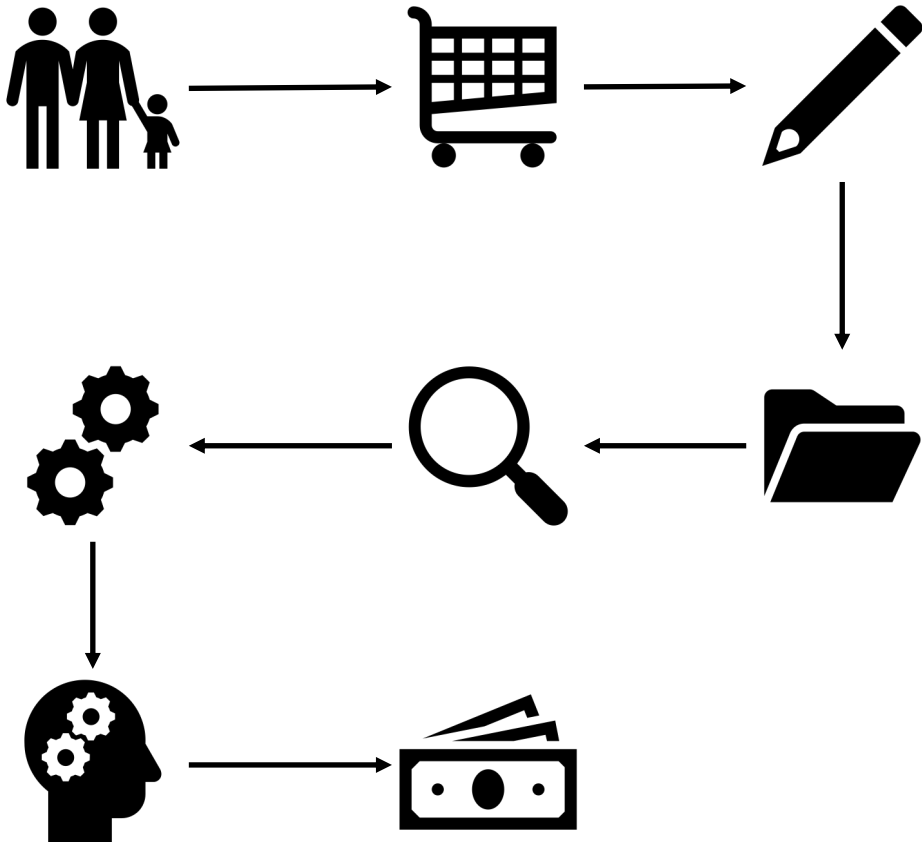


Технологии SAP для Data Science проектов в Банках

Антон Ноздрин, SAP
May 23, 2019

PUBLIC

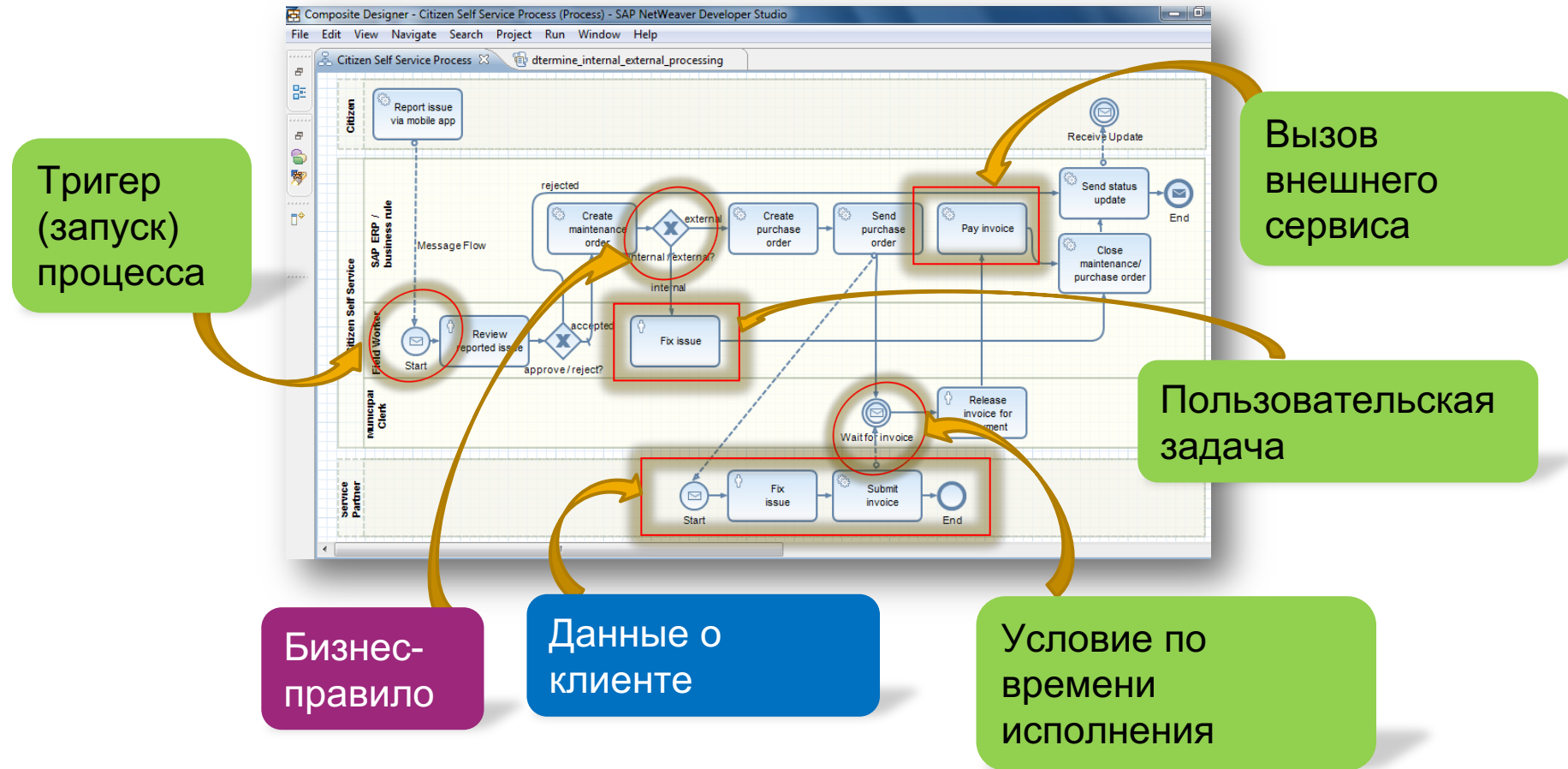
Что может быть проще кредитного конвейера



- База клиентов
- Подбор продукта
- Заполнение заявки
- Работа с внутренними и внешними источниками
- Верификация данных клиента
- Андеррайтинг
- Принятие решения
- Сделка

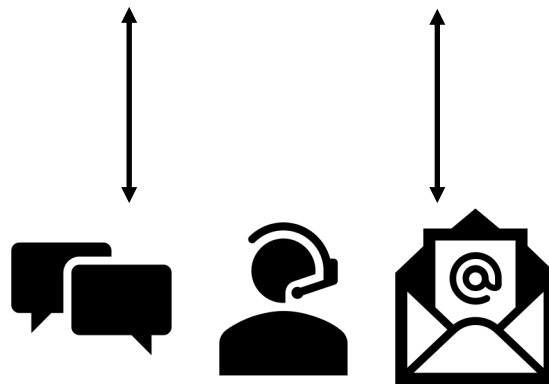
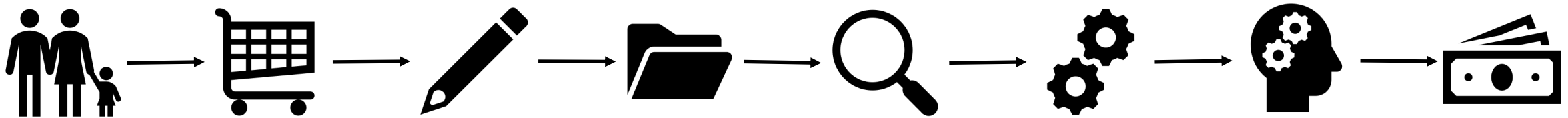
Бизнес-процесс кредитного конвейера

