

**Европейская экономическая комиссия****Конференция европейских статистиков****Семьдесят вторая пленарная сессия**

Женева, 20 и 21 июня 2024 года

Пункт 5 предварительной повестки дня

**Использование искусственного интеллекта  
и больших языковых моделей в официальной  
статистике и авторитетных  
геопространственных данных****Совместная работа по продвижению использования  
искусственного интеллекта в целях официальной  
статистики: обзор инициатив и результатов работы  
Группы высокого уровня по модернизации официальной  
статистики****Записка Группы по большим языковым моделям Группы  
высокого уровня по модернизации официальной статистики  
и секретариата***Резюме*

В настоящем документе представлен обзор инициатив в области искусственного интеллекта (ИИ), предпринятых Группой высокого уровня по модернизации официальной статистики (ГВУ-МОС) с 2019 года по настоящее время, и результатов, полученных в результате совместной работы. Приложение представляет собой выдержку из аналитического доклада «Использование больших языковых моделей в целях официальной статистики», подготовленного специальной целевой группой Группы высокого уровня по модернизации официальной статистики. Полный текст этого аналитического документа доступен [здесь](#).

Документ представляется для обсуждения на сессии Конференции европейских статистиков по теме «Использование искусственного интеллекта и больших языковых моделей в официальной статистике и авторитетных геопространственных данных».



## I. Введение

1. Искусственный интеллект (ИИ) обладает огромным потенциалом для статистических организаций. Он может сделать формирование статистики более эффективным, автоматизируя определенные процессы или помогая людям выполнять их. Кроме того, ИИ позволяет статистическим организациям использовать новые типы данных, такие как данные социальных сетей и изображения, что дает им возможность предоставлять обществу и политикам более подробные и своевременные сведения.
2. Многие статистические организации начали внедрять эту новую технологию, чтобы повысить актуальность и качество официальной статистики. Однако, как это часто бывает с новыми технологиями на ранних стадиях, каждая организация оказывается ограниченной в ресурсах, чтобы самостоятельно исследовать весь потенциал этой технологии. Это делает объединение опыта и знаний различных организаций бесценным для облегчения внедрения новой технологии.
3. Группа высокого уровня по модернизации официальной статистики (ГВУ-МОС) служит международной платформой для специалистов-практиков из национальных и международных статистических организаций для обмена опытом и извлеченными уроками, совместной разработки передовой практики и рамок, чтобы вместе находить оптимальный маршрут в этом быстро меняющемся ландшафте. Изначально основное внимание уделялось машинному обучению (МО), совокупности методов ИИ, которое предполагает обучение компьютерных алгоритмов извлекать информацию из данных и делать прогнозы или принимать решения без явного программирования, но впоследствии сфера его применения расширилась и стала включать в себя более современные разработки в этой области, такие как большие языковые модели (БЯМ).
4. Цель этой записки — представить обзор инициатив, предпринятых ГВУ-МОС с 2019 года по настоящее время, и материалов, разработанных в результате совместной работы.

## II. Работа по теме машинного обучения (2019–2022 годы)

5. Дискуссия о МО зародилась в сообществе ГВУ-МОС в середине 2010-х годов, например в контексте больших данных. Однако она получила значительный импульс после того, как в 2018 году был опубликован позиционный документ по МО, подготовленный Сетью передовых исследований и изысканий (СПИИ), группой ГВУ-МОС по перспективному обзору в контексте модернизации. Этот документ привел к созданию проекта ГВУ-МОС по МО, который впоследствии привел к созданию Группы по машинному обучению, возглавляемой Управлением национальной статистики (УНС) Соединенного Королевства Великобритании и Северной Ирландии.

- **Проект ГВУ-МОС по машинному обучению (2019–2020 годы):** проект был начат в рамках ГВУ-МОС в марте 2019 года и завершился в конце 2022 года. Проект состоял из трех рабочих модулей: 1) пилотные исследования; 2) качество; и 3) интеграция. Проект начался с небольшой группы из 11 участников, но постепенно превратился в крупное сообщество по интересам, в состав которого входили более 120 участников из 37 национальных и международных организаций. Более подробную информацию о проекте по МО можно найти на специально посвященной этому [вики-странице](#).
- **Группа УНС-ЕЭК ООН по машинному обучению (2021–2022 годы):** широкое сообщество по интересам, созданное в рамках проекта по МО, смогло продолжить свою работу благодаря УНС, которое выразило готовность координировать дальнейшее международное сотрудничество в области МО. По сравнению с проектом по МО, который больше фокусировался на экспериментах вокруг МО («Что мы можем делать с помощью МО?»), больше внимания в работе Группы по МО уделялось вопросам интеграции («Как интегрировать МО в повседневную работу?»). Помимо продолжения пилотных исследований с более разнообразными вариантами использования

(например, веб-скрейпинг данных, данные системы автоматической идентификации, оценка малых территорий, оптимизация маршрутов), группа по МО изучала такие темы, как проблемы перехода от экспериментов к производству, этические соображения, качество учебных данных и инфраструктуры. Группа по МО насчитывала в своем составе более 400 человек из 35 с лишним стран и 20 с лишним международных организаций. Более подробную информацию о работе Группы по МО в каждом году можно найти на [вики-странице ML Group 2021](#) и [вики-странице ML Group 2022](#).

6. За эти 4 года было подготовлено много материалов. Для удобства навигации по обширным ресурсам они структурированы в соответствии с типичными вопросами, задаваемыми о МО, в остальной части раздела.

## **A. Что такое машинное обучение?**

7. Использование машинного обучения в контексте официальной статистики все еще является относительно новым. Обучение тому, что такое машинное обучение и соответствующим навыкам, имеет первостепенное значение для внедрения этой технологии в статистических организациях. Многие учебные ресурсы, связанные с МО, существуют за пределами официального статистического сообщества. Несколько ознакомительных ресурсов, подготовленных под эгидой ГВУ-МОС, включают:

- Публикация ЕЭК ООН [«Использование машинного обучения в целях официальной статистики» \(2022\)](#) — глава 2.
- Ознакомительные учебные ресурсы, рекомендованные участниками проекта по МО, которые были собраны на [вики-странице «Обучение и тренинги»](#).

## **B. Где машинное обучение может быть использовано в статистических организациях?**

8. Машинное обучение может вносить свой вклад в работу статистических организаций различными способами. Оно позволяет автоматизировать процессы, которые раньше в основном выполнялись людьми, и дает организациям возможность использовать новые источники данных. Использование больших данных часто требует применения МО, поскольку предполагает эффективную обработку огромного объема данных. Основные области применения включают в себя:

- кодирование и классификацию: [тематический доклад о классификации и кодировании проекта по МО \(2020\)](#) и [тематический доклад о классификации текста группы по МО \(2022\)](#); [Тематический доклад Группы по ПО «Веб-скрейпинг» \(2022\)](#);
- редактирование и импутация: [Тематический доклад проекта по МО «Редактирование и интерполяция» \(2020\)](#);
- анализ изображений: [Тематический доклад проекта по МО «Изображения» \(2020\)](#).

9. Исследования и сопутствующие коды (при наличии) по этим областям применения в 2018–2022 годах собраны на [вики-странице «Исследования и коды»](#) вместе с другими примерами использования. На заседании первого [рабочего совещания ЕЭК ООН по МО \(2023\)](#) также представлены новейшие примеры применения МО. В [докладе «Передача знаний» \(2021\)](#) рассматриваются факторы, способствующие и препятствующие тиражированию примеров МО в различных статистических организациях.

### **С. Требующие учета последствия для качества и этические соображения**

10. Обеспечение качества является безусловным требованием для производителей официальной статистики. Однако так называемый «черный ящик» машинного обучения и его сильная зависимость от обучающих данных усложняют задачу статистических организаций по защите от предвзятости и ошибок. Поэтому качество и этические аспекты были одной из наиболее важных проблем с самого начала изучения МО. В последующих материалах рассматриваются вопросы качества и этические аспекты:

- Система качества для статистических алгоритмов ([публикация ЕЭК ООН «Использование машинного обучения в целях официальной статистики» \(2022 год\) — глава 4](#)) и ее применение (2021 год).
- [Качество обучающих данных \(2022 год\)](#).
- [Этические аспекты использования МО для исследований и статистики \(2021 год\)](#).

### **Д. Какие организационные проблемы возникают при переходе от экспериментов к производству?**

11. Несмотря на успешные пилотные исследования, внедрение моделей МО в производство часто оказывается сложной и трудоемкой задачей, и многие решения, найденные в ходе экспериментов, остаются на полке. Для обеспечения успешного развертывания необходимо заблаговременное планирование, учитывающее организационные и культурные проблемы. Основные материалы, посвященные этим проблемам, включают:

- [Переход от эксперимента к производству \(2021 год\)](#).
- [Организационные аспекты внедрения редактирования данных на основе МО в статистическое производство \(2024 год\)<sup>1</sup>](#).

