



ТИНЬКОФФ

AI Banking в обслуживании

Как его сделать, как он влияет на нашу жизнь

Петр Савостин



КТО Я?

2019 - 2020 Tinkoff Business ML Lead

2020 – 2021 AI Bank in Service Lead

2021 AI Bank Products Lead



План

- ▶ От простого к сложному
- ▶ Точки роста автоматизации
- ▶ Тренды
- ▶ Другие подходы
- ▶ AI Bank

От простого к сложному

Задача



Обучить бота отвечать человеку

Пример диалога

Добрый день!



Здравствуйте! Поменял сотового оператора, хочу изменить контактный номер

Хорошо, скажите девичью фамилию матери



Иванова

Скажите новый номер



8 985 999 99 99

Номер изменили, еще есть вопросы?



Нет, спасибо!

Бот

Клиент



От простого к сложному

Шаг 1



Выделить диалоги, где вопрос решается в одно касание (т.е. вопрос-ответ).



Делаем таблицу с парами вопросов пользователей и ответов бота.

Пример диалога

Добрый день!

Здравствуйте! Чем могу помочь?

<Интент:
"Приветствие">

Как поменять ПИН?

Для того, чтобы поменять ПИН-код, зайдите в раздел карты в мобильном приложении, далее нажмите "шестеренку", после кнопку сменить "ПИН-код"

<Интент:
"Смена ПИНа">

Как поменять номер?

Клиент

Перевожу на профильного специалиста, он Вам поможет

<Перевод
на оператора>

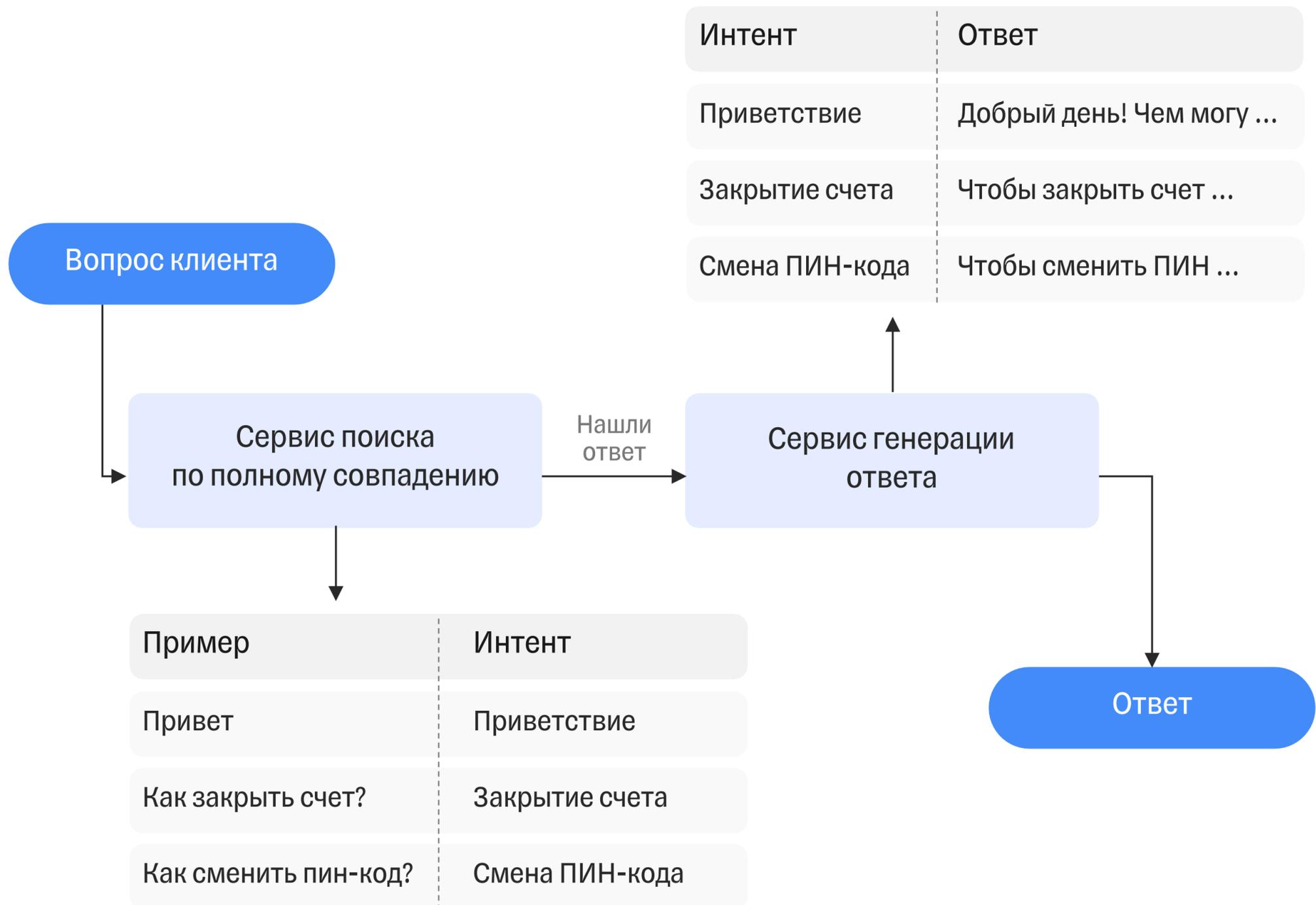
Бот

Полное совпадение

+ Преимущества

Быстрый ответ бота, так как ищем заранее заданным примерам, не запросы из больших моделей.

Можно подключить elastic-search, чтобы учитывать хотя бы одну ошибку.



Полное совпадение

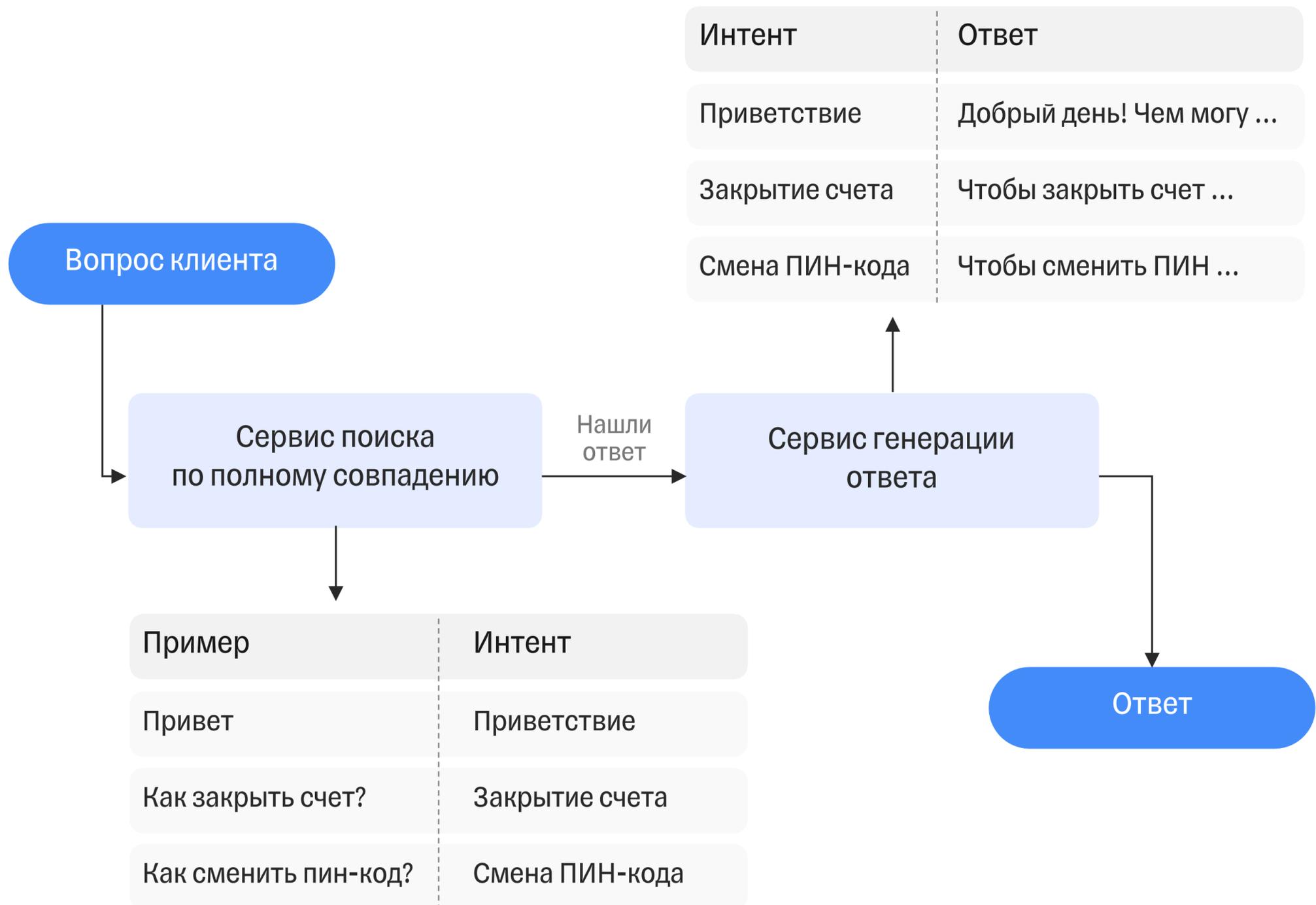
+ Преимущества

Быстрый ответ бота, так как ищем заранее заданным примерам, не запросы из больших моделей.

Можно подключить elastic-search, чтобы учитывать хотя бы одну ошибку.

- Недостатки

Не учитываем морфологию.
Низкий уровень автоматизации



Регулярные выражения

+ Преимущества

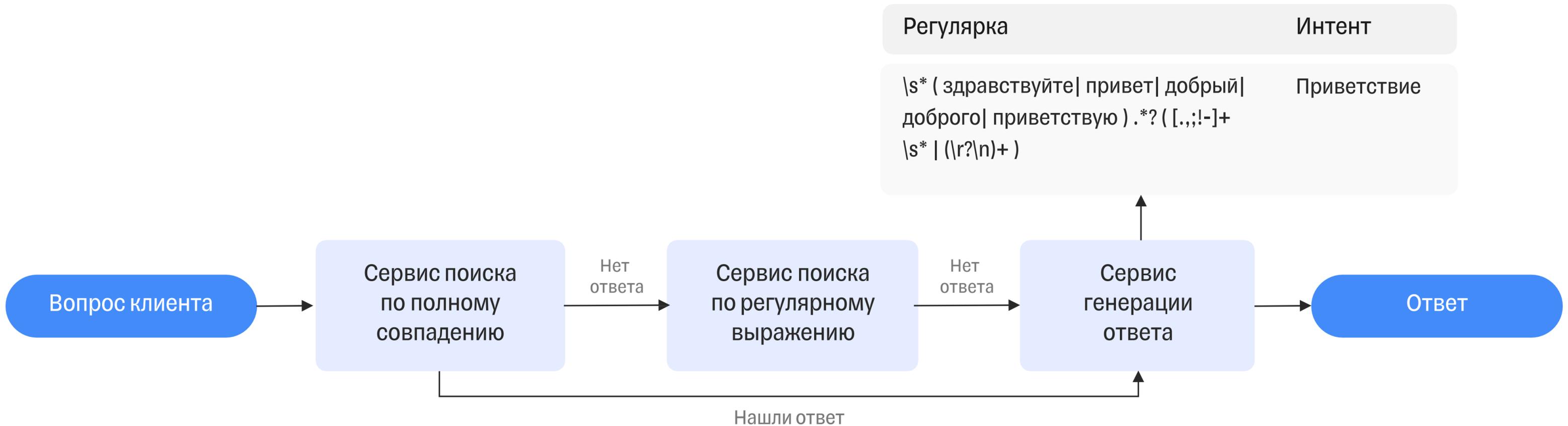
Можем не перечислять примеры, а просто задать шаблон.

Больше охватываем вопросов.

Незначительные проблемы с точностью.

- Недостатки

Не учитываем морфологию.



Сценарии

Шаг 2



В вопросах, где нужно у пользователей что-либо уточнять, необходимо использовать сценарии, чтобы не терять контекст диалога.



При использовании сценариев необходимо выделять сущности, для движения по диалогу и запоминать значения этих сущностей.

Пример диалога

Здравствуйте! Поменял сотового оператора, хочу изменить контактный номер

Хорошо, скажите девичью фамилию матери



Иванова

Скажите новый номер



8 985 999 99 99

Клиент

Номер изменили, еще есть вопросы?



