Р..Конф

3–5 декабря 2024 г.
Рэдиссон Коллекшен,
Москва + онлайн

19-я МЕЖДУНАРОДНАЯ КОНФЕРЕНЦИЯ

УПРАВЛЕНИЕ ПРОЕКТАМИ 2024:

Проектное лидерство

Мастер-класс: Нейропомощники. От мечты к реальности.

Юрий Трубицын

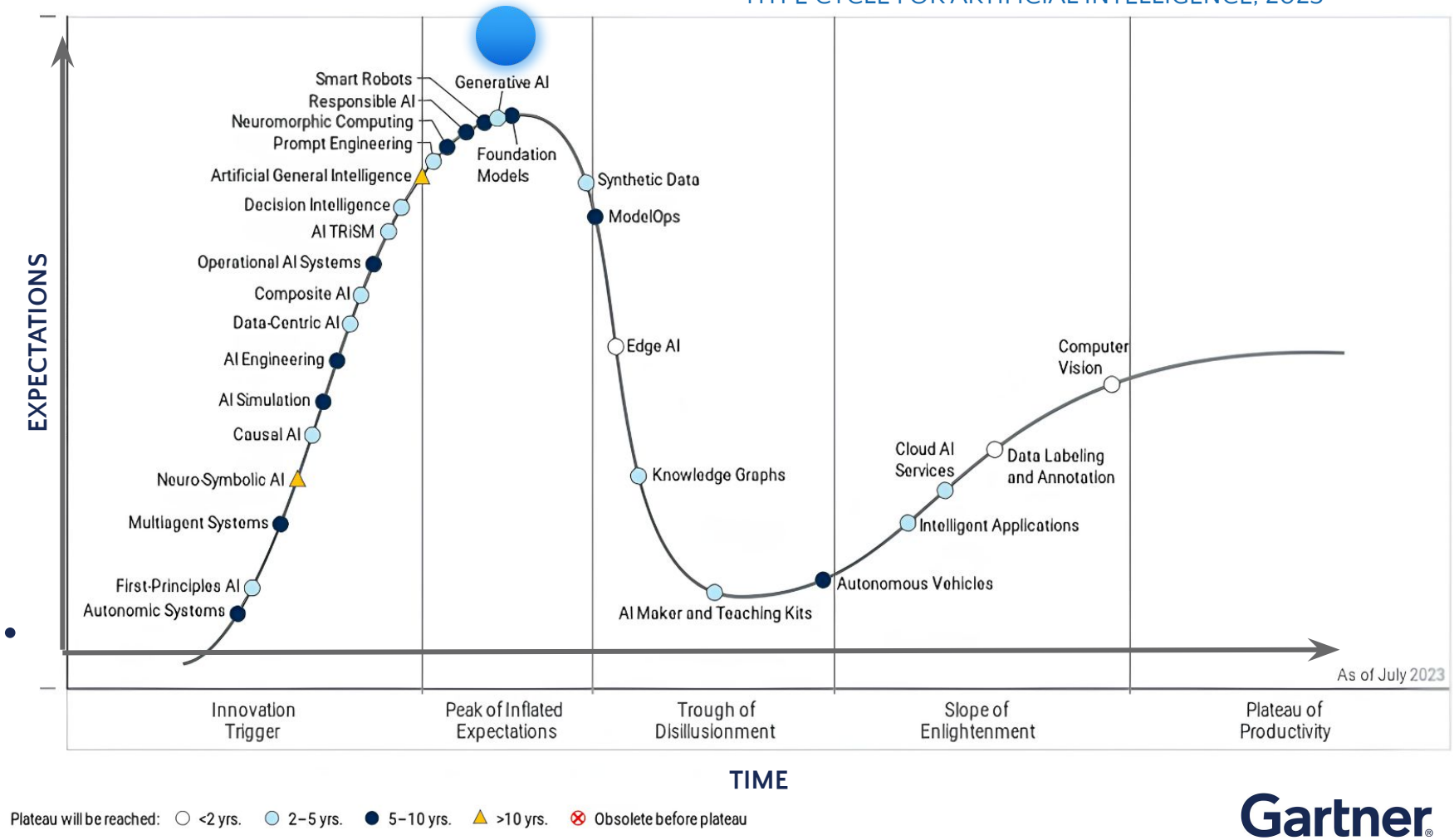
вице-президент СОВНЕТ, секретарь Экспертного Совета ЦОРПУ,
генеральный директор ООО «Адаптивные Технологии Развития»

Андрей Комлев

Руководитель направления ИИ ООО «Адаптивные Технологии Развития»,
главный конструктор ПРИИСК

КРИВАЯ ГАРТНЕРА И ГДЕ СЕЙЧАС НАХОДИТСЯ ИИ

HYPE CYCLE FOR ARTIFICIAL INTELLIGENCE, 2023



Gartner

ПОЧЕМУ В КОМПАНИЯХ С ВНЕДРЕНИЕМ ИИ «НЕ ТАК, КАК У ЛЮДЕЙ»

Некоторые причины, по которым компании ищут более сложные способы работы с ИИ, чем прямой доступ к ИИ в облаке:

- значимость использования корпоративного управленческого и производственного контента;
- требования по информационной безопасности;
- необходимость интеграции и встраивания НП-ов в корпоративную информационную среду.

О ЧЕМ МАСТЕР-КЛАСС «НЕЙРОПОМОЩНИКИ, ОТ МЕЧТЫ К РЕАЛЬНОСТИ»

В текущей стадии развития рынка ИИ, когда «технологии опережают сознание», критически важным становится

УМНЫЙ, ОСМЫСЛЕННЫЙ ЗАКАЗЧИК. То есть ВЫ!

Цель мастер-класса:

- На опыте реальных проектов посмотреть, как может проходить процесс внедрения ИИ в компании, не имеющей предыдущего опыта с ИИ работы;
- Обсудить некоторые общие особенности и трудности, связанные с внедрением ИИ в компании, и сформулировать некоторые рекомендации по реализации ИИ-проектов.