### **Е. Каким образом возможности машинного обучения могут быть интегрированы в статистические организации?**

12. Для использования и масштабирования МО требуется не только базовый код МО. После первых экспериментов и проверки концепции важно создать структуру, которая позволит расширить использование МО за пределы небольшой группы экспертов. Одним из ключевых аспектов создания устойчивого потенциала для МО является фокус на MLOps — практике интеграции моделей МО в производственную среду систематическим и эффективным образом. Это предполагает создание инфраструктуры и платформ, необходимых для развертывания, мониторинга и поддержки моделей МО в масштабе. Эти вопросы рассматриваются в следующих материалах:

- Доклады о переобучении моделей за [2021](#) и [2022](#) годы.
- [Построение экосистемы МО в статистических организациях \(2022 год\)](#).

---

<sup>1</sup> Этот материал подготовлен группой ГВУ-МОС по применению аналитики данных и современных методов», но он включен в настоящий документ по причине его высокой релевантности. Более подробную информацию о Группе см. в разделе IV.

### III. Работа над большими языковыми моделями (2023 год)

13. За последние несколько лет возможности ИИ значительно расширились благодаря появлению больших языковых моделей (БЯМ), и в статистическом сообществе растет признание преобразующего потенциала БЯМ.

14. В середине 2023 года две группы ГВУ-МОС по модернизации — Сеть передовых исследований и изысканий (BSTN) и Группа по применению аналитики данных и современных методов (ADSaMM) — подготовили аналитический документ «Использование БЯМ в целях официальной статистики».

15. В документе дается краткое введение в БЯМ, обзор того, где они могут быть использованы в статистическом производстве, и связанных с этим рисков, приводятся конкретные примеры использования в пяти статистических организациях и излагаются основные соображения, которые следует принимать во внимание для продвижения вперед в использовании этой новой технологии. С полным текстом документа можно ознакомиться [здесь](#), а выдержки из него приведены в приложении.

### IV. Проводимая в настоящее время работа (2024—)

16. В 2024 году ГВУ-МОС начала несколько направлений работы, направленных на интеграцию приложений ИИ в официальную статистику.

#### A. Проект Группы высокого уровня по модернизации официальной статистики по генеративному ИИ в целях официальной статистики

17. С опорой на аналитический документ по БЯМ данный проект направлен на дальнейшее изучение потенциала генеративного ИИ — более широкой категории передовых систем ИИ, включающих в себя БЯМ. Наряду с конкретными случаями использования генеративного ИИ в статистических организациях (например, генерация ответа, дополненная результатами поиска, генерация кода), в рамках проекта будут изучены следующие аспекты: управление проектами и путь развития (организационные аспекты), оперативное проектирование (лучшие практики и обеспечение качества), архитектура и стек программного обеспечения (инфраструктура, инструменты), а также управление и этика (анализ безопасности, правовых и этических рисков и их снижение).

#### B. Группа по применению аналитики данных и современных методов

18. Группа по применению аналитики данных и современных методов (ADSaMM) была создана в начале 2022 года, что отражает растущее значение новых источников данных и методов для составления официальной статистики. Группа стремится выйти за рамки концептуальных основ аналитики данных и современных методов, выявляя конкретные возможности для дальнейшей модернизации бизнес-процессов статистических организаций. В 2024 году в рамках группы ADSaMM будут изучаться следующие темы:

- Продвижение ответственного ИИ в статистических управлениях: опираясь на «Рамочную основу ответственного ИИ/МО для официальной статистики»<sup>2</sup> 2023 года, целевая группа будет работать над продвижением этичного внедрения ИИ в статистическую практику, согласуясь с принципами добросовестности и прозрачности. Эта работа будет включать в себя разработку руководящих принципов и инструментария для этичного применения ИИ, а также повышение уровня понимания и внедрения ответственного ИИ

<sup>2</sup> В стадии окончательной доработки.

статистическими организациями путем проведения тренингов и рабочих совещаний.

- Количественная оценка неопределенности: результаты, основанные на МО, часто представляются без учета неопределенности, что вызывает сомнения в их надежности. Строгие методы количественной оценки неопределенности могут помочь статистическим организациям обеспечить большее доверие при информировании пользователей и широкой общественности об использовании МО/ИИ. Целевая группа проведет исследование традиционных подходов (например, байесовских, инициализации) и подходов конформного предсказания.

### **С. Группа по развитию потенциала и коммуникации**

19. Группа по развитию потенциала и коммуникации (CapComm) является группой ГВУ-МОС по модернизации, которая занимается организационными изменениями и задачами в области коммуникации, необходимыми для поддержки модернизации в статистических организациях. Группа создала целевую группу «Использование ИИ в целях официальной статистики с точки зрения коммуникации», чтобы изучить, как повысить производительность экспертов по коммуникации с помощью ИИ и как обеспечить коммуникацию с использованием ИИ для поддержания доверия к официальной статистике.

## Приложение

### **Аналитический доклад «Использование больших языковых моделей в целях официальной статистики», подготовленный Группой высокого уровня по модернизации официальной статистики**

Обратите внимание, что данное приложение включает в себя только Раздел 1 («Введение в большие языковые модели»), Раздел 2 («Последствия и возможности для официальной статистики») и Раздел 5 («Соображения по поводу продвижения статистических организаций к большим языковым моделям») из документа по БЯМ в связи с ограничениями в отношении объема текущего документа. Остальные разделы — Раздел 3 («Примеры использования в статистических организациях») и Раздел 4 («Риски и меры по их снижению») — см. в [полном тексте документа](#).

### **Резюме**

Большие языковые модели (БЯМ) — это класс технологий искусственного интеллекта, способных понимать, интерпретировать и генерировать тексты. Благодаря длительному обучению на огромных массивах данных с миллиардами параметров БЯМ способны понимать и генерировать тексты на уровне, неотличимом от человеческого. Это отличает их от традиционных моделей машинного обучения, применение которых направлено в первую очередь на помощь человеку в решении задач прогнозирования, а не на создание контента.

Нет сомнений в том, что БЯМ будут играть важную роль в деятельности статистических организаций и в будущем. Как и в любых учреждениях во многих секторах и областях, в статистических организациях существуют регулярные рабочие задачи, такие как написание электронных писем и записок о совещаниях. БЯМ могли бы помогать сотрудникам в выполнении этих рутинных, но трудоемких обязанностей. Более того, БЯМ могут быть использованы для повышения эффективности на различных этапах процессов статистического производства и других связанных с ними работ при условии человеческого контроля и тщательной проверки существующих методов. Эти возможности не просто теоретические, а вполне реальные. Это подтверждают примеры внедрения в различных национальных и международных организациях, такие как перевод из SAS в R, обновление системы статистической классификации, создание отчетов, поиск данных на естественных языках и редактирование метаданных.

Однако, существуют и риски, связанные с БЯМ, такие как этические проблемы, юридические последствия (например, авторское право) и общий недостаток осведомленности и грамотности. Кроме того, из-за способности генерировать очень хорошо написанные и контекстуально релевантные тексты пользователи могут быть введены в заблуждение фактологически ошибочными, устаревшими и даже полностью сфабрикованными (так называемые «галлюцинации») данными. Вопросы конфиденциальности и безопасности, связанные с возможной утечкой данных через БЯМ, вызывают серьезную озабоченность и у статистических организаций. Эти риски часто зависят от типов сценариев использования БЯМ, однако существуют общие меры по их снижению, такие как обеспечение человеческого контроля, использование протоколов языкового тестирования, локальная тонкая настройка и применение принципов и требований конфиденциальности.

По мере продвижения статистических организаций во внедрении БЯМ следует учитывать несколько основных моментов. К ним относятся: как создать структуру управления, как взаимодействовать с технологическими компаниями, предоставляющими БЯМ, услуги на основе БЯМ и облачные вычисления, а также как выбрать БЯМ с разным уровнем открытости. С учетом повышенного общественного интереса и пристального внимания, с которым сталкиваются государственные

организации, информирование об ответственном использовании БЯМ — о том, что статистические организации используют их целенаправленно, когда имеются очевидные выгоды, учитывая при этом риски и необходимые меры по их снижению, — имеет жизненно важное значение. Использование БЯМ статистическими организациями все еще находится на начальном этапе, но имеется несколько практических рекомендаций:

- проводить обучение по БЯМ на всех уровнях организации (техническом, оперативном и управленческом);
- приступать к внедрению БЯМ с реализации небольших пилотных проектов, чтобы ознакомиться с технологией и понять ее потенциальную полезность;
- разработать всеобъемлющую стратегию БЯМ после того, как осведомленность о них и их знание достигнут достаточного уровня; и
- уделять постоянное внимание тому, чтобы быть в курсе постоянно меняющегося ландшафта БЯМ.

В связи с динамичным и быстро развивающимся характером этой области тесное сотрудничество между статистическими организациями будет по-прежнему иметь решающее значение для совместного изучения различных приложений и обмена знаниями и опытом на пути их внедрения.