Легко реализуется продуктами SAP



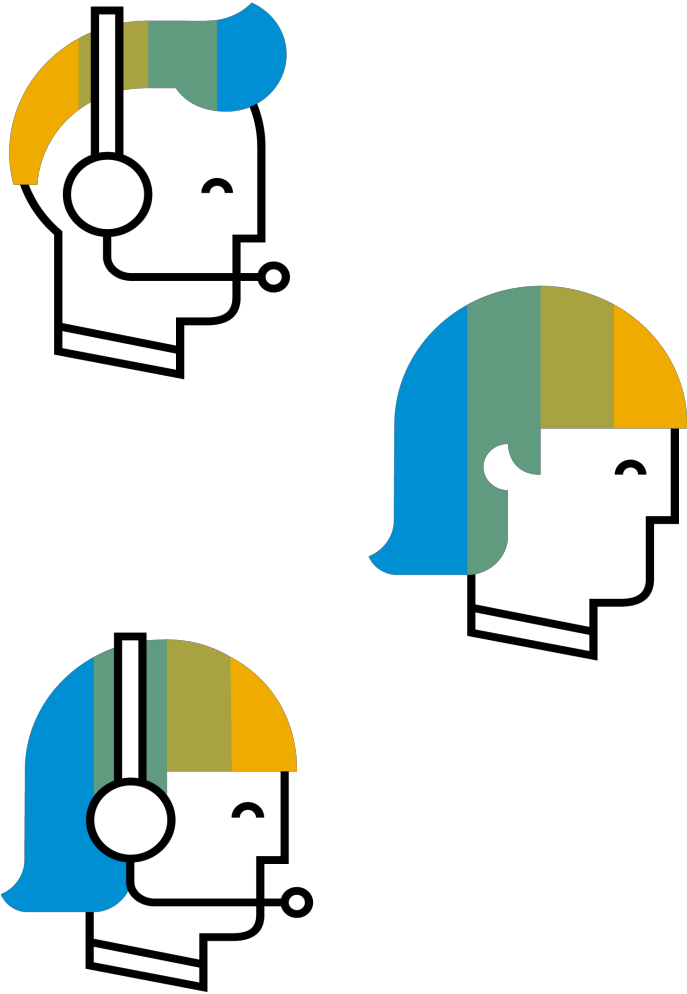
Кредитный конвейер

Работа с источниками
Верификация данных



CRM-система

Что не так?



Все правильно, но

- Люди – работают медленно и нестабильно
 - Люди неодинаковы
 - Люди подвержены эмоциям
 - Люди склонны к коррупции
-
- В схеме кредитного конвейера ответственность часто размывается

Один из наших клиентов (европейский банк)



Покупатель сканирует QR-код в офлайн-магазине в мобильном приложении Банка

Банк в течение 1-2 минут одобряет/не одобряет кредит

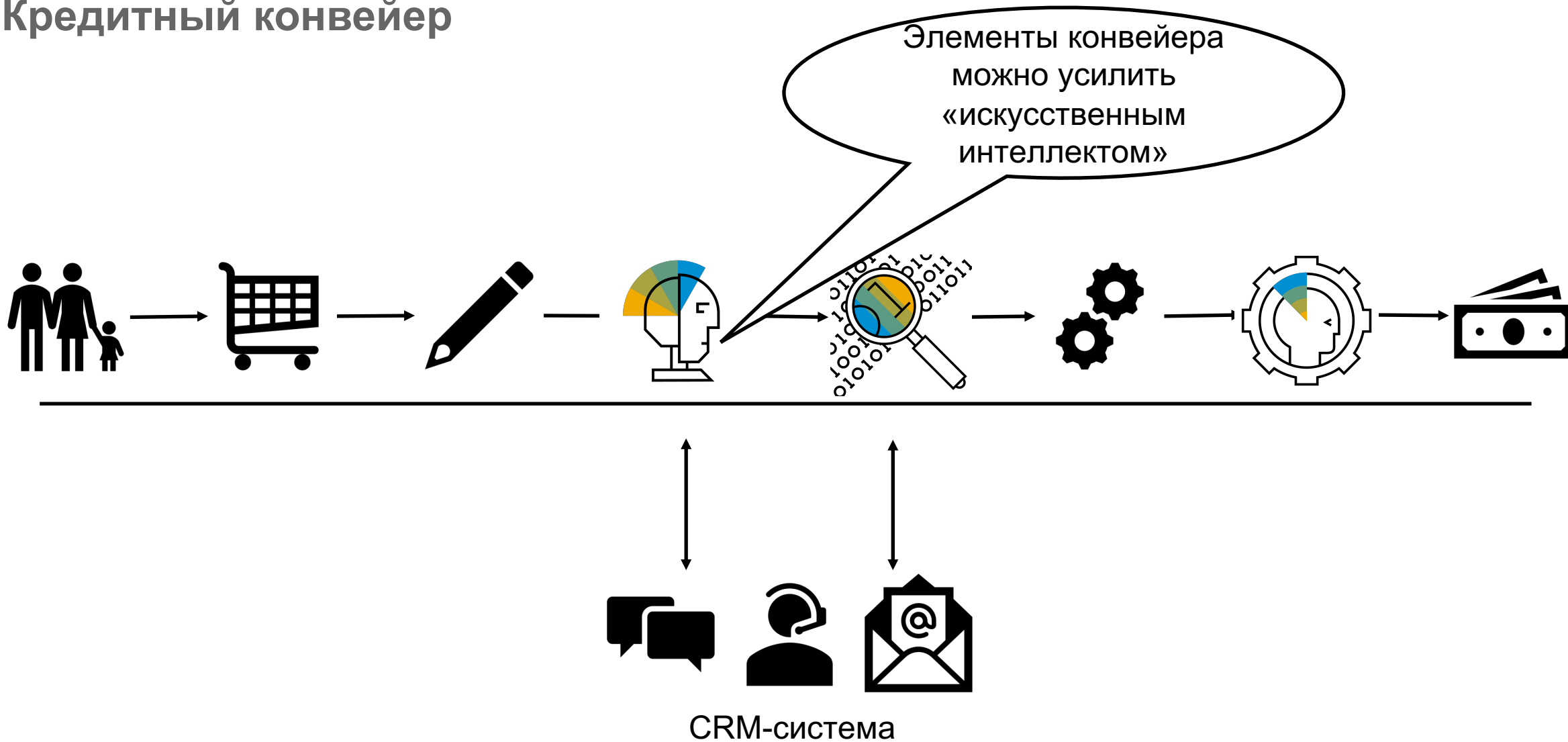
Банк переводит деньги сразу на счет магазина

Покупатель уезжает с покупкой

Можно ли это
сделать
конвейером
из людей?



