



**Open Source Python-библиотека  
для Process Mining исследований SberPM**

**Андрей Бугаенко**  
/ СберБанк

# Process Mining связывает управление процессами и науку о данных



# Зачем нужен Process Mining?



## Снижение расходов

- Устранение лишних этапов
- Оценка трудозатрат, высвобождение / перераспределение сотрудников
- Автоматизация ручных операций
- Оптимизация использования ТМЦ



## Рост доходов

- Увеличение конверсии
- Упрощение клиентского пути, повышение удовлетворенности
- Увеличение эффективности рекламных кампаний



## Контроль соответствия

- Контроль выполнения требований законодательства
- Процедуры внутреннего контроля и аудита
- Контроль исполнения SLA

# Python-библиотека SberPM

Актуальная версия 2.4.1 от 25.05.2023

NEW

Внешняя	open source	Базовый функционал	<code>pip install sberpm</code>
Внутренняя		Полный функционал	



Ближайший релиз 2.5.0 - 25.07.2023

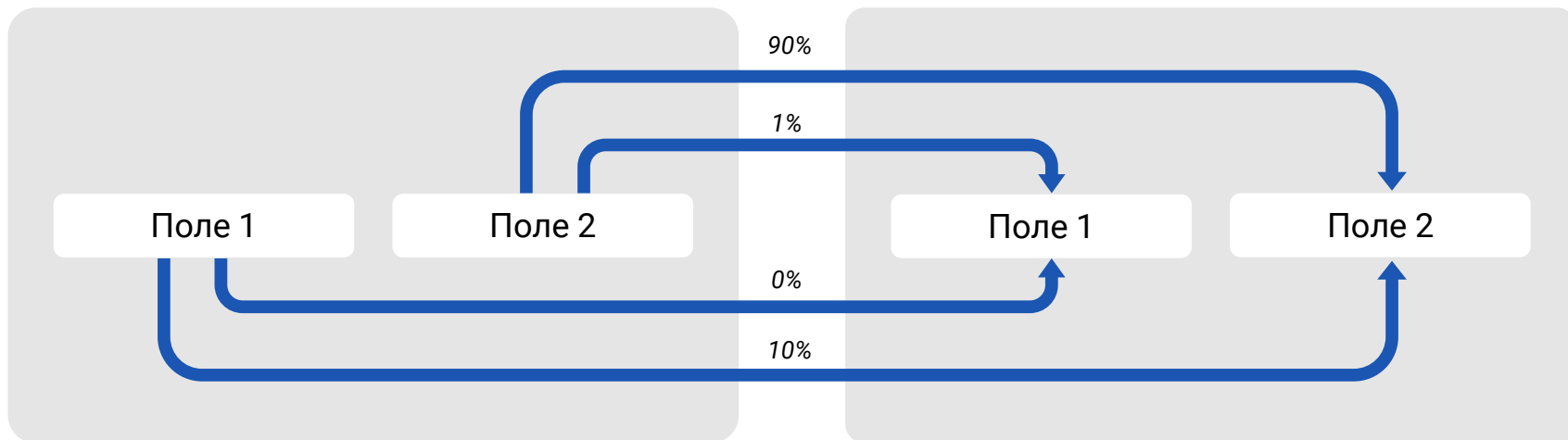
## В чём разница?

\* <https://pypi.org/project/sberpm/>

# Предобработка данных: автоматический мэтчинг таблиц

Таблица 1

Таблица 2



**Модуль автоматичинга таблиц**  
позволяет автоматически определять  
каким полям из одной таблицы  
соответствуют поля другой таблицы

# Предобработка данных: проставка синтетических id

Время	Действие	id экземпляра процесса
11:15	Вход в систему	???
11:20	Заведение заявки	???
11:25	Согласование заявки	???
11:30	Выход из системы	???
12:15	Вход в систему	???
12:16	Заведение заявки	???
12:20	Согласование заявки	???
12:25	Выход из системы	???
14:40	Вход в систему	???
14:45	Заведение заявки	???
14:50	Выход из системы	???