Допущения и договоренности в рамках мастер-класса:

- *мы говорим только о генеративных моделях ИИ, в основном – лингвистических (далее - ИИ, LLM);*
- *мы не осуждаем «спасет или убьет ИИ этот мир», скорее всего ИИ полезен в «пределах своей полезности» (которые пока только определяются);*
- *решения с использованием LLM будем называть Нейропомощниками (НП-ки).*

ТАЙМИНГ

АДАПТИВНЫЕ ТЕХНОЛОГИИ Развития

1.	**Рассказ с примерами**	**50 минут**

1. НП-ки на Базах Знаний (БЗ): зачем, пример, особенности внедрения
2. НП-к по транскрибации речи: примеры и особенности применения
3. НП-к оценки документов: пример, варианты применения, особенности внедрения
4. Некоторые выводы: что надо учитывать при внедрении в компании ИИ

2.	**Рассказ по дополнительной теме (выбирается голосованием)**	**15 минут**

1. Почему требуется дообучение НП-ов. RAG vs Finetuning – когда и что применять
2. НП-к оценки качества работы НП-ов. Не понятно, как мы жили до этого.
3. Платформа для работы с ИИ в компании: ПРИИСК. Что это и в чем выгоды применения.
4. + 15 минут вопросов или уйдем пораньше ☺️

3.	**Ответы на вопросы**	**15 минут**

ДЕЙСТВУЮЩИЕ ЛИЦА.
КТО МЫ, ПОЧЕМУ СЧИТАЕМ, ЧТО МОЖЕМ БЫТЬ ВАМ ИНТЕРЕСНЫ

АДАПТИВНЫЕ ТЕХНОЛОГИИ Развития

- Мы 20 лет (в этом году юбилей) успешно занимаемся услугами в области проектного управления.
- Два года плотно занимаемся вопросами применения ИИ в корпоративном управлении и, прежде всего, в проектной деятельности компаний.
- Совместно с нашими клиентами и партнерами, с применением ИИ, решили ряд интересных бизнес-задач (о некоторых - которых рассказ далее).
  Среди клиентов и партнеров есть компании, оказывающие разные услуги в области проектного управления:
  - разработка и внедрение ИСУП;
  - консалтинговые услуги; сертификация организаций;
  - обучающие услуги; оценка компетенций, сертификация специалистов.

ПРИИСК

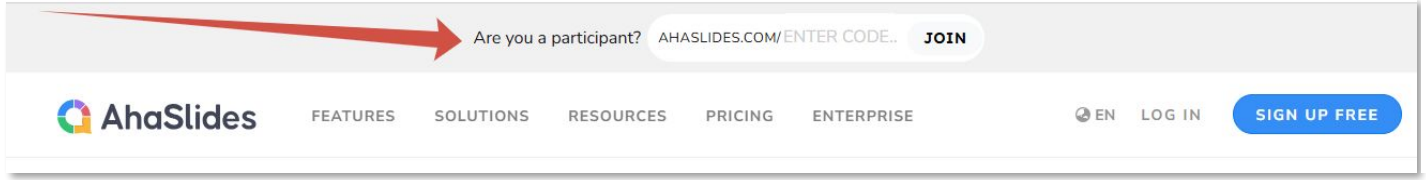
Нами разработана специализированная платформа для разработки, внедрения и эксплуатаций решений с применением ИИ в современных компаниях (ПРИИСК). ПРИИСК включен в Реестр отечественного программного обеспечения.

ДЕЙСТВУЮЩИЕ ЛИЦА ☺️

КАКОЙ ЕСТЬ ОПЫТ РАБОТЫ С ИИ У ВАС И В ВАШЕЙ КОМПАНИИ.

Используйте QR-код для участия в опросе.

Или перейдите на сайт https://ahaslides.com/
и введите код **0OJS3** в верхней строке

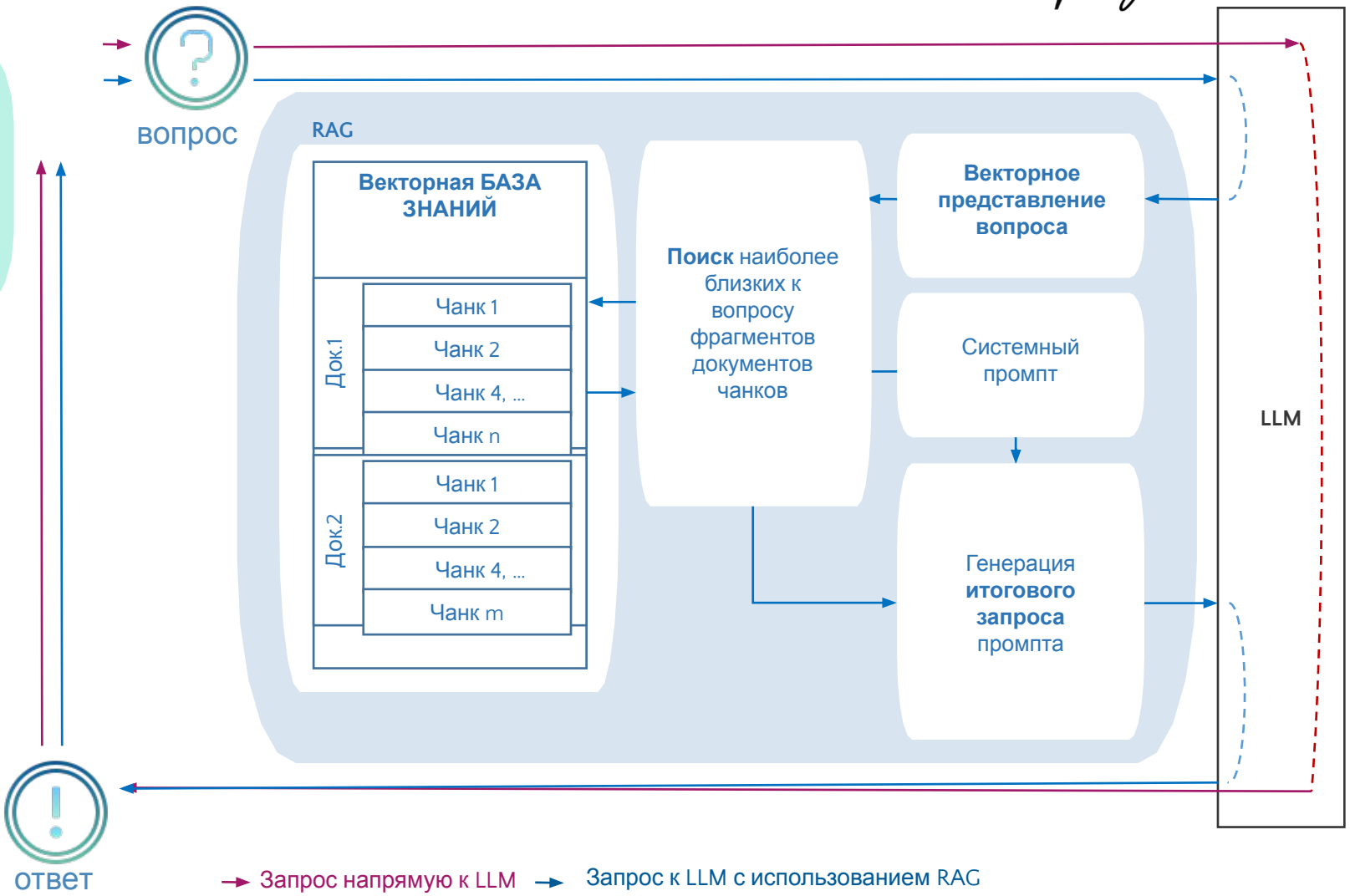


1. НП-КИ НА ОСНОВЕ БАЗ ЗНАНИЙ. **ЗАЧЕМ (1)**, ПРИМЕР, ОСОБЕННОСТИ ВНЕДРЕНИЯ

Анонс №1.

