



unesco

Серия публикаций ЮНЕСКО по журналистскому образованию

Пособие для  
преподавателей  
журналистики

# Освещение тематики искусственного интеллекта

Опубликовано в 2023 г. Организацией Объединённых Наций по вопросам образования, науки и культуры (ЮНЕСКО)  
7, place de Fontenoy, 75352 Paris 07 SP, France  
и Бюро ЮНЕСКО в Алматы.

© UNESCO 2023

ISBN 978-92-3-400070-3



Данная публикация предлагается в открытом доступе под лицензией Attribution-ShareAlike 3.0 IGO (CC-BY-SA 3.0 IGO) (<http://creativecommons.org/licenses/by-sa/3.0/igo/>). Используя содержание данной публикации, пользователи соглашаются с правилами пользования Репозитория открытого доступа ЮНЕСКО (<https://www.unesco.org/ru/open-access/cc-sa>).

Название на языке оригинала: «Reporting on artificial intelligence: a handbook for journalism educators». Опубликовано в 2023 г. Организацией Объединённых Наций по вопросам образования, науки и культуры.

Изображения на Shutterstock не подпадают под действие лицензии CC-BY-SA и не могут быть использованы или воспроизведены без предварительного разрешения правообладателей.

Использованные названия и представление материалов в данной публикации не являются выражением со стороны ЮНЕСКО какого-либо мнения относительно правового статуса какой-либо страны, территории, города или района или их соответствующих органов управления, равно как и линий разграничения или границ.

Ответственность за взгляды и мнения, высказанные в данной публикации, несут авторы. Их точка зрения может не совпадать с официальной позицией ЮНЕСКО и не накладывает на Организацию никаких обязательств.

**Редактор:** Маарит Яаккола

**Основные авторы:** Чарли Бекетт, Эдвард Финн, Фредрик Хайнц, Фредерик Хейманс, Сурен Джаясурия, Саяш Капур, Сантош Кумар Бисвал, Арвинд Нараянан, Агнес Стенбом и Дженни Вийк

**Дополнительные авторы:** Дженни Бергенмар, Аммина Котари, Бернхард Дотцлер, Теему Роос, Николас Кайзер-Бриль, Стив Вулгар

**Управление проектами ЮНЕСКО, рецензирование и редактирование:**

Гай Бергер, директор по стратегиям и политике в секторе коммуникации и информации, ЮНЕСКО

Ванесса Дрейер, младший специалист по коммуникации и информации, ЮНЕСКО

Максим Карлюк, программный специалист, ЮНЕСКО

**Внешнее рецензирование:**

Серкан Булут, доцент, факультет журналистики, Университет Чукурова

Брюс Муцвайро, кафедра ЮНЕСКО по вопросам дезинформации, данных и демократии, Утрехтский университет

Манджу Роуз Мэтьюз, заведующая кафедрой медиаисследований, колледж Крест Нагар, Тривандрум

Камилла Нигматуллина, заведующая кафедрой цифровых медиакоммуникаций, Санкт-Петербургский университет

**Координация со стороны ЮНЕСКО:**

София Веллек, консультант по работе со СМИ

Вероника дель Карриль, МПРК / консультант по журналистскому образованию

**Графический дизайн:** Аурелия Мазойер

**Дизайн обложки:** Niso Bang Bang (Николя Шатен)

**Перевод на русский язык и вёрстка русской версии:** Александр Кочетков

Отпечатано в Казахстане

## Знакомство с концепцией журналистского освещения тематики искусственного интеллекта

Развитие искусственного интеллекта (ИИ) и задача осуществления контроля над ним оказывают влияние на общество в целом. Это означает, что освещение темы ИИ должно информировать аудиторию о последствиях применения самой технологии, выходя за рамки журналистики. Например, освещение динамики сил в условиях меняющихся отношений между компаниями, властями, гражданами и компьютерными чипами, а также между данными и алгоритмами. Хотя многие разработки в области ИИ служат общественным интересам, журналистам также необходимы знания и компетентность, чтобы предупреждать общественность о таких аспектах, как риски социальной эксклюзии, неравное распределение благ и нарушения прав человека.

Данная публикация является частью серии публикаций по журналистскому образованию, и подготовлена совместно Международной программой развития коммуникации ЮНЕСКО (МПРК) и Всемирным советом по образованию в области журналистики. Целью данного пособия является вдохновение и расширение возможностей преподавателей журналистики, чтобы они могли помочь начинающим и работающим журналистам разобраться в одной из важнейших проблем современности.

В пособии рассматриваются следующие темы:

- выработка общего представления о машинном интеллекте и определение различных типов ИИ;
- рассмотрение потенциала ИИ, а также его сильных и слабых сторон;
- формулирование сценариев будущего в реалиях ИИ путём выявления распространённых популярных установок, формирующих общественное сознание;
- понимание роли журналистики в опосредовании и формировании дискурса об ИИ;
- поиск способов взвешенного, реалистичного и ответственного освещения проблематики ИИ;
- увязка с существующими жанрами журналистики — от новостного репортажа до журналистики данных.

Повышение уровня журналистского образования — один из ключевых результатов, к которым стремится МПРК, уникальная межправительственная программа в системе ООН, специализирующаяся на развитии СМИ.

**60%**  
**новостей в области технологий**  
приходится на отраслевые продукты; самая большая категория источников новостей — представители бизнеса (30%)



*«Мысли о войне возникают в умах людей, поэтому в сознании людей следует укоренять идею защиты мира»*



unesco

Серия публикаций ЮНЕСКО по журналистскому образованию

Пособие для  
преподавателей  
журналистики

# Освещение тематики искусственного интеллекта

Под редакцией Маарит Яаккола








# Содержание

<b>Вступление</b>	<b>6</b>
<b>Предисловие: проблема многообразия технологий искусственного интеллекта</b>	<b>9</b>
<b>Предисловие: необходимость знания журналистами новых технологий</b>	<b>13</b>
<b>Введение: методические аспекты преподавания темы технологий искусственного интеллекта</b>	<b>15</b>
Искусственный интеллект как предмет программы обучения	20
Задачи настоящего пособия	23
Структура настоящего пособия	25
Благодарности	26
<hr/>	
<b>Модуль 1</b>	
<b>Определение понятия искусственного интеллекта</b>	<b>27</b>
<i>Очерк Агнес Стенбом</i>	
<hr/>	
<b>Модуль 2</b>	
<b>Культурные мифы и представления об искусственном интеллекте</b>	<b>37</b>
<i>Очерк Эдварда Финна и Сурена Джаясурии</i>	
<hr/>	
<b>Модуль 3</b>	
<b>Нормативная база и рекомендации в области искусственного интеллекта</b>	<b>49</b>
<i>Очерк Фредрика Хайнца</i>	
<hr/>	
<b>Модуль 4</b>	
<b>Журналистика в алгоритмических культурах</b>	<b>61</b>
<i>Очерк Фредерика Хейманса</i>	
<hr/>	
<b>Модуль 5</b>	
<b>Освещение вредоносного использования технологий искусственного интеллекта</b>	<b>69</b>
<i>Очерк Сантоша Кумара Бисвала</i>	






**Модуль 6****Модели подачи материала об искусственном интеллекте 81***Очерк Маарит Яаккола***Модуль 7****Искусственный интеллект в журналистике и журналистской практике 94***Очерк Дженни Вийк***Восемнадцать «ловушек» в ИИ-журналистике 109***Саяш Капур и Арвинд Нараянан***Послесловие: подготовка журналистов, компетентных в вопросах ИИ 115****Словарь терминов 120****Авторы 124****Библиографический список 127****Список иллюстраций 135**

## Как пользоваться настоящим пособием

*Разделы модулей*

- |  |  |
|--|--|
|  <b>Краткое содержание</b>      |  <b>Конспект</b>                  |
|  <b>Вопросы для размышления</b> |  <b>Предлагаемые задания</b>      |
|  <b>Цели модуля</b>             |  <b>Дополнительная литература</b> |
|  <b>Результаты обучения</b>     |  |

*Категории предлагаемых заданий*

- |   |   |
|---|---|
|  <b>ПРОЧИТАТЬ И ПОРАЗМЫШЛЯТЬ</b> |  <b>ПОИСКАТЬ И ИССЛЕДОВАТЬ ИСТОЧНИКИ</b> |
|  <b>ОБСУДИТЬ</b>                 |   |
|  <b>ВЫРАБОТАТЬ ИДЕЮ</b>          |  <b>СОЗДАТЬ</b>                          |

# Вступление

*Первый текст, приведённый ниже, был написан роботом GPT-3 компании OpenAI (<https://openai.com/ari/>). GPT-3 означает Generative Pre-trained Transformer 3 – это авторегрессионная генеративная языковая модель третьего поколения, использующая механизмы глубокого обучения для создания текстов, похожих на человеческие (см. Floridi & Chiriatti, 2020). Инструкция для робота была следующей: «Написать вводный текст по теме “журналистика и искусственный интеллект”. Привести аргументы “за” и “против”. Дать обзор темы с точки зрения принципов демократии».*

Журналистика долгое время была областью, в которой изменения и адаптация происходили медленно. Однако с появлением систем искусственного интеллекта (ИИ) журналистика как отрасль начинает меняться. Искусственный интеллект начинает использоваться в журналистике различными способами – от помощи в написании статей до выявления фальшивых новостей. Это даёт определённые преимущества, но в то же время вызывает определённые опасения.

**Одно из преимуществ применения искусственного интеллекта в журналистике заключается в том, что он может помочь авторам создавать более качественные сюжеты** ИИ может помочь выявить закономерности и тенденции, которые могут быть не видны репортёрам. ИИ также может помочь в сборе данных и информации для репортажей. Это помогает журналистам экономить время и силы, которые можно направить на написание более качественного материала.

Однако существуют и некоторые опасения по поводу применения ИИ в журналистике.

Одно из опасений связано с тем, что ИИ может использоваться для создания фальшивых новостей. Дело в том, что средства ИИ способны создавать реалистичные сюжеты, которые не соответствуют действительности. Это может привести к тому, что люди будут дезинформированы или поверят ложной информации. Другая проблема заключается в том, что ИИ может быть использован для контроля за содержанием новостей. Это связано с тем, что с помощью технологий ИИ можно определить, какими историями активнее всего делятся в социальных сетях и что больше всего нравится людям, а затем вынести такие истории на первые места. Это может привести к формированию предвзятого и одностороннего взгляда на новости.

В целом, в использовании технологий искусственного интеллекта в журналистике есть как преимущества, так и вызывающие опасения аспекты. Важно знать все плюсы и минусы при рассмотрении вопроса о том, как технологии ИИ меняют сферу журналистики.

*Приведённые ниже тексты – это два варианта (из нескольких), предоставленных сервисом ChatGPT в декабре 2022 года. Это бета-версия сервиса от компании OpenAI. Это ответы на вопрос «Как в рамках преподавания журналистики научить студентов освещать тему искусственного интеллекта?»*

1. Начать с объяснения основ концепции ИИ и его применения в различных областях, таких как здравоохранение, финансы, транспорт и т. д.
2. Обсудить потенциальное влияние технологий ИИ на общество, их преимущества и недостатки.
3. Подчеркнуть важность проведения исследований и проверки фактов при освещении темы ИИ, поскольку эта технология окружена большим количеством дезинформации и шумихи.
4. Побуждать студентов к поиску экспертных мнений и точек зрения при освещении тем, связанных с ИИ, а также к рассмотрению этических последствий применения этой технологии в своих репортажах.
5. Обсудить, как проводить интервью со специалистами по ИИ и как задавать правильные вопросы, чтобы получить информативные и увлекательные ответы.