## 1. Введение в большие языковые модели

Большие языковые модели (БЯМ) все еще являются относительно новой технологией. Поэтому важно понять, что это такое и как они работают, прежде чем углубляться в вопрос о том, как БЯМ могут использоваться в официальной статистике. В этом разделе мы расскажем о возможностях БЯМ, их источниках в более широком ландшафте искусственного интеллекта и их преобразующей силе в обработке естественного языка (NLP). Мы кратко опишем динамическую эволюцию языковых моделей, от сложности нейронных сетей-трансформеров до адаптивности базовых моделей, таких как двунаправленные кодирующие представления модели на основе трансформеров (BERT) и генеративный предварительно обученный трансформер (GPT). Затем мы кратко обсудим концепции, важные для БЯМ, такие как тонкая настройка моделей и настройка подсказок, которая улучшает возможности БЯМ без необходимости переучивать их с нуля, а также открытый исходный код в БЯМ.

### 1.1 Что такое большие языковые модели?

БЯМ — это класс технологий искусственного интеллекта (АИ), способных понимать, интерпретировать и генерировать тексты. Благодаря длительному обучению на огромных массивах данных БЯМ способны понимать и генерировать тексты на уровне, неотличимом от человеческого. БЯМ становятся все более популярными благодаря их исключительной способности решать широкий спектр задач NLP и задач понимания естественного языка (NLU), таких как перевод и резюмирование текста.

В сервисах, разработанных на основе БЯМ (например, ChatGPT), пользователи могут взаимодействовать с БЯМ с помощью естественных языков, называемых «подсказками» (инструкции, генерирующие ответы от БЯМ), например, как показано ниже:

**Пользователь:** Не могли бы вы дать мне функции excel, которые генерируют случайные целые числа от 1 до 10?

**Сервис БЯМ:** Конечно! Вы можете использовать функцию `RANDBETWEEN`. `RANDBETWEEN(1, 10)` генерирует случайное целое число между заданными минимальным и максимальным значениями.



*Пользователь: А если мне нужно действительное число от 0 до 10?*

*Сервис БЯМ: Если вам нужно действительное число (включающее десятичные знаки) в диапазоне от 0 до 10, вы можете использовать функцию RAND(), а затем масштабировать результат.*

### **Взаимосвязь с искусственным интеллектом, машинным обучением и генеративным искусственным интеллектом**

БЯМ — это не непредвиденная, новая технология, появившаяся из ниоткуда; она является кульминацией непрерывного развития и эволюции ИИ. Чтобы лучше понять суть БЯМ, важно понять контекст их создания и различия между различными технологиями и определениями. Искусственный интеллект, машинное обучение, большие языковые модели и генеративный ИИ — все это взаимосвязанные понятия, но между ними имеются принципиальные различия. Прежде чем сосредоточиться на БЯМ, мы рассмотрим тесно связанные с ними понятия<sup>3</sup>.

- **Искусственный интеллект (ИИ)** — это широкая область компьютерной науки, которая занимается созданием систем и машин, способных выполнять задачи, для которых обычно требуется человеческий интеллект. Эти задачи включают в себя решение проблем, обучение, аргументацию, восприятие, понимание языка и многое другое.
- **Машинное обучение (МО)** — это подмножество технологий искусственного интеллекта, которое включает в себя использование алгоритмов и статистических моделей, позволяющих компьютерам улучшать свою производительность при выполнении конкретной задачи путем обучения на основе данных, без явного программирования. Другими словами, речь идет о том, чтобы научить компьютеры учиться на примерах и делать прогнозы или принимать решения на их основе. Многие приложения ИИ используют методы МО для достижения своих целей.
- **Глубокое обучение** — это подмножество МО, в котором используются искусственные нейронные сети с большим количеством взаимосвязанных слоев (глубокие нейронные сети). Эти сети могут автоматически обнаруживать и учиться представлять закономерности или тенденции в больших объемах данных. Глубокое обучение оказалось весьма успешным в таких задачах, как распознавание изображений и речи. Оно особенно хорошо подходит для задач, связанных со сложными неструктурированными данными, такими как изображения, аудио и текст. Это специализированный инструмент в наборе инструментов МО.
- **Генеративный ИИ** относится к системам ИИ, которые могут генерировать новый контент или данные, не полученные явно из существующих примеров. Это может быть генерация текста, изображений, музыки и многого другого. Генеративный ИИ часто использует такие методы, как генеративно-состязательные сети (GAN) и вариационные автокодировщики (VAE).

БЯМ, такие как GPT-3, являются конкретным применением глубокого обучения в области ИИ. Они способны понимать и генерировать естественный язык (т. е. использовать алгоритмы и модели, способные точно интерпретировать человеческий язык) и используются в широком спектре приложений. Современные БЯМ появились в 2017 году и используют нейронные сети- трансформеры, обычно называемые трансформерами. Благодаря использованию большого количества параметров и модели-трансформера БЯМ способны быстро понимать и генерировать точные ответы, что делает технологию ИИ широко применимой в различных областях.

<sup>3</sup> URL: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8259629>.

БЯМ можно рассматривать как подмножество базовых моделей, которые фокусируются на задачах, связанных с языком<sup>4</sup>. Базовые модели представляют собой большие нейронные сети глубокого обучения, обученные на больших наборах данных, и служат базовым структурным элементом для различных приложений. Они могут производить широкие и разнообразные выходные продукты или вообще применять к набору данных задачу предварительного обучения, чтобы базовая модель стала хорошо справляться с этой задачей (например, создание изображения). Базовые модели могут использоваться в качестве «основы» для других моделей, которые могут быть построены поверх базовой модели. Базовая модель настолько велика и влиятельна, что служит основой для дальнейшей оптимизации и конкретных сценариев использования.

## 1.2 Как работают большие языковые модели?

Большие языковые модели основаны на архитектурах глубокого обучения, при этом особое внимание уделяется так называемым моделям-трансформерам. Трансформеры — это архитектуры нейронных сетей, которые используют механизмы самовнимания для обработки входных данных, что позволяет им эффективно обрабатывать дальние зависимости в языке. В следующих разделах подробно описаны компоненты и процесс обучения больших языковых моделей.

### Компоненты больших языковых моделей

- (1) **Параметры:** Основными компонентами большой языковой модели являются ее параметры, которые включают веса и непреднамеренные предпочтения. Эти параметры настраиваются в процессе обучения, чтобы минимизировать разницу между предсказаниями модели и реальными значениями.
- (2) **Слои:** Большая языковая модель состоит из нескольких слоев, каждый из которых отвечает за извлечение и обработку различных уровней информации из входных данных. Эти слои обычно включают входной и выходной слои, а также несколько скрытых слоев.
- (3) **Механизм внимания:** механизм внимания — важнейший компонент больших языковых моделей, позволяющий им избирательно фокусироваться на соответствующих частях входных данных. Этот механизм помогает модели улавливать зависимости между словами и фразами, даже если они находятся далеко друг от друга в тексте.

Большие языковые модели обучаются на массивном наборе данных, обычно содержащем миллиарды слов из различных источников. Этот процесс самоконтролируемого обучения позволяет модели изучать структуру и закономерности языка. Обучение такого крупномасштабного БЯМ, как GPT-3, имеющего 175 млрд параметров, — очень дорогостоящий процесс, который может обойтись в десятки млн долларов только за счет стоимости оборудования и электроэнергии. Однако предварительно обученные модели можно точно настроить на небольшом наборе данных для конкретной задачи. Этот процесс «тонкой настройки» улучшает понимание моделью конкретной задачи, помогая ей лучше обобщать и достигать более высоких результатов в решении этой задачи. Этап тонкой настройки по-прежнему требует достаточной вычислительной мощности для заданной

---

<sup>4</sup> Базовая модель определяется как «любая модель, которая учится на данных и средствах наглядности (обычно с использованием самоконтроля) и может быть адаптирована (например, тонко настроена) для широкого спектра последующих приложений» (Центр исследований фундаментальных моделей (CRFM) Стэнфордского института человекоцентрированного искусственного интеллекта (HAI) <https://hai.stanford.edu/news/reflections-foundation-models>). Уникальность фундаментальных моделей заключается в их общем характере и размере, что отличает их от традиционных моделей машинного обучения. Они могут быть использованы в качестве основы для разработки последующих специализированных приложений.

модели и задачи, но он менее ресурсоемкий по сравнению с предварительным обучением моделей с нуля. Тонкая настройка описана в следующем разделе.

### 1.3 Модели тонкой настройки

БЯМ часто используются в готовом виде (т.е. они поставляются предварительно обученными с полным набором весов). Тем не менее их можно настраивать, используя ряд приемов, включая настройку подсказок и тонкую настройку, которые позволяют улучшить результаты модели без необходимости полностью обучать ее с нуля.

Настройка подсказок — это «облегченный» метод, который предполагает разработку конкретных входных данных, определяющих выход модели. Эта настройка выполняется без изменения параметров модели. Настройка подсказок использует существующие знания и возможности модели, просто изменяя способ запроса.

Тонкая настройка — это более интенсивный процесс, который подразумевает обучение модели после первоначального обучения на специализированном наборе данных. Это направлено на улучшение производительности модели применительно к целевой задаче или области. В процессе тонкой настройки параметры модели обновляются, чтобы лучше соответствовать целевой задаче или области. Это позволяет модели производить более релевантные и конкретные выходные данные для специализированных приложений. Однако для получения эффективных результатов тонкая настройка требует дополнительных вычислительных ресурсов и специально сформированного набора данных.