Бот

Отличие сценария оператора от бота

Пример диалога

Здравствуйте! Поменял сотового оператора, хочу изменить контактный номер

Хорошо, скажите девичью фамилию матери

<Интент: "Смена номера" запускаем ветку идентификации>

Иванова

<Заполняем слот с помощью NER, сравниваем с истинным значением>

Скажите новый номер

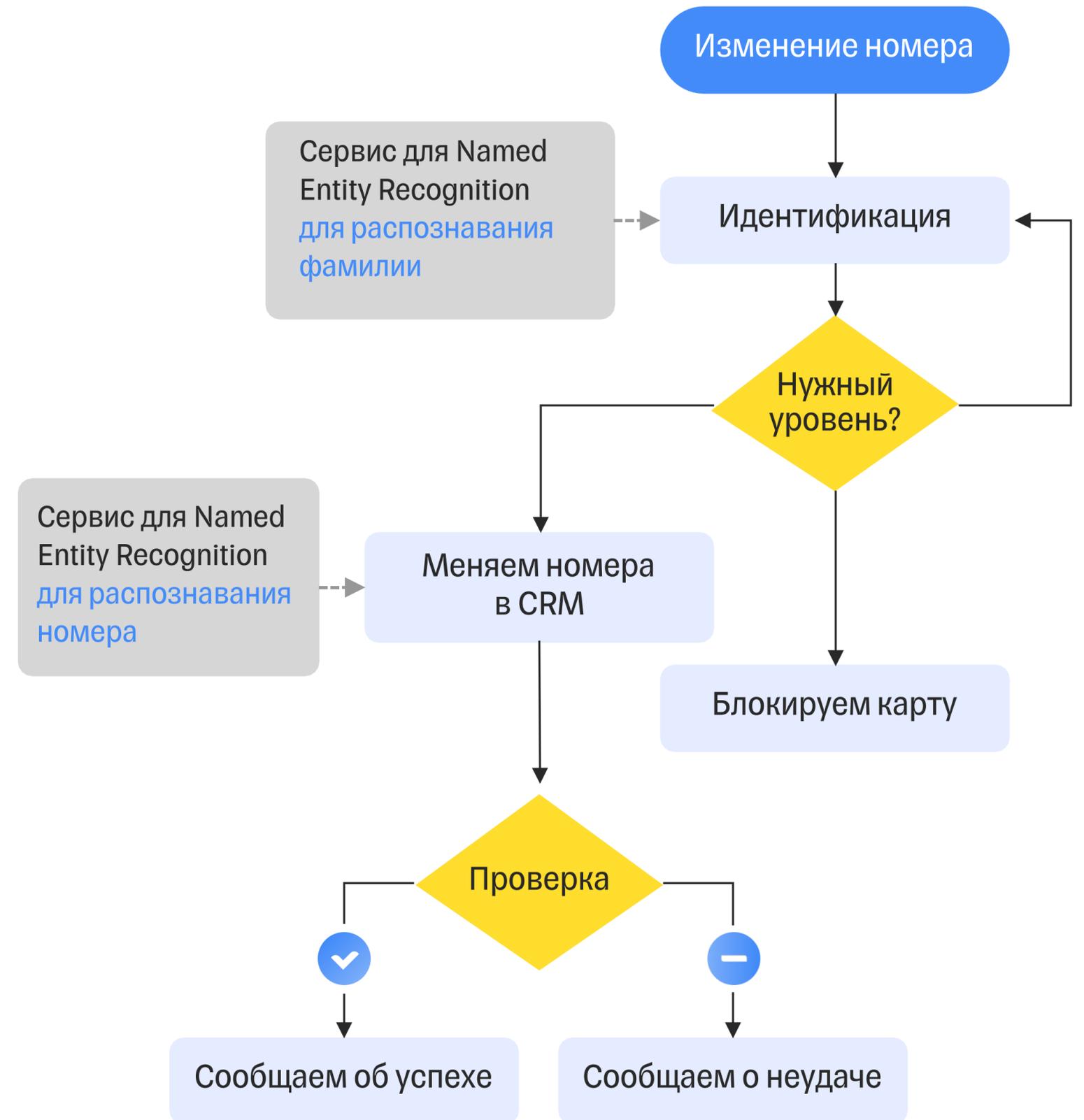
<Удачное сравнение, запускаем ветку уточнения номера>

8 985 999 99 99

<Заполняем слот с новым номером, вызываем сервис смены номера>

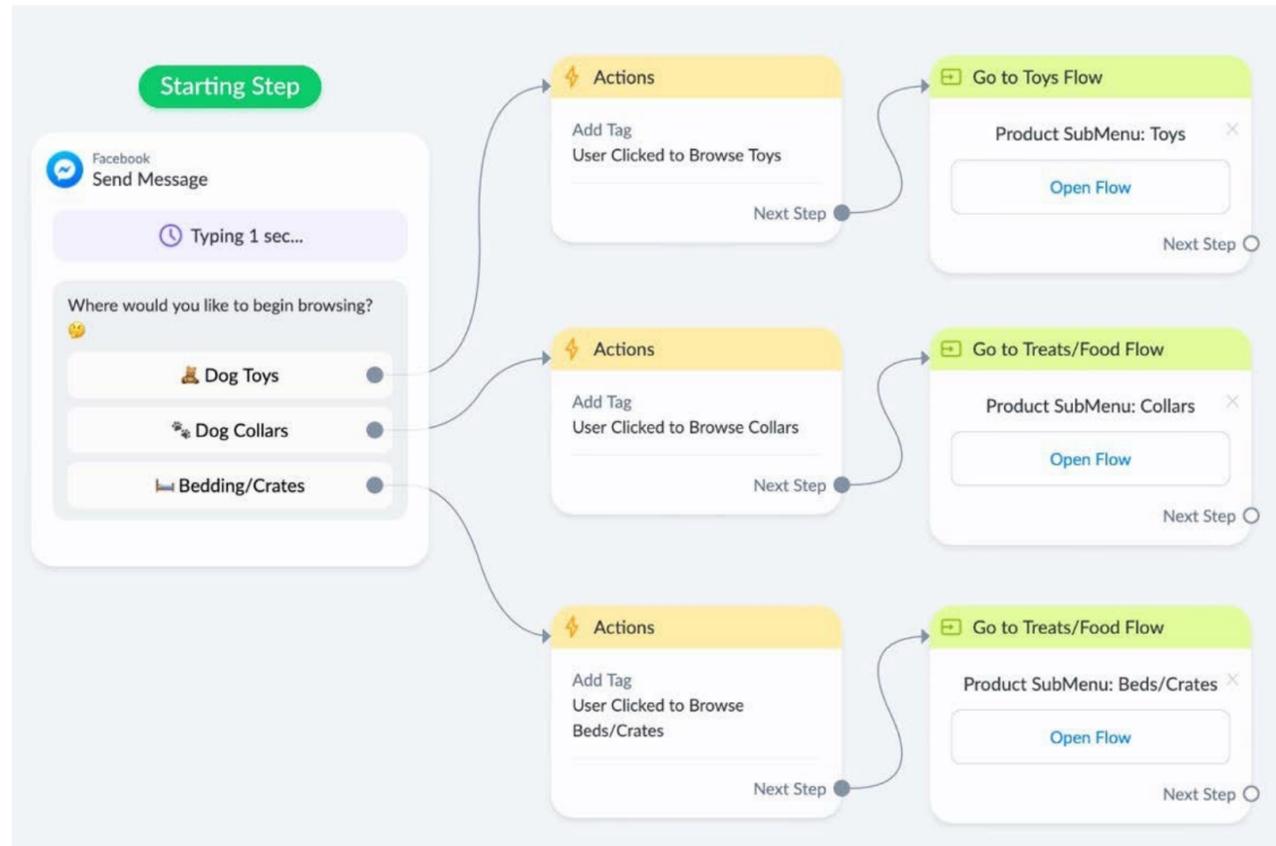
Номер изменили, еще есть вопросы?

Бот

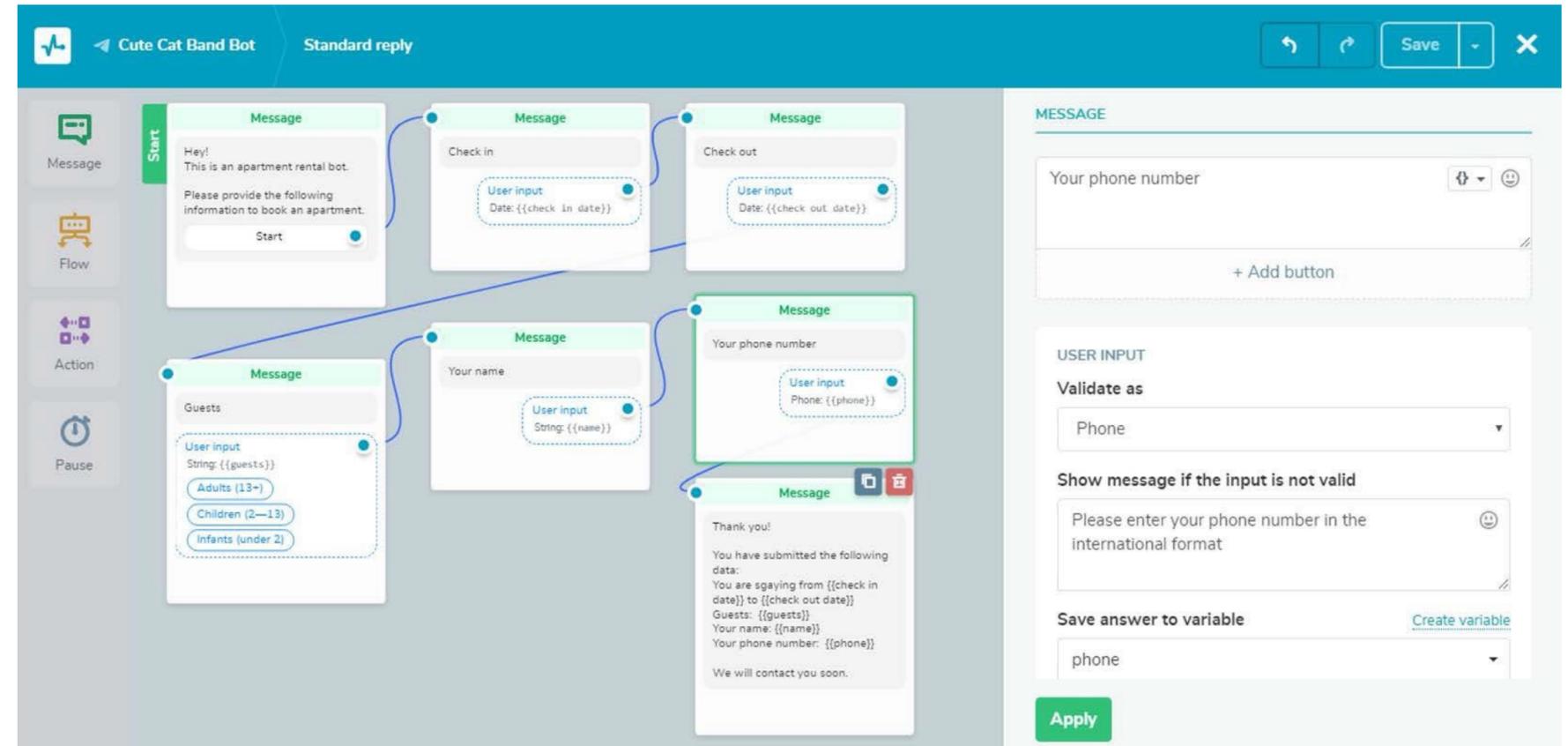


Примеры конструкторов сценариев

ManyChat



SendPulse



Обучение модели

Шаг 3



Много диалогов, следовательно много данных. Создаем эмбединги (BERT, GPT-3), поверх них модель для классификации (kNN, logreg).



Для разметки можно использовать собственных операторов либо сервисы а-ля

[Толока](#)

Вопрос

Как сформулировать задание для разметчиков, если у нас порядка 100 интенгов?

Вариант разметки №1

Если есть эмбединги, можем подать на вход классификатору пример без разметки и получить список близких интенгов. Добавить вариант “Другое”, если не подходит



Преимущества

Быстро. Дешево.



Недостатки

Не факт, что нужная тема будет среди примеров. Также сложно интерпретировать «Другое»

Обучение модели

Вариант разметки №2

Конвертировать “плоские” темы в иерархию тем. Плоские тематика == листья иерархии

+ Преимущества

Точнее, чем предыдущий вариант. Открывает новые возможности для аналитики диалогов.