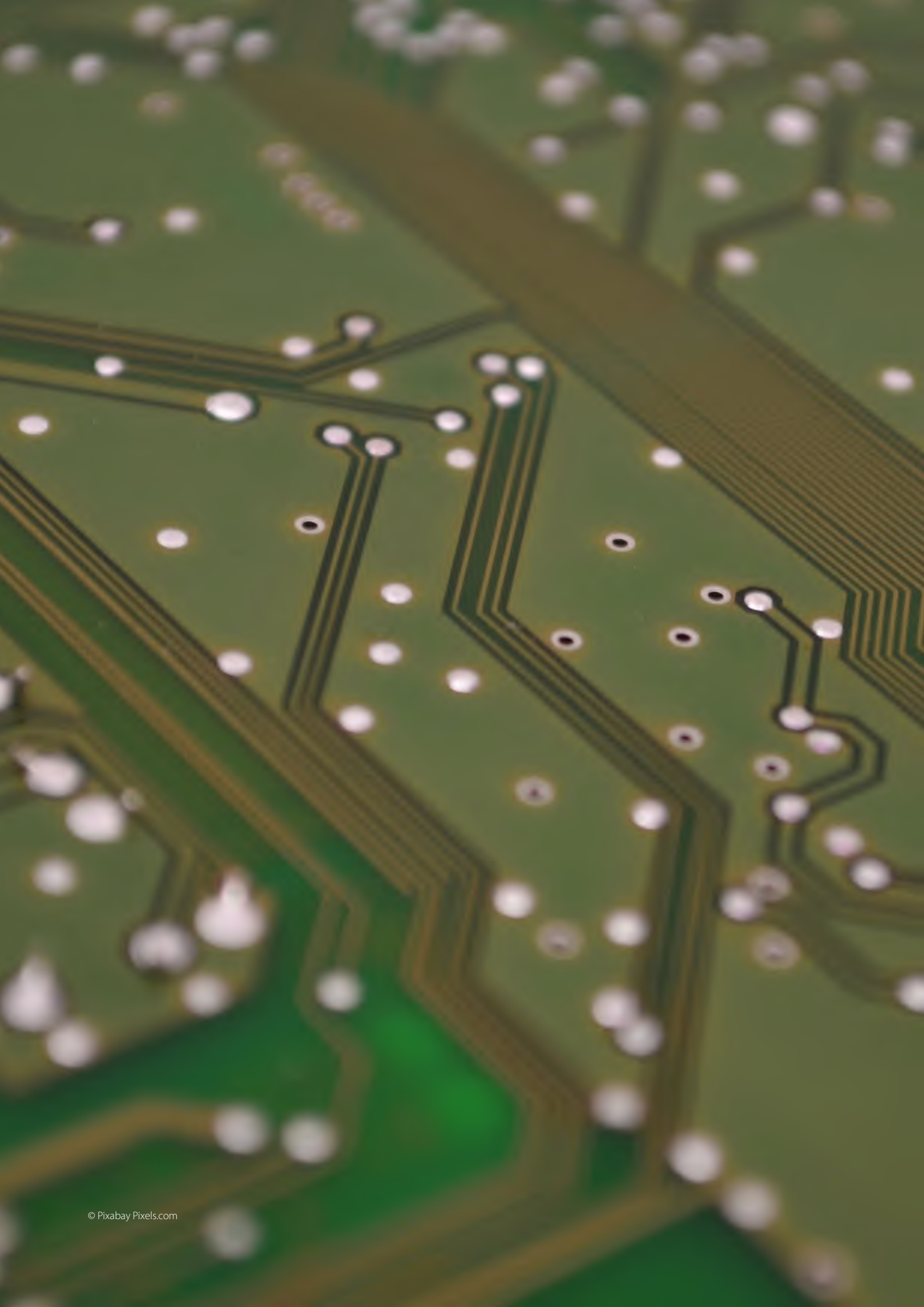
6. Поощрять студентов к использованию различных источников и форматов при освещении темы ИИ, таких как интервью, визуализация данных и экспертная аналитика.
7. Предоставить рекомендации по освещению темы ИИ таким образом, чтобы она была доступна и понятна широкой аудитории, но при этом сохраняла высокий уровень достоверности и глубины.
8. Помогать студентам развивать критический взгляд при освещении темы ИИ, побуждая их подвергать сомнению принятые допущения и критически осмысливать последствия применения технологии.
9. Предоставить студентам возможность отработать навыки журналистики в области ИИ, выполняя задания на занятиях и делая репортажи в реальных условиях.
10. Поощрять постоянное обучение и повышение квалификации в области журналистики по теме ИИ, поскольку эта область постоянно развивается и меняется.

При освещении тематики искусственного интеллекта преподаватели журналистики должны проинструктировать студентов по следующим ключевым моментам:

1. Важность поддержания собственной осведомлённости о последних достижениях и разработках в области ИИ.
2. Необходимость понимания этических и социальных последствий применения технологий ИИ и их влияния на различные сектора и отрасли человеческой деятельности.
3. Важность поиска различных точек зрения и экспертных мнений по вопросам ИИ для сбалансированного и взвешенного освещения этой темы.
4. Необходимость критической оценки утверждений и заявлений экспертов и лидеров отрасли об ИИ, а также проверки фактов и верификации информации перед её публикацией.
5. Важность осознания потенциальных ошибок и ограничений технологий ИИ, а также учёта этих факторов при освещении использования и применения ИИ.
6. Необходимость изложения сложных технических концепций и разработок в области ИИ в ясной и доступной для широкой аудитории форме.
7. Важность использования различных источников и форматов, таких как интервью, анализ данных и визуализации, для всестороннего и увлекательного освещения темы ИИ.

*Эти три примера текстов, созданных искусственным интеллектом, содержат моменты, которые пересекаются с последующими главами. Тем не менее, читатель найдёт в этом пособии множество идей, сформулированных человеком, а также различные примеры и творческие упражнения. В настоящее время преподаватели журналистики имеют возможность внести свой уникальный и значительный вклад в развитие этой темы.*





# Предисловие: проблема многообразия технологий искусственного интеллекта

**Гай Бергер,**

бывший секретарь Международной программы развития коммуникации ЮНЕСКО (2011–2022 гг.)

То обстоятельство, что в создании этого пособия участвовало множество авторов, объясняет разнообразие взглядов и определений, представленных на его страницах. В одних главах говорится об **«ответственном ИИ»**, в других — о необходимости обеспечения **«этичности»** ИИ. Часто встречается спорное выражение «четвёртая промышленная революция». Такое разнообразие говорит нам о том, что должно стать отправной точкой в преподавании темы искусственного интеллекта (ИИ), а именно: существует множество трактовок, и не следует воспринимать предложенную семантику как общепринятую.

Соответственно, с самого начала студенты-журналисты должны чётко осознавать, что сам термин «ИИ» для многих людей означает разные вещи. Их задача как журналистов — избегать создания у аудитории впечатления, что существует однозначный смысл этого квазимагического словосочетания.

Вопреки маркетинговым утверждениям, использование алгоритмов не равнозначно ИИ как таковому (хотя ИИ предполагает применение алгоритмов). Данные являются неотъемлемым компонентом функционирования ИИ, но не все проблемы и разработки, «основанные на данных», конгруэнтны ИИ.

Общественность, а значит, и СМИ, должны понять, что различные определения и интерпретации концепции ИИ отражают различные контексты и интересы. Особенно когда термин «ИИ» используется как «сцена» для «раскрутки» предложений различных консультантов, компаний и директивных органов.

*Все сведения (и статистические данные), касающиеся новостей, связанных с ИИ, должны подвергаться критическому журналистскому анализу, независимо от того, кто является источником — научные круги, регулирующие органы, крупные корпорации или предприниматели. То же касается и известий об использовании средств ИИ как составной части деятельности компаний, занимающихся компьютерными играми, служб электронной безопасности, продавцов невзаимозаменяемых токенов [англ. non-fungible token, NFT — прим. перев.], криптовалютных бирж, здравоохранения, сельского хозяйства, военных действий и т. д. Необходимо анализировать и разбирать все упоминания об ИИ, к какой бы области деятельности они ни относились.*

Пожалуй, самое интересное применительно к искусственному интеллекту зачастую заключается не в технических аспектах, связанных с программным обеспечением или аппаратной частью. Сфера ИИ — это **тесное переплетение различных аспектов жизнедеятельности человека**, включая корыстные интересы, всегда связанные с формированием ландшафта и продвижением технологий ИИ, разработкой программного обеспечения, контролем и доступностью аппаратных средств. И, конечно, вопросы наличия, владения и условий доступа к данным.

На самом деле каждая микроистория — например, достижение в медицине с использованием средств ИИ — является частью более общей системной истории человечества, которая даёт множество богатых ракурсов для освещения конкретных событий. Для того чтобы понять, насколько это важно, необходимо осознать, что совокупность технологий, известных под общим названием ИИ, неразрывно связана с обществом, его динамикой сил и проблемами. В свою очередь, это наблюдение должно привлечь внимание

тех, кто не специализируется на технологиях ИИ, к предоставляемым журналистикой возможностям проследить множество «человеческих зацепок» когда ИИ становится очевидным компонентом репортажей на темы, которые поначалу кажутся далёкими от технологий.

Внимание к человеческому аспекту наводит, например, на множество ярких сюжетов о труде человека. Одна подборка посвящена роли Глобального Юга в разметке данных для обучения ИИ. В другой подборке — истории об (ограниченном) демографическом составе программистов, занимающихся обработкой данных и взаимодействием с ними. Действительно, журналистика может выявить проблемы многообразия и инклюзивности, а также утечки мозгов. Она может пролить свет на то, как эти факторы определённым образом формируют ландшафт ИИ и как они влияют на одни возможности (например, использование ИИ в беспилотных автомобилях), но не влияют на другие (например, использование ИИ в управлении общественным транспортом).

Человеческий аспект также включает людей как объекты функционирования искусственного интеллекта — будь то распознавание лиц или определение размеров ипотечных займов или вынесение решений суда с помощью искусственного интеллекта. **Повсеместное извлечение (а также продажа и использование) данных о людях для развития технологий искусственного интеллекта** — ещё один сюжет, требующий постоянного мониторинга, как, например, новости 2022 года об опасениях по поводу приложений для отслеживания менструального цикла, которые позволяли третьим сторонам получать информацию о состоянии беременности. И, как показали такие авторы, как Кейт Кроуфорд и Шошана Зубофф, данные — это не просто ресурс, который можно просто взять — и это ценное наблюдение, которое журналисты должны изучить при подготовке материалов об ИИ.

*Можно добавить, что, хотя «пользователи» ИИ могут быть активными субъектами процесса, они также часто одновременно являются «сырьём» для производства «товаров» для продажи клиентам, готовым платить, и, не в последнюю очередь рекламодателям и другим лицам, заинтересованным в манипулировании, а не в поддержке независимости человека.*

Человеческие аспекты данных включают в себя вопросы потенциальной предвзятости, и простираются дальше. Они побуждают нас задаваться вопросами об осведомлённости и согласии отдельных лиц на сбор, использование и владение их информационными активами, особенно (хотя и не только) когда речь идёт о персональных данных. Так, развитие технологий ИИ вызвало интерес к **альтернативным моделям управления данными** — например, к идее «доверительных фондов данных» для облегчения законного обмена между владельцами и пользователями данных — и это ещё одна история, которую журналистам предстоит исследовать и рассказать.

Рассматривая человеческие аспекты технологий ИИ, мы, конечно же, не должны оставлять без внимания экологические аспекты использования ИИ. Например, такие технологии как «облака» и «биткойн» прочно обосновались в массивных центрах обработки данных, потребляющих значительное количество энергии, что неизбежно приводит к выбросам углекислого газа, и эта история освещается сравнительно мало.

При всем этом **централизация власти над разработкой и внедрением технологии ИИ**, а также её сетевые эффекты и зависимость от траектории предшествующего развития требуют **более пристального внимания со стороны СМИ**. Речь идёт о тенденции, которая имеет поистине глубокое значение для всего человечества с точки зрения прав человека, демократии, экологической устойчивости и инклюзивного развития.

Парадокс заключается в том, что при освещении темы ИИ, а также касаясь тем ИИ в рамках других сюжетов студентам-журналистам необходимо заглядывать далеко за пределы самой технологии. Вся цепочка взаимосвязей требует внимания... и в то же время она является настоящим рогом изобилия сюжетов, возможно, больше, чем многие другие разработки.

Общим для всех начинающих репортёров должно стать рассмотрение того, как каждая новая разработка и внедрение средств ИИ соотносятся с вопросами прав человека и соответствующими нормами.

*Это означает, что необходимо знать, как задавать вопросы о наличии (или отсутствии) и уровне должной заботы о правах человека со стороны разработчиков технологий ИИ, а также заказчиков и клиентов. В журналистских*

*интервью не следует пренебрегать вопросами о предварительных и последующих оценках воздействия на права человека и соответствующих мерах по снижению рисков.*

Развивая эту мысль, следует напомнить, что оценка того, какие права человека могут быть ущемлены, является частью общего призвания журналиста. В случае с ИИ, чтобы не ослепнуть от обещаний прогресса, важно учитывать, прежде всего, соблюдение прав на уважение человеческого достоинства, неприкосновенность частной жизни, равенство и справедливость. В этом смысле новостные сюжеты могут поставить на политическую повестку дня вопрос об обеспечении независимого аудита воздействия ИИ, а также о непредвиденных последствиях внедрения ИИ. Журналисты также могут сыграть свою роль в привлечении внимания к деятельности по раскрытию нарушений, что приведёт к изменениям в политике компаний и правительств в отношении использования ИИ и его многочисленных компонентов.

При подготовке журналистских материалов, связанных с ИИ, можно также задавать следующие вопросы:

- Какие этические регламенты использовались при создании, хранении, продаже и использовании данных и какие проверки проводились при внедрении соответствующих процессов?
- Какие существуют (или отсутствуют) системы отчётности и отслеживания инцидентов, связанных с причинением вреда, и какие выводы можно сделать для улучшения ситуации?
- Соответствует ли система глобальному стандарту, согласно которому «техническая и организационная концепция должна обеспечивать возможность проверки и отслеживания (работы)...» (цитата взята из п. 43 Рекомендации ЮНЕСКО об этических аспектах искусственного интеллекта 2021 г., согласованной 193 государствами-членами)?