Начало процесса

Окончание процесса

1

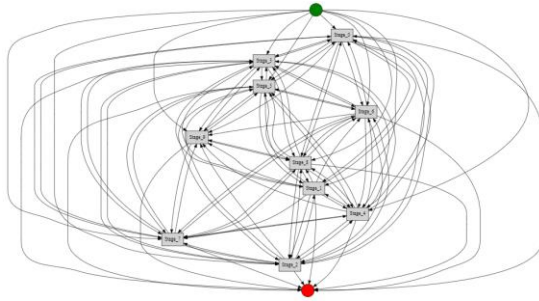
2

3

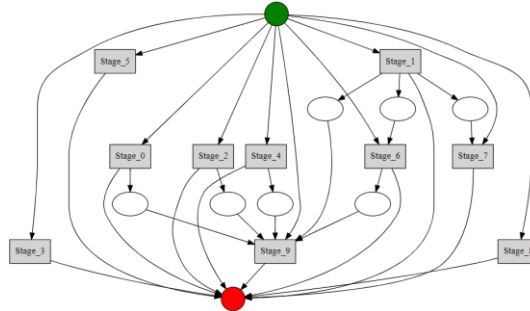
# Майнеры

**NEW** II-майнер

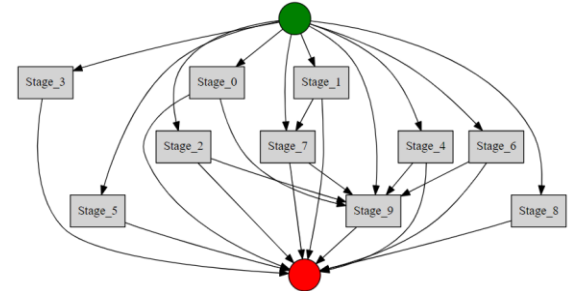
**NEW** ML-майнер, NLP-майнер



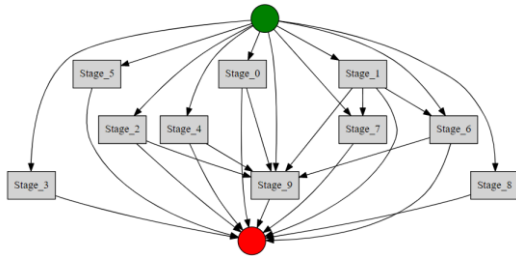
Simple майнер



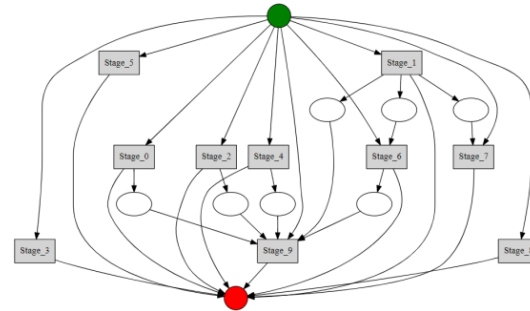