Кредитный конвейер



Таких «конвейеров» в банке много

Основной бизнес:

- Маркетинг и продажи
- Сервис и колл-центр
- Риски
- Скоринг

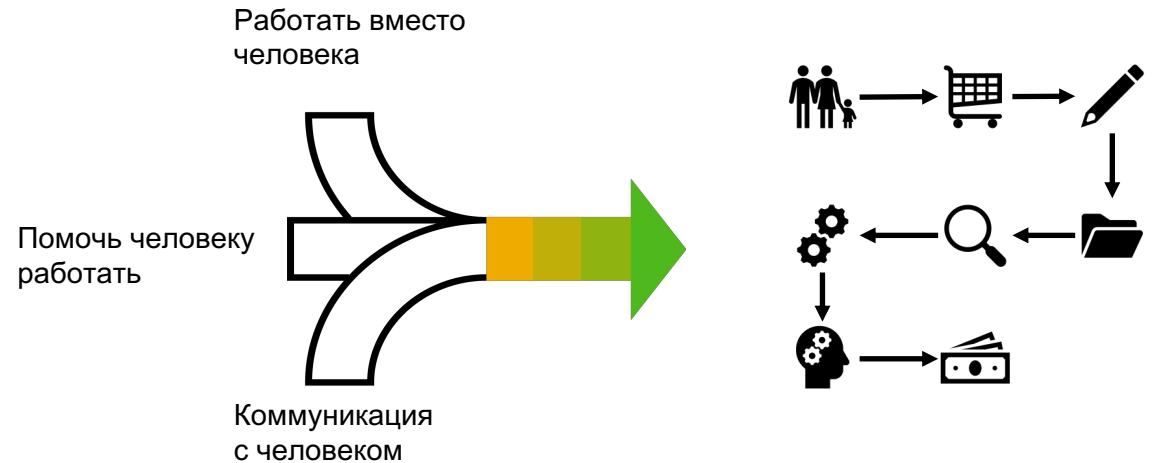
Общехозяйственная деятельность:

- Платежи (дебиторы/кредиторы)
- Учет ТМЦ
- Управление мастер-данными
- Закупки

HR

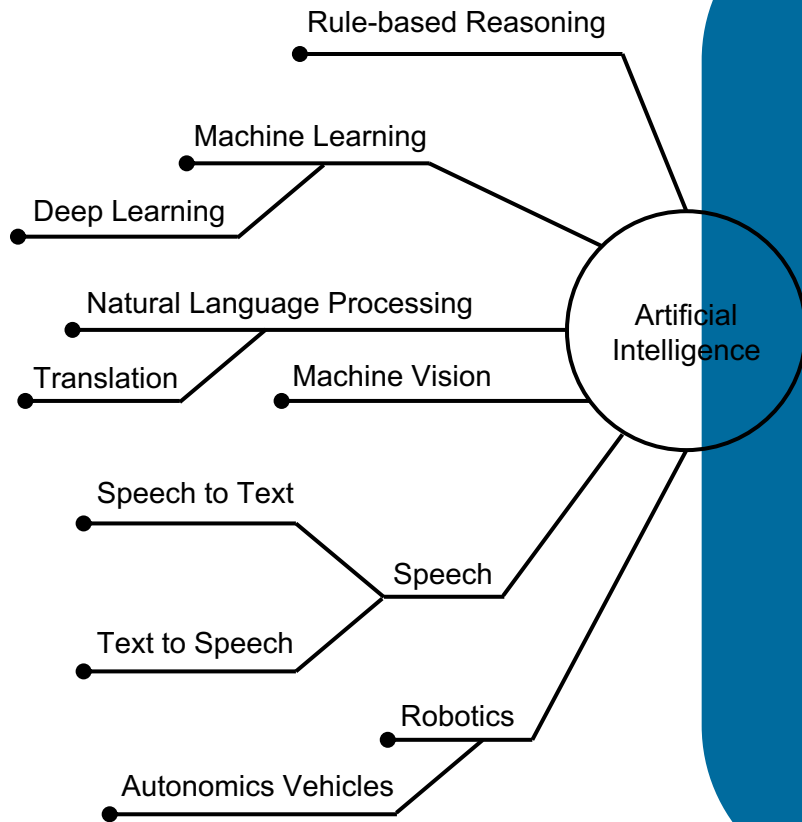
- Найм и увольнение
- Развитие сотрудников

...



Определения: AI, ML и DL

ИИ области исследования



Искусственный Интеллект (AI)

- Интеллект, демонстрируемый машинами
- Любая симуляция человеческого интеллекта
- Область исследований, разработок и инвестиций
- Роботизация, принятие решений по правилам, обработка естественного языка и т.д.

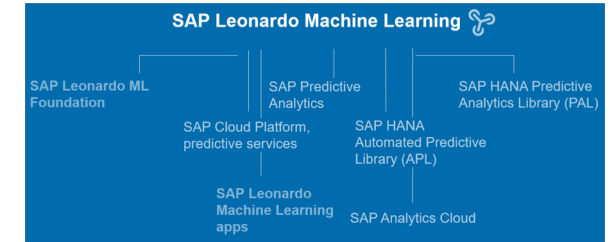
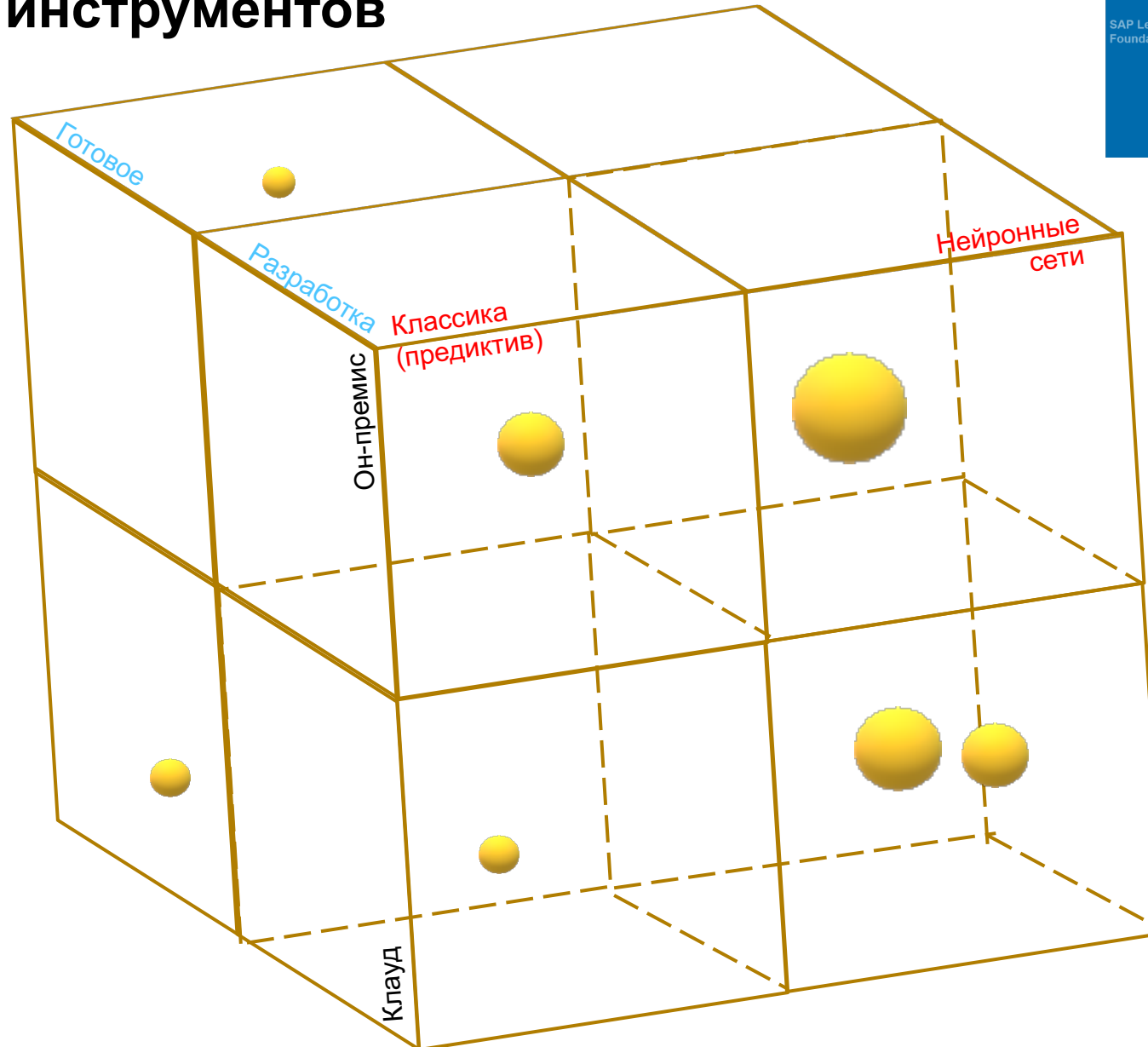
Машинное обучение (ML)