Без информированной и независимой журналистики об ИИ общество в целом будет испытывать огромный дефицит **прозрачности и подотчётности**, которые абсолютно необходимы для управления столь влиятельной технологией. С другой стороны, именно качественное освещение в прессе позволяет понять, какие пакеты мер государственного **регулирования** и саморегулирования являются оптимальными и каков потенциал таких альтернатив, как совместное регулирование и регулирование с участием многих заинтересованных сторон.

Журналисты могут с пользой применить тот же критический подход к процессу интеграции систем искусственного интеллекта в свои рабочие задачи и работу редакции в целом. И, конечно, как и в других областях, для всех материалов, затрагивающих тему ИИ, также необходима прозрачность относительно случаев, когда деятельность журналистов или их материалы так или иначе спонсируются.

Вся эта динамичная ситуация и дебаты вокруг ИИ открывают захватывающие возможности для преподающих и изучающих журналистику, и работающих журналистов. В этой связи ЮНЕСКО благодарит редактора Маарит Яаккола, авторов и Всемирный совет по образованию в области журналистики под руководством Верики Рупар, которая курировала инициативу, приведшую к созданию данного пособия.

Повышение уровня журналистского образования — один из ключевых результатов, к которому стремится комитет ЮНЕСКО, состоящий из 39 государств-членов, входящих в Международную программу развития коммуникации (МПРК). При поддержке фондов МПРК данное пособие входит в серию публикаций ЮНЕСКО по журналистскому образованию.

Некоторые публикации этой серии были встречены с таким энтузиазмом, что читатели предложили их добровольный перевод. Если кто-либо из читателей данного пособия заинтересован в добровольном переводе этого ресурса на другие языки, мы будем рады помочь вам в этом. Кроме того, у вас могут быть идеи по устранению пробелов в работе. Эксперты, рецензировавшие данную публикацию, настоятельно рекомендовали уделить больше внимания мнению студентов и опыту стран Глобального Юга при обучении журналистике в области ИИ.

Всех, кто заинтересован в осуществлении перевода и/или проведении дальнейших исследований, могущих дополнить данное пособие (возможно, для публикации в качестве сопутствующего сборника), мы просим обращаться по адресу [IPDC.Secretariat@unesco.org](mailto:IPDC.Secretariat@unesco.org)

А пока мы рекомендуем вам ознакомиться с этим пособием. Вы будете щедро вознаграждены. Поделитесь с ЮНЕСКО своим мнением. Ваши усилия помогут журналистскому образованию сохранить свою актуальность в быстро меняющемся мире.



FIRE ESCAPE

# Предисловие: необходимость знания журналистами новых технологий

**Чарли Бекетт,**

практикующий профессор факультета СМИ и коммуникаций Лондонской школы экономики, директор международного аналитического центра по журналистике Polis.

Наберите в Интернете словосочетание «искусственный интеллект», и на экране появятся изображения сияющих роботов или светящихся компьютерных схем. Это выглядит захватывающе и футуристично, но в то же время довольно пугающе. То же самое чувствуют и журналисты.

*! Заберут ли роботы наши рабочие места? Будут ли алгоритмы писать статьи и выпускать новости?*

Реальность же одновременно и более тревожащая, и более захватывающая.

После более чем трёх лет работы по всему миру с журналистами, использующими средства искусственного интеллекта, я убеждён, что эти технологии могут помочь нам сделать нашу работу более эффективной, результативной и увлекательной. Но ИИ не собирается «спасать» журналистику или полностью её трансформировать. На практике он выполняет довольно скучные, повторяющиеся, простые задачи. Его необходимо тщательно программировать, адаптировать и управлять им в соответствии с редакционными требованиями.

ИИ не так прост в использовании и не так «интеллектуален», как можно было бы ожидать. Это влечёт за собой **проблемы предвзятости и ответственности за технологию**. Это означает, что новостным организациям придётся вкладывать средства в развитие и внедрение новых навыков и методов работы. Многие из них не имеют для этого ресурсов и могут остаться позади. Однако технологии ИИ уже способны существенно изменить ситуацию. Есть надежда, что они смогут дополнить работу журналистов, предоставляя им больше времени и ресурсов для использования своих человеческих навыков, таких как креативность, эмпатия и способность рассуждать, чтобы повысить ценность своей журналистской работы.

Независимо от того, оптимистично или пессимистично вы относитесь к этой технологии, крайне важно, чтобы мы все уделяли ей больше внимания. В ближайшем будущем все аспекты нашей жизни будут в большей степени опираться на данные и алгоритмы, а искусственный интеллект будет формировать наше общество. Как журналисты мы должны понимать, что это за технология и на что она способна. Отчасти потому, что она будет влиять на процессы создания и потребления новостей, но также и потому, что мы должны будем осуществлять журналистскую деятельность с тем, чтобы **информировать граждан о влиянии ИИ на наш мир**.

Как будет показано в данном пособии, ИИ трудно дать определение. На практике **чистого ИИ не существует**. Не существует роботов, которые думают сами за себя. На практике это сочетание технологий машинного обучения, обработки естественного языка, автоматизации и обработки данных. Вы уже используете ИИ при поиске в Интернете или прокладывании пути домой в приложении-навигаторе. На самом деле работа искусственного интеллекта может быть сложной, но это не повод не понимать возможности его применения и потенциал. Данное руководство является отличным источником информации об этом.

В рамках проекта Journalism AI («ИИ в журналистике») в Лондонской школе экономики мы изучили глобальный спрос на более подробную информацию и присоединились к дебатам вокруг технологий ИИ в журналистике. Мы разработали тренинги, информационные ресурсы и ряд совместных инновационных программ при участии журналистов по всему миру. Они рассказали нам, что ИИ может улучшить все аспекты их работы. Он может помочь персонализировать новостные материалы, чтобы люди получали актуальные для них новости в удобном для них виде и в удобное для них время. Он может помочь найти скрытые в данных истории, или выявить **новые мнения**. Он может помочь в создании контента, особенно простых, функциональных, фактологических материалов, таких как прямые финансовые, спортивные или погодные сводки. Он может использоваться для масштабной проверки фактов и даже противостоять человеческой предвзятости журналистов. В настоящее время адаптацией средств ИИ занимаются крупные новостные организации, такие как BBC, Bloomberg или Wall Street Journal. Но если вы зайдёте на сайт JournalismAI, то увидите прекрасные примеры того, как небольшие новостные редакции разрабатывают рациональные методы использования ИИ для повышения эффективности своей журналистической работы.

**! Это уже третья волна технологических изменений для СМИ, а также для других секторов нашей становящейся всё более цифровой «метавселенной».**

Сначала мы все ушли в Интернет, перенесли наши традиционные процессы на веб-сайты. Затем, около 15 лет назад, появились социальные сети, и журналистика стала интерактивной, многоформатной, со множеством источников и платформ. Теперь третья волна технологий искусственного интеллекта обещает взять на себя большую часть нынешнего журналистского труда. Кроме того, она открывает возможности для развития новых форм, таких как **дополненная реальность, аудио- и видеоматериалы по запросу, гиперперсонализация и ещё большее разнообразие за счёт автоматизированного перевода, транскрипции и редакторской правки**.

Журналисты также отметили, что необходимо провести дискуссию о том, кто производит и контролирует эти технологии. Какие **искажения присущи** создаваемым **наборам данных или программным продуктам**? Как использовать эти инструменты и системы для улучшения журналистики, а не только для создания материалов-приманок или дезинформации? Очень важно, чтобы преподаватели и студенты факультетов журналистики участвовали в обсуждении того, как эти технологии могут повлиять на работу новостных СМИ. Также важно, чтобы в дискуссии участвовали технические специалисты, регулирующие и директивные органы и граждане. Мы живём в мире, где вся наша жизнь — от покупок до политики — становится всё более медиатизированной, всё более онлайн-овой, всё более социальной. Это даёт огромные удобства и захватывающие возможности для самовыражения, развлечения и образования. Но мы также знакомы с проблемами, которые таит в себе цифровой мир, такими как дезинформация и экстремистские высказывания.

Как человек, более двух десятилетий проработавший журналистом в аналоговую эпоху, я прекрасно понимаю, как много всего изменилось. И это меня волнует. Журналистика, как и многие другие профессии, стала более конкурентной, а в некоторых местах в мире — более опасной. На новое поколение ложится огромная ответственность за решение этических дилемм, связанных с новыми технологиями. Но у них могут появиться и новые возможности.



# Введение: методические аспекты преподавания темы технологий искусственного интеллекта

**Маарит Яаккола**

Разработка рекомендаций для будущих журналистов по теме ИИ как общественной и, соответственно, журналистской проблематике предполагает большую и сложную работу. Несмотря на то что технологии, связанные с ИИ, существуют уже давно, многие футуристические идеи прошлых эпох начинают воплощаться в жизнь только сейчас. «ИИ-защия» общества — это долгий, медленный процесс, который проникает во все сферы жизни. Поэтому, когда речь идёт о сборе рекомендаций и лучших практик по освещению этой темы в журналистике, возникают открытые вопросы и опасения, которые необходимо учитывать.



Многие аспекты, о которых говорится применительно к искусственному интеллекту, возможно, появятся в будущем, но ещё не доступны нам. Научно-техническое развитие, необходимое для достижения **«сверх-разума»** и **«сингулярности»**, — это будущий процесс, связанный с развитием вычислительной техники, а что именно он означает и как может повлиять на различные сферы жизни — это вопросы, ответы на которые невозможно полностью и однозначно предугадать. В последние годы общественные дискуссии и развитие нормативно-правовой базы в этой области активизировались, однако в обществе до сих пор нет общих, устоявшихся и прошедших широкую апробацию основ подхода к ИИ с точки зрения соответствия демократическим принципам, с точки зрения прав и интересов граждан и потребителей. Нормативно-правовая база и сопутствующая передовая практика только формируются, и не в последнюю очередь это касается использования технологий ИИ в журналистике и для журналистики. Одновременно всё больше исследований в области социальных и гуманитарных наук пытаются выяснить, как ИИ можно увязать с существующими направлениями исследований, и постоянно публикуются новые работы. Такие всеобъемлющие работы, как *«The Routledge Social Science Handbook on AI»* («Социологический справочник издательства Рутледж в области искусственного интеллекта») (Elliott, 2021), были опубликованы лишь недавно, а ряд из них всё ещё находится в стадии подготовки.

**Следовательно, нам в определённой степени не хватает понятийно-терминологического аппарата, чтобы осмыслить онтологию и эпистемологию, а также влияние и последствия распространения ИИ как общественного явления. Какие же вопросы должны задавать журналисты?**

Очевидно, что **ИИ — это результат процесса социально-технического развития**, который ставит перед журналистикой и журналистской практикой ряд задач — и как **тема для освещения, и как совокупность методов, приёмов и подходов в рамках самой журналистики**. Журналистика должна доносить до граждан достоверную, объективную информацию, выполняя свою роль как общественного блага, работающего как источник познания. Это относится и к технологиям, и понятие «искусственный интеллект», вошедшее в повседневный обиход, многим кажется трудно уловимым. Термин «интеллект» может вызывать ассоциации и образы, далёкие от повседневности, в то время как, как это ни парадоксально, ИИ в значительной степени является глубоко практическим явлением, оказывающим всё большее влияние на нашу повседневную жизнь. Между пугающими антиутопиями и праздничной шумихой бывает трудно сформировать разумные ожидания.

Эксперты выделяют несколько задач, которые журналисты должны решать в публичной сфере (см., например, Hansen et al., 2017, p. 7). Говорят, что журналисты должны находить иголки в стогах сена. Они также должны уметь выявлять тенденции, а также отклонения от основных тенденций. Они должны посвятить себя изучению применения ИИ или вычислительных систем в качестве темы журналистского материала. Некоторые учёные даже утверждают, что журналисты должны активно участвовать в разработке новых технологий, интегрировать журналистский взгляд в такие структуры и сотрудничать с директивными органами, определяющими этические рамки будущего ИИ. Хотя многие могут не согласиться с активной ролью журналистов в отрасли, возможно, самая важная задача журналистики во всём мире — участвовать в осмыслении технологического феномена и активно соотносить его с различными контекстами, чтобы иметь возможность формировать, информационно насыщать, уравновешивать и актуализировать общественный дискурс. Что касается журналистского образования, то чем раньше преподаватели журналистики смогут участвовать в дискуссиях и принимать активное участие в формировании будущего журналистского мышления, тем больше шансов, что будущие журналисты будут обладать глубоким пониманием особенностей будущих обществ.