Модели с точной настройкой могут быть эффективным подходом для:

- уточнения стиля и выражения: настройка выходных данных модели в соответствии с определенными стилями, тонами, форматами или другими желаемыми качественными аспектами (например, статистический чат-бот);
- выполнения сложных инструкций: эффективная работа с многосложными и подробными подсказками, даже теми, которые включают в себя несколько шагов;
- повышения согласованности: повышение надежности модели в плане производства запланированных выходных данных, снижения вероятности получения неожиданных или нежелательных результатов;
- овладение новыми навыками: способность модели выполнять новые задачи или применять новые навыки, которые может быть сложно определить в явном виде в формате подсказок.

С технической точки зрения, набор данных, необходимый для точной настройки модели, часто структурируется в виде файла JavaScript Object Notation (JSON). Он организован в виде тройки ролей (определенных как «система», «пользователь» и «помощник»), каждая из которых связана с контентом. Тройки определяют поведение модели в конкретной ситуации; каждая тройка «система/пользователь/помощник» используется для перекалибровки модели.

В примере ниже показан фрагмент набора данных, используемого для тонкой настройки чат-бота с использованием ChatGPT 3.5; представлены три тройки (для краткости подсказка «Система» показана только один раз, поскольку она одинакова для всех трех случаев).

```
{«сообщения»: [{«роль»: «система», «контент»: «Вы — агент поддержки, чья основная задача — помогать пользователям. Вы дружелюбны и лаконичны. Вы предоставляете только фактические ответы на запросы и не даете ответов, не связанных с официальной статистикой».}, {«роль»: «пользователь», «контент»: «Какое самое лучшее место в Австралии?»}, {«роль»: «помощник», «контент»: «Я прошу прощения, но я не могу высказывать свое личное мнение. Моя основная задача — помочь вам с запросами, связанными с официальной
```

статистикой. Имеется ли у вас какой-либо требующий ответа вопрос, касающийся официальной статистики?»}}}

{«роль»: «пользователь», «контент»: «Где я могу найти статистику по национальным счетам?»}, {«роль»: «помощник», «контент»: «Чтобы найти статистику по национальным счетам, вы можете подключиться к браузеру данных и найти национальный счет в выпадающем меню. Вот ссылка: —. -/—»}}}

{«роль»: «пользователь», «контент»: «У меня проблемы с подключением к браузеру данных»}, {«роль»: «помощник», «контент»: «Не волнуйтесь, давайте проведем предварительную проверку. Какой браузер вы используете?»}}}

## 1.4 Открытый исходный код

Термин «открытый исходный код» относится к тому, что люди могут изменять и распространять, поскольку ее структура находится в открытом доступе<sup>5</sup>. Слово «открытый» в термине «открытый исходный код» присутствует не просто так. БЯМ с открытым исходным кодом — это модель, код которой находится в открытом доступе под лицензией с открытым исходным кодом, что позволяет любому человеку использовать, адаптировать и распространять модель. Такие модели с открытым исходным кодом обычно сопровождаются подробной документацией, содержащей ценную информацию о структуре модели, методах обучения, конфигурациях модели и наборах данных, использованных при обучении и оценке. Такая документация способствует более глубокому пониманию внутренних механизмов и возможностей модели, что способствует прозрачности и сотрудничеству в сообществе ИИ и МО.

Благодаря такой открытости пользователи могут не только непосредственно использовать модель, но и изучать ее структуру, адаптировать и настраивать код, тем самым улучшая модель. Это может стать одной из возможностей для международного сотрудничества между национальными производителями статистических данных. Ряд моделей с открытым исходным кодом доступен через Hugging Face — компанию и популярную платформу в области обработки естественного языка (NLP) и искусственного интеллекта.

Однако, как подробно рассматривается в разделе 5.3, пользователям следует внимательно проверять лицензию, с которой предоставляется БЯМ, чтобы понять, соответствует ли ей их сценарий использования. Например, несколько БЯМ были выложены в открытый доступ с лицензиями, ограничивающими коммерческое использование. Другие лицензии, напротив, могут обязывать пользователей публично распространять производные работы на тех же условиях, что и оригинальные БЯМ, или требовать от пользователя явного указания на автора оригинала. В общем, тот факт, что БЯМ может находиться в открытом доступе, не обязательно означает, что не существует правил, регулирующих ее использование.

## 2. Последствия и возможности для официальной статистики

Возможности использования БЯМ впечатляют, но не безграничны, поэтому важно понимать, что БЯМ могут делать, а что нет. В этом разделе мы даем общий обзор того, как БЯМ могут повысить эффективность рутинных задач в статистических организациях, от коммуникации до управления проектами, подчеркивая их роль в оптимизации операций. Будет обсуждаться потенциал БЯМ в повышении эффективности процесса статистического производства, начиная с разработки плана обследования и заканчивая распространением данных. Мы также внимательно рассмотрим меняющийся информационный ландшафт и то, как БЯМ могут повлиять на то, как люди получают доступ к статистической информации.

<sup>5</sup> URL: <https://opensource.com/resources/what-open-source>.

## 2.1 Что статистические организации могут и не могут делать с большими языковыми моделями

БЯМ обучаются на огромных массивах информации и содержат миллиарды параметров для создания статистических прогнозов. Алгоритмы, используемые в коммерческих БЯМ, редко сообщаются, в результате чего они считаются «черными ящиками». Кроме того, обучающие данные не имеют четкой идентификации и могут содержать непреднамеренные предвзятости. К сожалению, эти предвзятости могут быть воспроизведены в результатах БЯМ. Кроме того, поскольку цель БЯМ — предсказать следующее слово, они могут выдавать неверную или бессмысленную информацию, которую принято называть галлюцинациями. Однако, поскольку выходные данные написаны очень хорошо, человеческая природа заставляет людей верить им как фактам.

Несмотря на эти потенциальные недостатки, БЯМ имеют множество возможностей для использования в статистических организациях. Они отлично разбираются в текстовой информации, обобщают большие объемы информации и генерируют человекоподобные ответы, которые могут быть полезны для автоматизации многих задач в статистических организациях. В этом разделе мы опишем несколько идей в отношении использования БЯМ, включая задачи, необходимые в любой организации, такие как составление электронных писем и предварительных отчетов, обобщение информации для мозговых штурмов, управление проектами и перевод на несколько языков. А также задачи, которые особенно актуальны для статистических организаций, такие как классификация текстов, визуализация данных и их распространение.

Более подробная информация об этих и других возможных вариантах использования представлена далее в разделе.

Несмотря на то, что БЯМ способны изменить работу статистических организаций, **они требуют тщательного контроля**. Электронные письма и отчеты, составленные БЯМ, должны проверяться людьми, чтобы убедиться, что контекст верен и не представляет собой предвзятую точку зрения. Это важно, так как БЯМ умеют создавать хорошо написанные тексты, но они не предназначены для проверки того, что их содержание является основанным на фактах или обязательно лучшим выбором. Если данные, на которых обучается БЯМ, неверны или подходят лишь в некоторой степени, она будет использовать эту информацию при формулировании своего ответа.

Например, Группа по применению аналитики данных и современных методов Группы высокого уровня по модернизации официальной статистики (ГВУ-МОС) задала несколько методологических вопросов несколькими БЯМ и проверила результаты. В целом, ответы были правильными, но не всегда наиболее подходящими. На вопрос о замене отсутствующих значений часто они, как правило, отвечали, что используют импутацию средними значениями. Хотя импутация средними значениями не является некорректной, она имеет ряд недостатков, таких как искажение распределения данных и неиспользование вспомогательной информации, которая может быть доступна. Вопросы, заданные Группой, иллюстрируют тот факт, что БЯМ является «механизмом рассуждений» и, в отличие от людей-экспертов, не задает вопросов, чтобы собрать больше информации для поиска более подходящих ответов. Ответственность за правильные подсказки лежит на лице, обращающемся к БЯМ с запросом.

Если пользователь не разбирается в предмете, БЯМ может не дать высококачественных ответов. Одна из основных задач консультанта — работать с клиентом, чтобы выяснить его реальные потребности. В контексте консультанта по статистике это сводится к пониманию потребностей в данных и конечного использования данных для восполнения информационных пробелов. Эта информация необходима для того, чтобы убедиться, что применяемые методы позволяют данным удовлетворять потребности клиента. Не собрав эту дополнительную информацию, БЯМ могут предложить методы, которые могут не подойти. Если человек,

взаимодействующий с БЯМ, обладает определенными знаниями в предметной области, он сможет предоставить дополнительную информацию для выбора подходящего метода. Однако, если у человека нет знаний, и он следует советам БЯМ, метод, который он использует, может не решить проблему должным образом.

Это подчеркивает важность **инженерии подсказок по цепочке рассуждений**, которая требует определенных знаний по обсуждаемому предмету и понимания того, как получить наилучший результат от БЯМ. Другими словами, БЯМ не смогут заменить человеческое взаимодействие, необходимое для четкого определения потребностей или исследуемого вопроса, который требуется для выбора наиболее подходящего статистического метода. В руках человека, который может не разбираться в предмете, слепое применение советов БЯМ может привести к не лучшим результатам.