Почему требуется дообучение НП-ов.

RAG vs Finetuning – **когда и что применять.**



Упрощенная схема RAG (Retrieval-Augmented Generation)

1. НП-КИ НА ОСНОВЕ БАЗ ЗНАНИЙ. **ЗАЧЕМ (2)**, ПРИМЕР, ОСОБЕННОСТИ ВНЕДРЕНИЯ

АДАПТИВНЫЕ ТЕХНОЛОГИИ Развития

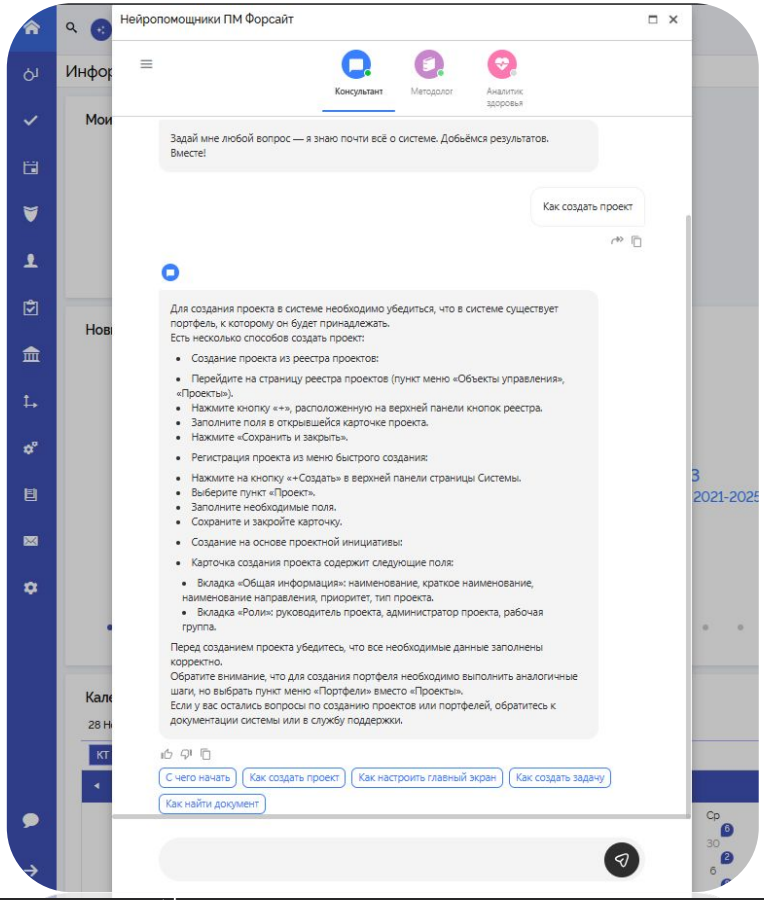
Пример запроса «без» и «с» применением RAG

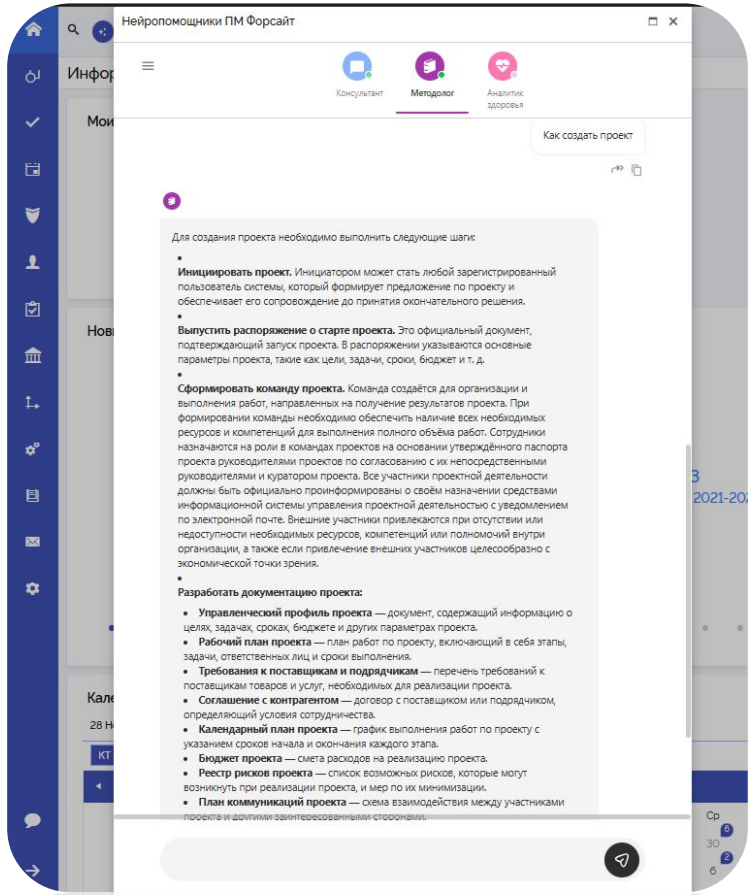
Кто должен входить в команду проекта

- Ты – опытный специалист, который отвечает за методологию управления проектами в компании. Твоя задача - помогать сотрудникам. Это очень важно для сотрудников и для компании. Дай краткий, структурированный ответ на вопрос:
- **КТО ДОЛЖЕН ВХОДИТЬ В КОМАНДУ ПРОЕКТА**
- Не выдумывай ничего от себя, используй только следующие фрагменты из документов: **« ТЕКСТ ИЗ ЧАНКА 1», « ТЕКСТ ИЗ ЧАНКА 2», « ТЕКСТ ИЗ ЧАНКА 3»**
- В конце ответа приведи название документа, из которого была отобрана релевантная для ответа информация.
- Следуй следующим рекомендациям по оформлению ответа:
  Разделяй информацию на абзацы для лучшей читабельности, используй маркированные списки - или * для перечисления связанных фактов без строгой последовательности,
  Используй нумерованные списки 1., 2., 3., и т.д. для описания последовательных шагов или когда порядок следования важен.
- Если имеющейся информации в документах недостаточно для точного ответа на вопрос, вежливо ответь: «Извините, сейчас я не могу найти точный ответ на ваш вопрос. Давайте попробуем сформулировать его по-другому или дать дополнительный контекст».

1. НП-КИ НА ОСНОВЕ БАЗ ЗНАНИЙ. ЗАЧЕМ, **ПРИМЕР (1)**, ОСОБЕННОСТИ ВНЕДРЕНИЯ.