Если немного утрировать, то такого жанра, как «репортаж об искусственном интеллекте», не существует. Очень редко журналисты пишут о самой технологии, как о чистом изолированном явлении, и очень редко журналистам требуется знать о технических процессах нейронных сетей или систем глубокого обучения. Вместо этого журналисты чаще пишут о демократических процессах, поведении людей или, например, структурных факторах несправедливости — и во всё это интегрирован ИИ, и на всё это влияют технологии ИИ. В связи с возрастающей ролью ИИ в общественных процессах журналистам необходимо знать об основных этических нормах, дискуссиях об этических аспектах и правилах, связанных с ИИ, а также осознавать страхи и надежды, которые ИИ может вызывать в представлениях людей. Журналистам также необходимо осмыслить возможности ИИ в своей области деятельности и быть готовыми к тому, чтобы

ответственно подходить к использованию технологий и методов ИИ для сохранения устойчивых отношений со своей аудиторией и того, что лежит в основе журналистики, — **доверия**.

Прежде чем обсуждать ИИ как тему учебной программы, я хотела бы остановиться на двух основных наблюдениях, имеющих отношение к журналистскому мышлению и практике освещения ИИ. Во-первых, ИИ в определённой степени подвержен влиянию самой идеи **проекции будущего**, и при его освещении необходимо балансировать между крайностями апокалипсиса и торжества технологий. Во-вторых, ИИ — это зонтичный термин, и информацию о нём всегда приходится синтезировать из различных источников, представляющих разные области. Наконец, до сих пор нет единого мнения о том, следует ли, в каком объёме и, главное, каким образом включать такие темы, как ИИ, в официальные учебные программы по журналистике. Ясно лишь то, что столкновение с новым явлением, сопряжённым с **проблемами этического характера**, такими как **взаимодействие человека и компьютера и человекоподобное поведение машин**, подводит нас к самой сути журналистики: журналистика должна оставаться журналистикой, придерживаясь своих основных ценностей, которые отличают её от сфер выработки политических решений, лоббирования и пропаганды, маркетинга и влияния на умы. Далее я рассмотрю ценностно-ориентационные и междисциплинарные аспекты будущего, соответственно.

«Скоро наступит сингулярность, и машины захватят власть!»  
 «Разумные машины спасут мир!»  
 «Наша повседневная жизнь будет похожа на научную фантастику!»

**Тема ИИ требует сбалансированного освещения.**

Общественный дискурс об искусственном интеллекте, как и дискурс обо всех новых технологиях, формируется под влиянием неопределённости, связанной с реализацией прогнозов на будущее. Как отмечается в модуле 2 «Культурные мифы и представления об искусственном интеллекте», эта идея сопровождается появлением большого количества литературы, фильмов, телепередач и других культурных продуктов, изображающих «гонку роботов». Такой подход вызывает интерес к теме и повышает её общественную значимость, поскольку угрозы и тревоги всегда являются точкой входа в поле зрения общественности. В то же время гибриды, наподобие оборотней и вампиров, повсеместно завораживают человеческое сознание, вызывая волнение, наполненное стремлением к отождествлению и инаковости.

Однако за акцентированной новизной ИИ можно забыть о том, что эта тема не является новой. В 1987 году, когда Брайан Б. Блумфилд опубликовал свою фундаментальную книгу «*The Question of Artificial Intelligence*» («*Вопрос об искусственном интеллекте*»), на заре революции в информационных технологиях велись жаркие споры о разумных машинах.

! *Всё большее проникновение информационно-коммуникационных технологий (ИКТ) в жизнь общества привело к появлению концепции информационного общества (Hofkirchner & Burgin, 2017) и, далее, производных от неё, таких как виртуальное общество (Woolgar, 2003), общество, управляемое данными, или общество данных, и автоматизированное общество (Bloomfield, 1995).*

Ожидается, что люди и организации будут сталкиваться с огромными объёмами и всё более сложными структурами данных, что в конечном итоге приведёт к появлению самостоятельно работающих компьютерных систем, к которым и относится термин ИИ. Для достижения «общества ИИ» — по сути, учёные и политики часто используют термин **«общество хорошего ИИ»**, чтобы подчеркнуть использование ИИ ради социального блага (Cath et al., 2018), — журналисты должны стать посредниками в распространении знаний об этих процессах, в то время как от них самих всё чаще требуется умение анализировать и курировать данные, имея хотя бы базовое представление об информатике как необходимом инструменте для освещения событий.

Как и многие современные проблемы, с которыми журналистам приходится сталкиваться в своей работе, такие как изменение климата, планетарный кризис ресурсов и пандемии, ИИ представляет собой с точки

зрения журналистики, склонной к освещению конкретных явлений, которые могут быть чётко очерчены и идентифицированы, абстрактный конструкт с глобальными последствиями. ИИ вездесущ, но его трудно обозначить; он порождает разнообразные последствия, которые трудно предугадать. Для его глубокого понимания необходимы экспертные знания, а для полного осмысления часто требуется триангуляция, то есть интеграция и учёт различных точек зрения. И, наконец, сам феномен ИИ ставит вопрос о том, должна ли журналистика оставаться зеркалом, отражающим происходящие процессы и весь замысел в целом в ходе его становления, или же занять более интервенционистскую, направленную на поиск решений позицию, исследуя альтернативы и предлагая ответы, а не формулируя вопросы.

Кроме того, современная информационная повестка, касающаяся ИИ, проникает повсеместно и будоражит воображение, поскольку очень тесно связана с научной фантастикой и фэнтези, изображающими утопические горизонты будущего как потенциально катастрофические. Как теоретики, так и практики ИИ продолжают утверждать, что **идея о захвате машинами власти является чрезмерно доминирующей**. Во многих случаях более тонкие, приземлённые повседневные вопросы, связанные с машинным обучением и вездесущими вычислениями, отходят на второй план из-за увлечения фантазией о захвате человечества машинами, а алгоритмические культуры, формирующие нашу повседневную жизнь, отбрасываются как скучные и неинтересные. Журналистов не должны вводить в заблуждение научно-фантастические фантазии, хотя их понимание крайне важно. Журналисты должны быть более любознательными, задаваться вопросами о том, как интеллектуальные системы влияют на нас различными способами, и искать альтернативные точки зрения.

На момент написания данного пособия всё ещё остаётся много вопросов без ответов и много места для спекуляций, когда речь идёт о применении и возможностях, рисках, вреде и угрозах, связанных с технологиями ИИ. Это означает, что журналисты сталкиваются с особой ситуацией, когда им приходится сочетать различные области знаний, уметь справляться с факторами беспокойства и в то же время не вызывать напрасных ожиданий и надежд, ненужных страхов и антиутопических опасений. Данное пособие призвано подготовить преподавателей журналистики к работе в обществе, в котором мы живём и которое всё больше подвержено влиянию технологий. В частности, мы хотим ответить на следующие вопросы:

*Как должны обращаться с темой искусственного интеллекта журналисты, выступающие в роли «фильтров» информации и формирующие общественное мнение? На что следует обращать внимание при освещении тем, связанных с ИИ? Какие источники следует использовать?*

Преподавание темы искусственного интеллекта в рамках журналистского образования предполагает гибкую позицию по отношению как к освещаемому обществу, так и к журналистской работе. При освещении ИИ и, возможно, с помощью технологий ИИ журналистам необходимо задавать масштабные вопросы о прошлом, настоящем и будущем ИИ. То, как дело обстоит сейчас, может отличаться от того, как оно будет обстоять в будущем, и журналисты могут оказаться в том числе и в том положении, когда они могут влиять на это будущее. Как правило, в **обучении журналистике** есть стремление к достижению баланса между академической и отраслевой сферами (см., например, Zelizer, 2004); постоянно возникает вопрос о том, должна ли подготовка журналистов основываться **на статус-кво отрасли или на инновационных способах освещения событий**, обеспечивая тем самым выход за рамки текущего положения дел. При освещении тематики ИИ это становится особенно заметным.

В нашем понимании ИИ представляет собой междисциплинарную инфраструктуру, в которой необходимо выявлять экспертные источники информации в различных областях знаний. Научные исследования, способные обеспечить глубокое понимание этого явления, предполагают теоретические разработки, использующие весьма специфические концепции для отражения соответствующих аспектов. Технические же разработки, опирающиеся на инновационное применение технологий, не всегда легкодоступны. Повседневный опыт неспециалистов и обычных пользователей часто не опирается на знания об этих инфраструктурах и инфраструктурных разработках, поскольку пользователи не знают, как всё работает, а понимают лишь, что оно работает. Наши способы обсуждения феномена ИИ основаны на циркулирующих дискурсах. Задача журналистики — сбалансировать эти области знаний, соединить различные аспекты, чтобы получить представление о том, что волнует общество. Для этого журналист должен обладать знаниями в этих областях и уметь фильтровать и синтезировать информацию, которая иногда может быть противоречивой и сложной. Тем не менее, задача создания общественного дискурса и понимания явле-

ний, которые всё сильнее влияют на нашу жизнь, — это задача журналистов, служащих общественному благу. Для обучения этому нужны соответствующие методические подходы.

«Заберут ли машины наши рабочие места?»  
 «Можно ли предвидеть все последствия развития ИИ?»  
 «Когда идеи станут реальностью?»

### Журналистика по ИИ должна раскрывать и освещать факторы беспокойства.

ИИ представляет собой **междисциплинарную, или «трансдоменную»** (Zhang & Peréz Tornero, 2021), **область исследований**, в рамках которой необходимо отслеживать и объединять различные области знаний для создания правильного понимания технических, этических, экономических, социальных и культурных аспектов использования систем ИИ в жизни общества. Как область исследований, ИИ опирается на компьютерные науки, социальные науки, а также на многие междисциплинарные области научных исследований, такие как взаимодействие человека и компьютера (дизайн), изучение игр и политические исследования. Во всех этих областях и подобластях приоритет отдаётся различным аспектам, выдвигающим — например, с точки зрения затрат, — роль технологий на первое место по сравнению с изучением социальных последствий. Таким образом, освещение темы ИИ в публичном дискурсе в ясной, понятной, а иногда и дидактической форме — не всегда простая задача, и журналистам, освещающим подобные темы, необходимо подумать, как лучше это сделать. Многие понятия, такие как машинное обучение, Интернет вещей, робототехника, глубокое обучение, да и само понятие ИИ, могут быть не знакомы широкой аудитории. В силу междисциплинарного характера этой темы ИИ представляет собой **область общественного сотрудничества**. Даже научно-исследовательские и политические круги могут интерпретировать результаты функционирования процессов машинного обучения и других автоматических процессов противоречивым образом.

*Компьютерные системы, даже если о них говорится, что они «интеллектуальны» и способны имитировать человеческое поведение и перенимать человекоподобные черты, не опираются на человеческие ценности и этику и не способны принимать соответствующие решения, требующие соблюдения баланса идеологий и культур, и даже необходимости в переговорах и компромиссах, которые могут показаться нерациональными с точки зрения логики.*

Даже если компьютер говорит «нет», человек должен интерпретировать, контекстуализировать и проблематизировать ответ, и эта ответственность не должна перекладываться на «умные» неживые машины. В обширной умственной работе, связанной с такой контекстуализацией, журналисты должны выступать партнёрами других участников, определяющих будущие направления развития технологий ИИ.

Для журналистов пока не существует такого устоявшегося жанра или профессиональной ниши, как «журналистика ИИ». Журналистам любого жанра может потребоваться умение задавать вопросы о том, почему и каким образом важны прорывные технологии и как рассматривать их, **не впадая в технологический детерминизм**. Кому-то может показаться, что ИИ — это тема, которую освещают специализированные виды журналистики, такие как научная журналистика, медицинская журналистика или общественная журналистика. По мере того как общество становится всё более цифровизованным, автоматизированным, наполненным данными и медиатизированным, необходимость понимания сути стоящих за этим технологий становится всё более актуальной. Журналисты, начиная с тех, кто занимается политической журналистикой и заканчивая журналистикой искусства и культуры, освещающие события как международного, так и местного уровня, независимо от того, в какой среде они работают (газеты, журналы, онлайн-информационные агентства, радио, телевидение), должны понимать, как технологии, основанные на искусственном интеллекте, влияют на все стороны жизни общества и проникают в них.

Одним из решений на столь раннем этапе является приглашение специалистов в области ИИ к участию в занятиях по журналистике. В современной цифровой среде организовать видеолекции достаточно просто, а обеспечить международный состав докладчиков — задача не из лёгких. Сейчас информация широко доступна в Интернете, и, например, Европейский союз следит за быстро развивающейся нормативной практикой на международном уровне. Большую часть работы по переводу информации из научно-технической сферы в социальную и демократическую общественную сферу, которую порождает журналистика, ещё предстоит проделать преподавателям журналистики вместе со своими студентами. Поэтому можно предположить, что если бы данное пособие было написано или пересмотрено через десять лет, то его содержание было бы иным. Некоторые аспекты могут перерасти в более сложные комплексы вопросов, а другие — обрести большую точность.

В процессе выработки учебных программ по теме ИИ университеты должны сохранять независимость и, безусловно, сотрудничать с отраслевыми организациями, даже если они сами в значительной степени обладают самыми свежими представлениями и информацией о развитии технологий, а также учитывать, что такие учреждения, как правило, лоббируют свои собственные цели. Необходимо занимать критические позиции относительно того, каким должен быть ИИ и в каких направлениях он должен развиваться.