- Область AI которая заставляет машины решать задачи без императивного программирования
- Использует статистические методы, включая нейронные сети

Глубинное обучение (DL)

- Область ML изучающая специализированные алгоритмы, как правило, многослойные нейронные сети
- Распознавание разных сущностей (форма-объект-сцена и т.д.)

Классификация инструментов



Самая лучшая классификация

Машинное обучение

Citizen Data Scientist

Короткий проект

Достаточное качество модели

Простые инструменты

Сложные инструменты

Профессиональный Data Scientist

Длинный проект

Высокое качество модели

Сложные инструменты

«Сложные» инструменты

Персона: «Профессиональный» data scientist

- предпочитает opensource
- имеет богатый опыт применения моделей
- нередко не интересуется бизнес-результатом

ПО:

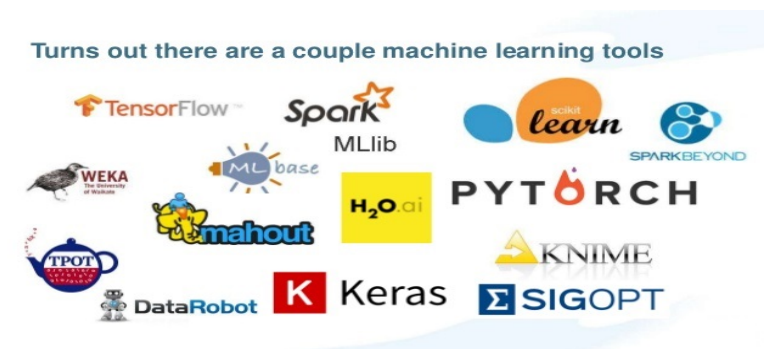
- opensource
- платные инструменты экспертного уровня

Результат:

- можно сделать все, что вообще возможно
- качество высокое и сверхвысокое

Трудозатраты:

- проект (одна модель в 2-3 месяца)



Сложные инструменты

1. **SAP HANA PAL, R, Tensorflow (on-premise и cloud)**
2. **SAP Predictive Analytics (Expert Modeler)**
3. **Opensource: Jupyter Notebooks/TensorFlow/Keras – нейронные сети**
4. **Opensource: R – предиктивная аналитика**

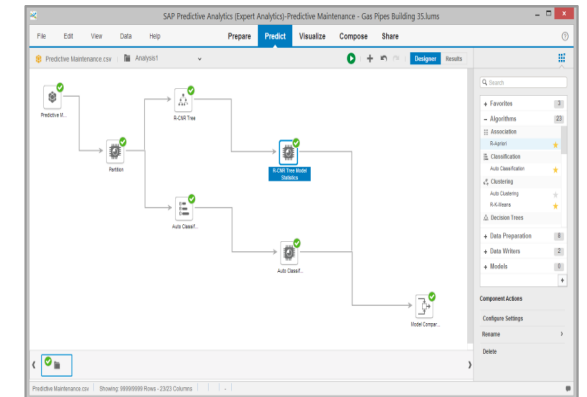
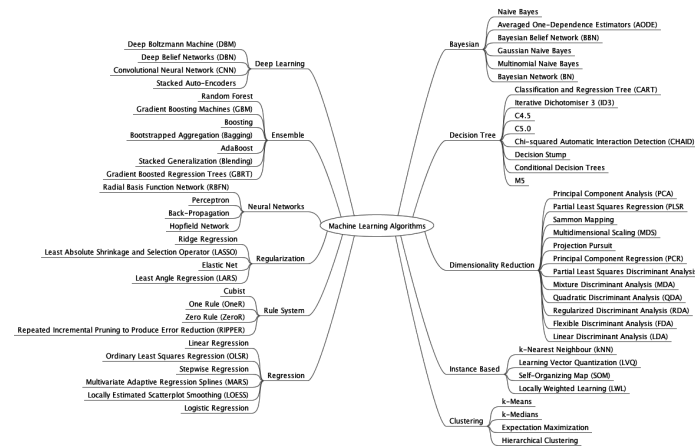
```

jupyter Vanilla+CNN-on-MNIST+dataset Last Checkpoint: 03/26/2019 (Unsaved changes)
File Edit View Insert Cell Kernel Help Python 3.0
In [4]: x_train = x_train.reshape(x_train.shape[0], 1, img_rows, img_cols)
x_test = x_test.reshape(x_test.shape[0], 1, img_rows, img_cols)
input_shape = (1, img_rows, img_cols)

In [5]: x_train = x_train.astype('float32')
x_test = x_test.astype('float32')
x_train = x_train / 255
print(x_train.shape, 'x_train.shape')
print(x_test.shape, 'x_test.shape')
x_train_shape = (x_train.shape[0], 1, 28, 28)
10000 train samples

In [6]: y_train = np.zeros((x_train.shape[0], num_classes))
y_test = np.zeros((x_test.shape[0], num_classes))

Build the CNN
In [7]: model = Sequential()
model.add(Conv2D(32, kernel_size=(3, 3), kernel_size=(3, 3),
border_mode='valid',
input_shape=input_shape))
model.add(MaxPooling2D(pool_size=(2, 2), border_mode='valid'))
model.add(Conv2D(64, kernel_size=(3, 3), kernel_size=(3, 3)))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Flatten())
model.add(Dense(100))
model.add(Dense(10))
model.compile(loss='categorical_crossentropy',
optimizer='adam',
metrics=['accuracy'])
    
```



SMART РЕКРУТИНГ



Статус проекта
Proof-of-Concept



Заказчик
Крупный банк

Система позволяет автоматически ранжировать поступающие отклики на вакансии. При ее использовании рекрутеры не упустят релевантного кандидата в **83%** случаев



SMART КАРЬЕРА



Статус проекта
Proof-of-Concept



Заказчик
Крупный банк

На основе анализа данных модель выдает потенциальные должности в порядке вероятности успешного перехода и рекомендации по «зонам роста» для каждого конкретного кандидата

Предиктивная аналитика в действии



Новый продукт: SAP Data Intelligence

Components



SAP Data Hub

Data sharing, pipelining, and orchestration. Including data preparation and cleaning.