- Недостатки

Среднее выполнение задание увеличивается. Выходит дороже.



Итоговый пайплайн

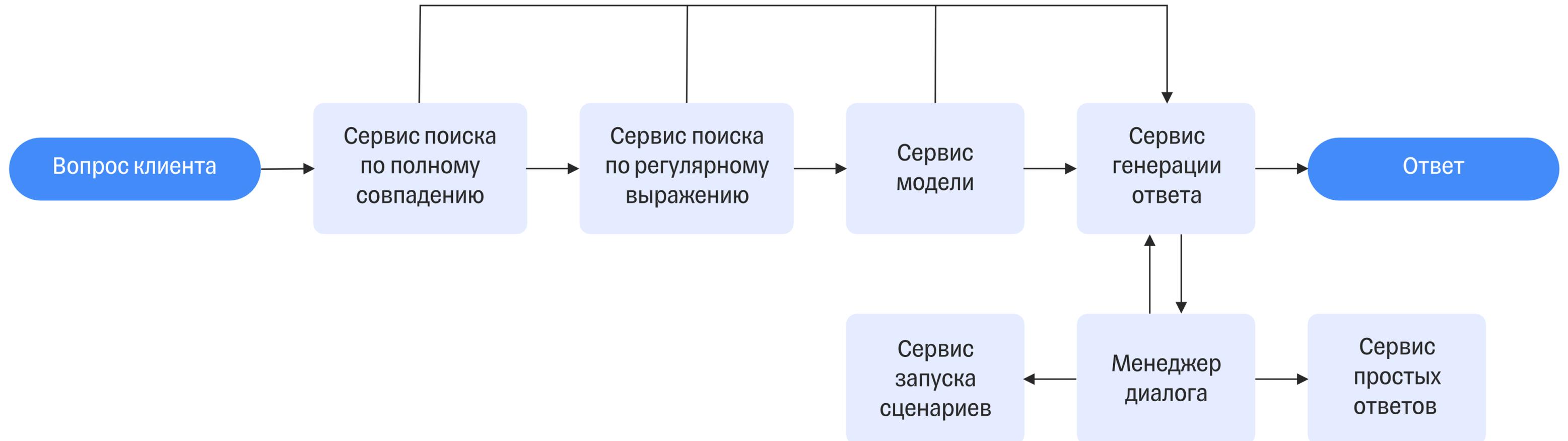
Появился целый продукт,
который можно улучшать
данными

Вопрос

Как искать точки роста
продукта?

Ответ

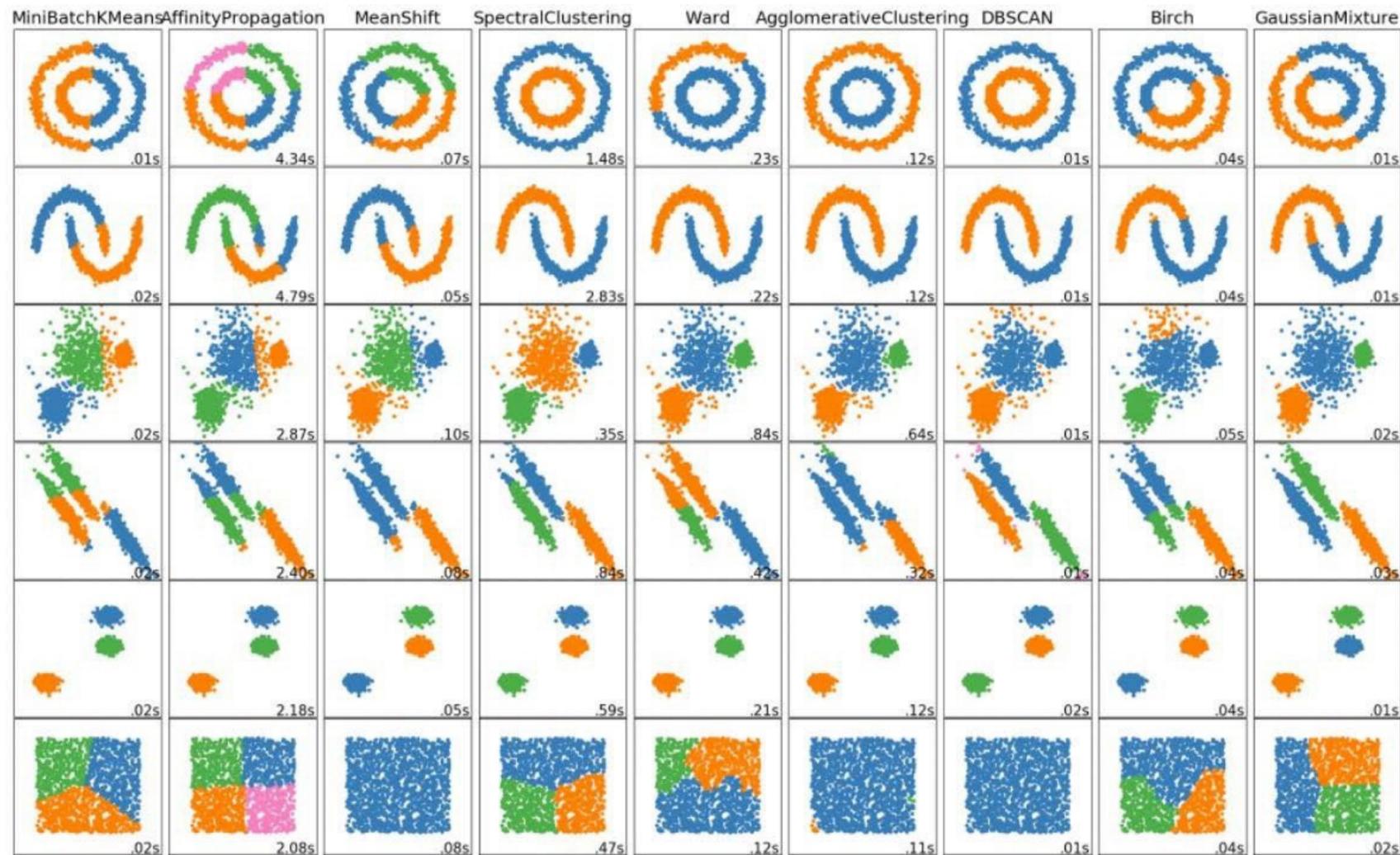
- ▶ искать новые темы
- ▶ искать больше примеров
- ▶ исправлять ошибки старых тем или сценариев



Точки роста автоматизации

Оценка числа тем

Кластеризация

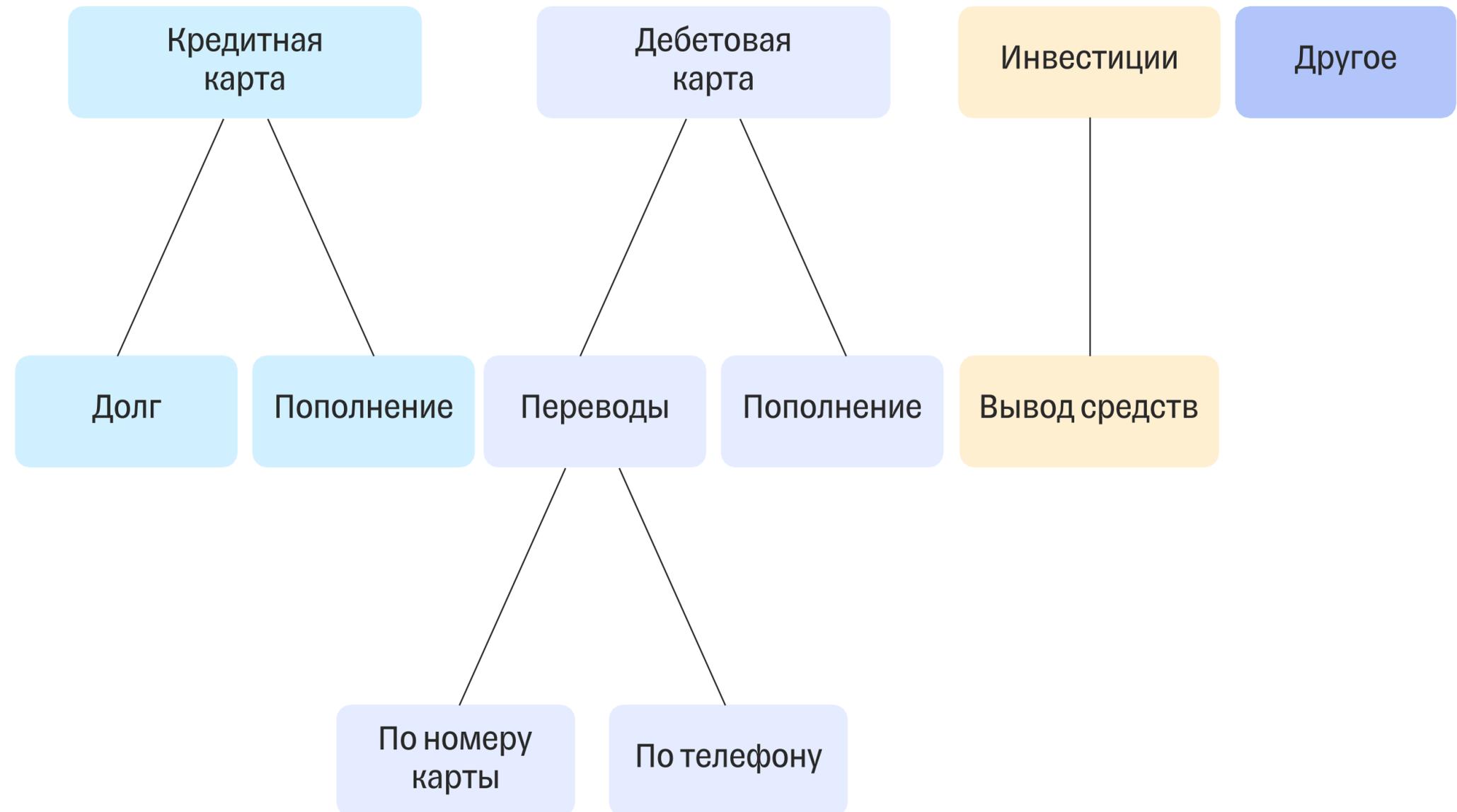


Тематическое моделирование



Разметка

- ✓ Можно использовать данные разметки для поиска точек роста.
- ✓ Поиск новых примеров можно поставить на поток, размечая постоянно по иерархии вопросы пользователей.
- ✓ Поиск новых тем и подтем можно также частично автоматизировать, если добавлять на каждом уровне иерархии вариант «Другое».



Модель уточнения



Раньше модель использовала листья иерархии, при классификации вопроса.



Можно построить модель, определяющую какую-то часть вопроса, например продукт и действие с ним.



В случае, если мы что-то определили одно, можно доспросить и определить точную тему.

Пример диалога

Здравствуйте! Поменял сотового оператора, хочу пополнить карту

Хорошо, скажите какую? Дебетовую или кредитную?

<Интент: “Действие: пополнение” уточняем продукт>

Дебетовую

<Заполняем слот с помощью NER>

Клиент

Пополнить карту можете с помощью ...

<Интент: “Пополнение дебетовой карточки”>

Бот

Точки роста по длине обращения



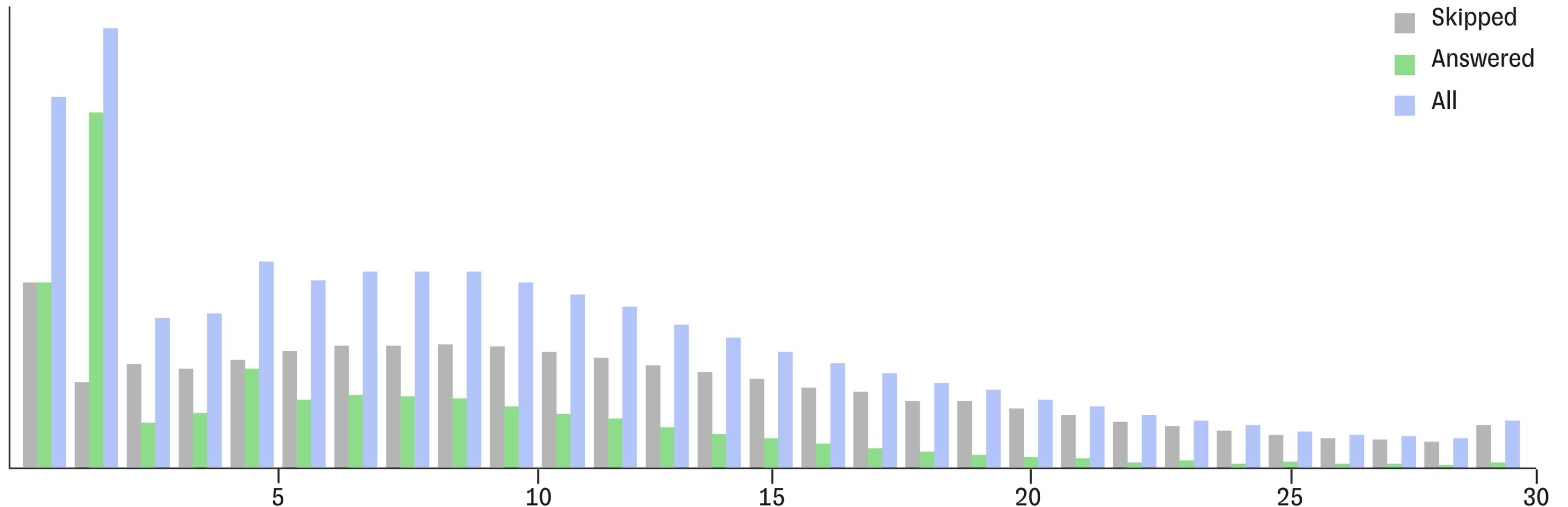
Можно построить гистограмму обращений клиента по длине (в словах) запроса.