Alpha майнер



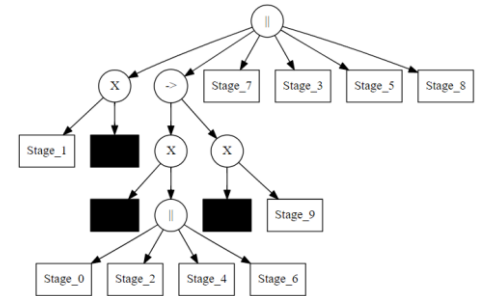
Эвристический майнер



Causal майнер



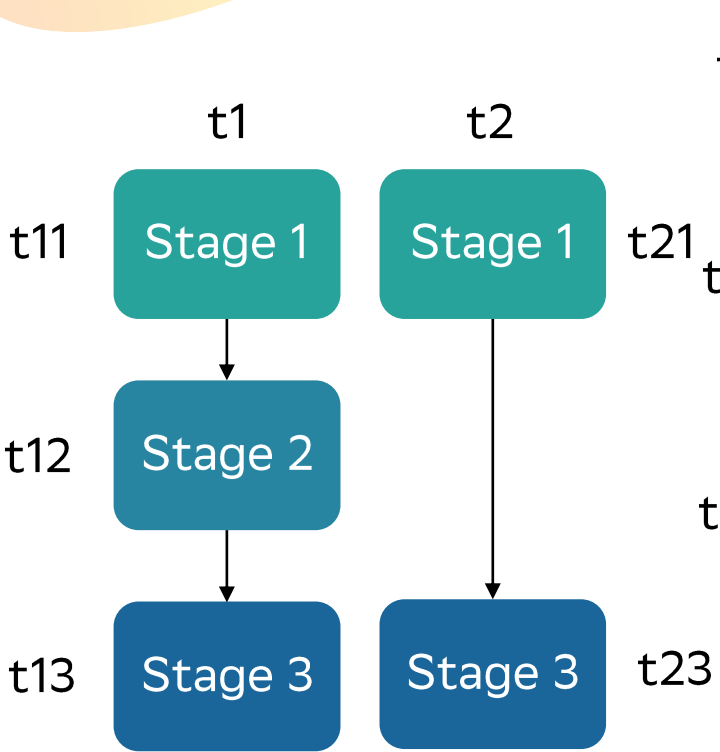
Alpha + майнер



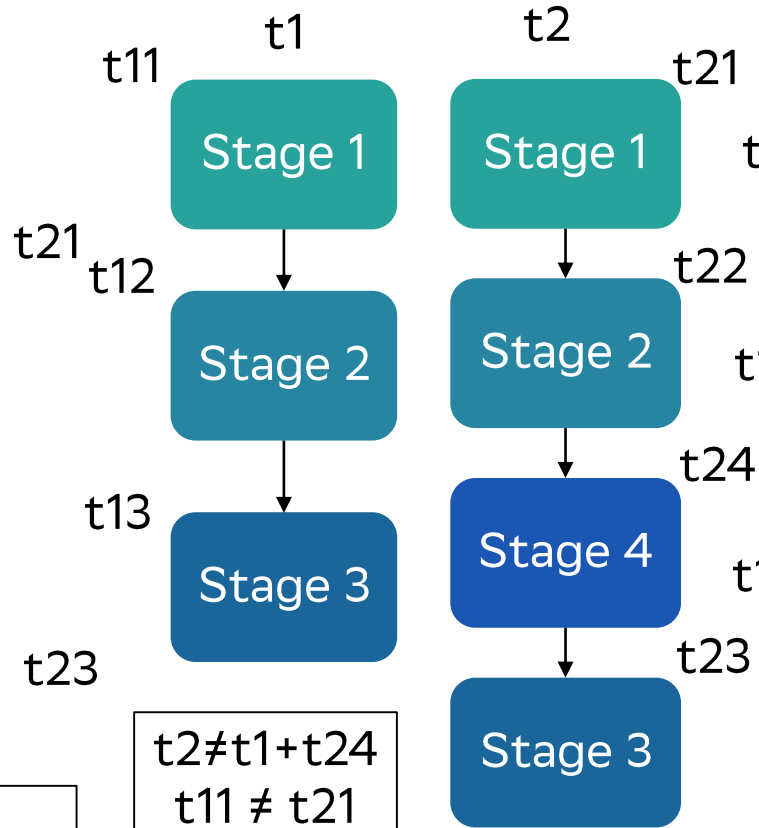
Inductive майнер

+ преобразование в BPMN

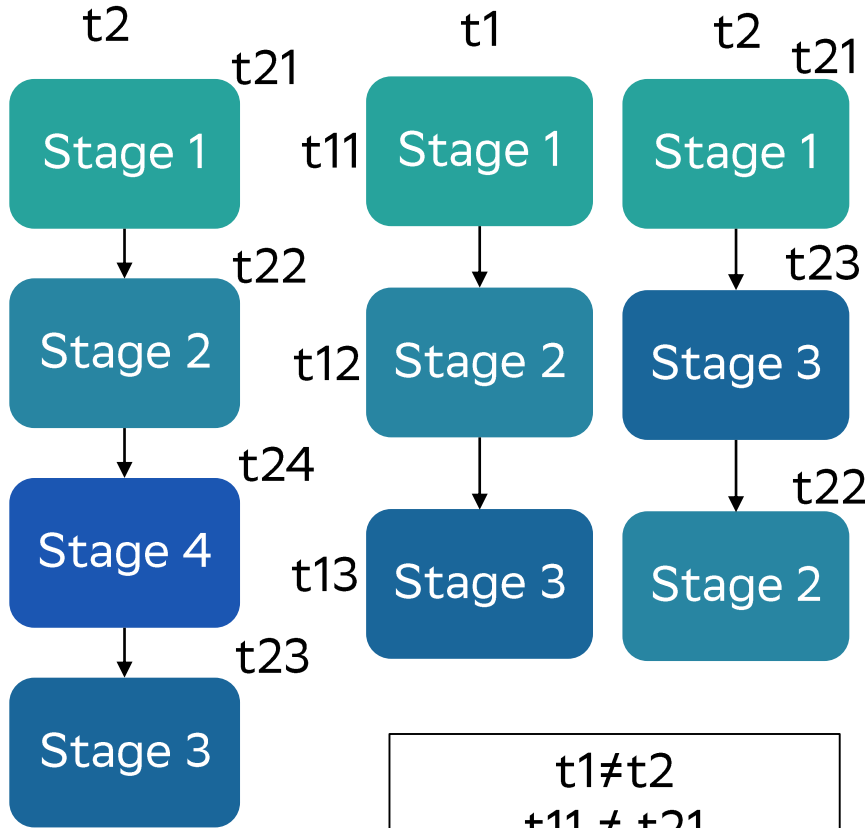
# What-if анализ



$$\begin{aligned}
 t2 &\neq t1 - t12 \\
 t11 &\neq t21 \\
 t13 &\neq t23
 \end{aligned}$$



$$\begin{aligned}
 t2 &\neq t1 + t24 \\
 t11 &\neq t21 \\
 t12 &\neq t22 \\
 t13 &\neq t23
 \end{aligned}$$



$$\begin{aligned}
 t1 &\neq t2 \\
 t11 &\neq t21 \\
 t12 &\neq t22 \\
 t13 &\neq t23
 \end{aligned}$$