SAP Predictive Analytics

Operationalization and automation



SAP Leonardo ML

Deep Learning (text, image, audio)



SAP HANA ML

In-database Machine Learning

Solution

Data Science платформа

- One integrated offering
- One data science front end
- Full lifecycle management
- Integrated with SAP

Простые инструменты

«Простые» инструменты

Персона: Citizen data scientist

- имеет бизнес-задачу или задачу по разработке ПО
- не интересуется математикой

ПО:

- Специализированные инструменты для неData Scientist

Результат:

- можно сделать многое (но не все)
- качество высокое

Трудозатраты:

- нанопроjekt (десятки моделей в день)

«Простые» инструменты

SAP Predictive Analytics (Automated modeler)

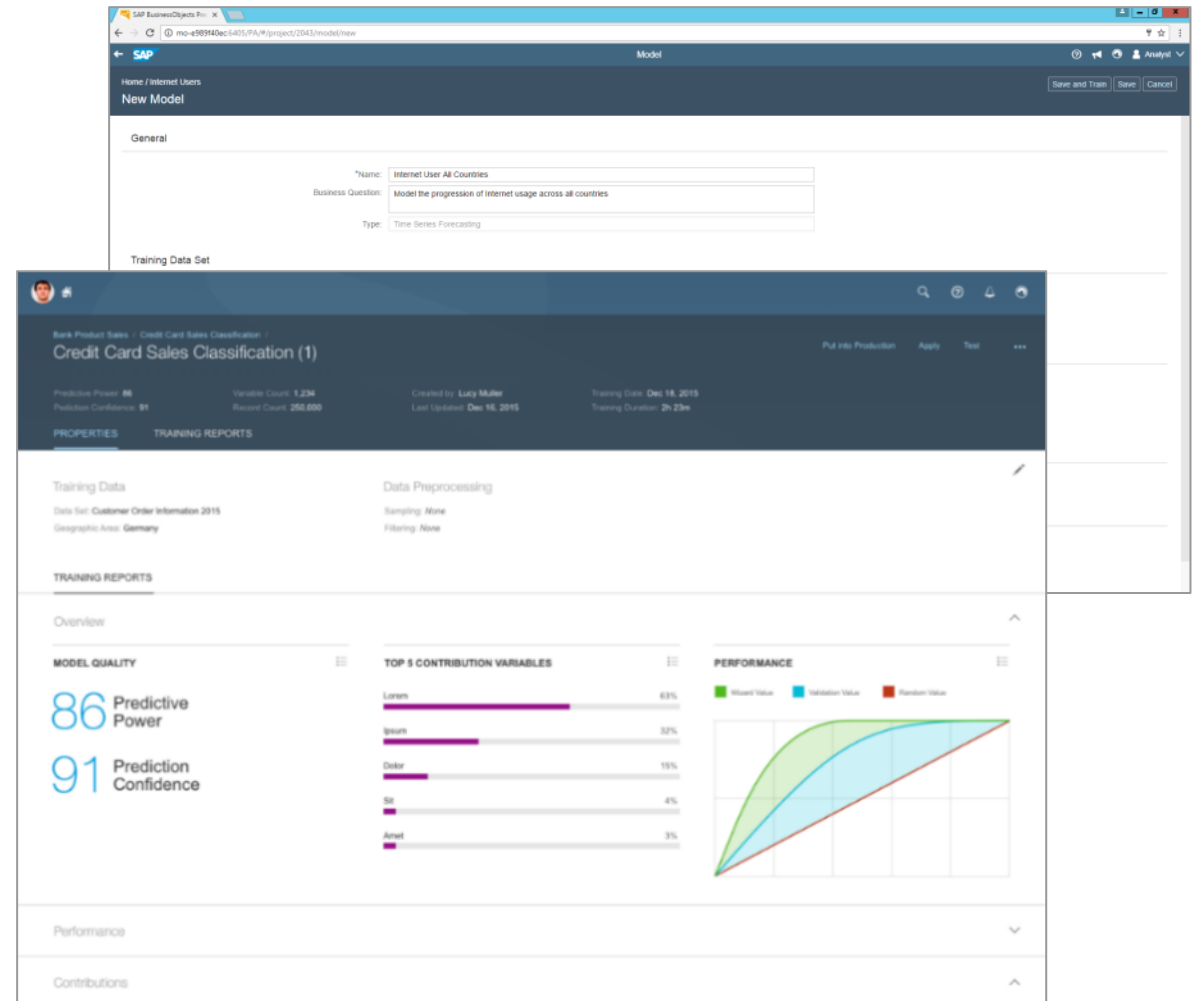
SAP Leonardo ML Foundation

SAP Analytics Cloud

1. SAP Predictive Analytics Automated Modeler



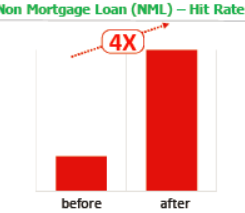
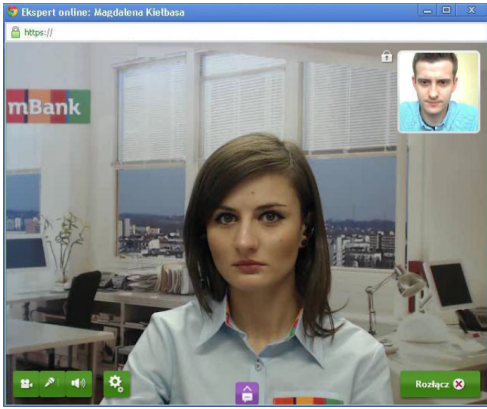
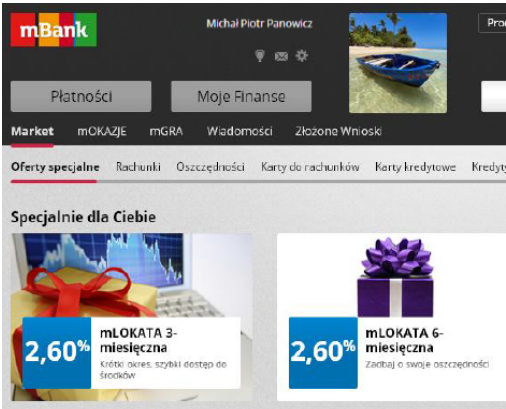
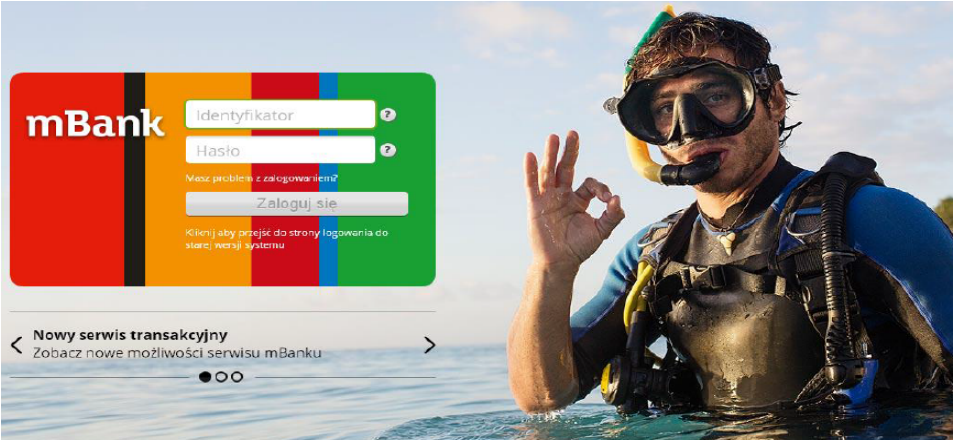
- Более новая математика
- Автоматическая оптимизация
- Автоматическая подготовка данных
- Любые источники данных