Далее к графику добавить количество **отвеченных** и **пропущенных** запросов.



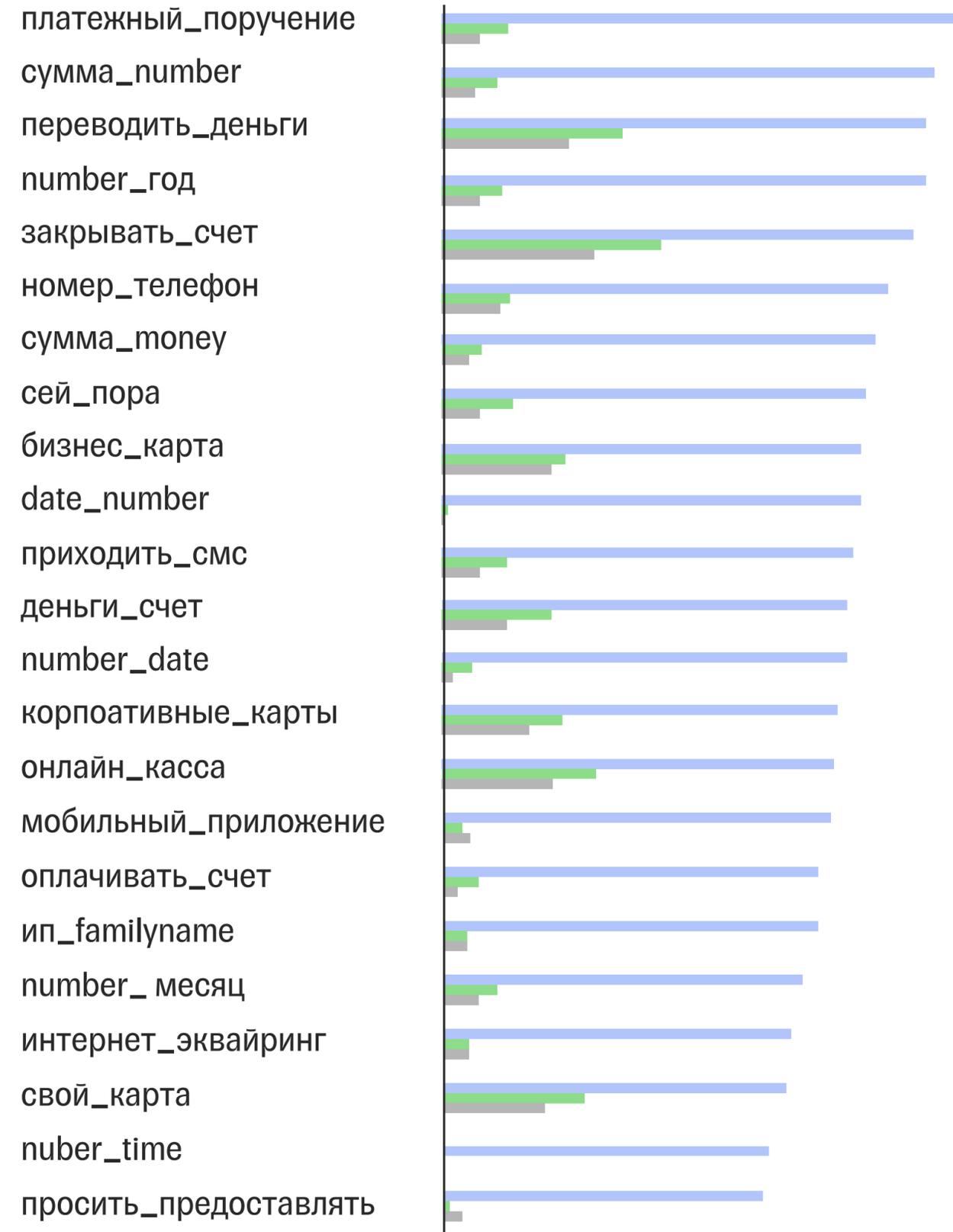
Благодаря этому можно искать простые примеры вопросов, на которые мы не отвечали.



Частотность биграмм/ униграмм



Чтобы локализовать проблемные тематики, можно посмотреть сколько биграмм/униграмм от общего количества обрабатывает модель



Точки роста удачные и не очень



Сегментирование вопросов по каналам веб/мобильное приложение

Выяснили, что лучше работаем в МП, в МП вопросы задают проще.



Развитие сценариев

Смотрели на переводы на операторов искали ветки, в которых чаще всего это происходит.



Ансамбль моделей

При обучении мы где-то приобретаем, где-то теряем, попытались использовать ансамбли. Поняли, что мы увеличили число ошибок



Деградация данных

Со временем наблюдается эффект снижения процента автоматизации, необходимо постоянно дообучать модели примерами с прода.

Тренды

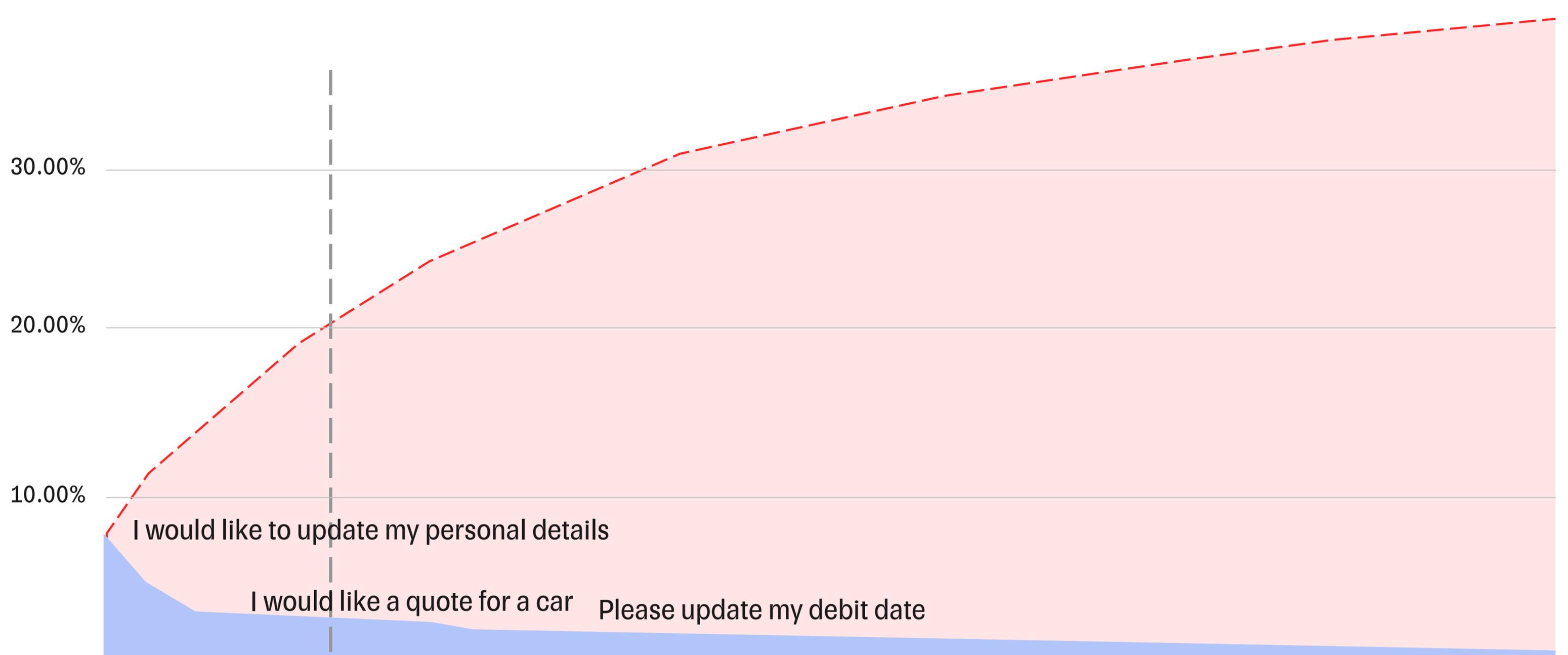
Пределы обучения



С увеличением числа тем покрытие
растет логарифмически

■ Per-intent coverage - - - Aggregated coverage

Intebts/Coverage



Баланс



Бот помогает снизить нагрузку, отвечая на несложные вопросы.



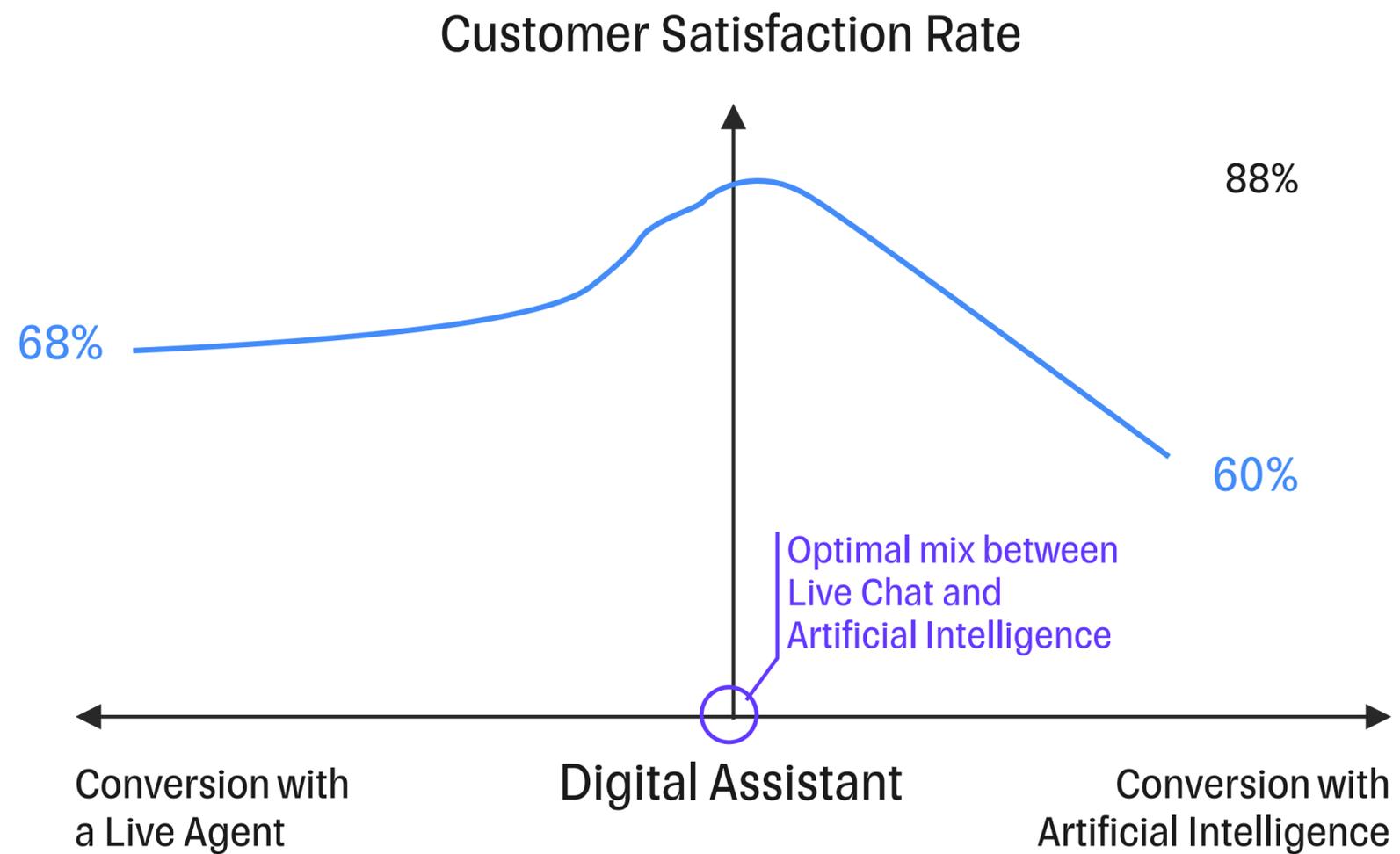
Чувствительные случаи обрабатываются человеком

Live Chat Limitations

- Knowledge Attrition
- Waiting Time (32% wait more than 1 minute)
- Low Level of Expertise
- Expensive

Artificial Intelligence Limitations

- Sensitive Cases
- Isolated Cases



Последствия на рынке труда

Искусственный интеллект не отберет работу у людей, по крайней мере, если профессия не поддается легкому алгоритму.