## 2.2 Повышение эффективности выполнения обычных рабочих задач

Как и у любой другой организации, у статистических организаций есть регулярные задачи, которые весьма похожи на те, что встречаются как в государственном, так и в частном секторе. Эти задачи включают в себя такие действия, как управление электронной почтой, создание отчетов и презентаций, а также ведение записей о совещаниях. Хотя эти рутинные обязанности жизненно важны для эффективного функционирования организации, они требуют значительного количества времени и усилий от преданных своему делу сотрудников.

БЯМ/ChatGPT могут помочь оптимизировать работу статистических организаций и повысить производительность имеющихся ресурсов. Таким образом, управление сможет более эффективно распределять свои ресурсы на выполнение основных задач и способствовать достижению своей цели — предоставлению точной и своевременной статистической информации. В следующем разделе будут приведены примеры того, как БЯМ/ChatGPT могут быть использованы для повышения эффективности статистической организации, позволяя ей более эффективно достигать своих основных целей.

(1) **Коммуникация** — Одним из наиболее распространенных применений БЯМ является непосредственное использование их возможностей в целях коммуникации. БЯМ доказали, что они могут помочь в составлении электронных писем, планов и отчетов, предлагая контент, оказывая помощь в форматировании и создавая сам текст<sup>6</sup>. Это экономит время и повышает качество письменных материалов. В отчетах БЯМ обобщают объемные документы, предлагают варианты визуализации данных, выявляют ошибки и предлагают рекомендации.

(2) **Мозговые штурмы и генерирование идей** — БЯМ могут способствовать проведению мозговых штурмов, предлагая творческие предложения, изучая различные аспекты проблемы и генерируя новые идеи на основе представленных входных данных. Это может быть особенно полезно в плане изучения различных точек зрения, изучения проблемных аспектов, постановки вопросов для углубления анализа, оценки идей, формирования выводов и экономии времени.

(3) **Управление проектами и планирование** — БЯМ можно эффективно использовать для выполнения рутинных задач, необходимых на различных этапах процесса управления проектами, автоматизируя планирование задач и управление зависимостями, оптимизируя распределение ресурсов на основе исторических данных и требований проекта, оценивая длительность задач для планирования сроков, а также упрощая ведение протоколов совещаний путем

<sup>6</sup> Обратите внимание, что письма, сгенерированные БЯМ, легко распознать; дружеский совет — не вставляйте сгенерированный текст из ChatGPT напрямую без форматирования стилей, так как это сохранит оригинальный шрифт и серый фон.

расшифровки и составления резюме, обеспечивая эффективное документирование и обобщение важной информации.

(4) **Перевод с других языков и на них** — БЯМ могут переводить документы и тексты с одного языка на другой, облегчая доступ к информации на разных языках. В целом, на текущем этапе развития БЯМ/ChatGPT предлагают значительные преимущества при решении переводческих задач, поскольку они более чувствительны к контексту. Однако традиционные системы автоматического перевода все еще сохраняют свои преимущества в сценариях с большими массивами данных, требованиями к скорости и эффективности, а также четко определенными областями.

(5) **Презентации** — БЯМ можно использовать для создания презентаций от базовых до продвинутых слайдов с макросами. Они могут использоваться не только для создания контента слайдов, позволяя настраивать стиль, структуру и количество слайдов, но и для разработки тезисов для более удобной для восприятия презентации.

(6) **Образовательные цели** — БЯМ могут использоваться в образовательных и учебных целях в организации. Они могут давать пояснения, создавать викторины и помогать в разработке материалов для электронного обучения, чтобы повысить уровень знаний и навыков сотрудников.

(7) **Генерирование изображений** — шаблонные изображения часто используются в отчетах и продуктах статистических организаций. Вместо того чтобы приобретать шаблонные изображения, статистические организации могли бы использовать БЯМ для создания изображений, сопровождающих статистическую продукцию.

Разумное использование БЯМ/ChatGPT может высвободить человеческие ресурсы для решения более стратегических и сложных задач, позволяя сотрудникам быть более креативными, продуктивными и сосредоточиться на более приоритетных областях.

### 2.3 Повышение эффективности статистического производства и качества предоставления услуг

БЯМ могут использоваться в широком спектре приложений для повышения эффективности на различных этапах процесса статистического производства, например при условии человеческого контроля и тщательной проверки существующих методов и опыта, накопленного в организациях;

- Разработка плана сбора данных (подпроцесс 2.3 ТМПЦИ): БЯМ могут внести свой вклад в разработку обследований и вопросников, предлагая вопросы, форматы и формулировки, которые с большей вероятностью дадут точные ответы<sup>7</sup>.
- Классификация и кодирование (подпроцесс 5.2 ТМПЦИ): БЯМ обладают способностью автоматически сортировать текстовые данные по заранее определенным категориям или меткам. Статистические организации могут использовать их для организации ответов на опросы и других текстовых данных в соответствующие категории в системах статистической классификации.
- Валидация и редактирование данных (подпроцесс 5.3 и 5.4 ТМПЦИ): БЯМ могут оптимизировать задачи по очистке и предварительной обработке данных, выявляя и устраняя ошибки в данных, недостающие значения и несоответствия.
- Создание продуктов распространения (подпроцесс 7.2 ТМПЦИ): БЯМ могут генерировать текстовые описания из таблицы или ряда чисел (см. пример использования в разделе 3.4. Генерирование отчетов с использованием БЯМ

<sup>7</sup> Типовая модель производства статистической информации  
(URL: <https://statswiki.unece.org/display/GSBPM>).

(Статистическое управление Канады)), которые могут быть адаптированы для различных сегментов аудитории, включая политиков, журналистов и широкую общественность. Это может значительно упростить работу аналитиков и специалистов по коммуникации, предоставляя первоначальные проекты, над которыми могут работать люди-эксперты. БЯМ также могут помочь в автоматизации создания графиков и диаграмм, хотя эта область находится в стадии изучения.

- Метаданные играют важнейшую роль в статистическом производстве, и редактирование метаданных может быть облегчено БЯМ (см. пример использования в разделе 3.5. Редактирование метаданных с использованием GPT (Банк международных расчетов)).

Помимо применения в процессе статистического производства, БЯМ могут оказать поддержку в нескольких сквозных областях, имеющих решающее значение для статистических организаций:

- помощь в кодировании и переводе с одного языка программирования на другой: БЯМ могут работать не только с естественными языками, но и с языками программирования, которые статистические организации широко используют для многих частей своего производства, в частности для обработки и анализа. БЯМ могут значительно повысить эффективность и результативность работы программистов и аналитиков, помогая упростить и оптимизировать разработку кода, предоставляя фрагменты кода и осуществляя перевод между различными языками программирования (см. пример использования в разделе 3.2. Перевод и объяснение кода (с SAS на R) с помощью БЯМ (Центральное статистическое управление Ирландии)).
- Обновление и поддержка статистических стандартов: генерирование проектов текстовых описаний для помощи экспертам в обновлении систем статистической классификации (см. пример использования в разделе 3.1. Обновление определений статистической классификации (Австралийское бюро статистики)) и методологических документов.
- Генерирование синтетических данных: конфиденциальность и использование данных являются ключевыми вопросами при тестировании статистической методологии. БЯМ могут использоваться для создания синтетических текстовых данных, что позволяет тестировать методологию без использования реальных данных в тестовых средах.

В частности, способность БЯМ быстро обрабатывать огромный объем текстовой информации и взаимодействовать с людьми на естественных языках может значительно повысить удобство работы с платформами для распространения статистических данных. В настоящее время платформа распространения данных большинства статистических организаций структурирована по областям и темам. Чтобы найти нужную статистику, пользователям приходится переходить по нескольким страницам, а в более неудачном случае — проходить через несколько раундов «назад-вперед». Кроме того, такая структура может быть обременительной для пользователей, которые ищут и интегрируют данные из нескольких областей и тем. Хотя статистические организации стремятся предоставлять продукты в форматах, предназначенных для разных аудиторий (например, основные цифры для журналистов, необработанные данные для исследователей, аналитические отчеты для политиков), пользователи, не знакомые с тем, как их можно получить на веб-сайте, могут столкнуться с трудностями. БЯМ могут помочь смягчить эти проблемы и повысить качество предоставления данных пользователям — конечную цель производителей официальной статистики — например, путем

- интерактивных запросов: предоставление возможности БЯМ вступать в диалог с пользователями для уточнения их информационных потребностей и уточнения запросов может привести к получению более точных и релевантных ответов (см. пример использования в разделе 3.3. StatGPT (Международный валютный фонд)).



- Предоставление оптимизированной под требования пользователей информации: статистические организации могут позволить пользователям настраивать формат получения статистической информации от БЯМ. Некоторые пользователи предпочитают обобщенные отчеты, в то время как другим нужен глубокий анализ или исходные данные.
- Помощь в интерпретации данных: БЯМ могут помочь пользователям интерпретировать сложные статистические данные, предоставляя объяснения, визуализации и контекст. Это помогает пользователям понять значение и последствия запрашиваемой статистики.