Предиктивная аналитика в банке



Выдача кредитов за 30 секунд



2. SAP Cloud Platform ML Foundation (нейронные сети)



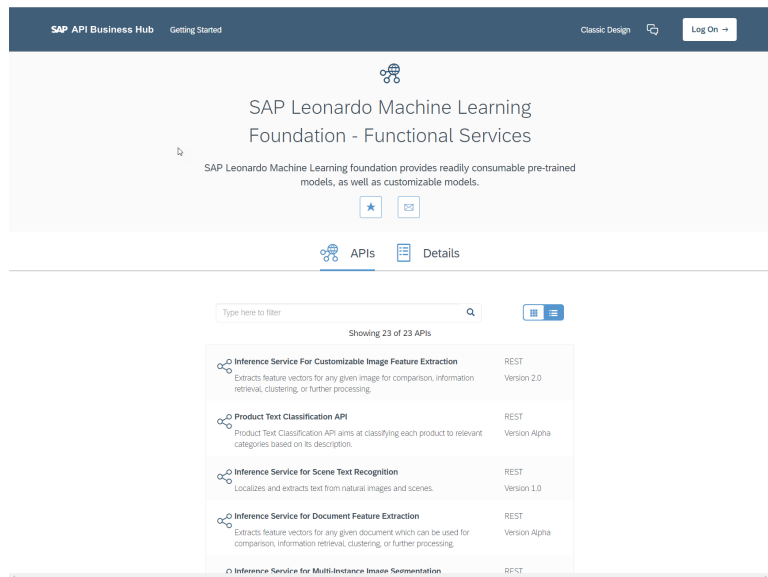
Image



Speech



Text



SAP API Business Hub Getting Started Classic Design Log On

SAP Leonardo Machine Learning Foundation - Functional Services

SAP Leonardo Machine Learning foundation provides readily consumable pre-trained models, as well as customizable models.

APIs Details

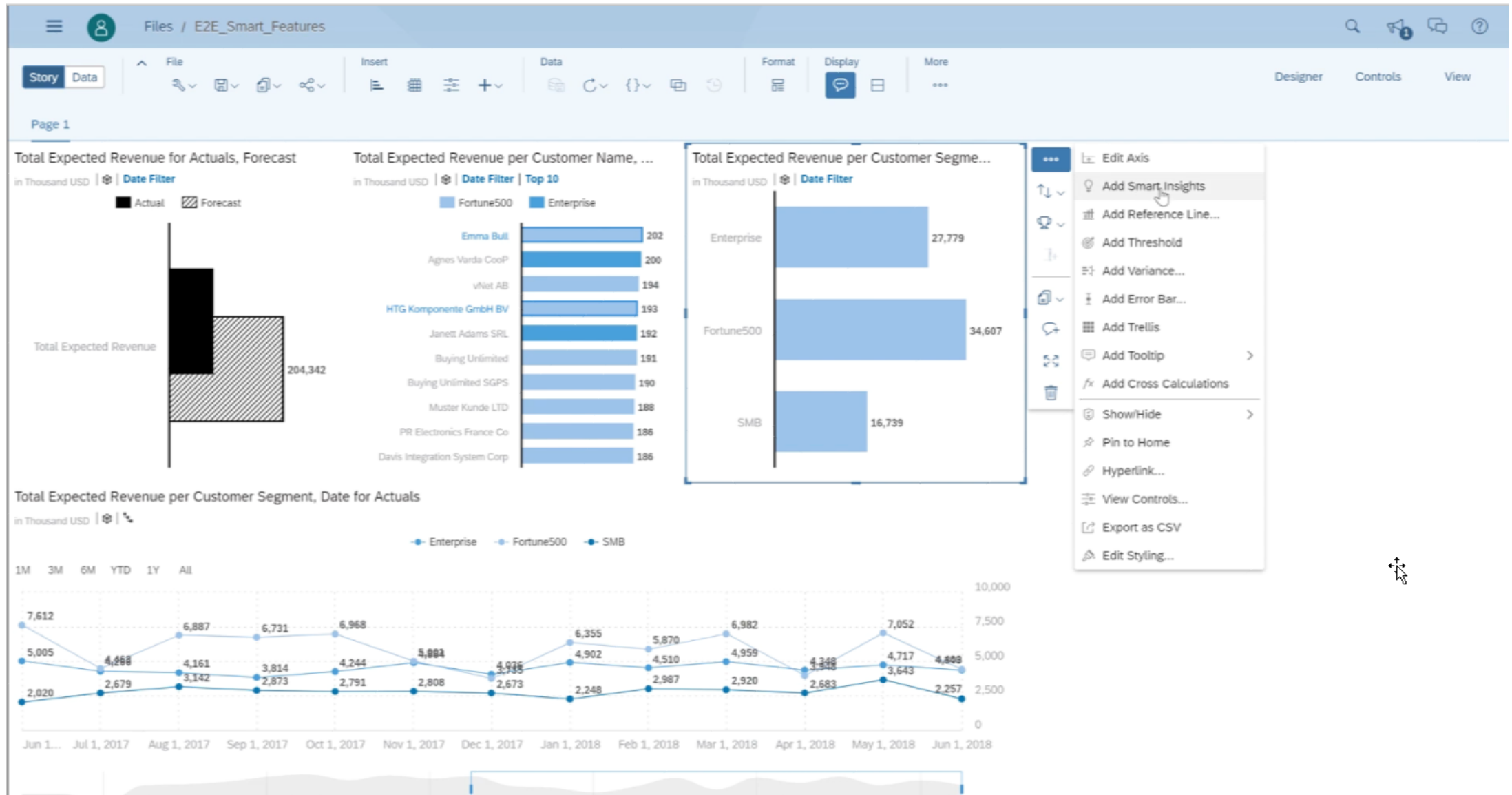
Type here to filter Showing 23 of 23 APIs

Inference Service For Customizable Image Feature Extraction	REST
Extracts feature vectors for any given image for comparison, information retrieval, clustering, or further processing.	Version 2.0
Product Text Classification API	REST
Product Text Classification API aims at classifying each product to relevant categories based on its description.	Version Alpha
Inference Service for Scene Text Recognition	REST
Localizes and extracts text from natural images and scenes.	Version 1.0
Inference Service for Document Feature Extraction	REST
Extracts feature vectors for any given document which can be used for comparison, information retrieval, clustering, or further processing.	Version Alpha
Inference Service for Multi-Instance Image Segmentation	REST