ИИ расширяет возможности человека. В некоторых задачах не может заменить его полностью (по крайней мере сейчас)

Человек сможет заниматься более творческими задачами



Sustainability of AI

- Sustainable ML — более эффективное обучение и выполнение моделей
- Лаборатория Тинькофф — обучение компактных моделей

Другие подходы

Как расти дальше?

Вопрос

Как искать точки роста продукта?

Ответ



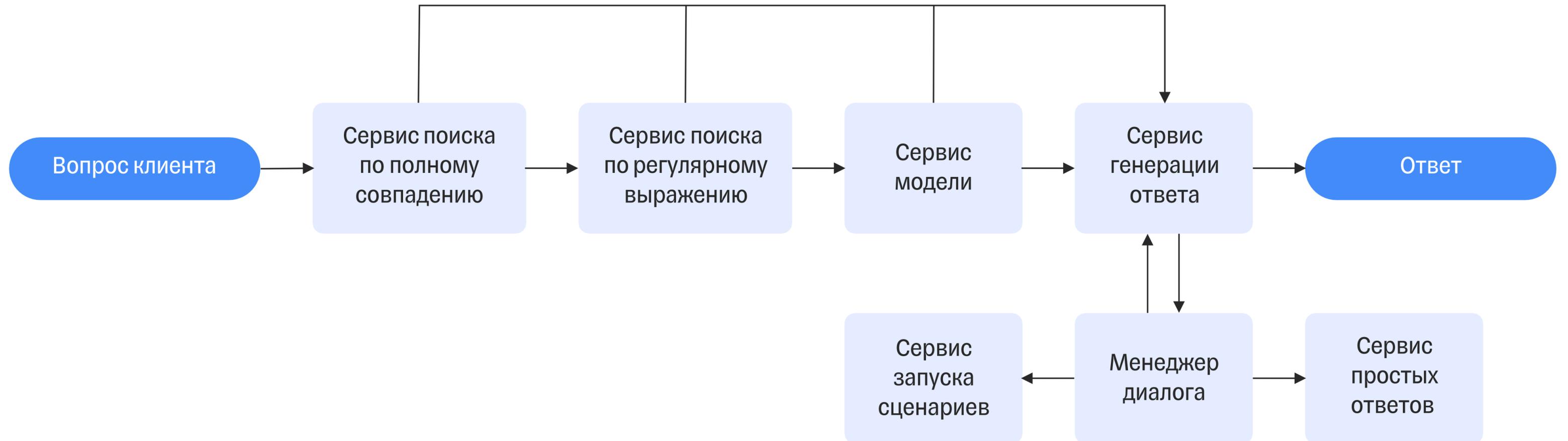
искать новые темы



искать больше примеров



исправлять ошибки старых тем или сценариев



Как расти дальше?

Вопрос

Как искать точки роста продукта?

Ответ



искать новые темы



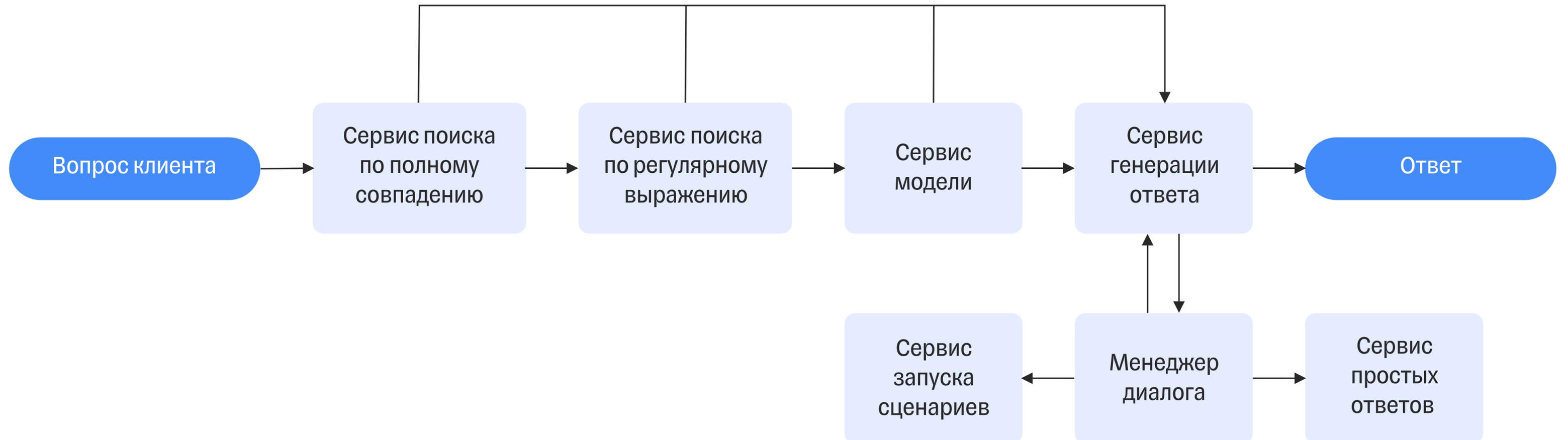
искать больше примеров

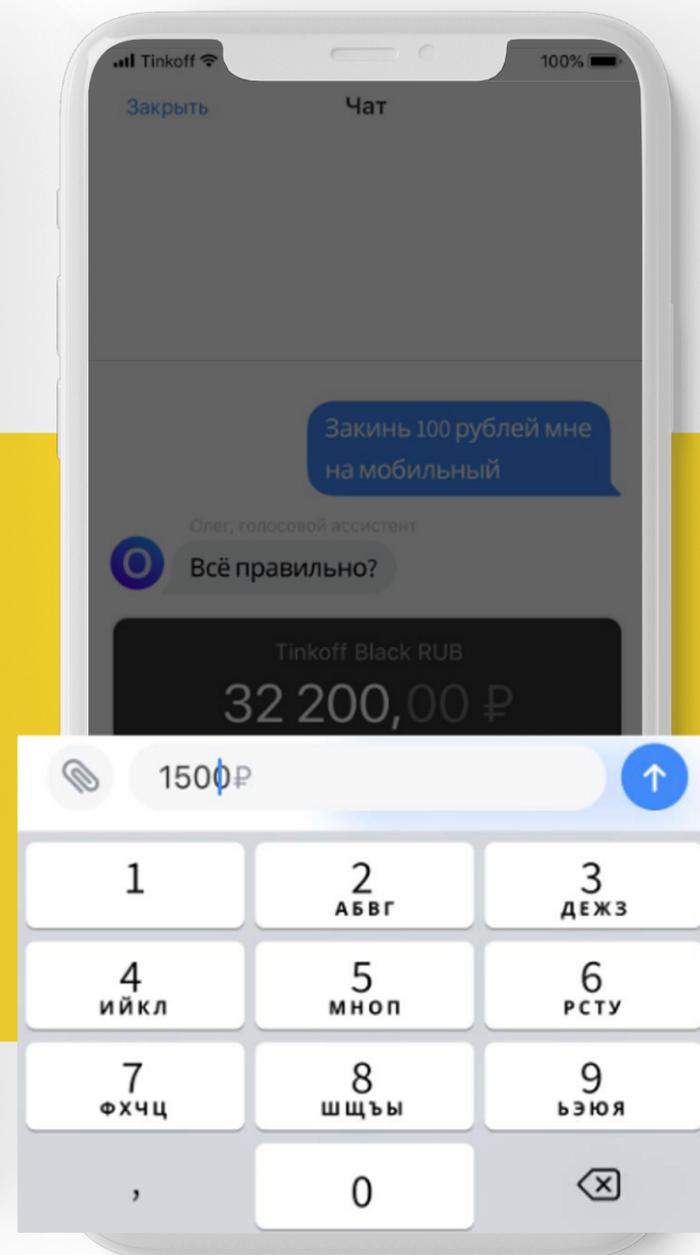
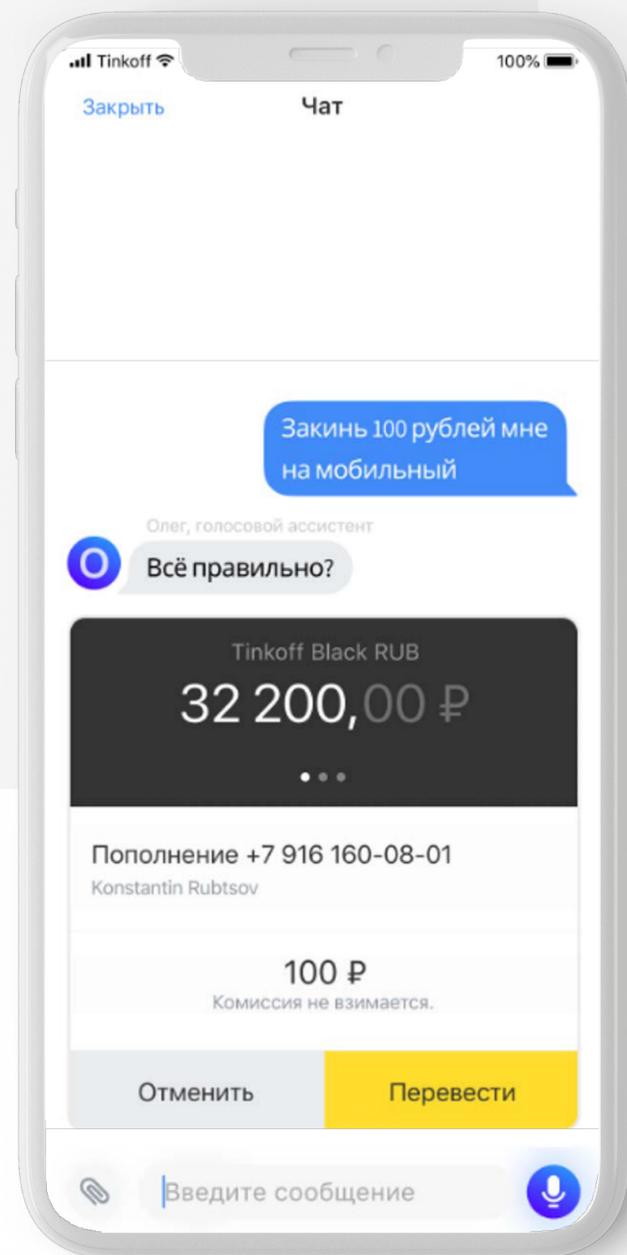