## 2.4 Изменение способа поиска информации и знаний

Статистические организации адаптируются к меняющемуся ландшафту распространения информации, диверсифицируя свои каналы, чтобы максимально охватить пользователей и аудиторию данных. За последнее десятилетие способы поиска людьми информации значительно изменились. Они редко напрямую посещают сайты статистических организаций для получения официальной статистики; часто люди начинают поиск на таких платформах, как Google.

Эти поисковые системы и цифровые платформы используют алгоритмы (например, поисковый индекс Google), чтобы просеивать огромные массивы информации в Интернете и предоставлять пользователям релевантную информацию. Например, при поисковом запросе «уровень инфляции в стране X в году Y» эти платформы могут отображать официальную статистику соответствующей национальной статистической организации, но также могут включать данные из других источников. Хотя точные процедуры работы этих алгоритмов остаются нераскрытыми, появились стратегии, направленные на повышение видимости и раскрытия контента на этих платформах, к которым приспособились многие статистические организации.

Однако с появлением и ростом популярности удобных для пользователей сервисов, построенных на основе БЯМ (например, ChatGPT), парадигма информационного поиска вновь начинает меняться. Уже сейчас БЯМ могут извлекать историческую статистику из своих обучающих данных с помощью пользовательских подсказок, не прибегая к помощи официальных статистических организаций. Однако в зависимости от возраста и источника обучающих наборов, используемых БЯМ, будут возникать проблемы со своевременностью и качеством выходных данных. Проблемы своевременности и точности не всегда очевидны для рядового пользователя БЯМ, как не очевидно и то, что БЯМ в настоящее время не могут создавать актуальную значимую статистику.

Признавая риски использования БЯМ, организации официальной статистики должны понимать возможности, которые предоставляют БЯМ, и их потенциальное влияние на предоставление официальной статистики и пробные варианты использования статистики.

Для того чтобы официальная статистика оставалась актуальной в эпоху БЯМ, статистические организации должны предоставлять услуги, которые БЯМ не могут оказывать сами по себе, предоставляя варианты высококачественных, точных и своевременных статистических «предпочтительных источников» пользователям официальной статистики.

Организации официальной статистики могут заниматься этим в рамках своей страны или организации, а могут работать совместно с поставщиками БЯМ, чтобы создавать комбинированные статистические продукты, недоступные сегодня, используя возможности БЯМ. БЯМ должны рассматриваться в качестве одного из ключевых факторов, способствующих более своевременному и эффективному предоставлению статистических данных как на национальном, так и на международном уровне.

## **5. Соображения, которые необходимо учитывать статистическим организациям при внедрении больших языковых моделей**

БЯМ открывает множество возможностей для статистических организаций, однако при интеграции БЯМ в организациях необходимо действовать осторожно, принимая во внимание различные факторы. В этом разделе мы рассмотрим основные моменты, связанные с изучением БЯМ, такие как управление, взаимодействие с технологическими компаниями, модели открытого доступа и связи с общественностью. Хотя эта тема быстро развивается, в конце этого раздела мы постараемся дать краткие практические рекомендации.

### **5.1 Управление**

Чтобы воспользоваться преимуществами, которые сулят БЯМ/GPT, описанными в разделах 2 и 3, организации должны создать новые показатели управления или интегрировать свою собственную систему внутреннего управления для ограничения рисков, описанных в разделе 4. К числу сопряженных с риском областей, обсуждаемых в этом документе, относятся этика и предвзятость, точность, конфиденциальность и безопасность, судебные разбирательства по авторским правам и юридические вопросы, а также потенциальное ненадлежащее использование из-за недостаточной грамотности и непонимания. Там же были описаны возможные стратегии уменьшения рисков.

В этом разделе мы рассмотрим, как мы можем управлять БЯМ посредством реализации этих стратегий уменьшения рисков в контексте современных статистических организаций, работающих в среде, уже определенной национальным законодательством, международными рамками и соглашениями, существующими и меняющимися техническими ландшафтами с доминирующими игроками, а также существующей организационной культурой.

#### **Управление большими языковыми моделями**

В тех случаях, когда управление будет касаться внедрения или использования БЯМ, заинтересованные стороны проекта должны установить разумные и соответствующие цели проекта, согласующиеся с основными ценностями организации и принципами официальной статистики, а также с учетом национального контекста. Мы отмечаем, что управление всегда будет ограничено тем фактом, что самые мощные БЯМ/GPT в конечном итоге принадлежат третьим сторонам и контролируются ими, а из-за своих размеров чаще всего должны работать на сторонних облачных платформах, которые также контролируются извне.

Поэтому рекомендация заключается не в том, чтобы полномасштабно внедрять ответственный ИИ, а в том, чтобы обратить внимание на проблемы и конфликты, которые генеративный ИИ (в частности, сервисы БЯМ) создает с точки зрения ответственного ИИ. При внедрении БЯМ/GPT в рабочие процессы организации (будь то в составе готовых продуктов сторонних производителей, через обращение к API-интерфейсу или путем доработки базовой модели и внедрения ее во внутренний разработанный и развернутый продукт) мы должны рассмотреть проблемы и конфликты, возникающие при использовании БЯМ/GPT в предполагаемом рабочем процессе/приложении и определить соответствующие меры по их снижению.

#### **Управление большими языковыми моделями/генеративными предварительно обученными трансформерами в текущем техническом ландшафте**

БЯМ и GPT редко обучаются полностью на локальных или других общедоступных наборах данных. Они часто обучаются, размещаются и работают на сторонних платформах, например предоставляемых Amazon (AWS), Google (GCP) или Microsoft (Azure). Организации будут заключать соглашения с поставщиками технологий, чтобы обеспечить защиту ключевых национальных интересов и

соблюдение соответствующих законов (например, хранить данные на местных серверах). Тем не менее некоторые части конвейерных последовательностей и продуктов БЯМ/ГРТ, используемых организациями, не будут находиться под нашим контролем и, более того, могут быть не совсем видны сотрудникам организации.

Таким образом, характер и уровень управления БЯМ/ГРТ в статистических организациях будет зависеть от того, насколько БЯМ/ГРТ входят в сферу компетенции организации. Управление проектом, касающимся разработки (например, тонкой настройки) БЯМ/ГРТ, будет отличаться от управления, связанного с использованием стороннего приложения с закрытым исходным кодом. В каждом случае управление потребует описания рисков и определения соответствующих мер по их снижению. Более подробная информация о классах рисков и возможных мерах по их снижению представлена в разделе 4.

Ниже приведены некоторые примеры управления БЯМ/ГРТ.

**Пример А:** Лицензионное соглашение на установку стороннего приложения на базе БЯМ/ГРТ

Компания Microsoft встроит свой инструмент искусственного интеллекта CoPilot в пакет Office365, который, как она утверждает, должен повысить производительность рабочих мест. Определенный уровень управления будет осуществляться на юридическом уровне — например, требование о размещении данных на территории страны. Однако после установки и ввода в использование CoPilot некоторые вопросы управления придется решать с помощью более мягких мер. Например, сотрудники статистических организаций, запрашивающие информацию с помощью ИИ, могут быть чрезмерно уверены в точности результатов и публиковать/распространять потенциально ошибочную информацию или принимать решения на основе неверной или неполной информации. Для дальнейшего обсуждения см. пример в разделе 3.1, касающийся ошибок в различении сгенерированных человеком и БЯМ списков профессиональных задач, а также общую дискуссию о ненадлежащем использовании в разделе 4.5. Статистические организации не могут устранить риск ненадлежащего использования, но могут принять меры, описанные в разделе 4.5, в том числе повысить грамотность в области данных и ИИ, разработать четкие протоколы использования и установить технические препятствия, предотвращающие ненадлежащее использование.

**Пример В:** Разработанная внутри компании конвейерная последовательность или продукт, в котором используется предварительно обученная БЯМ/ГРТ

Все чаще внутренние разработчики, знакомые с целями организации, наборами данных и вариантами использования, такие как аналитики данных или инженеры машинного обучения, используют предварительно обученные модели (также называемые базовыми моделями). Организация будет ограничена в своей способности полностью управлять продуктом или конвейерной последовательностью, использующими базовую модель.

Например, статистическим организациям будет сложно гарантировать, что в продукте или конвейерной последовательности не используются компоненты (наборы данных или код), маркированные или разработанные в условиях, где применяются низкие стандарты человеческого труда. Будет трудно доказать, что точность данных приемлема, а модель не страдает предвзятостью, как описано в разделе 4.2, или что не произошло «отравления» данных, как описано в разделе 4.3.

Даже когда сторонние разработчики этих базовых моделей публикуют обучающий код или обучающие данные через публичный репозиторий и/или предлагают пользователям менее строгую лицензию на открытый доступ или открытый исходный код, прозрачности все равно не хватает. Индекс прозрачности базовых моделей, публикуемый Стэнфордским центром базовых моделей, оценивает многие известные базовые модели по 100-бальной шкале, присуждая по одному баллу за каждый критерий, по которому компания предоставила достаточно информации на каждый вопрос. Модель Llama 2 от Meta получила высшую оценку 54/100. Это

означает, что по 46 критериям Мета не предоставила достаточно информации, чтобы исследователи сочли этот критерий прозрачности удовлетворенным<sup>8</sup>.