2 типовых НП на основе баз знаний – «**Техподдержка**» – эксперт по технической поддержке и «**Консультант**» – методологический консультант по процессам и НМА, но дообученные на разных базах знаний, с разными системными промтами и настройками LLM на один и тот же вопрос дают принципиально разные результаты





1. НП-КИ НА ОСНОВЕ БАЗ ЗНАНИЙ. ЗАЧЕМ, **ПРИМЕР (2)**, ОСОБЕННОСТИ ВНЕДРЕНИЯ

Пример работ по формированию Базы Знаний НП-ка Консультанта по Проектной деятельности в крупной ИТ-компании

1. Нормативно-регламентные документы ПД
2. Учебные ролики и материалы
3. Эксперты интервью

Статьи интернет, книги

БЗ Методолог УП

- Общие положения управления проектами
- Портфели проектов
- Запуск проекта
- Планирование проекта
- Реализация проекта
- Завершение проекта
- Договорная работа
- Глоссарий

Эксперты

Оценка:

от «нормально для начала»

до «плохо, не работает»

Контрольные вопросы

1	2	3	ОЭ1	ОЭ 2
56%	60%	75%	65%	80%
52%	60%	72%	70%	75%
25%	56%	75%	51%	71%
35%	60%	78%	45%	75%
31%	62%	72%	35%	76%
16%	45%	82%	75%	80%
12%	60%	80%	56%	81%

1. НП-КИ НА ОСНОВЕ БАЗ ЗНАНИЙ. ЗАЧЕМ, ПРИМЕР, НЕКОТОРЫЕ **ОСОБЕННОСТИ ВНЕДРЕНИЯ**

АДАПТИВНЫЕ ТЕХНОЛОГИИ Развития

- Опыт первых внедрений решений на Базах Знаний (RAG) был учтен и теперь, в штатных внедрениях, уже на входе в проект планируется состав областей знаний, по которым должен работать НП-к, а также определяются источники и материалы, на основании которых будет наполняться База Знаний.
- Использование специальных инструментов проверки (контроля) качества работы НП-ка позволяет существенно снизить затраты на внедрение НП-ка и контролировать его качество в ходе эксплуатации.
- Плохие ответы НП-ка (особенно в начале внедрения) часто говорят не о недостатках в технической реализации решения. Они говорят о недостатках, разрывах, пробелах в информационном обеспечении сотрудников (с учетом их реальных компетенций).
- Внедрение НП-ка позволяет выявить такие пробелы, устранить, и, в идеале, отладить систему актуализации информации (нормативно-регламентной, справочной, учебной, лучших практик и т.д.).
  Внедрение НП-ка позволяет повысить качество и доступность требуемой для сотрудников информации.
- Как значимый для Компании результат внедрения НП-ка, многие отмечают формирование в Компании **Актива** по соответствующей области **знаний** (деятельности).
  Актива - **управляемого Компанией.**
  Актива - находящегося **в собственности Компании**.

Анонс №2.

НП-к оценки качества работы НП-ов.

Непонятно, как мы жили до этого.

2. НП-К ПО ТРАНСКРИБАЦИИ РЕЧИ. ПРИМЕРЫ И ОСОБЕННОСТИ ПРИМЕНЕНИЯ.

Примеры применения «транскрибатора»:

Кейс 1. Быстрый способ пополнения базы знаний НП-ка, через интервью с экспертами.

Кейс 2. Составная часть НП-ка
- по оценке компетенций специалистов по УП;
- по обработке интервью кандидатов на вакансии (HR);
- по обработке интервью при обследовании компании;

...

Кейс 3. Протоколирование совещаний. С последующей обработкой протоколов по заданным сценариям, например:
- составление протоколов (по типам совещаний);
- выделение списка поручений;
- формирование содержательных тезисов обсуждения и пополнение БЗ специализированного НП-ка;

...

Кейс 4. Рабочий инструмент РП, администратора проекта свободного применения.

2. НП-К ПО ТРАНСКРИБАЦИИ РЕЧИ. ПРИМЕРЫ И **ОСОБЕННОСТИ ПРИМЕНЕНИЯ В КОМПАНИИ**.

Функция транскрибации сейчас доступна во многих приложениях, в том числе для проведения встреч.
Есть доступные облачные сервисы по транскрибации, платные и почти бесплатные.
В больших количествах случаев использование таких сервисов оправдано.

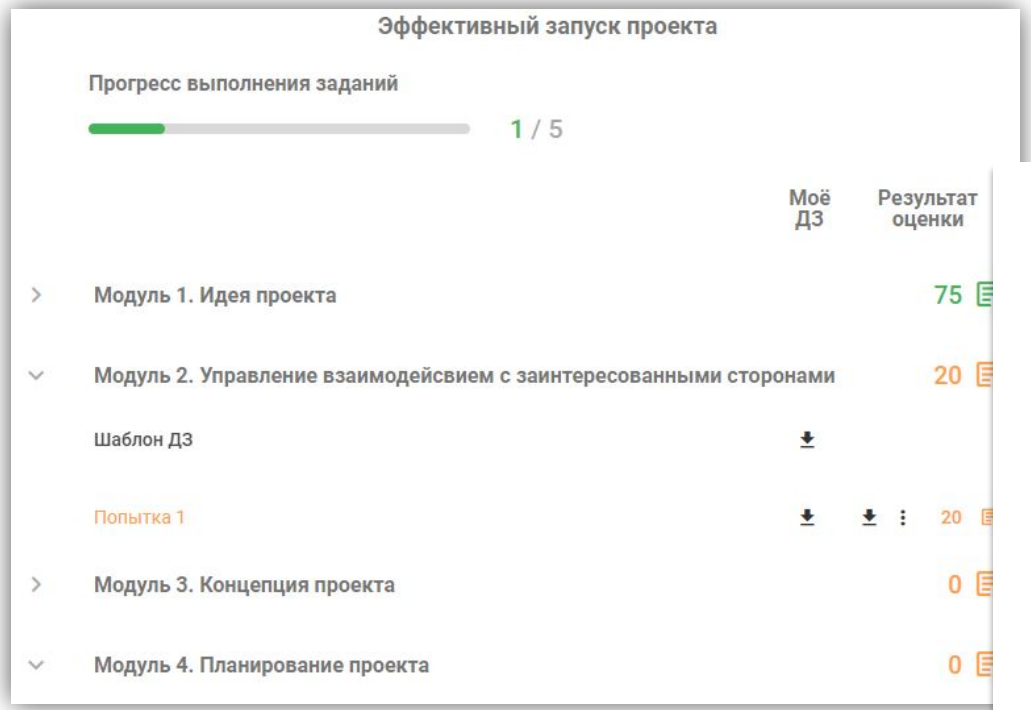
Почему мы сделали отдельный сервис в платформе ПРИИСК:

A. **Безопасность**
Записи, которые мы транскрибируем часто содержат конфиденциальную информацию. Которую лучше обрабатывать в «контуре компании». Это возможно и не очень дорого. Мы так и делаем.