*«Как следует освещать использование технологий ИИ в журналистской работе?»  
«В какой степени аудитория должна быть осведомлена о процессах с использованием технологий ИИ в журналистике?»  
«Как смягчить нежелательные последствия освещения темы ИИ? Существуют ли неэтичные аспекты применения ИИ, которые не должны освещаться в журналистике?»*

**Журналисты также должны задумываться и проявлять прозрачность в отношении использования средств ИИ в своих репортажах**

## ⇒ Искусственный интеллект как предмет программы обучения

**Журналистика с использованием компьютерных технологий** является предметом дискуссий в журналистской среде и в сфере журналистского образования на протяжении многих десятилетий. Для журналистского образования ИИ — это относительно новая область знаний, которая обычно рассматривается в категории новых навыков и компетенций, наряду с онлайн-журналистикой, мобильной журналистикой и журналистикой данных. Несмотря на то что в компьютерных науках ИИ как предмет обсуждается и изучается уже несколько десятилетий, в журналистском образовании тема ИИ стала приобретать актуальность благодаря интеграции инструментов ИИ в работу новостных организаций. Вопрос об искусственном интеллекте связан с обсуждением новых методов, таких как автоматизированный сбор больших данных, и потребностью в новых профессиональных навыках, таких как программирование. Для разработчиков учебных программ это означает вопрос о том, какие области необходимо исключить из учебного плана, чтобы включить в него новые направления обучения.

Тем не менее тема ИИ вызывает озабоченность у многих заинтересованных сторон, и очевидно, что вопросы о роли журналистики в демократических обществах, управляемых данными, должны обсуждаться в упреждающем режиме. Крайне важно рассмотреть, в какой степени журналистика и учебные программы по журналистике должны быть подвержены влиянию научно-технического развития и политических решений. Традиционно отношения между журналистским образованием и отраслью носят двойственный характер: при всей важности отслеживания и отражения происходящих изменений и соответствия ком-

петенций, формируемых в рамках журналистского образования, внешнему миру, журналистское образование всегда должно быть на шаг впереди, даже в отношении систем, которые сами по себе не могут сформулировать цели, поскольку эти цели будут сформулированы людьми, которые создают эти системы.

Вопрос об **учебных программах по журналистике** включает в себя не только проблему того, чему и как учить, но и, по выражению Марка Дойзе (2006), **«структурирование, переосмысление и создание институтов, школ или факультетов журналистики»**. Вопрос о том, каким образом предмет ИИ будет окончательно интегрирован в учебные программы школ журналистики, будет решаться в ближайшие годы, и этот вопрос также в определённой степени зависит от будущего развития отрасли и нормативно-правовой базы. Подобные дебаты велись в области флэш-журналистики, документальной журналистики и программирования, и до сих пор ведутся, например, в области журналистики данных, журналистики погружения (или иммерсивной журналистики) и подкастов. Это «новые компетенции», о которых ведутся переговоры и которые в конечном итоге могут быть объединены с более традиционными методами и подходами, связанными с подачей информации и проверкой фактов.

Журналистам, как специалистам широкого профиля, освещающим политику и другие общественные проблемы, лежащие в основе демократии, не требуется глубоких знаний во всех областях жизни общества, которые они отслеживают. Некоторые области, которые считаются актуальными для общества и требуют более глубоких знаний при их освещении, становятся областями журналистской специализации. В значительной степени это зависит от **организационной структуры редакции**: создаются ли подразделения или группы, в которые входят специалисты. Научная журналистика — это сложившееся направление журналистики, в котором работают специалисты с углублёнными навыками и научными знаниями, прежде всего, в области естественных наук. Если одни направления специализации, такие как политическая, экономическая и культурная журналистика, давно и прочно занимают свои позиции в деятельности редакций, то другие, такие как научная, медицинская, потребительская и экологическая журналистика, позиционируются по-разному.

Данное пособие опирается на подход, основанный на убеждении, что журналисту общего профиля необходимо получить базовые знания из различных областей, связанных с ИИ: основные концепции, связанные с ИИ, типичные мифы и дискурсы, влияющие на общественные представления и, соответственно, надежды и опасения в отношении технологий ИИ, существующие нормативно-правовые рамки и этические рекомендации, а также передовой опыт, включая потенциальные «ловушки» и самоанализ того, что представляет собой ИИ в самой журналистике. Сильные стороны журналистов широкого профиля заключаются в их способности задавать правильные вопросы и излагать знания в доступной для широкой аудитории форме, а также преодолевать разрыв между специализированными областями жизни общества, такими как наука и промышленность. Эта сложная задача зиждется на способности журналистов утверждать себя в качестве авторитетов посредством стратегических действий риторического характера, создающих **толкательный авторитет**, который заключается не в том, что журналисты знают — они не являются «носителями» знаний, — а в том, как они подают свои знания (Zelizer, 1993). На рис. 1 показаны основные области исследования технологий ИИ: наука, начиная с информатики и гуманитарных наук, с чётким акцентом на различные подобласти информатики, и заканчивая прикладной наукой, а также промышленностью, образованием, политикой и жизненным миром пользователей. Журналистам необходимо приобрести навыки грамотности в этих областях, чтобы обращаться к ним как к источникам информации, овладев отдельным понятийно-терминологическим аппаратом, набором основных концепций и теорий, а также традициями ведения риторики и дискурса. Необходимым условием такой грамотности является определение основных действующих лиц и заинтересованных сторон, а также инфраструктур, и понимание того, как работают соответствующие культурные и политические институты. Например, применительно к науке журналистам необходимо знать, о каких научных дисциплинах идёт речь, как происходит производство научных знаний, как можно получить доступ к научным публикациям и как их интерпретировать, как работает наука в целом. Если речь идёт о политике, то журналистам необходимо определить центральные политические структуры, понять характер их работы и иметь возможность получать доступ к их программным документам и отслеживать их динамику. Что касается промышленности, то журналистам необходимо понимать коммерческую логику производства технологий ИИ и осознавать интересы производителей ИИ и соответствующих заинтересованных сторон, поддерживая с ними регулярный диалог. Подготовка студентов к освещению темы ИИ как области журналистской деятельности не имеет существенных отличий от других областей журналистики. Журналисты должны использовать свои обычные

стратегии, как и в случае с любой другой темой, включая такие журналистские процессы, как контекстуализация, проверка данных, анализ последствий и рисков, релятивизация, выявление нормативности и отклонений, занятие критической позиции по отношению к объектам исследования. Журналистика по теме ИИ должна быть информативной, объективной и нейтральной, а также независимой, автономной и представительной, и, как и любая другая форма журналистики, она должна оставаться ответственной, заслуживающей доверия и актуальной для своей аудитории.

*Журналистика и журналистское образование могут использовать существующий опыт и знания, полученные при освещении других развивающихся технологий, таких как нано- и биотехнологии, наука и другие сложные этические сферы жизни общества, начиная от сегрегации и заканчивая клонированием.*

Какие же компетенции необходимы для освещения сферы ИИ? На рис. 1 показаны области компетенций, которые имеют центральное значение для будущих журналистов и, таким образом, должны стать частью журналистского образования в интересах качественного освещения тематики ИИ и связанных с ним новых технологий. Сфера 1 — внутренняя зона круга — представляет собой жизненный мир человека, сфера 2 — мир журналистики, а сфера 3 — мир ИИ, которые соответственно включают в себя повседневные знания обывателя, профессиональные журналистские знания и экспертные знания, связанные с ИИ. Журналист должен опираться на все эти сферы при использовании своих компетенций, связанных с обществом, профессией, технологиями, языком и человеческим бытием.

**Рисунок 1.** Области журналистской компетенции, связанные с освещением тематики ИИ



Источник: оригинальный материал авторов

При освещении новых технологий, таких как искусственный интеллект, студент-журналист выстраивает отношения между этими основными областями через **призму повседневных, профессиональных и экспертных знаний**:

- ▶ **общество:** представления об автоматизированном и управляемом данными обществе в условиях современной демократии — опыт граждан в отношении ИИ, связь журналистики с обществом, научное знание об обществе;
- ▶ **профессия:** представления о журналистских идеалах, нормах, практике и процедурах, «журналистика ИИ» (ИИ как объект исследования) и «ИИ в журналистике» (использование ИИ в целях журналистской деятельности) — личный профессиональный опыт, знания о журналистском профессионализме, журналистские исследования;

- ▶ **технологии освещения событий:** понимание и владение технологиями, связанными с журналистским производством — владение инструментами подготовки репортажей, общей практикой редакционной работы, передовой практикой использования технологий;
- ▶ **язык:** смыслообразование, владение экспрессией, лексикой и навыками коммуникации с целевой аудиторией — владение техниками смыслообразования, общие модели смыслообразования в журналистике и редакционной работе, исследование средств подачи информации;
- ▶ **человеческая сущность:** онтология бытия человека в постчеловеческую эпоху, когда расширяется человеко-машинное взаимодействие и происходит объединение людей и компьютеров — личный опыт человека, онтология человечности, философия ИИ.

Ключевыми методическими аспектами при работе над освещением тематики ИИ для преподавателей журналистики являются общепринятые понятия и установки — можно говорить о контролируемом развитии или изменении культурного климата. Как относиться к новым и возникающим явлениям? Как отслеживать их последствия и, прежде всего, связанные с ними социокультурные проблемы, которые оказывают решающее влияние на демократические общества? Как контекстуализировать возникающие явления? Как избежать живописания антиутопических сценариев или превращения в защитника ангажированных интересов? С методически-практических позиций это требует ознакомления с различными аспектами технологий и надлежащего выполнения журналистской роли.

Прорывной характер технологий ИИ имеет своим следствием то, что, хотя ИИ трансформирует общество в целом, при обсуждении освещения тематики ИИ мы имеем дело не только с определённым типом журналистской практики. Ожидается, что ИИ изменит характер отношений между людьми и машинами в различных сферах жизни общества, включая журналистику, и этот фундаментальный сдвиг требует **переоценки наших эпистемологических и онтологических концепций гражданской ответственности, общества и взаимодействия человека и машины.**

Таким образом, задачи понимания предмета ИИ в журналистском образовании, которые ставятся в данном пособии, заключаются в следующем (см. Long & Magerko, 2020):

- **распознавание ИИ** как явления — понимание того, что это такое, и проведение различий между технологическими артефактами, использующими и не использующими ИИ;
- понимание концепции **машинного интеллекта** и определение различных типов ИИ;
- **междисциплинарный характер** ИИ как области деятельности и зонтичного термина, а также разнообразие технологий, обеспечивающих его применение, — от когнитивных систем до робототехники и машинного обучения;
- рассмотрение **потенциала технологий ИИ**, а также их сильных и слабых сторон;
- формирование представлений о будущем в реалиях ИИ путём понимания и преодоления **распространённых установок**, на основе которых ведутся общественные дискуссии;
- понимание **роли журналистики** в опосредовании и формировании дискурса об ИИ;
- поиск путей сбалансированного, реалистичного и ответственного освещения проблематики ИИ;
- увязка с существующими жанрами журналистики — от новостного репортажа до журналистики данных;
- расширение знаний об использовании ИИ в журналистской работе и для целей журналистики;
- развитие **профессионального самоанализа** в отношении философских, этических, социальных и культурных аспектов технологий ИИ и журналистики.

## ⇒ **Задачи настоящего пособия**

Данное пособие связано с предыдущими пособиями по журналистскому образованию и журналистской практике, опубликованными ЮНЕСКО (Fengler et al., 2021; Hunter, 2011; Keith & Cozma, 2022; Impe, 2019; Ireton & Posetti, 2018; Muratova et al., 2021). Серия публикаций ЮНЕСКО по журналистскому образованию призвана выполнять функцию методического источника в виде типовой учебной программы, учебного



пособия или практического руководства. Кроме того, ЮНЕСКО опубликовала руководства по технологиям ИИ для лиц, ответственных за формирование политики в области образования (Miao et al., 2021). В этих пособиях освещаются проблемы реализации мер политики и собирается глобальный передовой опыт для информационного обеспечения специалистов-практиков по всему миру. В соответствии с предыдущими публикациями данная работа называется справочным пособием, но, если быть точным, она не соответствует академической традиции справочного пособия, а представляет собой методическое пособие или типовую учебную программу, в которой излагаются методические принципы, связанные с относительно новым, но глубоко укоренившимся общественным явлением.