Учитывая эти конфликты и факторы напряженности, мы не рекомендуем запрещать БЯМ/GPT, так как это создает риск появления теневого ИИ у статистических организаций. Вместо этого мы рекомендуем оценивать каждый проект или приложение на предмет рисков и принимать соответствующие меры по их снижению.

### **Управление на практике — оценка и мониторинг**

**Параметры оценки:** Если БЯМ используется для предоставления ответов на запросы или рекомендаций, производительность БЯМ должна оцениваться по таким критериям, как верность (например, является ли сгенерированный текст верным по отношению к исходному документу?), воспроизводимость (например, дает ли он одинаковые или похожие ответы на один и тот же или похожий запрос) и релевантность (удовлетворяет ли ответ запрос?). Оценка также охватывает вопрос о том, как результаты отражают ценности организации (например, могут ли результаты нанести репутационный ущерб?). Разработчикам также может потребоваться скорректировать параметры, чтобы тон результатов был непредвзятым, политически нейтральным и основанным на фактах, а результаты были ориентированы на соответствующую аудиторию (какой бы она ни была). Необходимо проверять результаты генерирования текстов, чтобы убедиться, что они не являются непреднамеренным плагиатом существующих публикаций — компьютерные тексты все еще требуют доработки, и негативная огласка и возможное влияние на общественное доверие не стоят того, чтобы рисковать.

**Мониторинг:** Степень, в которой конвейерная последовательность или продукт ИИ/БЯМ соответствует целям или создает риск, должна измеряться, контролироваться и правильно отражаться в отчетах в течение всего срока службы конвейерной последовательности или продукта. В качестве этапа мониторинга заинтересованные стороны могут провести пороговую оценку или оценку воздействия, в ходе которой реализация проекта и использование продукта оцениваются по соответствующим категориям риска. Для того чтобы системы ИИ оставались ответственными с течением времени, разработка должна включать план сопровождения, в том числе частоту обновления обучающих данных, а также обзоры методов и кодов для обеспечения актуальности модели ИИ. Вышеперечисленные соображения должны быть интегрированы в техническое обслуживание, чтобы учитывать изменения, происходящие с каждым из них с течением времени.

## **5.2 Взаимодействие с технологическими компаниями, предоставляющими сервисы больших языковых моделей**

Экосистема БЯМ — сложная и быстро развивающаяся область. Центральное место в этой экосистеме занимают такие крупные организации, как Google, OpenAI, Microsoft и Meta AI, которые играют ключевую роль в определении и развитии технологий БЯМ. В этом контексте статистическим организациям крайне важно также изучать и подчеркивать использование моделей и платформ с открытым исходным кодом. Такие компании, как Hugging Face и EleutherAI, построенные на идеологии открытого исходного кода, способствуют созданию более разнообразной и доступной среды. Взаимодействие с этими организациями требует соблюдения баланса между запатентованными технологиями и технологиями с открытым исходным кодом, чтобы стимулировать инновации и поддерживать этические стандарты.

Понимание различных ролей технологических компаний в экосистеме БЯМ имеет очень важное значение. Учитывая такие факторы, как основные предлагаемые продукты, роль в экосистеме и спектр предоставляемых услуг, статистические организации смогут эффективно ориентироваться в этом пространстве.

<sup>8</sup> URL: <https://crfm.stanford.edu/fmti/>.

## Роль поставщиков облачных услуг

Поставщики облачных услуг являются неотъемлемой частью работы и развития БЯМ. При взаимодействии с такими поставщиками статистические организации должны учитывать несколько ключевых факторов. Конфиденциальность и безопасность данных имеют первостепенное значение, равно как и масштабируемость и производительность сервисов. Управление расходами — еще одна важная область, требующая ясного понимания моделей ценообразования и возможных скрытых сборов. Важнейшими соображениями также являются обеспечение соответствия законодательству, например доступность услуг в определенных регионах (например, в Европе или Западной Европе) и техническая совместимость с существующими системами статистической организации.

На мировом рынке облачных технологий в основном доминирует «большая тройка»: Azure, AWS и Google Cloud. Однако альтернативные поставщики часто специализируются на нишевых услугах, предлагая конкретные интеграционные решения, потенциально более подходящие для определенных статистических организаций. Выбор поставщика облачных услуг для инфраструктуры или платформ ИИ требует соответствия конкретным потребностям и учета долгосрочных трендов развития. Также важно помнить о рисках зависимости от основных ключевых игроков в экосистеме БЯМ.

## Экосистема больших языковых моделей

В экосистеме БЯМ услуги, предлагаемые технологическими компаниями, часто охватывают несколько категорий, что подчеркивает взаимосвязанную природу этой области. Например, Azure Machine Learning от Microsoft позволяет пользователям получить доступ к моделям, разработанным OpenAI и Meta AI, а также к некоторым моделям Hugging Face. Аналогичным образом, Hugging Face отличается тем, что предлагает широкий спектр услуг практически по всем категориям экосистемы БЯМ.

Для статистических организаций признание и понимание этих многогранных ролей имеет решающее значение. Определив конкретную категорию или категории, в которых работает компания, статистические организации могут более эффективно выстраивать стратегию взаимодействия с технологическими компаниями. Эти знания позволяют им определить, какие компании предлагают наиболее актуальные и полезные услуги для их конкретных нужд, будь то использование передовых моделей искусственного интеллекта, доступ к различным наборам данных или использование эффективных обучающих платформ. Кроме того, понимание этих категорий помогает статистическим организациям предвидеть и учитывать потенциальное дублирование услуг и сотрудничества, обеспечивая более рациональный и эффективный подход к интеграции технологий БЯМ в их деятельность.

В категорию **разработчиков и поставщиков БЯМ** входят компании, специализирующиеся на исследованиях, разработке и внедрении БЯМ. В качестве примера можно привести OpenAI, Meta AI, Google DeepMind, а также игроков с открытым исходным кодом, таких как EleutherAI и Technology Innovation Institute (ТИ). Эти организации находятся в авангарде развития технологий БЯМ. С технической точки зрения решающее значение имеет обеспечение беспрепятственной интеграции БЯМ этих поставщиков в системы статистических организаций. Это подразумевает совместимость с существующей инфраструктурой и способность адаптироваться к конкретным техническим требованиям. Соответствие этическим стандартам статистических организаций имеет первостепенное значение. Очень важно, чтобы БЯМ соответствовали принципам ответственного ИИ, включая прозрачность, справедливость, конфиденциальность и подотчетность. Обеспечение этичности разработки и внедрения таких моделей соответствует более широкому общественным ценностям и нормативно-правовой базе.

**Поставщики инфраструктуры и платформ ИИ** — это вторая категория экосистемы БЯМ, в которую входят компании, предоставляющие необходимую аппаратную и программную инфраструктуру для обучения, развертывания и эксплуатации БЯМ, такие как Microsoft Azure ML, Google Cloud AI platform, AWS

SageMaker и другие. Для статистических организаций работа с такими поставщиками требует внимания к масштабируемости, производительности, технической совместимости и глубокого понимания структуры затрат, включая любые потенциальные скрытые расходы.

**Разработчики приложений БЯМ** — это технологические компании, которые занимаются разработкой приложений или сервисов, использующих БЯМ для конкретных функций, таких как разработка чат-ботов. Инновации в разработке приложений, ориентированная на пользователя структура и соблюдение стандартов конфиденциальности данных — важнейшие аспекты их вклада.

**Услуги по кастомизации и тонкой настройке ИИ** — это важный сегмент компаний, включающий стартапы и специализированные технологические фирмы в области ИИ, которые адаптируют существующие БЯМ для удовлетворения конкретных потребностей клиентов. Их адаптивность и способность интегрировать индивидуальные решения в существующие системы являются ключевыми соображениями для статистических организаций.

Не менее важны **исследовательские и инновационные лаборатории БЯМ**, которые включают в себя академические исследовательские лаборатории и отделы исследований и разработок. Эти организации расширяют границы возможностей БЯМ, уделяя особое внимание передовым исследованиям и этичным методам работы с искусственным интеллектом. Их работа вносит значительный вклад в расширение базы знаний в области ИИ и БЯМ. Сотрудничество с такими лабораториями может обеспечить статистическим организациям доступ к новейшим исследованиям и этичным практикам использования ИИ.

В этой экосистеме **сообщество БЯМ и инициативы с открытым исходным кодом** играют ключевую роль. Такие платформы, как Hugging Face и различные репозитории GitHub, посвященные исследованиям в области БЯМ, способствуют вовлечению сообщества, продвигая культуру открытого исходного кода в разработке БЯМ. Эти инициативы стимулируют инновации и обеспечивают доступность инструментов и ресурсов, что крайне важно для совместной и инклюзивной экосистемы БЯМ. Статистические организации должны сотрудничать с этими инициативами, чтобы получить доступ к богатому набору инструментов и ресурсов с открытым исходным кодом.