исправлять ошибки старых тем или сценариев



искать точки роста за рамками продукта





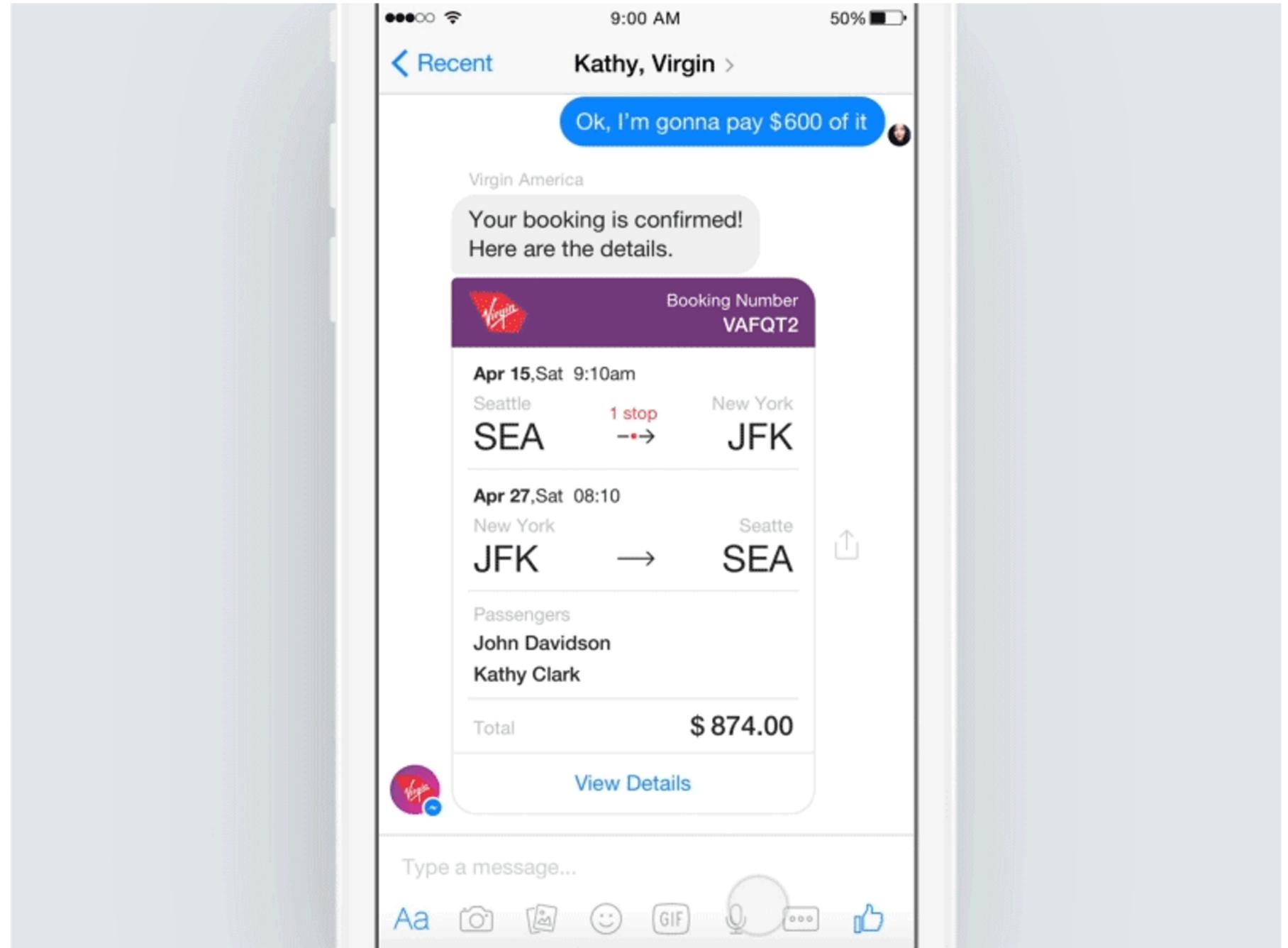
Другие подходы

В предыдущих подходах мы рассматривали чатбот в отдельности от его среды обитания — чата или мессенджера.

Следующий этап развития — развитие диалоговых интерфейсов:

- ▶ Подсказки
- ▶ Виджеты
- ▶ Кастомные клавиатуры

Интерфейсные точки роста. Подсказки во время набора



A solid yellow vertical bar is positioned on the left side of the image, extending from the top to the bottom edge.

AI Bank

AI Bank

Миссия

избавить пользователей от проблем, ожидания оператора

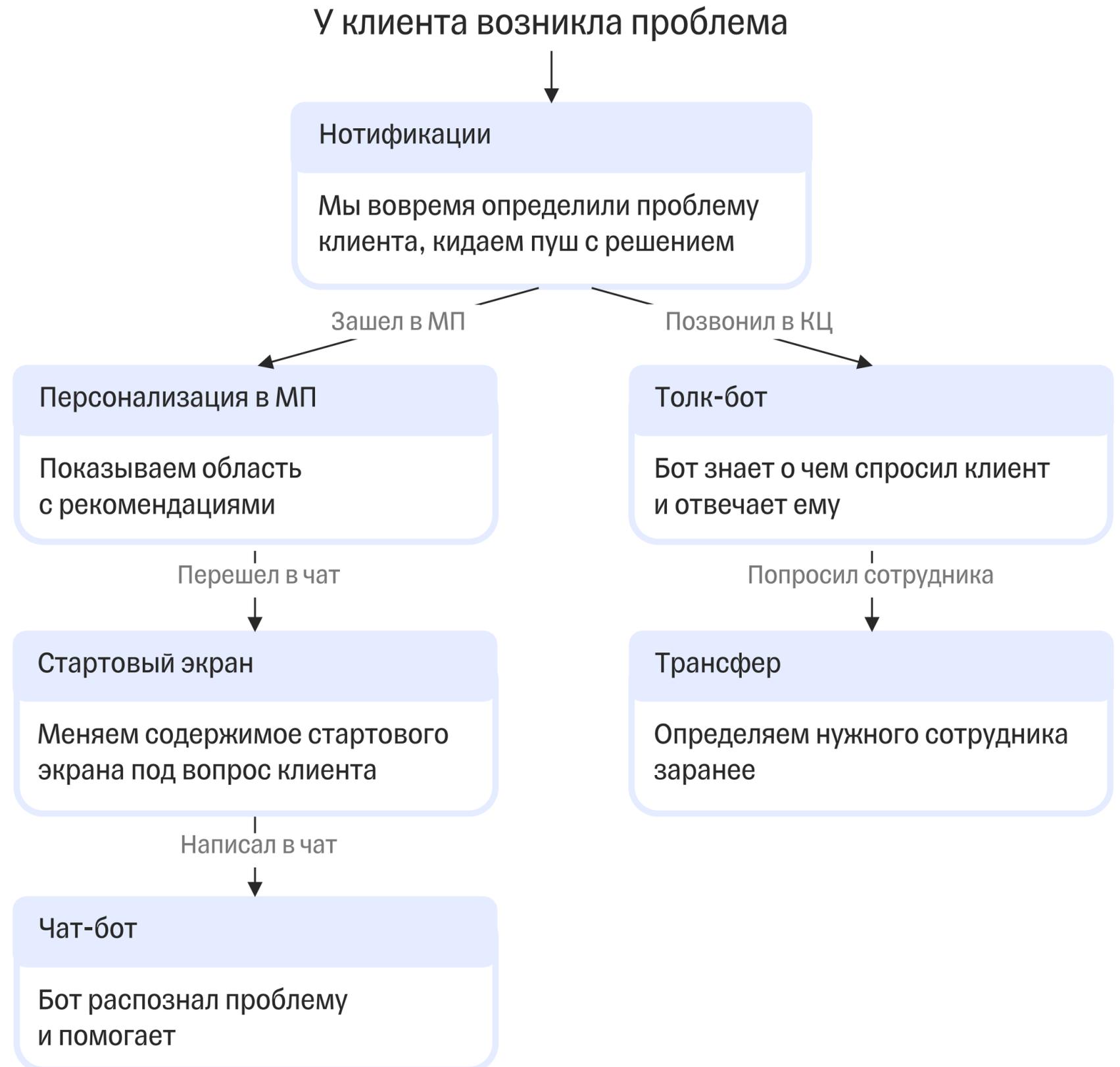
Три столпа сервиса в эпоху AI:

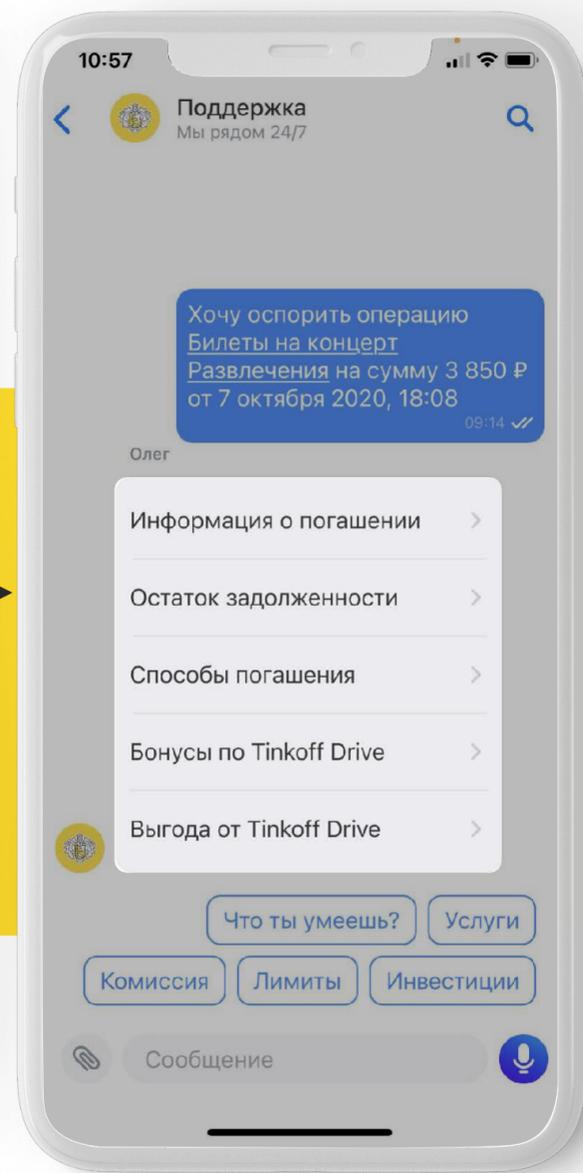
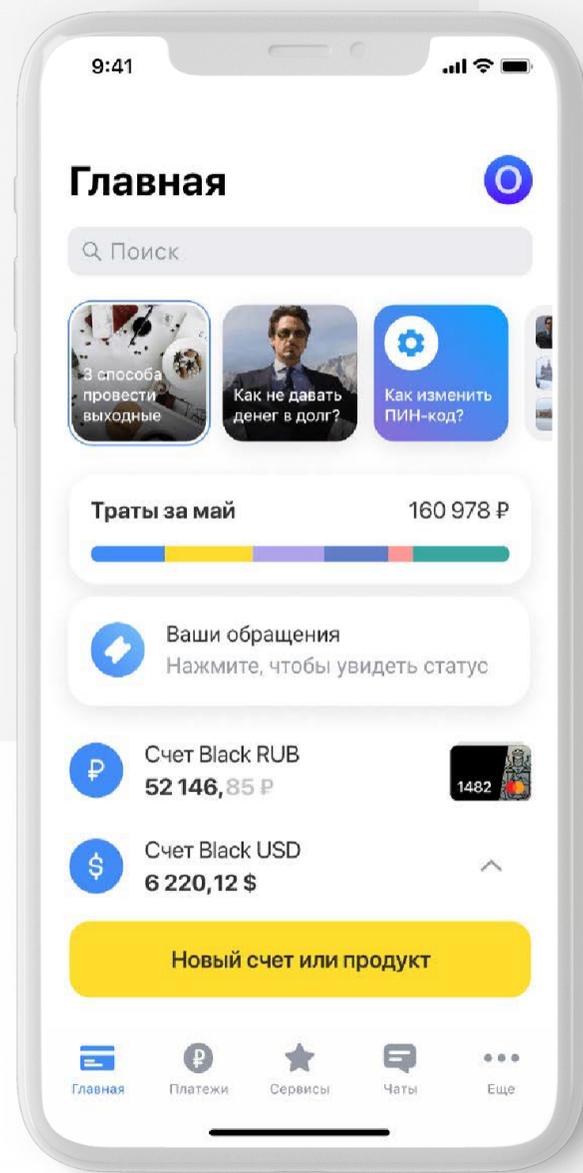
- ▶ **Проактивность** — решать проблему как можно скорее, в идеале до ее возникновения
- ▶ **Self Service** — пользователь должен иметь возможность сделать то, что ему нужно в приложении
- ▶ **Contactless service** — делать продукты, которые не требуют вопросов