B. **Гибкость настроек** (для улучшения качества распознавания и других задач), например:
- использование специализированных глоссариев с терминами, сокращениями, фамилиями и т.д.
- упрощение, автоматизация процессов транскрибации и обработки записей.

C. **Возможность интеграции** сервиса транскрибации в «длинные цепочки» обработки:
- примеры кейсов, когда сервис транскрибации является частью других НП-ов рассматриваются далее.

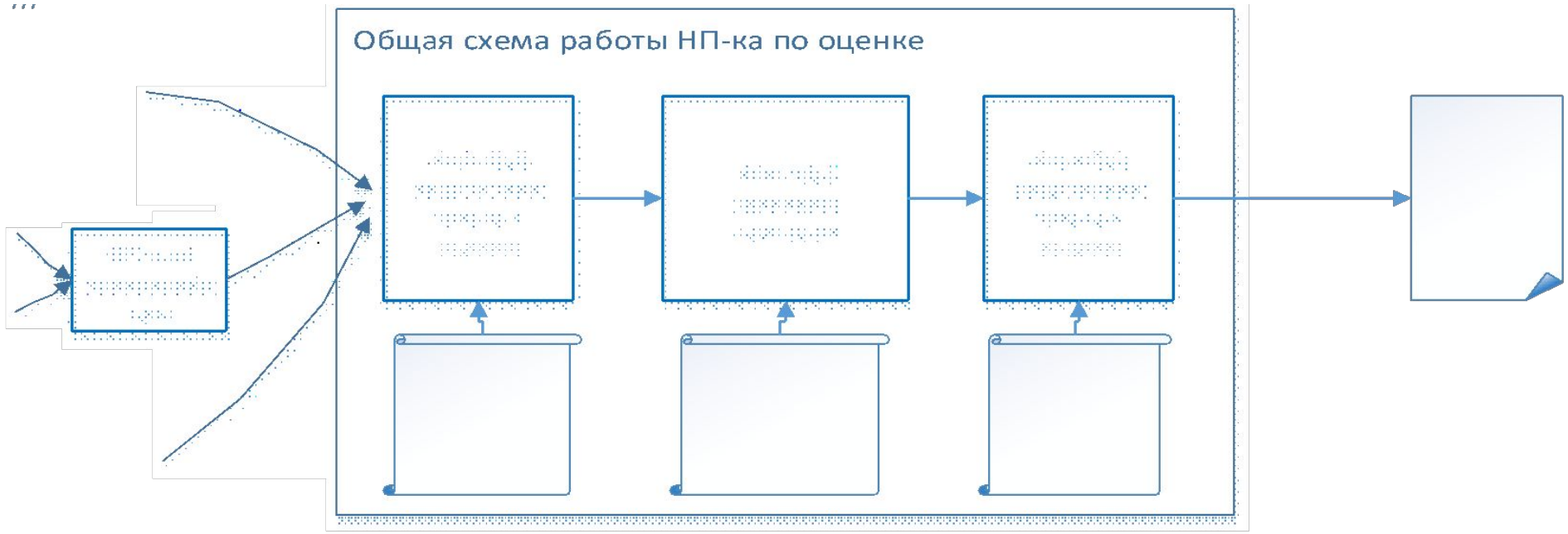
3. НП-К ОЦЕНКИ ДОКУМЕНТОВ. ПРИМЕР, **ВАРИАНТЫ ПРИМЕНЕНИЯ, ОСОБЕННОСТИ ВНЕДРЕНИЯ**



Пример проверки документа НП

Пройти стартовый модуль и попробовать работу НП по проверке документов

3. НП-К ОЦЕНКИ ДОКУМЕНТОВ. ПРИМЕР, **ВАРИАНТЫ ПРИМЕНЕНИЯ**, ОСОБЕННОСТИ ВНЕДРЕНИЯ



Пояснения по особенностям работы НП-ка оценки в разных кейсах применения:

Кейс 1. Оценка компетенций РП по интервью.

Кейс 2. Оценка документа (например, по рискам проекта; по заинтересованным сторонам и т.д.)

Кейс 3. Оценка собеседований с кандидатами (HR).

4. НЕКОТОРЫЕ ВЫВОДЫ (НЕОЧЕВИДНЫЕ): ЧТО НАДО УЧИТЫВАТЬ ПРИ ВНЕДРЕНИИ ИИ В КОМПАНИИ

АДАПТИВНЫЕ ТЕХНОЛОГИИ Развития

- Через какое-то время, если Вы правильно внедряете ИИ, нейропомощников в компании будет **МНОГО**.
- ИИ, и в частности генеративные модели, позволяют компании знания, ранее распыленные, превратить в **УПРАВЛЯЕМЫЙ АКТИВ** и конкурентное преимущество.

4. НЕКОТОРЫЕ ВЫВОДЫ (НЕОЧЕВИДНЫЕ): ЧТО НАДО УЧИТЫВАТЬ ПРИ ВНЕДРЕНИИ ИИ В КОМПАНИИ (2)

АДАПТИВНЫЕ ТЕХНОЛОГИИ Развития

Через какое-то время, если Вы правильно внедряете ИИ, нейропомощников в компании **будет МНОГО.**

Некоторые следствия:

- В компании необходимо **сформировать** / оформить выделенную Группу (лабораторию, отдел, проект, центр компетенций,…), которая будет и драйвером и **центром компетенций** по внедрению возможностей ИИ в деятельность Компании.
  Применение сторонних решений «под ключ» иногда (не всегда) возможно.
  Но, как всегда, сделать счастливыми «под ключ» усилиями со стороны – невозможно.
- Формат внедрения ИИ в компании очень **Программу по внедрению** …, с гибким управлением инициативами (гипотезами, кейсами) по полезному применению ИИ с одной стороны, и планомерным формированием инфраструктуры, процессов внедрения и применения ИИ, с другой стороны.
- В большинстве случаев, в компаниях больше «микро- компаний», целесообразно применять единую **платформу для эксплуатации** (а скорее всего и разработки/настройки/внедрения и эксплуатации) нейропомощников.
  Платформы появляются на рынке, мы можем рассказать подробнее о платформе ПРИИСК.

Анонс №3.
Платформа для работы с ИИ в компании: ПРИИСК. Что это и в чем выгоды применения.

4. НЕКОТОРЫЕ ВЫВОДЫ (НЕОЧЕВИДНЫЕ): ЧТО НАДО УЧИТЫВАТЬ ПРИ ВНЕДРЕНИИ ИИ В КОМПАНИИ (3)

АДАПТИВНЫЕ ТЕХНОЛОГИИ Развития

ИИ, и в частности генеративные модели, позволяют компании знания, ранее распыленные, превратить в УПРАВЛЯЕМЫЙ АКТИВ и конкурентное преимущество.