В самой последней категории — **сервисы по предоставлению данных и обучению БЯМ** — представлены компании, которые играют важную роль в предоставлении обширных и разнообразных наборов данных, необходимых для обучения БЯМ. Эти организации не только предоставляют данные, но и могут предложить важнейшие услуги, облегчающие процесс обучения БЯМ. В этой области выделяются такие компании, как EleutherAI и Hugging Face, которые предлагают ряд наборов данных и инструментов, необходимых для разработки надежных и эффективных БЯМ. Их вклад имеет решающее значение для обеспечения обучения БЯМ на разнообразных, обширных и высококачественных наборах данных, что имеет решающее значение для точности и надежности этих моделей. Кроме того, эти сервисы часто включают в себя инструменты и платформы, которые помогают в эффективном и результативном обучении БЯМ, что делает их незаменимой частью экосистемы БЯМ. Статистические организации должны сотрудничать с этими организациями для получения высококачественных, разнообразных источников данных и эффективных платформ для обучения.

Каждая категория в экосистеме БЯМ предлагает уникальные возможности для участия, способствуя общему росту и этичному использованию технологий БЯМ.

### 5.3 Соображения по поводу открытого доступа

Существует несколько аспектов, касающихся БЯМ и поставщиков, которые будут использоваться статистической организацией, которые требуют тщательного рассмотрения. Основные параметры, которые необходимо учитывать, включают в

себя доступность базовой структуры модели и обучающих данных, лицензионные условия, регулирующие использование модели, а также доступ к входным и выходным данным при использовании БЯМ. Как правило, оценка включает в себя анализ компромисса между преимуществами удобства и необходимостью контроля. Стоимость и доступ к компетенциям также являются важными моментами, поскольку для масштабной работы БЯМ требуются значительные ИТ-инфраструктура и специальные знания.

Доступность БЯМ характеризуется широким спектром возможных вариантов. Некоторые модели находятся в открытом доступе, что позволяет проверять и изменять их архитектуру и веса путем тонкой настройки. И наоборот, другие имеют статус активов ПИС, доступ к которым предоставляется только через API или другие интерфейсы. В некоторых случаях поставщики моделей с закрытым доступом могут предлагать возможности для тонкой настройки, позволяя пользователям адаптировать веса модели к своим конкретным данным и случаям использования. Хотя прямой доступ к весам модели может показаться несущественным, возможность настройки БЯМ под конкретные данные и случаи использования может оказаться очень важной для статистической организации.

Что касается моделей, находящихся в открытом доступе, то тщательное изучение условий лицензирования является императивом. Создатели БЯМ могут наложить особые условия на использование модели, которые могут быть непреднамеренно нарушены неосведомленными пользователями.

Прозрачность и доступность обучающих данных модели имеют первостепенное значение для оценки потенциального присутствия необъективных, вредных или защищенных авторским правом материалов, которые могут повлиять на результаты, генерируемые моделью (как обсуждается в разделе 4). В таких случаях полная прозрачность и доступность необходимы для снижения репутационных рисков, поскольку набор обучающих данных в значительной степени определяет результаты работы модели. Еще один аспект, требующий рассмотрения, касается доступности для БЯМ современной информации. Корпус знаний, используемый в процессе обучения модели, ограничивается датой учета данных, предшествующей началу процесса обучения модели. Устранение этого недостатка предполагает интеграцию обновленного контента, что часто называют генерацией ответа, дополненной результатами поиска (RAG). Кроме того, некоторые системы оснащены механизмами, позволяющими БЯМ получать доступ к данным из Интернета в режиме реального времени.

Наконец, следует тщательно продумать вопрос конфиденциальности входных и выходных данных. Во многих облачных сервисах входные и выходные данные могут храниться поставщиком услуг, чтобы облегчить будущие итерации БЯМ путем обучения и точной настройки. В связи с этим пользователи могут столкнуться с ограничениями в использовании конфиденциальной информации. Однако стоит отметить, что некоторые поставщики начинают предоставлять доступ к закрытым моделям в среде «песочницы», предлагая пользователям возможность сохранять полный контроль и конфиденциальность над входными и выходными данными.

В целом, статистические организации должны оценивать потенциальные преимущества и недостатки моделей открытого доступа при оценке БЯМ, в частности уровень прозрачности и способность сотрудничать с другими статистическими организациями, учитывая открытый характер этих моделей.

## 5.4 Связи с общественностью

Область БЯМ быстро развивается. Хотя эти модели обладают потрясающими возможностями, их стремительное развитие ставит технологию искусственного интеллекта в довольно серую зону, где общественное мнение и настроения могут быть неопределенными и подверженными изменениям.

Будучи государственными учреждениями, чьи продукты оказывают значительное влияние на политику и принятие решений в масштабах страны, статистические организации несут большую ответственность за ответственное использование БЯМ, а также за прозрачное информирование общественности. Сам факт того, что основная деятельность статистических организаций (т. е. формирование официальной статистики и предоставление услуг по обработке данных) в значительной степени зависит от доверия общества, требует, чтобы статистические организации уделяли еще больше внимания и усилий связям с общественностью, в частности с поставщиками данных, которые могут выразить опасения по поводу того, что их данные могут быть использованы ненадлежащим образом при взаимодействии статистических организаций с БЯМ. В конце концов, общественность имеет фундаментальное право понимать, как могут быть использованы ее данные, и быть уверенной в том, что приняты меры по их защите.

Сообщая об использовании БЯМ, важно отметить, что статистические организации:

- целенаправленно используют БЯМ при наличии очевидных преимуществ: важно четко объяснить, почему БЯМ используются в статистических организациях, и подчеркнуть ощутимые преимущества этой технологии (например, повышение эффективности, экономия затрат, улучшение качества услуг), приведя конкретные примеры, когда использование БЯМ принесло успех. Например, чат-боты на основе БЯМ, помогающие общественности лучше понимать статистические данные и получать доступ к ним, являются одной из ролей, которую генеративный ИИ/БЯМ может играть совершенно автономно;
- отдают себе отчет в ограничениях и рисках: важно продемонстрировать, что статистические организации не используют БЯМ вслепую и осведомлены о потенциальных ограничениях и рисках, связанных с БЯМ. Можно упомянуть области, в которых БЯМ не используются и не будут использоваться (например, для составления индивидуальных прогнозов, которые могут негативно повлиять на людей);
- принимают необходимые меры по снижению рисков: крайне важно объяснить, какие меры были приняты для смягчения ограничений и рисков (например, меры, принятые для защиты конфиденциальности и безопасности данных), подчеркнув при этом, что вмешательство человека осуществляется для контроля и руководства использованием БЯМД.

Что касается внутренней коммуникации, то важно учитывать, что тот или иной вариант использования может непреднамеренно рассказать сотрудникам о приоритетах организации. Например, БЯМ могут обеспечить эффективный способ подготовки нетехнических резюме, но это также можно рассматривать как передачу организацией этой задачи модели, вместо того чтобы развивать внутренние навыки написания нетехнических текстов для общения с заинтересованными представителями общественности.

## 5.5 Практические рекомендации и заключительные замечания

Использование БЯМ статистическими организациями все еще находится на начальном этапе, но их ландшафт быстро развивается. Передовые методы разрабатываются с течением времени и требуют постоянных усилий со стороны статистических организаций, чтобы поддерживать их в актуальном состоянии. Имеется несколько практических рекомендаций, которые, на наш взгляд, актуальны в краткосрочной перспективе и могут выдержать испытание временем.

*Первая* касается проведения обучения по БЯМ на всех уровнях организации — техническом, оперативном и управленческом — для повышения осведомленности и лучшего понимания возможностей и ограничений БЯМ.



*Во-вторых*, мы рекомендуем приступать к внедрению БЯМ с реализации небольших пилотных проектов, чтобы ознакомиться с технологией и понять ее потенциальную полезность. Эти небольшие проекты могут расширить возможности статистических организаций в этой области, предоставить результаты, которые могут оправдать дальнейшие инвестиции и уточнить их направленность, и в конечном итоге снизить риски, связанные с изучением использования БЯМ.

*В-третьих*, статистические организации должны разработать общую стратегию БЯМ, когда осведомленность о ней и ее освоение достигнут достаточного уровня, реализовав несколько небольших проектов, о которых говорилось выше.

*Наконец*, статистические организации должны прилагать постоянные усилия, чтобы идти в ногу с постоянно меняющимся ландшафтом БЯМ, как с технологической, так и со стратегической точки зрения.

Признавая стремительное развитие БЯМ, мы понимаем, что темпы прогресса находятся за пределами нашего полного понимания. Целью данного документа является сбор существующих примеров использования до сегодняшнего дня и глубокое изучение этой темы с различных точек зрения, актуальных для статистических организаций. По причине динамичного характера развития этой области совместная работа имеет решающее значение. Поэтому мы приглашаем экспертов к сотрудничеству, обмену мнениями и совместному изучению этого постоянно меняющегося ландшафта. Мы продолжаем изучать эту тему, и мы приветствуем постоянное участие в этом исследовании.

---