Цель

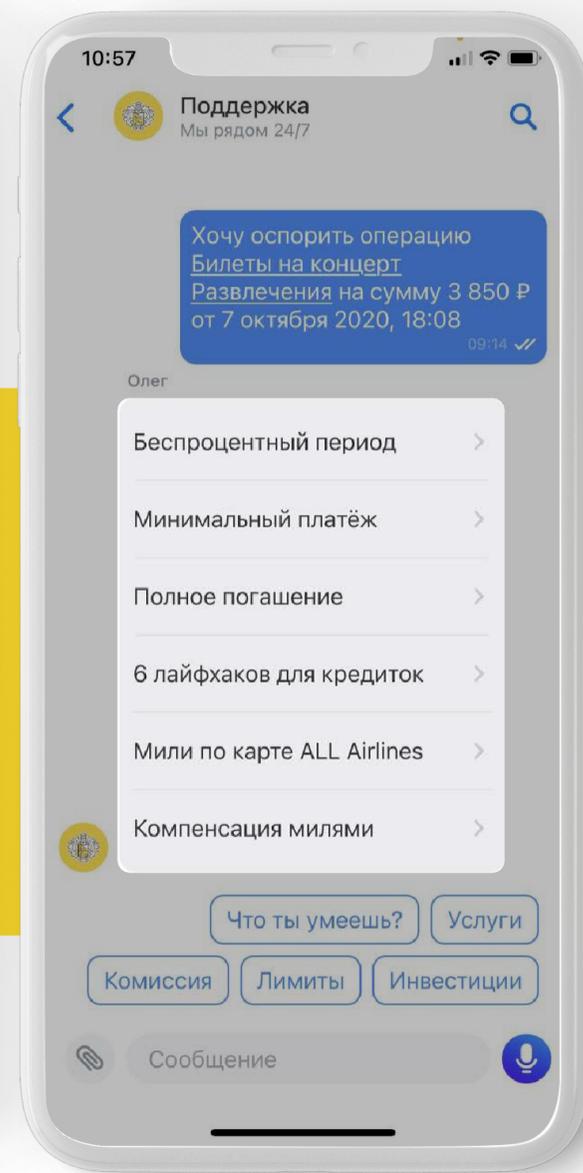
вау эффект от персонализации обслуживания:

- ▶ Сокращение времени до целевого действия
- ▶ Сокращение обращений





Перед
мин.платежом



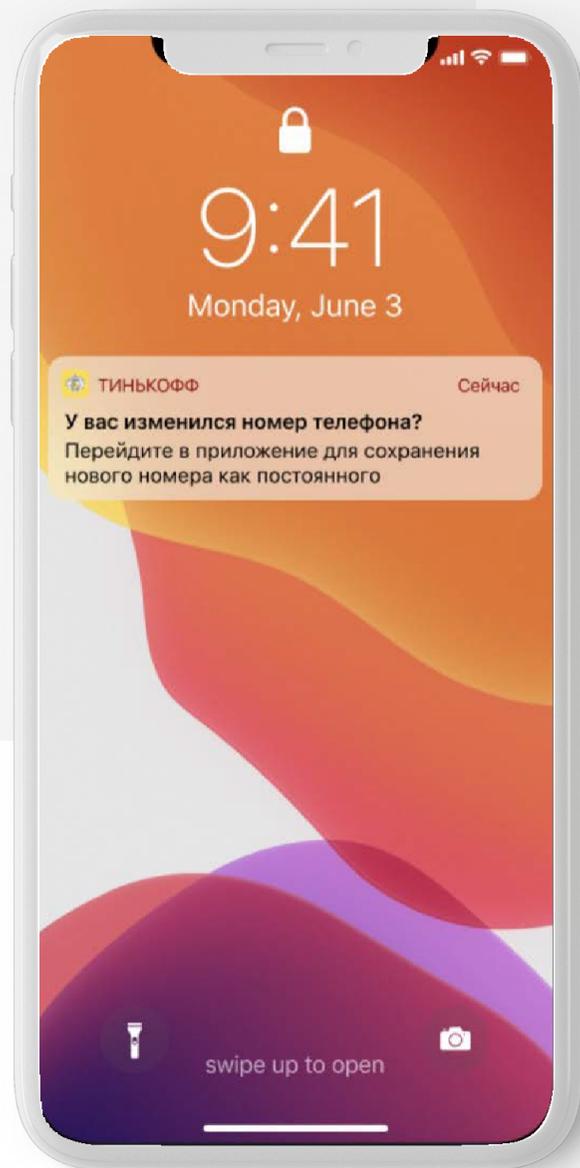
После встречи

Предиктивный интерфейс в чате

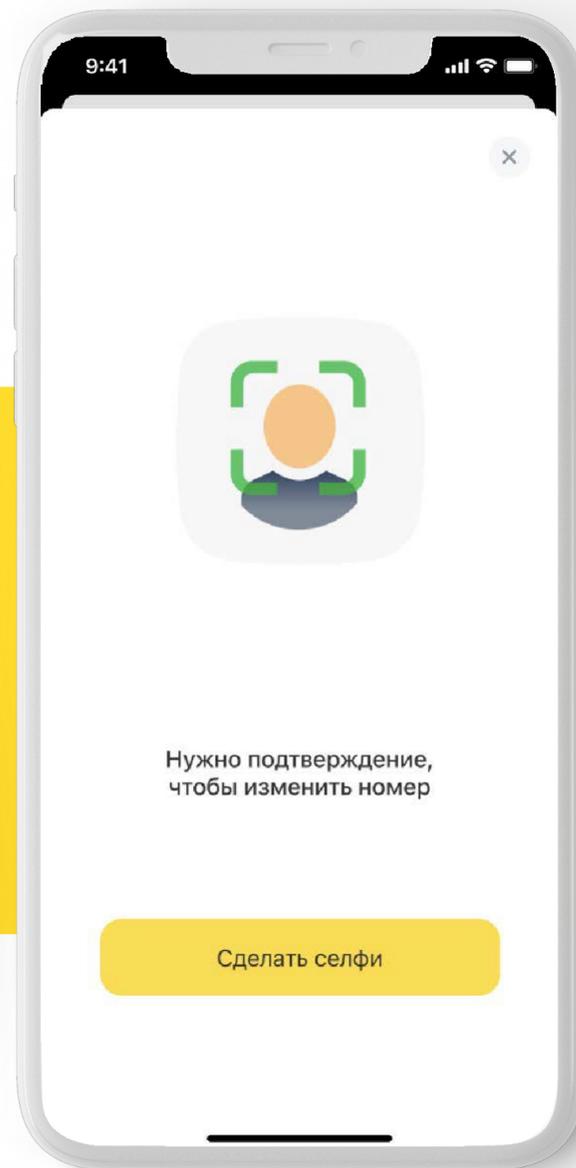
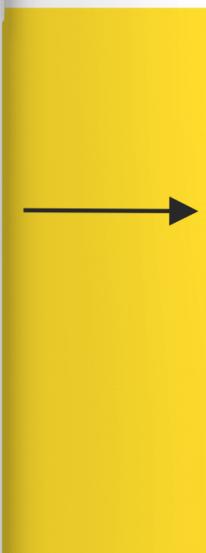
Персонализация

Меняем контент в зависимости от:

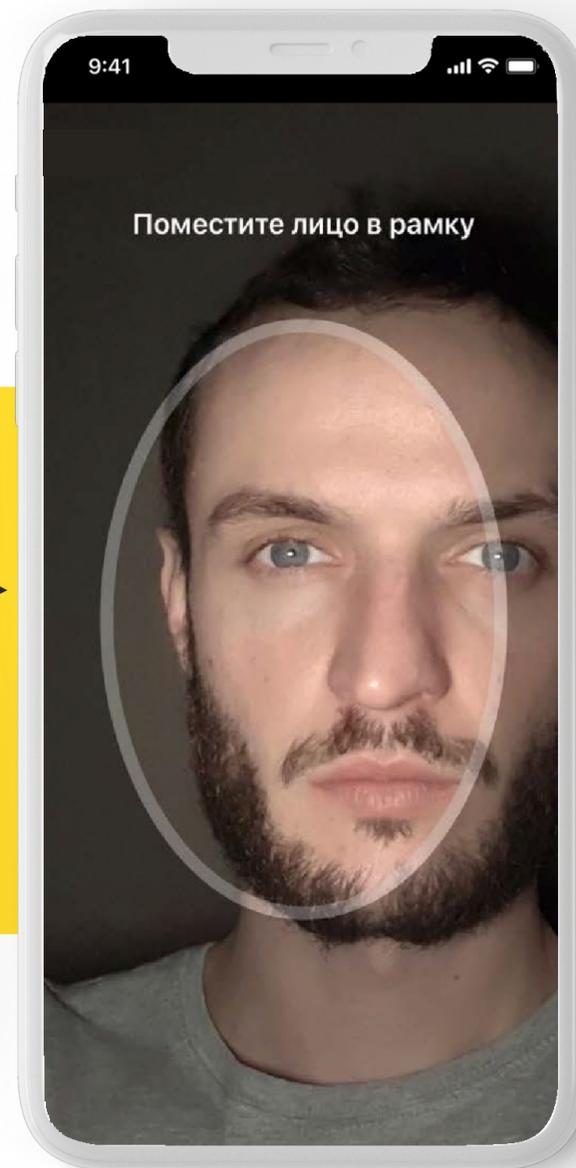
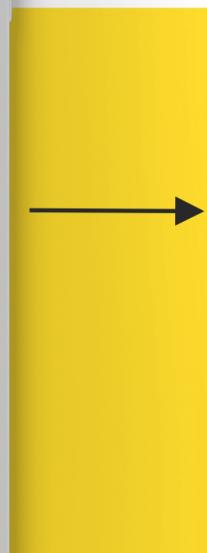
- ▶ Продукта
- ▶ Программа лояльности
- ▶ Некоторых событий (мин. платеж, выписка)



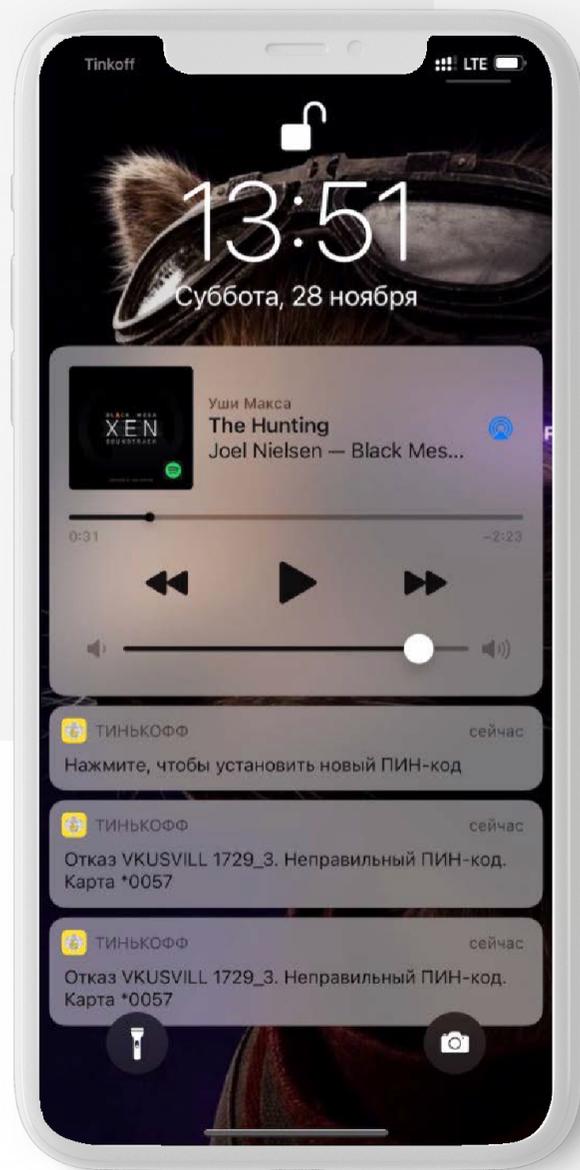
Пуш «изменить номер»