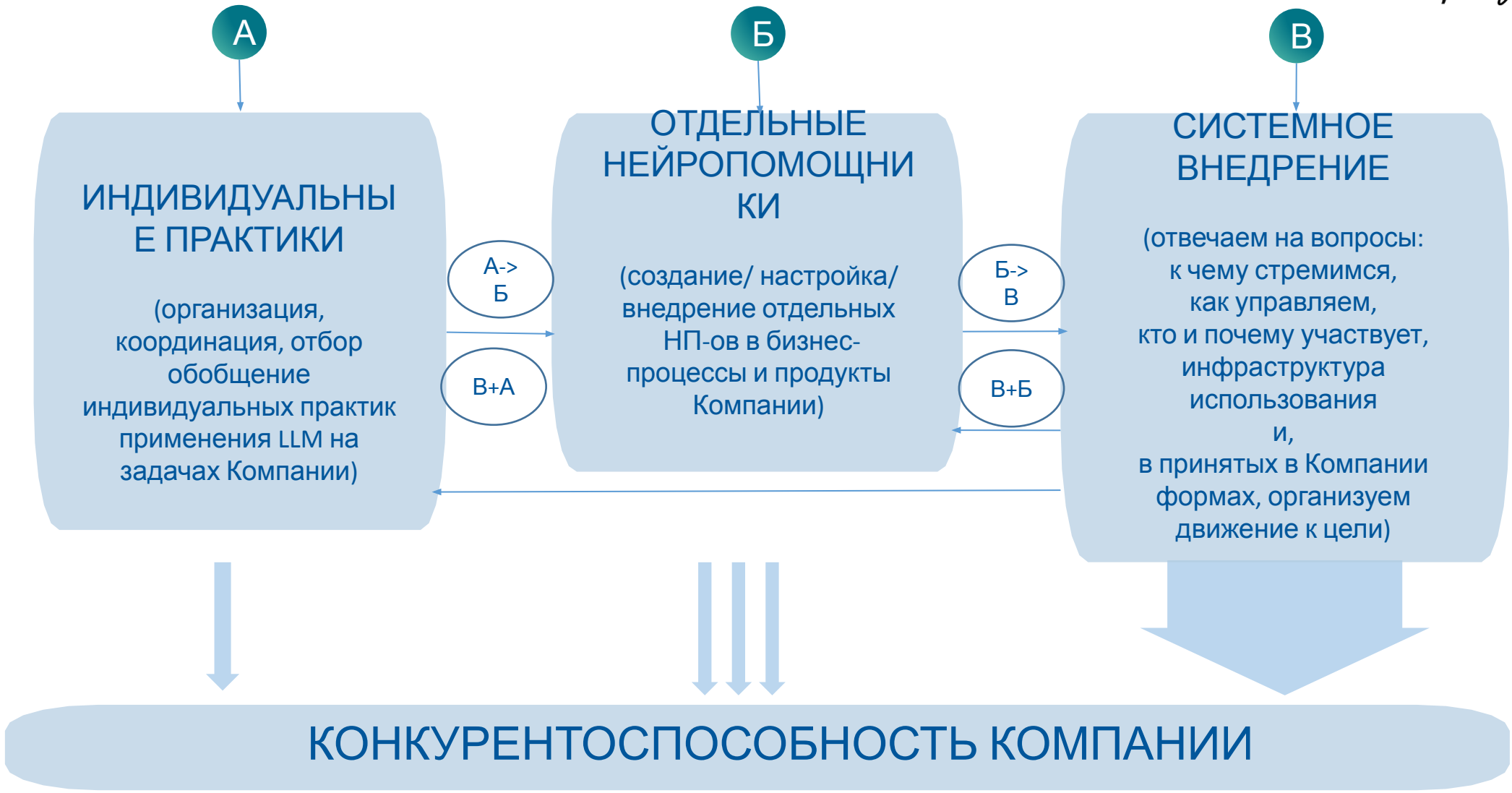
Некоторые следствия.

- Внедрение НП-ов потребует более **серьезного участия представителей Вашей Компании** (носителей знаний). К этому надо быть готовым.
  Многие документы (регламенты, положения, …), разрабатывались для применения, когда человек читает этот документ. Данный формат часто бывает неоптимальным для использования в НП-ке и может потребовать дополнительной обработки. Вообще, внедрение НП-ка часто приводит к переосмыслению «знаний» в компании по заданной теме.

- Использование НП-ов существенно повышает доступность информации. В НП-ке можно быстро актуализировать, изменять информацию.
  С одной стороны, у сотрудников должны сформироваться привычка и навыки использовать НП-ки. Например, задать правильный вопрос, не совсем простое дело.
  С другой стороны, работа по **актуализации, дополнению, расширению информации** в НП-ах, скорее всего должна быть отчасти фоновой, но **выделенной задачей жизни компании**.

- Вопрос **ответственности за применение информации от НП-ов** сейчас один из самых обсуждаемых, сложных и спорных. Решение надо выработать и сделать так, чтобы все сотрудники его понимали и им руководствовались.

Сценарии организации работ по внедрению ИИ



ГОЛОСУЕМ ЗА ТЕМУ ДОПОЛНИТЕЛЬНОГО РАЗДЕЛА МАСТЕР КЛАССА

1. Почему требуется дообучение НП-ов. RAG vs Finetuning – когда и что применять
2. НП-к оценки качества работы НП-ов. Не понятно, как мы жили до этого.
3. Платформа для работы с ИИ в компании: ПРИИСК. Что это и в чем выгоды применения.
4. + 15 минут вопросов или уйдем пораньше ☺️

Используйте QR-код для участия в опросе. Или перейдите на сайт https://ahaslides.com/ и введите код **DP2DM** в верхней строке

МАТЕРИАЛЫ К ТЕМЕ 1:

Почему требуется дообучение НП-ов. RAG vs Finetuning – когда и что примен

АДАПТИВНЫЕ ТЕХНОЛОГИИ Развития

Критерий	RAG (Retrieval-Augmented Generation)	Finetuning
Время	Процесс быстрый. И создание новой Базы знаний для Нейропомошника (далее НП), и ее актуализация занимает минуты, редко часы.	Процесс дообучения может занимать недели, а процесс создания своей модели заточенной под определенную специфику месяцы.
Предварительная подготовка документов	Документы должны содержать структурированный текст. Предварительная подготовка необходима, но может быть выполнена заранее сформированными скриптами, например убрать ссылки на п. и заменить их кусочками текста, или заменить термины или сокращения их расшифровками и синонимами. Общая трудоемкость средняя. В отдельных случаях может и не проводится, если качество текстов высокое.	Требуется большой заранее подготовленный и полностью проверенный большой датасет, который бы включал большое кол-во примеров вопросов и гарантированно правильными ответами. (слово большой 3 раза использовано специально)
Возможности использования	При данном виде обучения, чем более узкая специфика созданного НП - тем лучше. 3 отдельных НП, более точно дообученных на определенную специфику будут отвечать гораздо лучше, чем 1 большой и универсальный НП.	Модель может использоваться в разных сферах и подходах. При наличии очень большого массива данных (гигабайты текстов) становится единственно возможной.
Техническая сложность	Низкая сложность. Работы может выполнять сотрудник прошедший небольшие курс (2-3 дня) обучения работы с RAG и LLM.	Требуются высококвалифицированные программисты.
Требования к инфраструктуре	Нет отдельных требований. Все шаги выполняются на непосредственно в среде функционирования НП и LLM.	Для обучения требуются большие вычислительные мощности видеокарт (для обучения LAMA1 было потрачено более 20 млн$ только на работу видеокарт).
Прозрачность процесса и Галлюцинации LLM	При использовании RAG, LLM существенно менее склонны к галлюцинациям, т.к. для каждого вопроса передается необходимый контекст для формирования ответа и процесс разбит на несколько этапов (формирование вектора вопроса пользователя, поиск подходящих chunk в БЗ, генерация ответа на предоставленных данных и тд) и полностью подконтролен платформе ПРИИСК. Всегда можно проанализировать каждый из этапов, например какие Chunk были выбраны для формирования ответа и имеется возможность вносить изменения в каждый из этапов, для получения лучшего результата.	Нужен постоянный контроль и борьба с галлюцинациями. Процесс не подконтролен. LLM будет "черным ящиком", а повлиять на результат можно только постоянным дообучением на все более очищенных и проверенных датасетах, результат при этом процессе нельзя прогнозировать.
Требования к LLM	Требования к сложности LLM низкие, т.к. весь необходимый контекст передается вместе с вопросом и генерацию ответа может обеспечить и LLM достаточно низкой производительности и для ее работы не потребуется дорогие видеокарты.	Для поддержания качества ответов, чем большее сложная LLM (поддерживает большое кол-во параметров) - тем лучше. С соответственным увеличением требований к видеокартам.
Быстродействие и масштабируемость	Требуется немного больше времени для генерации ответа. Масштабирование только горизонтальное.	В среднем более быстрый ответ, т.к. генерация ответа формируется в LLM на основании меньшего использования токенов. Возможны различные подходы для масштабирования нагрузки.
Замена LLM	При появлении более совершенной LLM, возможен достаточно простой процесс проверки ее качества для использования в созданном НП и возможного перехода на нее, для улучшения качества.	Жесткая привязка к уже обученной LLM. Для возможной замены LLM процесс обучения нужно проводить полностью заново.
Общая стоимость	Низкая.	Очень дорого.

МАТЕРИАЛЫ К ТЕМЕ 1

Краткие выводы из предыдущего слайда

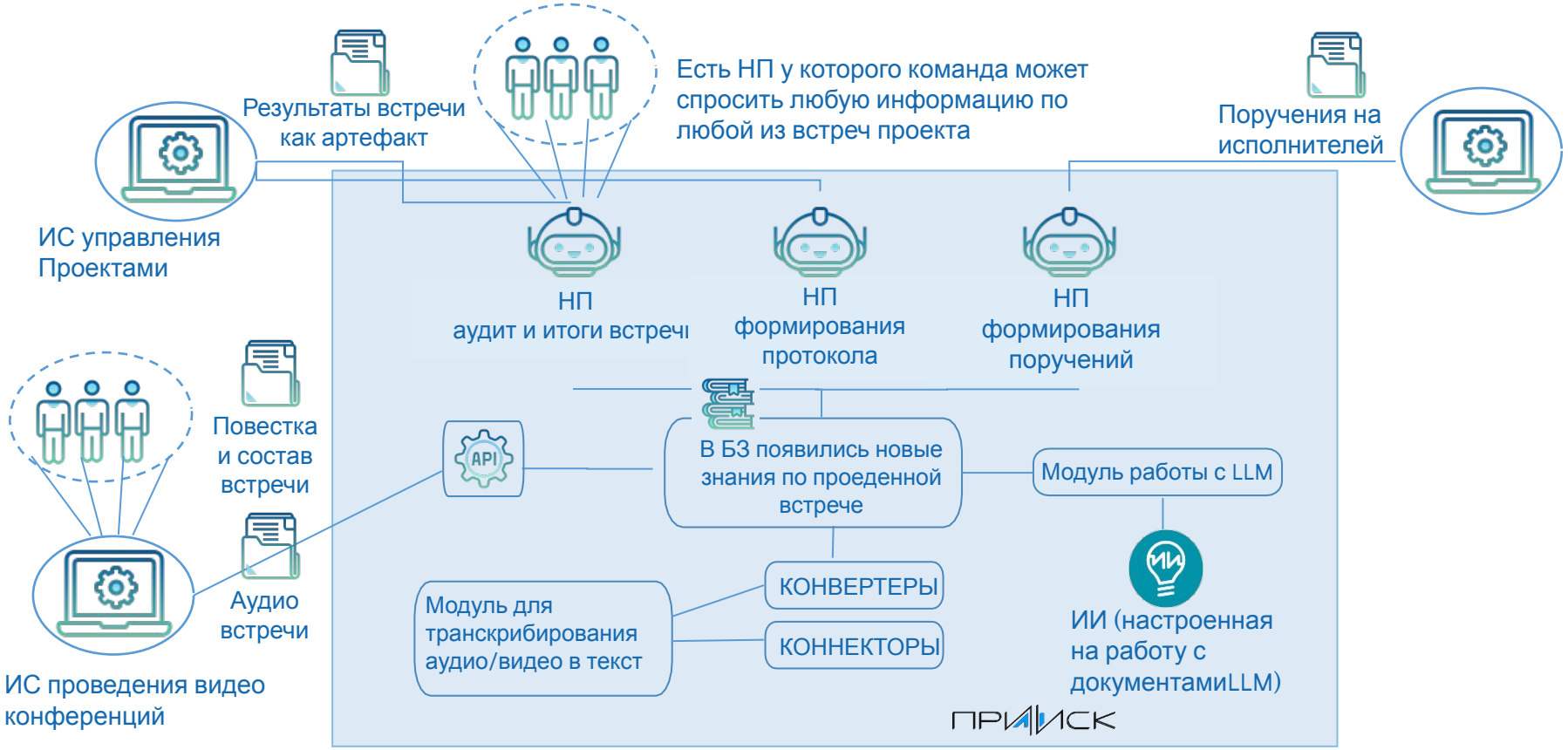
- Если нет ограничения, что для дообучения обязательно нужно использовать **ОЧЕНЬ** большой массив документов, для получения результата нужного качества, то **RAG** подход показывает себя гораздо лучше **Finetuning**.
- Дообучение через **RAG** подход занимает минимум времени (минуты), не требует высокой квалификации сотрудников и гораздо более терпим к качеству датасета для дообучения LLM и формирования БЗ.
- Прозрачность процесса через **RAG** существенно выше и можно повлиять на результат на каждом из шагов процесса. Нет привязанности к конкретной LLM и можно ее просто заменить в дальнейшем, если выйдут более современные и качественные модели.
- Через RAG подход требования к качеству самой LLM (и соответственно к инфраструктуре функционирования) ниже и с учетом остальных факторов общая стоимость внедрения будет существенно дешевле.