**Достаточно ли этих инструментов, чтобы проанализировать бизнес-процесс?**

**Ответ: ДА**

**... НО**

# Автопоиск инсайтов. Где в процессе неэффективности?

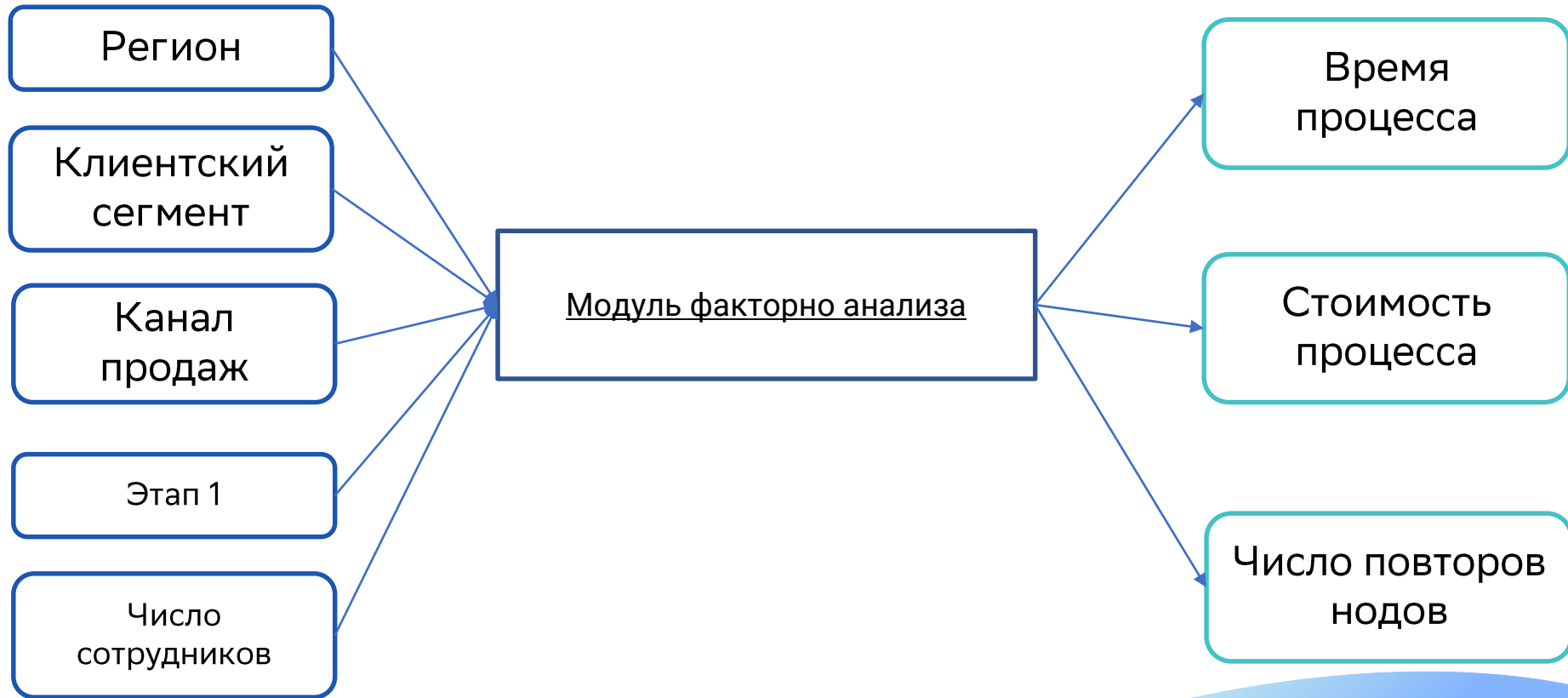


**Модуль автоинсайтов** позволяет идентифицировать **14 типов кейсов неэффективности**, предварительно оценить финансовый эффект от устранения кейсов и выдать текстовое заключение по эффективности

- ▶ Высокая длительность этапа
- ▶ Длительность этапа растет со временем
- ▶ Этап необязательный этап для успеха
- ▶ Зацикленность
- ▶ Bottle neck
- ▶ Ручной или не стандартизованный этап процесса
- ▶ Многократные инциденты, приводящие к замедлению процесса
- ▶ Разовые инциденты, приводящие к замедлению процесса
- ▶ Критический этап, отказ которого приводит к остановке всего процесса
- ▶ Ошибки системы на данном этапе, приводящие к замедлению процесса
- ▶ Критические ошибки системы на данном этапе, приводящие к неучеху процесса
- ▶ Структурные ошибки на данном этапе, приводящие к неучеху процесса
- ▶ Сторнирования, правки, изменения на данном этапе, приводящие к замедлению процесса
- ▶ Сторнирования, правки, изменения на данном этапе, приводящие к неучеху процесса