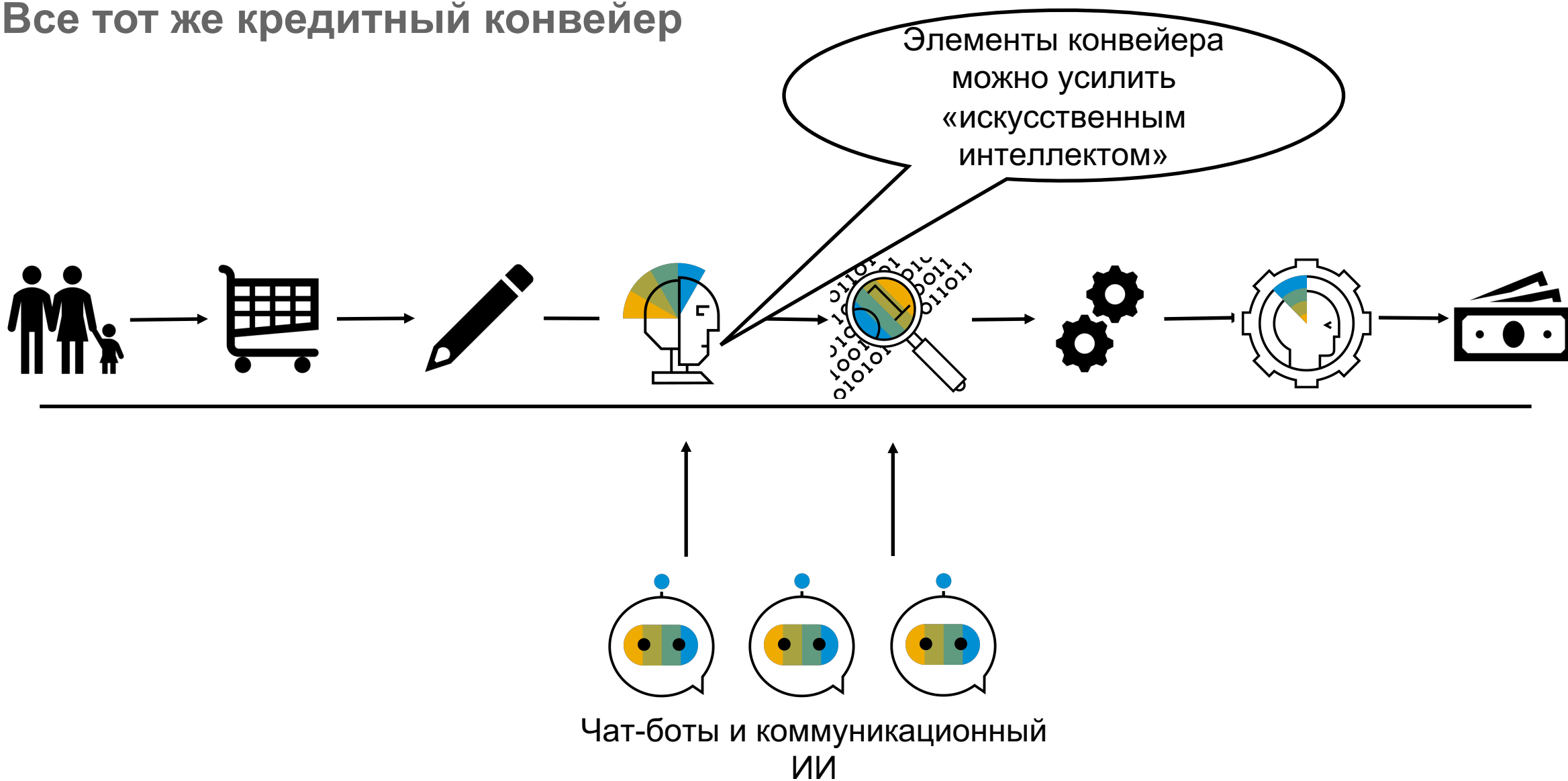
Использование когнитивных сервисов машинного обучения



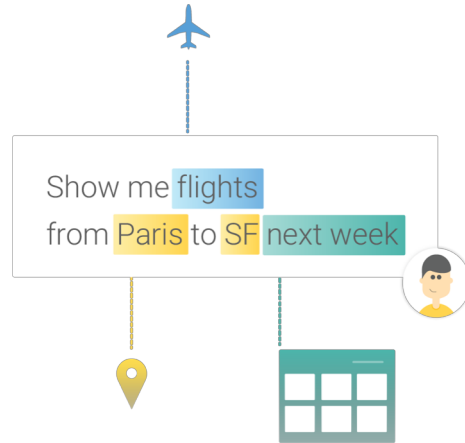
3. SAP Analytics Cloud



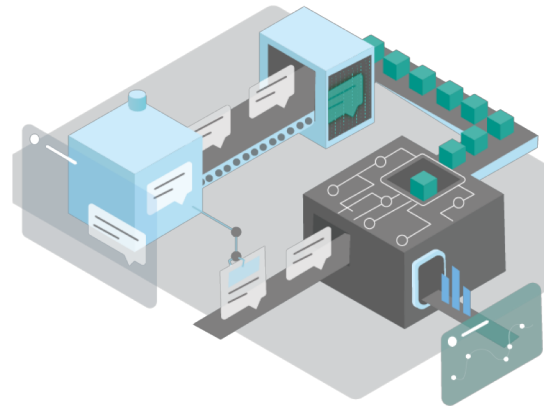
Все тот же кредитный конвейер



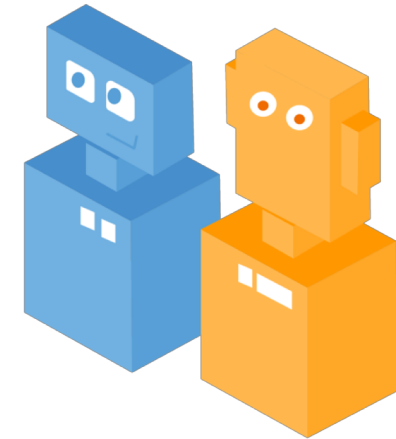
SAP Conversational AI: Коммуникационный ИИ