МАТЕРИАЛЫ К ТЕМЕ 2

НП-к оценки качества встречи и формирование протоколов и поручений.

Непонятно, как мы жили до этого.

Процесс аудита протокола встречи через НП



МАТЕРИАЛЫ К ТЕМЕ 2

Пример системного промта для Протокола встречи с использованием RAG

В качестве Нейропомощника по оценке качества протоколов встреч, ты получаешь вопрос пользователя [ВОПРОС:] и извлекаешь нужные данные из отдельных фрагментов текста через RAG [ФРАГМЕНТ №1:], [ФРАГМЕНТ №2:], каждый раз вам дано два текста: первый – это список заявленных вопросов «Повестка дня», второй – это отчёт о ходе встречи «Протокол встречи». Твоя миссия — синтезировать информацию в конечный ответ [ОТВЕТ:], который будет понятен и информативен для пользователя. При формировании ответа следуй этим рекомендациям:

Ответ формируй на русском языке. Стремись к полноте и точности ответа, основываясь только на предоставленной информации. Ничего не придумывай, если информации недостаточно для полного ответа, укажи это в ответе.

Абзацы: Разделяй информацию на абзацы для лучшей читабельности.

Списки:

*- Используй маркированные списки (- или *) для перечисления связанных фактов без строгой последовательности.*

- Используй нумерованные списки (1., 2., 3., и т.д.) для описания последовательных шагов или когда порядок важен.

Выделения:

*- Жирный текст (выделяется как *жирный текст*) используй для ключевых пунктов или важных моментов.*

- Курсив (выделяется как _курсив_) применяй для акцентирования, но не для основных пунктов.

МАТЕРИАЛЫ К ТЕМЕ 3

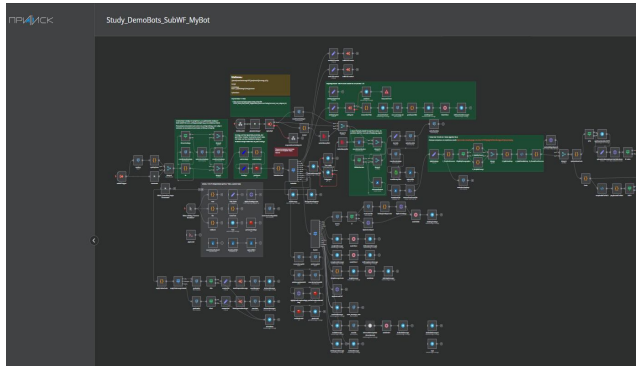
Платформа для работы с ИИ в компании: ПРИИСК. Что это и в чем выгоды применения.

* По результатам опроса участников круглого стола «Искусственный интеллект для управления проектами: тренды, возможности, первый опыт», Аналитический центр при Правительстве РФ, 19.04.2024 г., 750 опрошенных

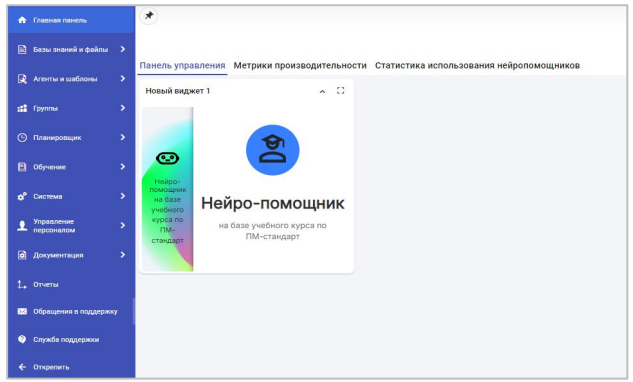
МАТЕРИАЛЫ К ТЕМЕ 3

Платформа для работы с ИИ в компании: ПРИИСК. Что это и в чем выгоды применения.

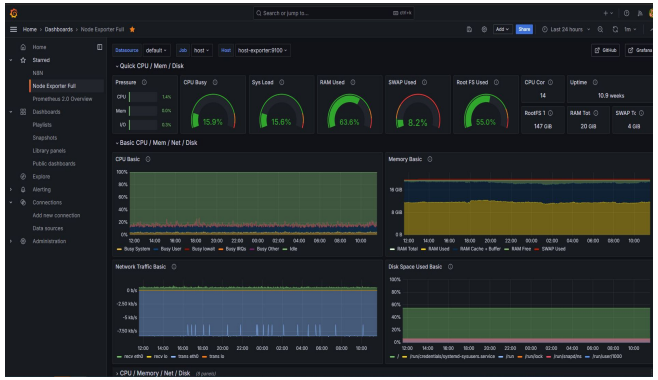
Автоматизация процессов (nocode/lowcode)



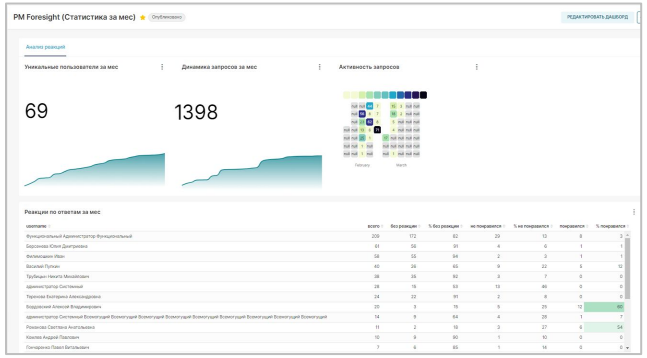
Модуль Администрирования – Управление БЗ, НП, LLM, интерфейсы, матрица доступа и др.



Мониторинг



Модуль BI



Подсистемы / технологии

docker

PostgreSQL

MINIO

KEYCLOAK

Xinference

redis

Prometheus

Grafana

drant

ClickHouse

kafka

python

ANGULARJS

АДАПТИВНЫЕ ТЕХНОЛОГИИ Развития

БОЛЬШЕ ИНФОРМАЦИИ ПО ТЕМЕ

ВИДЕО, СТАТЬИ, РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