

Кейс 3: Умный корпоративный ассистент для помощи в решении бизнес-задач



Oden's warriors



Жилкин Федор Игоревич Architecture & Engineering

> @fedoros +79312208944



Смирнов **Александр** Львович Research, Desgn & ML

@furiousteabag
+79119727982



Что ищут люди

Количество **тысяч** запросов по категориям в месяц*



Мы решили сосредоточиться на самых крупных категориях:

- Законодательство
- Бухгалтерский учет
- Личные налоги

^{*} Yandex.Wordstat май 2023



Где ищут люди

Внутри компании

AskGuru индексирует рабочие чаты, тикеты, внутреннюю базу знаний компании

Интернет, кодексы

Помимо универсальной модели знаний, AskGuru дополнен 24 кодексами и федеральными законами РФ

Проф. форумы

AskGuru дополнен 1160 статьями и вопросами с профессионального бухгалтерского сайта glavbukh.ru

Telegram

AskGuru дополнен информационными постами телеграм-канала о личных налогах @dulkarnaev

AskGuru использует **OpenAl GPT-4** для генерации ответов. Эта модель обладает широким универсальным знанием, но имеет пробелы в части современной меняющейся информации. На основании самых популярных тем запросов, мы отобрали **высокоинформативные ресурсы** для обогащения контекста модели и обеспечения **фактологической точности**



Функциональность AskGuru

Быстрый семантический поиск

АskGuru использует эмбеддинги текстовой информации и запросов, генерируемые нейронной сетью. Это позволяет поиску быть устойчивому к опечаткам и использованию синонимов. AskGuru использует специальную векторную базу данных для ускорения поиска релевантной информации.

Вы можете найти примеры поиска в предоставленном jupyter notebook

Фильтрация запросов

AskGuru позволяет задать набор правил для фильтрации по стоп-словам для предотвращения использования сервиса в недобросовестных целях

Автоматическая генерация ответа

AskGuru использует генеративные модели ОреnAI наряду с проприетарными языковыми моделями для балансировки нагрузки. Функция стриминга ответа позволяет сократить время ожидания, и дать пользователю возможность чтения во время генерации

Группы доступа

AskGuru позволяет использование групп доступа для пользователей и ресурсов. Пользователь сможет получить ресурс или ответ, опирающийся на него только в том случае, если имеет соответствующие настройки доступа.

Мы не демонстрируем эту функциональность в jupyter notebook, так как она не была заявлена в ТЗ, и чтобы не перегружать демо. Пожалуйста, дайте нам знать, если вы хотите увидеть эту функцию в действии.

Вы также можете найти информацию об использовании этой и других функций в swagger API



Логика обработки запросов





Архитектура AskGuru

Frontend



Core ML





Эмбеддинг данных

Генерирование ответа

OpenAI и проприетарные модели

Вычисления на GPU







Асинхронная обработка запросов

Логгирование

Семантический поиск

NoSQL база данных



Хранение логов запросов

Хранение правил фильтрации

Хранение вспомогательной информации

Векторная база данных



Хранение эмбеддингов данных

Быстрый поиск по заданной метрике

Архитектура AskGuru реализована в соответствии с микросервисной парадигмой. Каждая карточка, представленная на диаграмме, является отдельным сервисом. Это позволяет эффективнее масштабировать систему и осуществлять гибридное развертывание сервиса. Последнее означает, что часть AskGuru, отвечающая за хранение данных может быть развернута на серверах клиента для усиленной безопасности. Общение с backend осуществляется с помощью stateless запросов.

Примеры использования

[10]: query = "НДФЛ с иностранных граждан временно пребывающих на патенте, перечислять нужно отдельным платежом, как и в ФСС НС?" sources = api.get_ranking(query)

Самые релевантные документы:

- 1. НДФЛ с иностранных граждан временно пребывающих на патенте (Главбух: ответы на самые популярные вопросы за последний год)
- 2. Как дистанционному сотруднику платить российский НДФЛ и надо ли (Артур Дулкарнаев (налоговый юрист) Telegram-канал)
- 3. РБК пишет, что у переехавших за границу и ставших нерезидентами есть вариант зарегистрироваться самозанятыми. Я считаю, что это сработает только если вы не будете получать оплату по договору от нынешнего или бывшего работодателя (Артур Дулкарнаев (налоговый юрист) Telegram-канал)
- 4. НДФЛ с нерезидентов в 2023 году | Ставка НДФЛ для иностранных граждан (Главбух: лучшие статьи за последний год по налогам, страховым взносам и бухгалтерскому учету)
- 5. У меня часто спрашивают, мол, российский ИП на УСН с налоговой ставкой 6 процентов уехал в Армению и прожил там больше 183 дней в 2022 году. Надо ли теперь такому ИП пересчитать налоги и начать платить НДФЛ по ставке 30 процентов? Деятельность идет, деньги приходят на российский счет (Артур Дулкарнаев (налоговый юрист) Telegram-канал)

query = "Новые правила по отпускным 2023"
response = api.get_answer(query, user=USER, collections_only=COLLECTIONS_ONLY, stream=STREAM_ANSWER)
api.set_reaction(request_id=response["request_id"], rating=4)



В 2023 году утверждены новые правила предоставления отпусков. Изменения в Трудовом кодексе расширили категории работников, которые могут требовать предоставления отпуска в любое время. Вот некоторые ключевые изменения:

- 1. Сотрудникам с трех или более детьми до 18 лет предоставляется отпуск в удобное для них время до достижения младшим из детей 14 лет или одному из старших 18 лет. Как только младшему исполнилось 14 или одному из старших 18, сотрудник теряет право на льготу [1].
- 2. Сотрудникам, призванным на службу в рамках частичной мобилизации, предоставляется отпуск в любое удобное время в течение 6-ти месяцев с момента восстановления на работе (ст. 351.7 ТК) [2].

Также планируются изменения в сроке выдачи отпускных, установленном в статье 136 ТК. Госдума рассматривает проект, который касается ситуации, когда сотрудник подает заявление на отпуск за три дня или меньше до его начала. Для этого случая предусмотрят два срока в зависимости от способа выплаты денег [1].

Источники:

- 1. Новые правила предоставления отпусков в 2023 году: разъяснения
- 2. Новые правила предоставления отпусков в 2023 году

query = "штраф за неподачу отчета о движении денег"
response = api.get_answer(query, user=USER, collections_only=COLLECTIONS_ONLY, stream=STREAM_ANSWER)
api.set_reaction(request_id=response["request_id"], rating=5)



Штраф за неподачу отчета о движении денег и иных финансовых активов для физических лиц составляет от 2 000 до 3 000 рублей. Если повторно не подать отчет, то штраф составит 20 тысяч рублей [1].

Источники:

1. Максимальный штраф для физических лиц, если не подать отчет о движении денег и иных финансовых активов или подать такой отчет позже срока или подать с ошибками, составляет 3 тысячи рублей

Спасибо за возможность участвовать в хакатоне и интересный кейс!

Наша команда имела некоторые наработки по схожей задачи до хакатона, хотя многое пришлось реализовать с нуля для решения предложенного кейса.

Продукт AskGuru, разрабатываемый нами, планируется как коммерческий, и мы, внимательно изучив раздел об использовании проприетарного софта, считаем возможным заявить AskGuru как проприетарный софт, используемый в решении. Причина этого в том, что мнения в команде относительно возможности раскрытия кода, разделились.

Мы понимаем, что одной из целей tada.team является получение готового прототипа решение в виде кода, и мы готовы обсуждать передачу неисключительных прав и код всего продукта.

С нетерпением ждем вашей оценки!

Odens team