результату



Спасибо!

Контактная информация:

Денис Шувалов

Руководитель направления Database and Data
Management

denis.shuvalov@sap.com

Денис Савкин

Руководитель Центра Экспертизы SAP:
решения и технологии

denis.savkin@sap.com