Цель пособия — развить теоретические представления преподавателей и студентов о данном явлении (ИИ) в увязке с ценностями журналистики, определить соответствующие методы и практику журналистской работы, а также затронуть этические проблемы, представляющие общественный интерес и актуальные для журналистов. В работе выделено семь изучаемых областей:

- ▶ **Технология:** как дать определение ИИ и как проследить его развитие?
- ▶ **Концепции и развитие:** как общественные концепции и представления влияют на наше нынешнее и будущее понимание сферы ИИ?
- ▶ **Нормативно-правовое регулирование:** как регулировать сферу ИИ на уровне отдельно взятых государств и на международном уровне?
- ▶ **Инфраструктура:** каковы центральные структуры для производства, регулирования, разработки и потребления технологий ИИ?
- ▶ **Повседневная практика:** как технологии ИИ встраиваются в социальное поведение и культурные привычки граждан и потребителей, а также обществ в целом?
- ▶ **Злоупотребления:** как технологии ИИ могут использоваться во вредоносных целях и как можно предвидеть и минимизировать риски такого использования?
- ▶ **Практика освещения:** каковы лучшие практики и потенциальные «ловушки» при освещении тематики ИИ в журналистских материалах?
- ▶ **Журналистика:** как технологии ИИ используются в журналистской практике и как они могут способствовать повышению уровня профессионализма?

Пособие не ставит своей целью дать исчерпывающий и окончательный ответ на вопрос об искусственном интеллекте, а скорее стремится помочь преподавателям журналистики в создании учебных планов и соответствующих методических подходов, которые помогут студентам-журналистам принять сбалансированную позицию в рассмотрении искусственного интеллекта как общественной темы среди многих других областей политического, технического и культурного журналистского освещения. Таким образом, данное пособие не должно фокусироваться только на текущей картине, на том положении дел, которое сложилось на момент написания книги, а должно быть нацелено на формирование более целостной точки зрения, которая выдержит испытание временем. Одной из главных задач является также демистификация темы, которая для многих может показаться слишком технической и футуристической, и формулирование журналистского контекста, позволяющего обрисовать явление, имеющее социокультурные последствия, выходящие далеко за рамки технических аспектов.

*! Важнейшим шагом является осознание того, что никакие технологии не являются нейтральными: в основе технических структур и принципов работы различных платформ всегда лежит ряд интересов.*

Кроме того, разработка и использование технологий сопряжены с эмоциями и отношением. К технологиям подходят на основе **ценностного выбора**, а их использование оценивают по-разному, отдавая предпочтение и приоритет одним и игнорируя и отвергая другие.

Данное руководство построено на следующих принципах: **разнообразие мнений** (в силу многосторонности этой области существуют различные точки зрения и множество вопросов, на которые нет ответов, что делает целесообразным подход к теме с разных точек зрения, предоставление возможности высказаться различным экспертам); разные **жанры журналистики** (хотя средства ИИ могут быть интегрированы во все виды журналистской работы и их использование остаётся интересным и актуальным вопросом для всех жанров, а на сегодняшний день технологии ИИ в большей степени применяются в некоторых специ-

фических жанрах журналистики, таких как политическая, научная, техническая и спортивная журналистика, в то время как другие жанры, такие как художественная и культурная журналистика, ещё требуют освоения с точки зрения практического применения технологий ИИ в них; разные жанры подразумевают и разную практику и традиции освещения событий, а также разную профессиональную самобытность, по-разному подчёркивающую профессиональные ценности, поэтому развитие ИИ в журналистике должно учитывать жанровые различия журналистского поля); и **научно обоснованные данные** (развитие должно основываться на последних достижениях науки, в сочетании с установленными этическими рамками журналистской деятельности; журналисты должны непосредственно следить за последними и быстро развивающимися областями исследований, а также следить за теми областями, которые изучают саму журналистику, такими как исследования в области журналистики и исследования в области вычислительной журналистики).

## ⇒ Структура настоящего пособия

Пособие состоит из модулей-глав, посвящённых темам технологий, нормативно-правового регулирования, инфраструктуры, повседневной практики, злоупотреблений и журналистики. Главы имеют следующую структуру разделов:

- *Краткое содержание* — краткое изложение темы каждой главы.
- *Вопросы для размышления* — подготовительные вопросы для обдумывания студентами перед лекцией и небольшие задания, которые, как правило, представляют собой несложные задачи по наблюдению за окружением. Цель подготовительных заданий — стимулировать собственные размышления учащихся по теме и помочь им выработать личное отношение к теме, лучше овладеть ею и почувствовать причастность к ней. Вопросы и задания могут выполняться как индивидуально, так и в малых группах. Преподаватель может выбрать подготовительные задания для обучающихся, или предложить обучающимся самим выбрать предпочтительные или наиболее подходящие задания для введения в тему. Подготовительные задания хорошо подходят для выполнения дома перед лекцией или семинаром, а также в аудитории в начале общего занятия.
- *Цели модуля* — методические задачи, решаемые в каждой главе.
- *Результаты обучения* — цели занятия, которые должен достичь студент, ознакомившись с содержанием главы и выполнив задания, выбранные преподавателем.
- *Конспект* — основная часть главы, в которой излагается её ключевое содержание. Эта часть написана в соавторстве с экспертом по данной теме.
- *Предлагаемые задания* — виды учебной деятельности, которые могут быть организованы в различных условиях. Каждый раздел заданий содержит задания для выполнения индивидуально, в малых группах и всей группой. Предлагаются как задания, предназначенные для лекционных занятий (задания с теоретическим подходом), так и для учебных редакций для журналистской практики (задания с практическим подходом или с упором на создание журналистских материалов).
- *Дополнительная литература* — список литературы на английском языке по теме главы. В список включаются как научные книги и журнальные статьи, так и отраслевые доклады. Предпочтение отдавалось вариантам с открытым доступом, но, к сожалению, не всегда удавалось найти такую альтернативу. Здесь стоит отметить, что в такой быстро развивающейся области, как новые и перспективные технологии, исследования быстро устаревают. Поэтому следует поощрять студентов к самостоятельному поиску литературы, а преподавателям — дополнять список новейшими исследованиями. Ключевые слова, приведённые в конце краткого содержания главы, можно использовать в качестве поисковых терминов в базах данных учебной литературы (например, Google Scholar).

Если говорить более конкретно, то задания каждой главы составлены таким образом, чтобы вписаться в многообразие методических подходов в журналистском образовании, и включают следующие категории, обозначенные соответствующими символами:



**ПРОЧИТАТЬ И ПОРАЗМЫШЛЯТЬ.** Теоретическое задание, предполагающее чтение научной статьи и её краткий анализ. Цель этих заданий — познакомить студентов с практикой получения информации из материалов научных исследований.



**ОБСУДИТЬ.** Задание оптимальное для малых групп, как правило, связанное с конкретным случаем, направленное на формулирование мнений и взглядов, а также аргументов по теме. Цель таких заданий — сформировать мнение и стимулировать обсуждение по конкретной теме.



**ВЫРАБОТАТЬ ИДЕЮ.** Творческое задание, которое можно выполнять как индивидуально, так и в диалоге с коллегами, чтобы найти журналистский угол зрения на тему по одной концепции, которую необходимо подвергнуть всестороннему разбору. Цель этих заданий — развитие у студентов навыков журналистского мышления в рамках той или иной темы.



**ПОИСКАТЬ И ИССЛЕДОВАТЬ ИСТОЧНИКИ.** Задание на интервью, в рамках которого студенту необходимо получить информацию из стороннего источника. Цель таких заданий — найти подходящие методы и терминологию для сбора информации.



**СОЗДАТЬ.** Журналистское задание, направленное на подготовку журналистской презентации на заданную тему. Цель этих заданий — дать студенту возможность узнать что-то новое, познакомившись с темой через журналистскую работу.

Преподаватели сами выбирают типы заданий в зависимости от условий классной работы. Следует отметить, что задания всегда могут быть изменены или адаптированы к виртуальной среде. Иногда целесообразно использование цифровых инструментов.



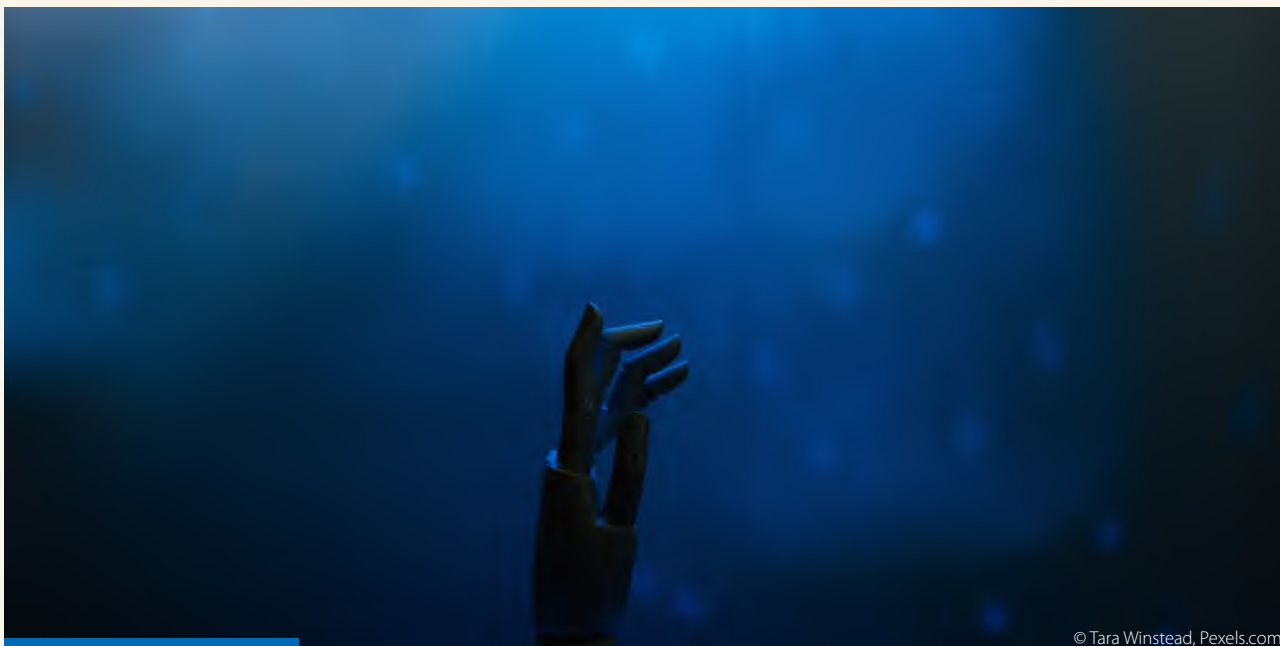
## Благодарности

Основные части каждого модуля — «конспекты» — писали приглашённые авторы. В остальном методические указания написаны редактором на основе опыта обучения журналистике в разных странах.

Собрать воедино знания и идеи из разных областей знаний — дело коллективное, и в составлении этого пособия мне помогли разные люди.

В процессе создания данного пособия я сотрудничала и консультировалась с учёными и практиками, занимающимися технологическими аспектами и журналистикой. Благодарю за отдельный вклад в создание этого пособия профессора Теему Рооса из Университета Хельсинки (Финляндия), почётного профессора Стива Вулгара из Линчёпингского университета (Швеция), профессора Дженни Бергенмар из Гётеборгского университета (Швеция), профессора Бернхарда Й. Дотцлера из Регенсбургского университета (Германия), Маттиаса Шпилькампа и Николаса Кайзер-Бриля из исследовательской группы AlgorithmWatch (Германия), а также профессора Аммину Котари из Род-Айлендского университета (США).

Отдельно я хочу поблагодарить ЮНЕСКО и Всемирный совет по образованию в области журналистики (WJES) за инициативу создания этого пособия. Кроме того, я хочу поблагодарить Гая Бергера, директора Управления ЮНЕСКО по вопросам свободы слова и развития СМИ, и Вериду Рупар, председателя WJES, за их широкую поддержку и ценные замечания на протяжении всего пути. Я также хочу поблагодарить всех корректоров этой книги, включая Кристин Клэй и Тимо Лийра, а также Веронику Инес Дель Карриль за руководство проектом.



© Tara Winstead, Pexels.com

## Модуль 1:

# Определение понятия искусственного интеллекта



## Краткое содержание

В данном модуле даётся определение понятия искусственного интеллекта (ИИ) как междисциплинарного зонтичного термина. Модуль знакомит читателя с некоторыми центральными понятиями, важными для понимания современных технологий ИИ в журналистике и вне её, и освещает конкретные примеры применения технологий ИИ на практике. В рамках модуля кратко описываются истоки и развитие ИИ, а также содержится призыв к взвешенному пониманию взаимодействия человеческого и искусственного интеллекта в современном и будущем обществе.

**Ключевые понятия:** прорывные технологии, прикладные технологии ИИ, общие технологии ИИ, машинное обучение, глубокое обучение



## Цели модуля

Цели данного модуля:

- ▶ выработка общего понимания понятия искусственного интеллекта;
- ▶ изучение основных концепций, необходимых для понимания предмета ИИ;
- ▶ обсуждение актуальности и последствий использования технологий ИИ для общественного дискурса, журналистского освещения и общества в целом.



## Результаты обучения

Ожидается, что по завершении модуля студент научится:

- ▶ давать определение и ориентироваться в различных определениях ИИ и технологий ИИ;
- ▶ знать основные контексты обращения к тематике ИИ в журналистике;
- ▶ выявлять проявления ИИ в технологически насыщенной среде, в которой мы живём, и иметь представление о том, как подходить к ним с позиций журналистики.