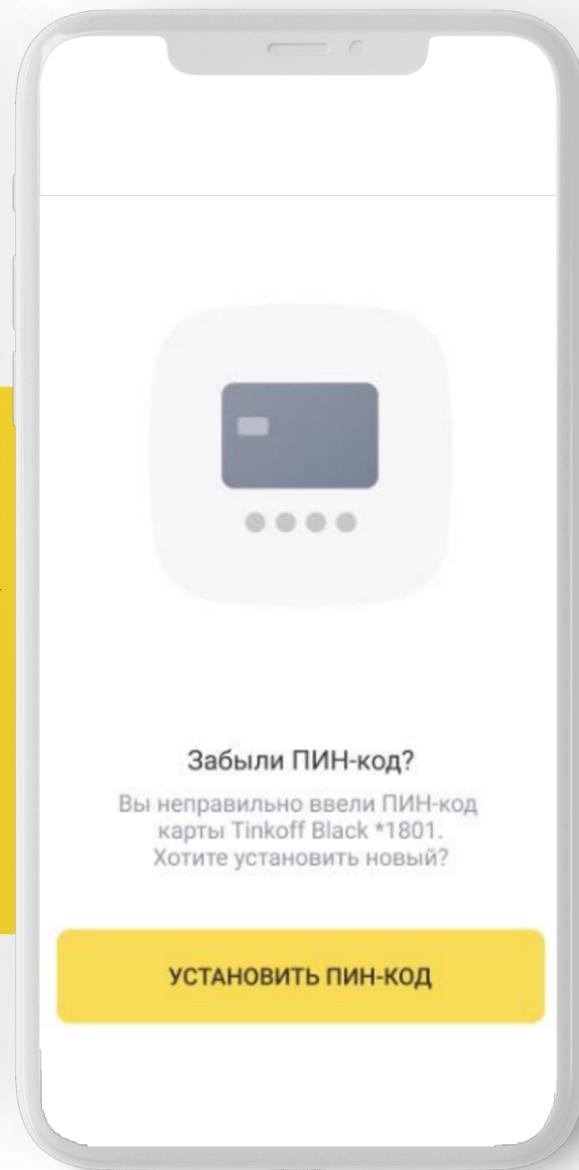
Сервис лицевой биометрии



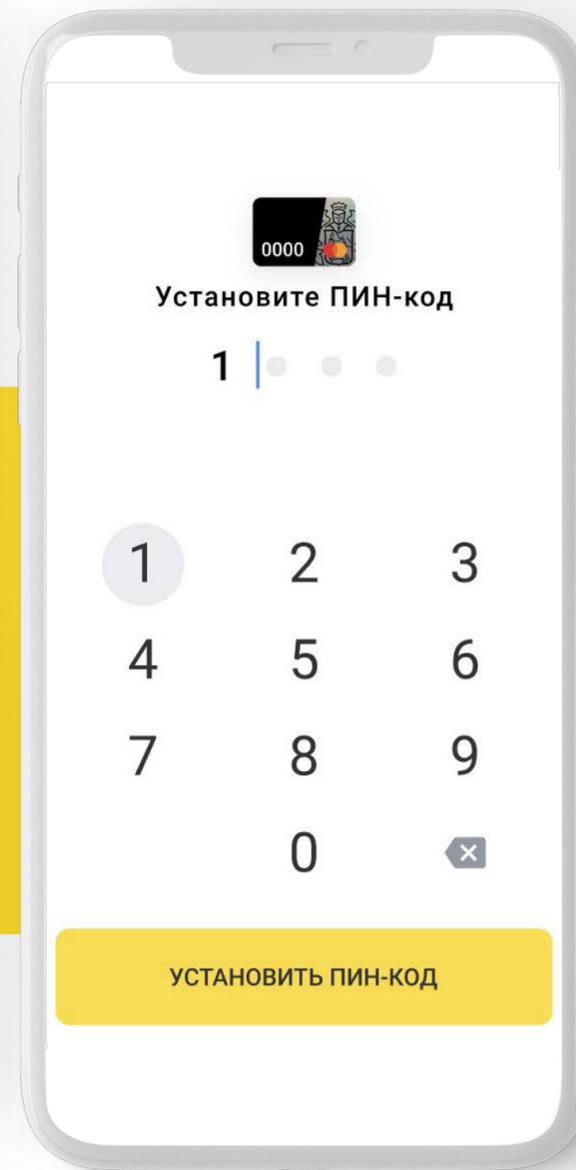
Проактивное изменение номера



Пуш при
неправильном вводе



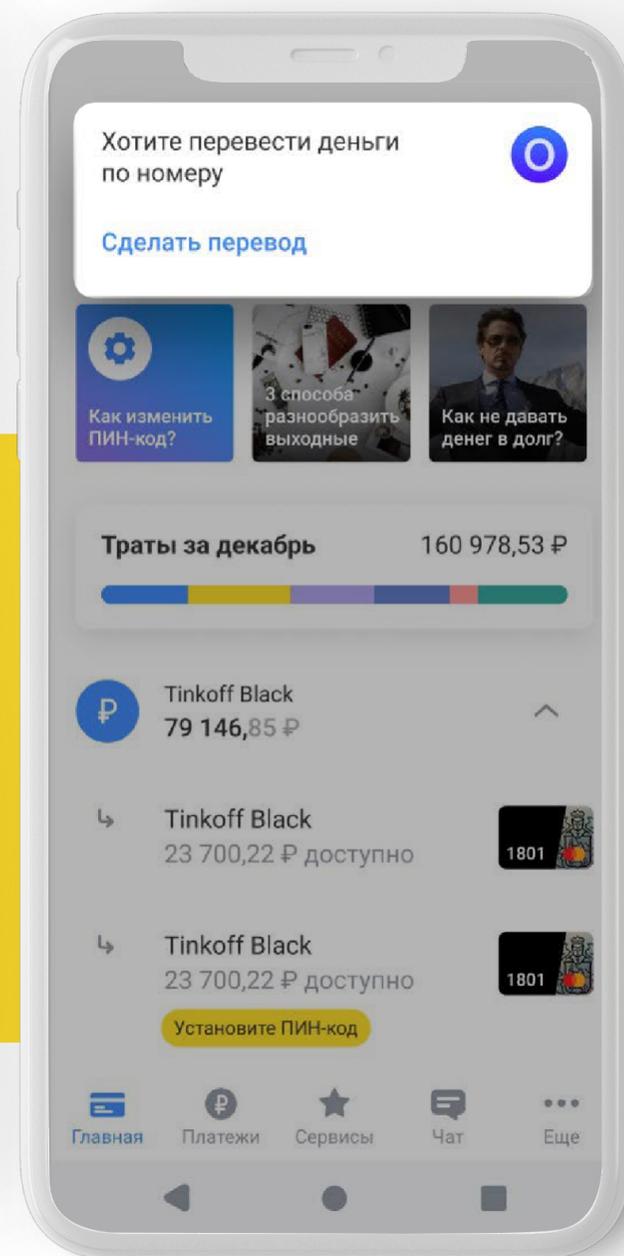
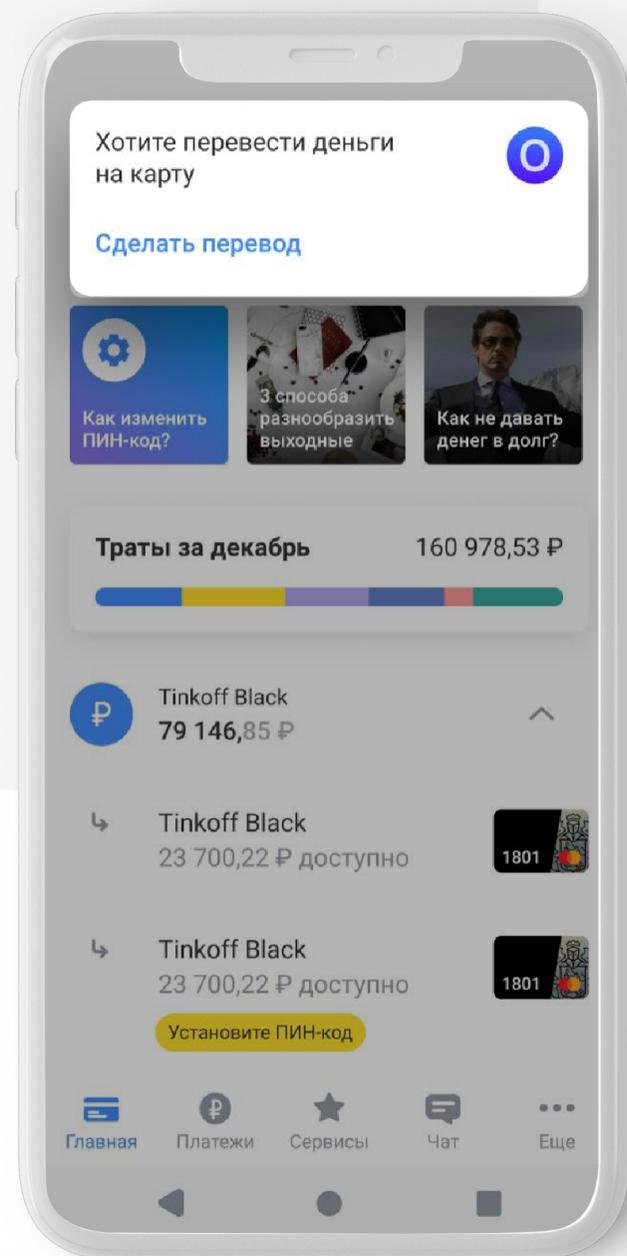
Онбординг
с предложением
сменить



Экран смены
ПИН-кода



Проактивное
изменение
пин-кода



Проактивные переводы

Пользователь скопировал номер карты или телефон

UX фишка года по версии Markswebb



Сейчас



Звонок



**Какой
вопрос?**



Будет



Звонок

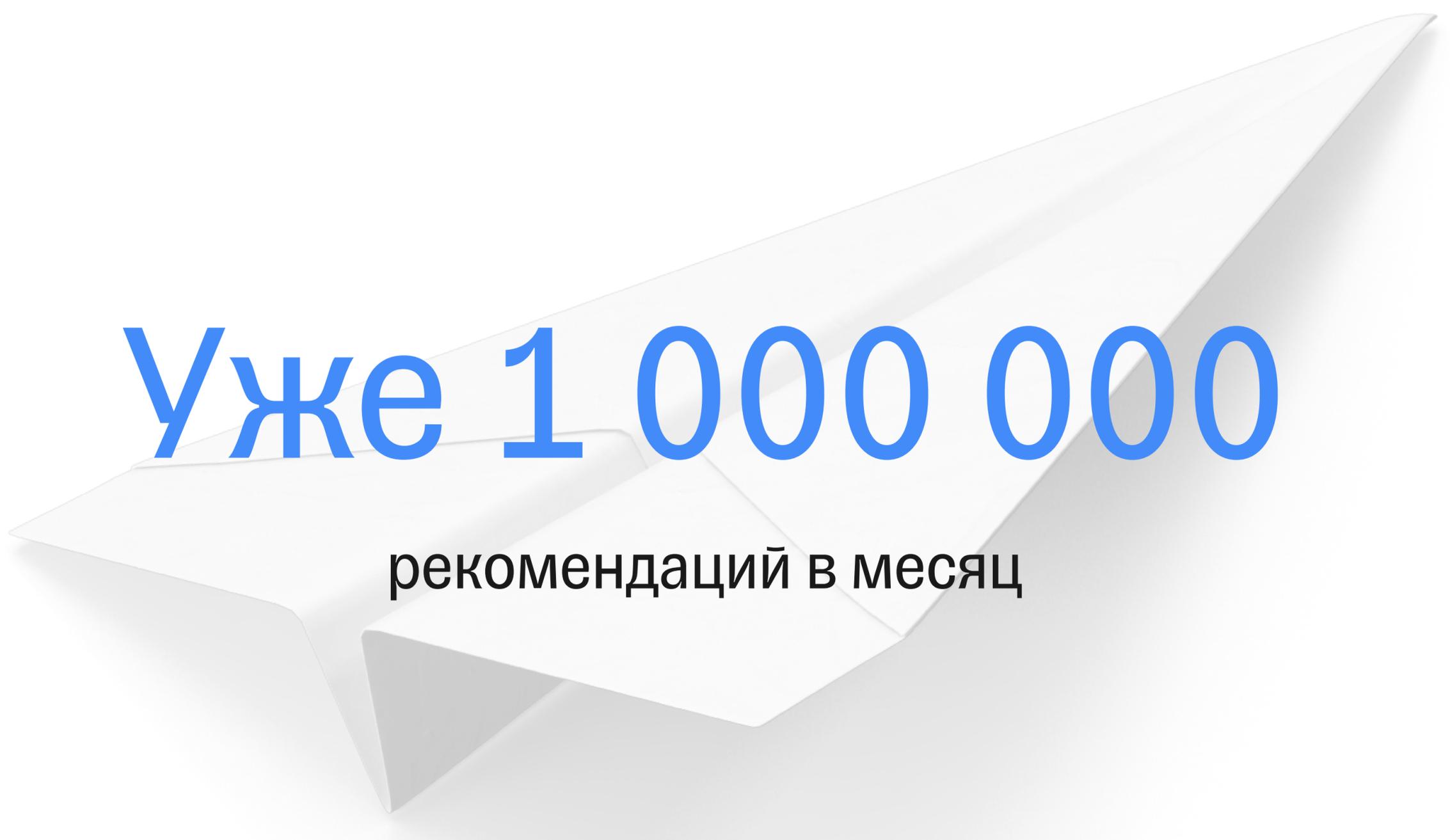


**Хотите
узнать
ПИН?**



Проактивный Бот

Пользователь скопировал
номер карты или телефон



Уже 1 000 000

рекомендаций в месяц



ТИНЬКОФФ

Спасибо за внимание!