## Факторный анализ.

*Какие факторы влияют на процесс ?  
По каким провести бэнчмаркинг?*



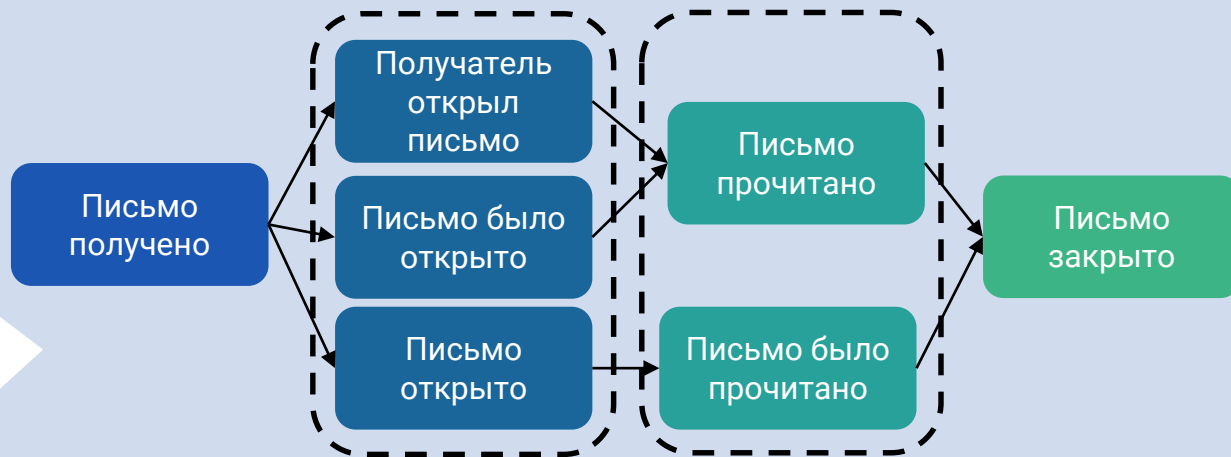
# Получение инсайтов из текста



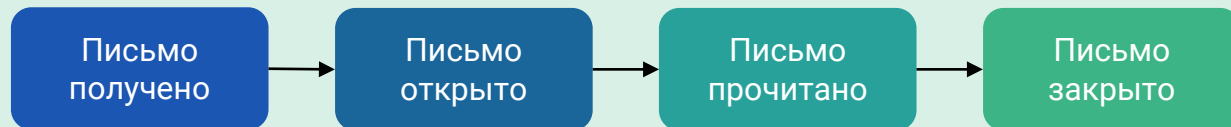
# Текстовая кластеризация этапов

## Граф из логов

*Модуль кластеризации текстовых описаний позволяет автоматически объединять этапы с похожими названиями*



## Модифицированный граф



# Happy path. Построение идеальной структуры бизнес-процесса

NEW

Разрезы: ТБ, ГОСБ, Сегмент

NEW

Start = const, Start = var

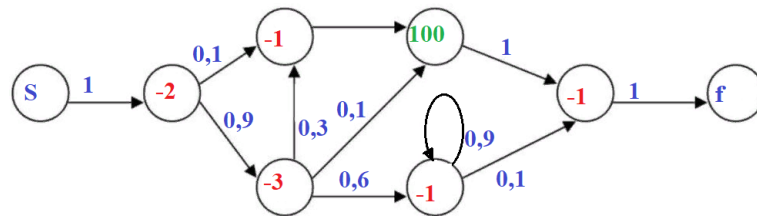
**Environment** – граф процесса

**Action** – переход

**States** – этап

**Loss** – время \ стоимость этапа

**Reward** – успех \ доход от 1 транзакции



## Ближайший релиз

Библиотека AutoRL

*Greed search*

Q-learning

$\alpha=...$ ,  $\epsilon=...$

*Greed search*

Cross-entropy

$\beta=...$ ,  $\epsilon=...$

*Greed search*

Genetic Algorithm

child=..., mutation=...

*Greed search*

Marcovic Algorithm

iteration =..., repeat=...