## Вопросы для размышления

Несколько вопросов, над которыми можно поразмыслить заранее:

1. Что приходит вам в голову, когда вы слышите термин «искусственный интеллект» (ИИ)? Перечислите свои ассоциации и сравните их с мыслями коллег. Возникли ли у вас схожие идеи? Как эти идеи могут быть отражены в доминирующих общественных дискурсах об ИИ?
2. Представьте себе научно-техническое развитие будущих трёх десятилетий в следующих средах (как вариант, выберите только одну из них): дом/семья, школа, здравоохранение. Какие процессы будут автоматизированы? Как автоматизация повлияет на поведение, социальное взаимодействие и опыт людей?
3. Чьи мнения слышны (или не слышны) в основных дебатах вокруг технологий ИИ? Как вы думаете, почему дело обстоит именно так и как это может повлиять на наше понимание того, чем является и чем может быть ИИ?

Несколько подготовительных заданий, которые необходимо выполнить перед тем, как приступить к изучению темы:

1. Поговорите с кем-то в вашем окружении, другом или членом семьи, и поинтересуйтесь его представлениями об ИИ. Какие открытые вопросы и сомнения возникают у вас, разделяете ли вы эти опасения?
2. Прочитайте ежедневный выпуск новостей в газете или другом информационном издании и выберите статьи, посвящённые технологиям. Опишите стиль изложения: какие закономерности прослеживаются?
3. Найдите в Интернете конкретные примеры применения технологий ИИ в перечисленных ниже областях.

<p><b>Предотвращение и урегулирование кризисов:</b> прогнозирование вспышек заболеваний, миграционных кризисов, природных и техногенных катастроф, поисково-спасательные работы</p>	<p><b>Расширение экономических возможностей:</b> качество сельскохозяйственной продукции, инициативы экономического роста и вовлечения, предложение рабочей силы</p>	<p><b>Образование:</b> доступность образования и его завершение, повышение успеваемости учащихся, производительность труда преподавателей и администрации</p>
<p><b>Окружающая среда:</b> охрана животных, растений, земли, воздуха и воды, борьба с изменением климата и адаптация к нему, энергоэффективность и экологическая устойчивость</p>	<p><b>Социальное равенство:</b> доступность, борьба с дискриминацией, интеграция маргинализированных групп, выявление предвзятости ради повышения разнообразия</p>	<p><b>Здравоохранение:</b> улучшение качества лечения и медицинского обслуживания, психическое благополучие</p>
<p><b>Информационный обмен:</b> выявление дезинформации, борьба с поляризацией общества</p>	<p><b>Управление инфраструктурой:</b> энергетика, инвестиции в недвижимость, транспорт, городское планирование, управление водными ресурсами и отходами</p>	<p><b>Безопасность, правосудие и мир:</b> недопущение нанесения вреда, помощь полиции и юристов</p>

Подробнее: Fosso Wamba et al., 2021; Chui et al., 2018



## Конспект

Автор: Агнес Стенбом

Искусственный интеллект (ИИ) представляет собой совокупность инструментов и технологий, которые изменяют характер деятельности и результаты в различных областях и секторах жизнедеятельности человека — здравоохранения, транспортных перевозок, сельского хозяйства и искусства и т. д. Применение ИИ всё больше влияет на повседневную жизнь людей во всём мире, в том числе на способы работы, организации и взаимодействия между людьми. Это растущая область научных исследований и внимания директивных органов. Но что такое ИИ?

Термин «искусственный интеллект», вероятно, был придуман в Дартмутском колледже летом 1956 г., когда математики, специалисты по вычислительным системам и нейробиологи собрались на семинар, где проверялась гипотеза о том, что «всякий аспект обучения или любое другое свойство интеллекта может в принципе быть столь точно описано, что машина сможет его симулировать» (McCarthy et al., 2006).

В настоящее время этот термин не обозначает какой-то один предмет или технологию, а должен пониматься как зонтичный термин, охватывающий различные области знаний и технологии. Это зонтичное понятие, ИИ, не имеет универсального определения, а описывается разными людьми и в разных дисциплинах по-разному. Приведённые ниже примеры показывают, как могут различаться эти определения:

- В Рекомендации ЮНЕСКО об этических аспектах искусственного интеллекта (2021), одобренной 193 государствами-членами, отмечается, что «системы на основе ИИ представляют собой технологии обработки информации, которые включают модели и алгоритмы, обеспечивающие способность обучения и выполнения когнитивных задач, с получением результатов в виде прогнозной оценки и принятия решения в материальной и виртуальной среде. ИИ-системы предназначены для работы с той или иной долей автономности посредством моделирования и представления знаний, а также использования данных и расчёта корреляционных зависимостей» (UNESCO, 2021). Стратегия Евросоюза в области искусственного интеллекта гласит: «[ИИ] относится к системам, которые демонстрируют интеллектуальное поведение, анализируя окружающую среду и предпринимая действия — с некоторой степенью автономности — для достижения конкретных целей» (European Commission, 2018).

## Определение понятия искусственного интеллекта

- В контексте журналистики ИИ описывается как «совокупность идей, технологий и методов, относящихся к способности компьютерной системы выполнять задачи, обычно требующие человеческого интеллекта» (Brennen et al., 2018).

Технологии искусственного интеллекта считаются **прорывными**, поскольку они способны изменить устоявшуюся практику и привести к преобразованиям. В отличие от поддерживающих технологий, прорывные технологии приводят к появлению новых способов ведения дел или даже позволяют осуществлять новые виды деятельности. Это не означает, что каждая инновация в области ИИ направлена на разрушение старых методов работы, скорее, эти технологии способны существенно изменить продукты и практику деятельности.

## ИИ узкого и общего назначения

В рамках более широкой формулировки понятия ИИ можно выделить несколько уникальных технологий и идей. Во-первых, важно различать следующие понятия:

- прикладной ИИ (также известный как ИИ узкого назначения), который представляет собой компьютерные системы, обладающие специфическим интеллектом, позволяющим им выполнять специализированные задачи;
- общий ИИ (также известный как искусственный интеллект общего назначения, ИИОН), представляющий собой компьютерные системы, которые обладают или должны обладать *общим* интеллектом, подобно человеку. ИИОН также связан с концепцией сверхразума, которая часто обсуждается как философами, так и писателями-фантастами.

Хотя в культуре существует множество иллюстраций общего ИИ, включая такие блокбастеры, как «Матрица», современные практические приложения ИИ действительно являются «прикладными». Тем не менее важно помнить, что для некоторых этот термин *может* представлять собой идеал интеллектуальных систем общего назначения.

Примечательно, что большинство современных сценариев использования технологий ИИ требуют применения *гибридных* подходов, в которых человек (помимо создания системы ИИ) необходим в качестве оператора и/или оценщика. Такие гибридные системы используют возможности алгоритмизации, такие как скорость и масштабирование, и опираются на возможности человека для выполнения поставленной задачи (Brynjolfsson & McAfee, 2014).

На восприятие людьми тех или иных определений ИИ влияют такие факторы, как образование, предыдущий технический опыт и даже личные убеждения. Лично я отношусь к ИИ как к комплексному термину, обозначающему компьютерные системы, которые стремятся имитировать человеческий интеллект, а иногда и превзойти его практические ограничения. Я привожу это определение, чтобы подчеркнуть точку зрения, с которой написаны следующие разделы, и побудить читателей обращаться за разъяснениями, когда другие используют этот термин.

## Искусственный — не значит автономный

Употребляя термин «*искусственный интеллект*», особенно когда речь идёт о прикладном ИИ, мы должны помнить, что ИИ — это не автономная сила или концепция, которая волшебным образом появится среди нас. ИИ, так же как и рукотворные инструменты, питается от физических ресурсов и управляется нашим выбором. Системы ИИ являются тем, чем мы их делаем, и для их функционирования необходим наш физический мир — не в последнюю очередь это касается огромного количества энергии, которую эти системы потребляют. Следовательно, мы должны помнить, что за технологиями стоят люди и что в функционировании цифрового мира также задействованы природные ресурсы.

В 1990-х годах Герберт Саймон (Herbert A. Simon, 1995) отметил, что термин «*искусственный*» имеет негативный оттенок. И хотя отношение к общей концепции искусственно созданного мира с тех пор, возможно, изменилось, общество по-прежнему беспокоит то, к чему могут привести современные искусственные системы. Следует помнить, что искусственный — не значит автономный; искусственная часть создаётся нами — людьми.

Идея *интеллекта* занимает центральное место в дискурсе об ИИ, но, как и ИИ, понятие «интеллект» не имеет универсального определения. Что такое интеллектуальная система? Когда инструмент следует считать «умным»? Использование термина *«интеллект»* может вызывать ожидания, что неживые системы будут похожи на людей, что не характерно для большинства современных прикладных ИИ. Влиятельные исследователи в области ИИ Стюарт Дж. Рассел и Питер Норвиг (2010) заявили, что предпочитают термин *«рациональный»*, а не *«интеллектуальный»*, указывая на то, что интерес представляет не человекоподобность, а оптимизация результатов.

ИИ не является ни хорошим, ни плохим, но он также не нейтрален. Человеческие ценности определяют развитие систем ИИ, например, путём постановки задач и создания обучающих данных.

## Краткий взгляд на историю технологий искусственного интеллекта

Люди давно представляли себе машины, обладающие человеческими возможностями, в том числе в литературных образах Древней Греции: в *«Илиаде»* Гомера описаны «автоматоны» («живые статуи»), самодвижущиеся кресла и другие полезные машины. Примерно в 1495 году, за четыре столетия до создания первого компьютера, Леонардо да Винчи нарисовал своего «робота-рыцаря» — человекоподобное механическое существо, способное сидеть, махать руками, двигать головой и челюстью.

Если отвлечься от «робота-рыцаря» да Винчи, то рождение того, что мы сегодня называем искусственным интеллектом, часто приписывают тесту Тьюринга, проведённому в 1950 году. Тест (который первоначально назывался «имитационной игрой») был предложен британским математиком Аланом Тьюрингом и представляет собой метод определения наличия у машины интеллекта. В ходе теста люди наблюдают за взаимодействием человека и машины, и если они не могут отличить их друг от друга (т. е. идентифицировать машину как машину), то считается, что машина прошла тест. Таким образом, практическое возникновение и изучение машинного интеллекта началось ещё до упомянутого выше Дартмутского семинара.

Со времени проведения первого теста Тьюринга в области ИИ появилось множество разнообразных инноваций — от создания Джозефом Вейценбаумом из Массачусетского технологического института одного из первых «болтунов» (виртуальных собеседников, ныне известных как чат-боты) «ELIZA» в 1966 году до победы экспертной системы Deep Blue компании IBM над тогдашним чемпионом мира Гарри Каспаровым в долгожданном матче 1997 года.

Однако история успеха технологий ИИ не была линейной. Были периоды, когда темпы развития замедлялись, и в этой области царил пессимизм. Термин *«Зима искусственного интеллекта»* используется для описания периодов, когда финансирование исследований сокращалось, а коммерческие инвестиции были скромными, что фактически представляет собой цепную реакцию на общее чувство пессимизма в сообществе ИИ.

## Основные подобласти сферы ИИ

ИИ можно рассматривать как самостоятельную область как исследовательской, так и практической деятельности, однако сфера ИИ фундаментально взаимосвязана с информатикой и наукой о данных, а также с другими областями. В двух словах эту связь можно описать так: информатика и наука о данных представляют собой комплексные области знаний, на которые опираются различные подобласти сферы технологий ИИ в создании ценностей посредством использования данных и вычислительных методов.

Кроме того, сфера ИИ часто рассматривается как область, взаимосвязанная с робототехникой. ИИ и робототехника, хотя иногда и рассматриваются как синонимы, могут быть, но не обязательно связаны между собой. Роботы не всегда используют технологии ИИ, а технологии ИИ не всегда нуждаются в физической форме (более того, чаще всего не имеют её). Однако ИИ в промышленности тесно связан с робототехникой, где сочетание этих двух областей помогает, например, в прогнозировании технического обслуживания, оптимизации цепочек поставок и обеспечении качества при производстве продукции различных размеров и типов — от автомобилей до компьютеров.



## Определение понятия искусственного интеллекта

Ниже я описываю две наиболее фундаментальные области, входящие в широкую концепцию ИИ.