Работа с естественным
языком



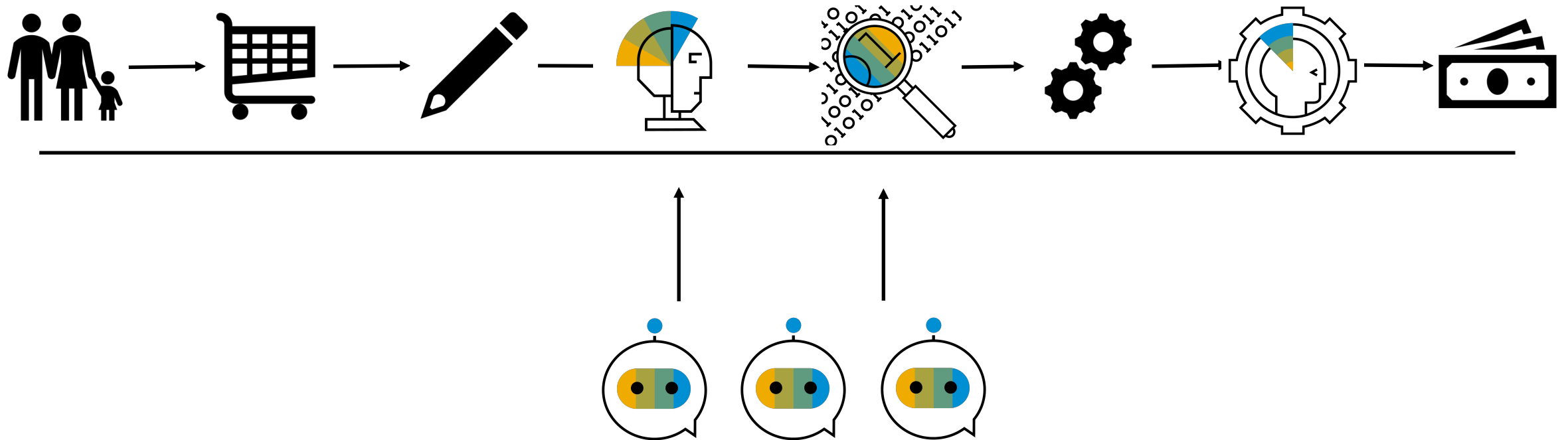
Платформа для
построения ботов



Автоматизированные
отраслевые решения

Smarter, Faster, Self-Learning

Кредитный конвейер

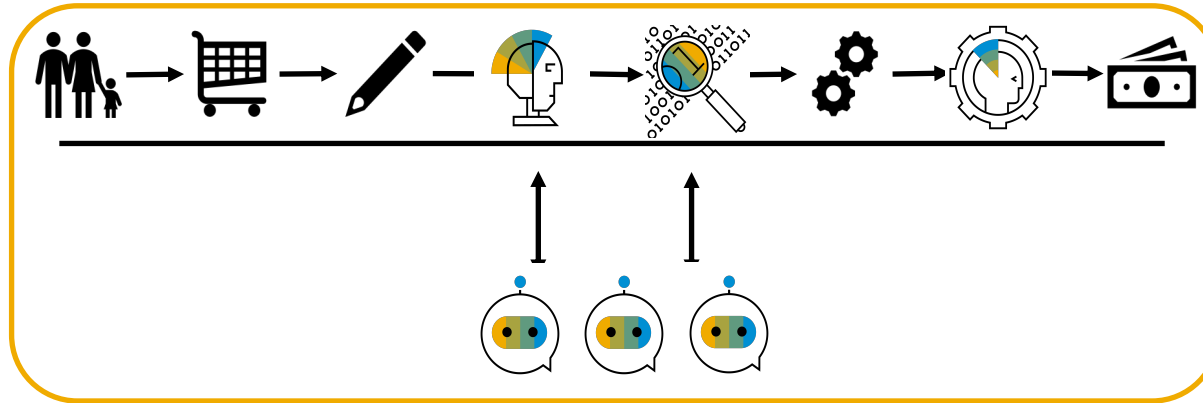


Автономное предприятие

МАРКЕТИНГ



КРЕДИТНЫЙ КОНВЕЙЕР



ПОСЛЕПРОДАЖНЫЙ
СЕРВИС И АНАЛИТИКА

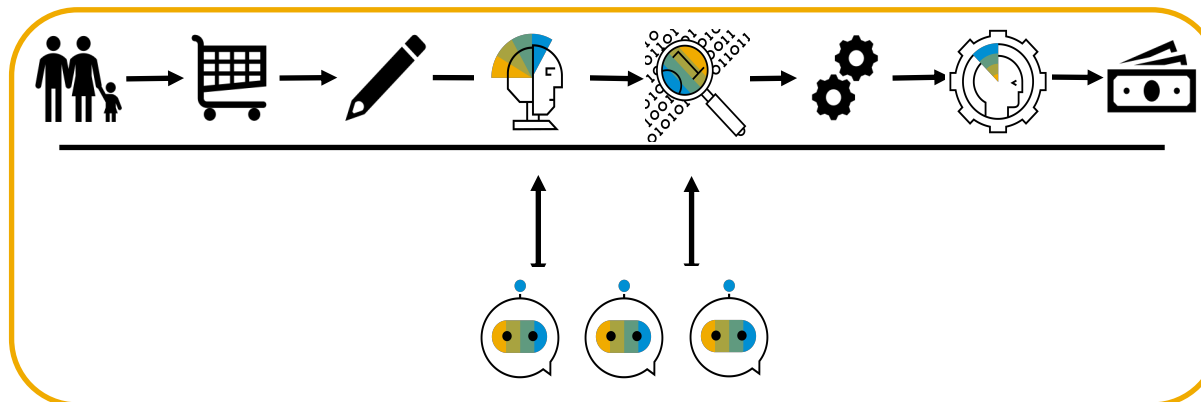


Автономное предприятие

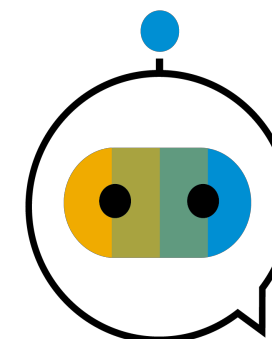
МАРКЕТИНГ



КРЕДИТНЫЙ КОНВЕЙЕР

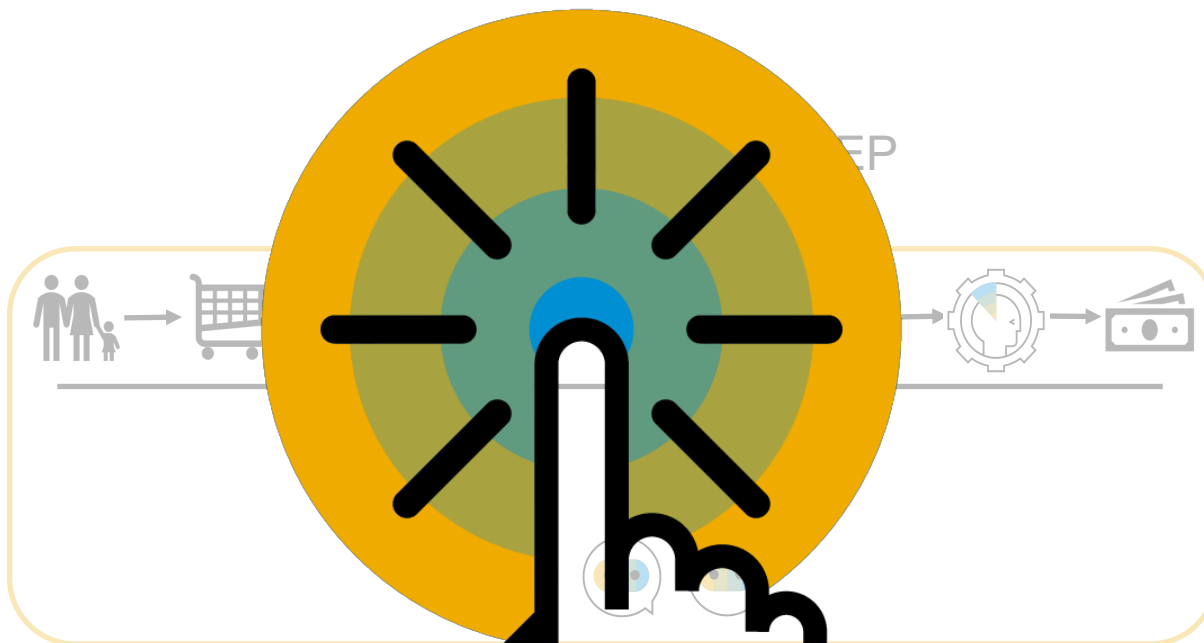


ПОСЛЕПРОДАЖНЫЙ
СЕРВИС И АНАЛИТИКА

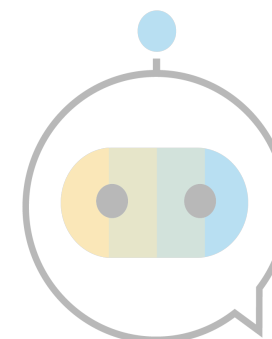


Автономное предприятие

МАРКЕТИНГ



ПОСЛЕПРОДАЖНЫЙ
СЕРВИС И АНАЛИТИКА



Фантастика?

60%

Человеческих задач будет автоматизировано к 2025

97%

Качество распознавания изображений (человек - 95%)

95.1 %

Распознавание речи (человек - 94.1%)

\$18B

Рынок машинного обучения для предприятий к 2020

Автоматизация процессов

Полная автоматизация бизнес-процессов, от HR до платежей, закупок и продаж

Интерфейсы следующего поколения

Пользовательские интерфейсы на базе голоса, компьютерного зрения, естественного языка

Новое знание

Проактивный анализ и выявление новых шаблонов в данных – путь к появлению нового знания



Вопросы?

Anton.Nozdrin@sap.com