*Greed search*

Loss + Reward  $\rightarrow$  max

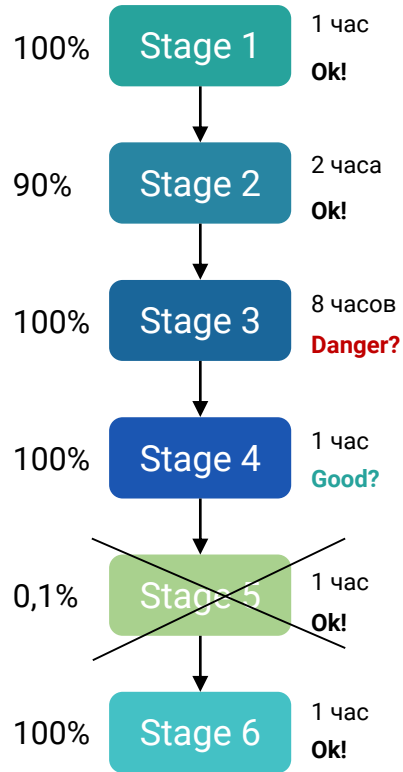
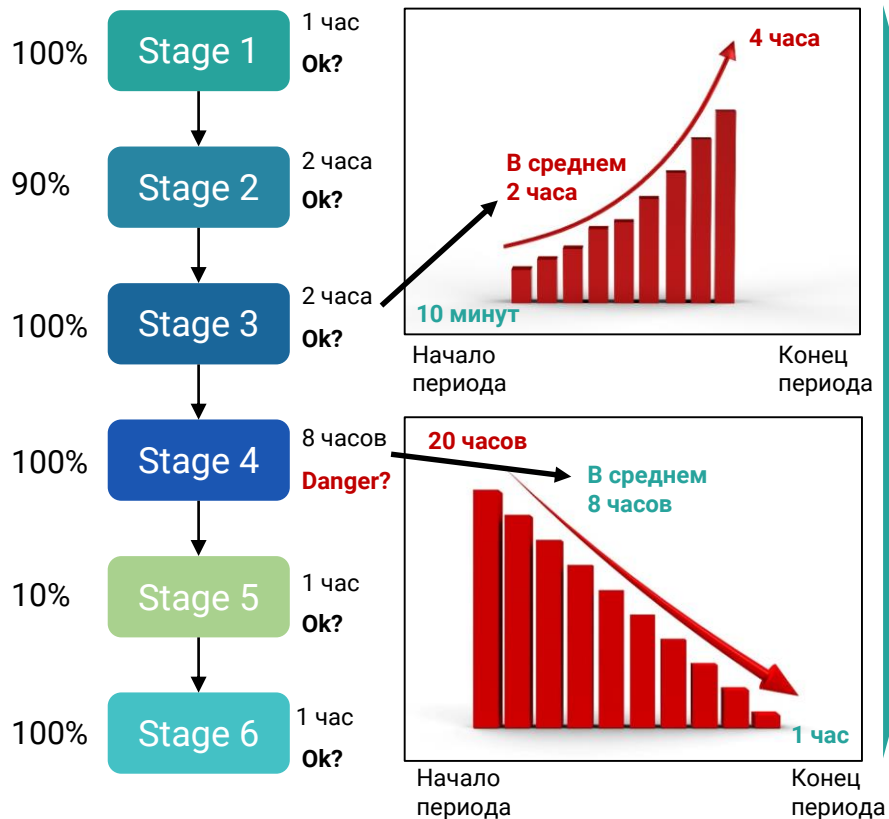
Dream Process \ Happy Path

Режим «Всегда успех»

Режим «Влияние стационарной среды»

Режим «Влияние динамической среды»

# Предсказание метрик и структуры процесса



*Модуль прогноза метрик позволяет спрогнозировать, какие будут метрики процесса и его структура при сохранении текущих трендов*

# Автоматическое исследование бизнес-процессов компании на основе искусственного интеллекта



Робот – «управленческий консультант»

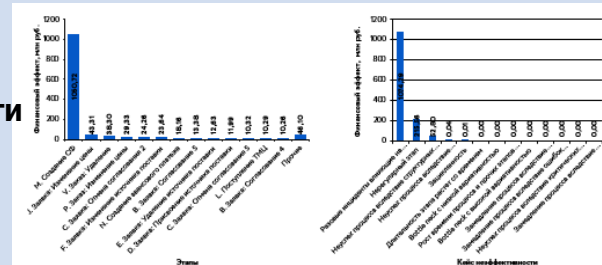
1

Выявление кейсов неэффективностей

	Длительность этапа растет со временем	Нерегулярный этап	Зависимость	Ветви кейс с низкой вариативностью	Разные инциденты возникают на длительности этапа	Результаты процесса и прерывание процесса из-за инцидентов, связанных с данным этапом	Ветви кейс с высокой вариативностью
U. Заказ: Восстановление заказа...	False	True	True	False	False	True	False
I. Заказ: Изменение источника...	False	True	True	False	False	True	False
F. Заказ: Изменение источника...	False	True	True	False	True	True	False
Q. Заказ: Изменение количества	False	True	True	False	False	True	False
M. Создание СД	False	False	True	False	True	True	False
V. Заказ: Удаление	False	True	True	False	True	True	False
G. Заказ: Изменение плана...	False	True	True	False	False	True	False
S. Заказ: Отмена согласования...	False	True	True	False	False	True	False
N. Создание авансового платежа	False	True	True	False	False	True	False
S. Заказ: Отмена согласования...	False	True	True	False	False	True	False

2

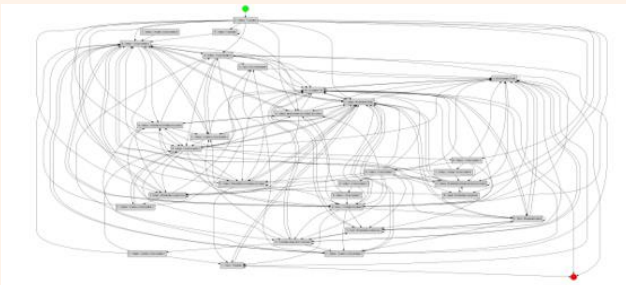
Бэнчмаркинги по самым значимым метрикам



Суммарный финансовый эффект от устранения обнаруженных кейсов неэффективности 1296599506 руб. Максимальный финансовый эффект достигается от устранения кейсов неэффективности на этапе «М. Создание СД». В размере 1050723623 руб., что составляет 81.037% от всего возможного финэффекта.

3

Построение идеальной структуры процесса



Используя Neuz miner видно, что 80% всех экземпляров состоит из 31 этапов и 185 переходов. Есть ограниченная возможность визуального анализа процесса.

4

Выводы

- Суммарный финансовый эффект от устранения обнаруженных кейсов неэффективности: **1342699280 рублей**.
- В следующих этапах процесса отменяется высокая длительности: «С. Заказ: Отмена согласования 7», «Е. Заказ: Удаление источника поставки», «F. Заказ: Изменение источника поставки», «M. Создание СД», «V. Заказ: Удаление». Максимальный потенциальный финансовый эффект от её снижения 1074290013 рублей.
- Следующие этапы являются нерегулярными (редкими) и не требуются для успешной реализации процесса: «B. Заказ: Согласование 7», «С. Заказ: Отмена согласования 5», «А. Заказ: Отмена согласования 6», «С. Заказ: Отмена согласования 7», «С. Заказ: Отмена согласования 4», «B. Заказ: Согласование 6», «С. Заказ: Отмена согласования 3», «H. Заказ: Изменение материала», «E. Заказ: Удаление источника поставки», «J. Заказ: Возстановления», «I. Заказ: Изменение количества», «F. Заказ: Изменение источника поставки», «Q. Заказ: Изменение количества», «V. Заказ: Удаление», «G. Заказ: Изменение плановой даты поставки», «С. Заказ: Отмена согласования 2», «N. Создание авансового платежа», «С. Заказ: Отмена согласования 1», «R. Заказ: Изменение цены», «B. Заказ: Согласование 5», «I. Заказ: Изменение цены», «B. Заказ: Согласование 4». Максимальный потенциальный финансовый эффект при отказе от данных этапов 215555336 рублей.
- В следующих этапах наблюдаются разовые инциденты, приводящие к замедлению процесса: «С. Заказ: Отмена согласования 7», «E. Заказ: Удаление источника поставки», «F. Заказ: Изменение источника поставки», «M. Создание СД», «V. Заказ: Удаление». Максимальный потенциальный финансовый эффект от их устранения 1074290013 рублей.





# AI в SberProcessMining

## Поиск аномалий

Майнеры

Автоматическое Process Mining исследование на основе AI

Автоматический поиск инсайтов

Факторный анализ

Приоритезация этапов

RootRoute анализ

Автопостроение моделей

## NLP

Кластеризация текстовых описаний и вывод заголовков

Поиск подтверждения гипотез

Сентиментный анализ

Объединение этапов процесса

## RL

Поиск Happy path

Генерация тестовых логов

What-if анализ

## Прогнозирование

Прогноз метрик и структуры графа в будущем

Прогноз следующего шага в процессе

## Предобработка данных

Фильтрация выбросов

Расчёт цифрового хронометража

Автомэтичнг таблиц

Простановка синтетических id

Преобразование графа в BPMN



Спасибо за внимание!



Python-библиотека SberPM

[aabugaenko@sberbank.ru](mailto:aabugaenko@sberbank.ru)



***Наша миссия в том, чтобы в России Process Mining стал основным инструментом для анализа бизнес-процессов.***