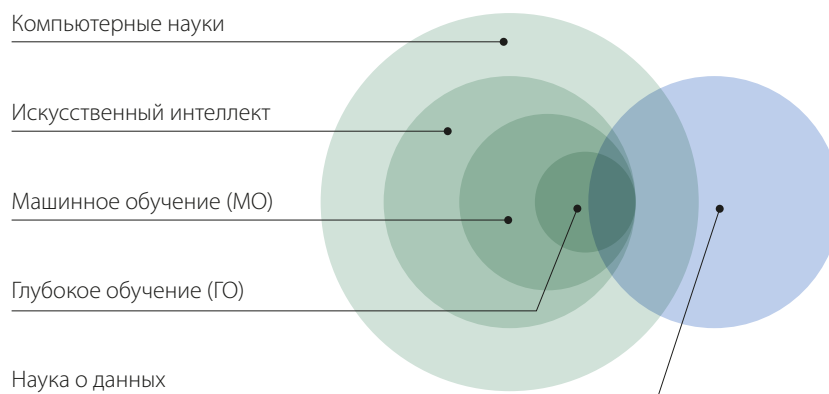
- **Машинное обучение**

Хотя термин «машинное обучение» (МО) иногда используется как взаимозаменяемый с термином «ИИ», его следует понимать как подобласть сферы ИИ. МО относится к алгоритмам, которые обучаются на примерах, а не кодируются с помощью явных инструкций. МО можно понимать как рецепты, полученные на основе данных; это подход к принятию повторяющихся решений путём алгоритмического выявления закономерностей в исторических данных и их использования для получения правильного (или желаемого) решения при поступлении новых данных.

- **Глубокое обучение**

Глубокое обучение (ГО) — это семейство методов МО, основанных на многослойных нейронных сетях, отсюда и термин «глубокое». Нейронные сети — это вычислительные системы, созданные для имитации человеческого мозга и составляющих его биологических нейронных сетей. Подходы ГО позволяют системам обучаться на основе исходных данных.

### Рисунок 2. Основные подобласти сферы ИИ



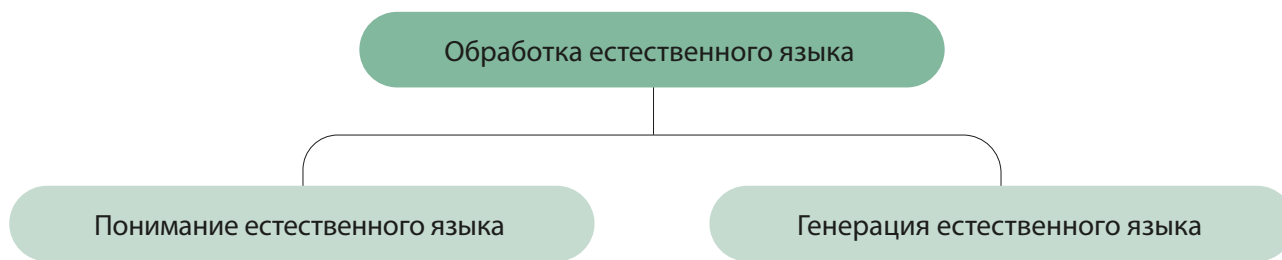
Источник: оригинальный материал авторов.

Эти концепции могут использоваться множеством способов для поиска различных возможностей создания ценности в самых разных областях и отраслях. Так, применение ИИ в механике или производстве существенно отличается от его применения, например, в журналистике. Примечательно, что эти термины, скорее всего, будут по-разному интерпретироваться и применяться в зависимости от сферы деятельности человека, а их определения (и значение) меняются с течением времени.

В следующих параграфах я описываю три области ИИ, которые должны представлять особый интерес для студентов и практикующих журналистов, а также привожу краткие примеры их применения в журналистике.

## Обработка естественного языка

Технология обработки естественного языка — сокращённо ОЕЯ — представляет собой различные подходы, направленные на анализ и генерацию текстов на *естественном языке*. Естественный язык представляет собой язык, используемый людьми для общения в письменной и устной формах. Это языки, сформировавшиеся в процессе использования и повторения, и поэтому они отличаются от конструктивных языков, например, языков, используемых для программирования компьютеров. В области ОЕЯ выделяют два основных направления — понимание естественного языка и генерация естественного языка. Как следует из терминов, эти области предлагают различные подходы к пониманию и генерации сообщений на естественном языке.

**Рисунок 3.** Иерархия методов обработки естественного языка (ОЕЯ)

Источник: оригинальный материал авторов.

Технологии ОЕЯ помогают обрабатывать огромные объёмы текстов, переводить с одного языка на другой и генерировать новые тексты. Повседневные примеры использования можно найти в программах перевода или проверки орфографии (как, например, та, которая использовалась при написании этого текста!), фильтрах спама в электронной почте, чат-ботах для обслуживания клиентов или виртуальных помощниках, таких как Amazon Alexa, Apple Siri, Microsoft Cortana и Google Assistant.

### ОЕЯ в журналистике

Ярким примером использования технологий ОЕЯ в журналистике является глобальное совместное расследование Международного консорциума журналистов-расследователей (англ. International Consortium of Investigative Journalists, ICIJ) и 36-ти СМИ-партнёров, в ходе которого выяснилось, как Изабель душ Сантуш, самая богатая женщина Африки и дочь бывшего президента Анголы, выводила сотни миллионов долларов государственных средств из экономики одной из беднейших стран планеты. В основу расследования «Луандские утечки» легли 715 000 документов, предоставленных ICIJ Платформой по защите осведомителей в Африке. Документы включали в себя электронные письма, внутренние служебные записки, контракты, отчёты консультантов, налоговые декларации, результаты частных аудиторских проверок и видеозаписи деловых встреч за десятки лет. Для их обработки ICIJ сотрудничал со студией искусственного интеллекта Quartz, которая создала систему, способную «прочитать» все документы. По словам представителей Quartz, цель была в том, чтобы «помочь журналистам из Quartz, ICIJ и других партнёрских организаций найти в массиве «утёкших» документов искомые, независимо от формата файла, орфографии, ошибок транскрипции, формулировок и даже языка документа» (Merrill, 2020).

### Компьютерное зрение

Если ОЕЯ понимается как технология, связанная с деятельностью человека, относящейся к языку (например, письмо, слушание или чтение), то компьютерное зрение следует понимать как технологию, связанную со зрением. В технологиях компьютерного зрения компьютерные системы обучаются анализу и классификации/сортировке визуальных материалов.

Существует подобласть подобласти (ИИ — это действительно зонтичный термин!) которая посвящена компьютерному зрению в применении к распознаванию человеческих лиц. Риски использования распознавания лиц в неблагоприятных целях и для контроля, в том числе в авторитарных режимах, вполне реальны и актуальны. Тем не менее, технология распознавания лиц ежедневно добровольно используется миллионами людей, которые с помощью собственных лиц выполняют разблокировку своих устройств, например, смартфонов.

Другой повседневный сценарий применения компьютерного зрения — это использование его в большинстве смартфонов для классификации изображений в зависимости от их содержания. Если люди набирают в своих библиотеках изображений запрос «кошка» или «ребёнок», они видят работу этой технологии на практике.

### Компьютерное зрение в журналистике

Компьютерное зрение может применяться в журналистике различными способами. Яркий пример — классификация и индексирование изображений в медиаархивах, где оно помогает структурировать большие базы данных. Один из примеров — газета The New York Times, которая использовала компьютерное зрение для создания цифрового архива печатных изданий, начиная с 1851 года. Распределив различные компоненты печатных статей по категориям, они собрали их в структуру, имитирующую современные цифровые статьи (Nordahl et al., 2021).

Как и ОЕЯ, компьютерное зрение может использоваться и для проведения журналистских расследований. В Аргентине газета La Nación использовала его для анализа избирательных бюллетеней, чтобы подтвердить правильность избирательного процесса и избежать фальсификаций. С помощью компьютерного зрения они проанализировали бумажные бюллетени для голосования, а затем проверили результат с помощью человеческой оценки через краудсорсинг. В 2015 году издательство реализовало аналогичный проект, в рамках которого вручную было проанализировано 16 000 бюллетеней. Благодаря использованию компьютерного зрения в издании 2021 года они смогли проанализировать 100 000 бюллетеней (Coelho et al., 2021).

### Рекомендательные системы

Ещё одним важным понятием в области ИИ являются рекомендательные системы. На самом фундаментальном уровне этот термин определяется как «описание любой системы, которая выдаёт на выходе индивидуализированные рекомендации или позволяет персонализированным образом направлять пользователя к интересным или полезным объектам в большом пространстве возможных вариантов. Такие системы имеют очевидную привлекательность в условиях, когда объем онлайн-информации значительно превышает возможности человека по её изучению» (Burke, 2002, p. 331).

Современные рекомендательные системы, как правило, используют средства машинного обучения для выработки предложений по информационным материалам, продуктам или деятельности различного типа. Известным примером является рекомендательный алгоритм Amazon, который предлагает пользователям товары, основываясь, например, на их предыдущих покупках и поведении в Интернете.

В журналистике рекомендательные системы всё чаще используются для персонализации и курирования новостных лент. Дизайн таких рекомендательных систем различен: общественные и коммерческие СМИ оптимизируют свои рекомендации для различных целей, таких как распространение новостей, представляющих общественный интерес, привлечение читателей в число платных подписчиков или максимизация времени, проведённого на новостных сайтах.

Рекомендательные системы также являются основой современных социальных сетей, поскольку они помогают создавать актуальные и увлекательные цифровые сервисы, объединяющие людей по всему миру. Многие СМИ отмечают отрицательные стороны этих систем, и критика небезосновательна. Facebook, YouTube и Twitter оказались в центре широких общественных дискуссий о распространении с помощью искусственного интеллекта языка вражды, угроз и дезинформации, причём рекомендательные системы обвинялись в том, что они направляют пользователей в метафорические «кроличьи норы» радикализации.

### Заключительные замечания

Как отмечалось во введении, ИИ — это не что-то одно, а многообразный зонтичный термин и набор инструментов, предоставляющих различные возможности для создания ценности в различных условиях и областях. Цель данного краткого обзора — дать читателям базовое представление об этом многообразии и о нередко сопровождающих его идеологических дискуссиях.

Технологии ИИ открывают широкие возможности, но в то же время создают новые серьёзные социальные и технические проблемы. Как уже говорилось, технологии ИИ могут носить прорывной характер и, следовательно, приводить к значительным изменениям как в процессах, так и в жизни людей. Освещение этих изменений и их социотехнических последствий является важной задачей современной журналистики.

### Определения понятия «искусственный интеллект»

«Под искусственным интеллектом понимаются машины или компьютерные системы, способные обучаться выполнению задач, для которых обычно требуется человеческий интеллект» (Bawack et al., 2019).

«Область искусственного интеллекта стремится к пониманию и созданию интеллектуальных существей» (Russell and Norvig, 2010).

«Способность системы правильно интерпретировать внешние данные, обучаться на их основе и использовать полученные знания для достижения конкретных целей и решения задач путём гибкой адаптации» (Haenlein and Kaplan 2019).

«Системы ИИ — это разработанные человеком программные (а возможно, и аппаратные) системы, которые, имея сложную цель, действуют в физическом или цифровом пространстве, воспринимая окружающую среду путём сбора данных, интерпретируя собранные структурированные или неструктурированные данные, рассуждая на основе имеющихся знаний или обрабатывая информацию, полученную из этих данных, и принимая решение о наилучшем действии (действиях) для достижения поставленной цели. Системы ИИ могут либо использовать символичные правила, либо обучаться на числовой модели, а также адаптировать своё поведение, анализируя то, как на окружающую среду влияют их предыдущие действия» (Samoili et al., 2020).

«Искусственный интеллект (ИИ) — это любая технология, позволяющая машинам работать, имитируя человеческие способности воспринимать, понимать и действовать. ИИ — это интеллект, демонстрируемый машинами, в отличие от естественного интеллекта, проявляемого людьми и животными, который включает в себя сознание и эмоциональность. ... ИИ — это технология общего назначения, которая может повлиять на всю экономику, как паровая машина, электричество или компьютерная сеть Интернет» — проект «ИИ для Африки» (Sedola, 2021).



### Предлагаемые задания



**ПРОЧИТАТЬ И ПОРАЗМЫШЛЯТЬ.** Зайдите на ведущую новостную ленту вашей страны, введите поисковый запрос «искусственный интеллект» и выберите одну интересующую вас статью. Проанализируйте статью в соответствии с приведёнными ниже категориями. Поразмышляйте, какое представление об искусственном интеллекте даёт статья аудитории и какие последствия может иметь. Насколько в статье удалось дать ясную и понятную картину сферы искусственного интеллекта?

Концепции ИИ	Применение	Терминология	Стиль	Источники	Потенциальные последствия
Как определяется понятие ИИ? О каком виде ИИ идёт речь?	Какой области применения посвящена статья?	Какие выражения и обороты речи используются?	Для какой аудитории предназначена эта статья?	Каковы источники? Чего не хватает?	Как потенциально может быть интерпретирована статья? Какие последствия она может иметь?



**ОБСУДИТЬ.** Какие, на ваш взгляд, существуют общие представления и заблуждения у людей об искусственном интеллекте? Перечислите свои идеи и поделитесь ими с коллегой. Обсудите, откуда берутся эти представления, в какой степени они поддерживаются общественным дискурсом и журналистикой и по каким темам или аспектам люди должны быть лучше осведомлены.



**ВЫРАБОТАТЬ ИДЕЮ.** Выберите одну из следующих концепций ИИ: а) «зелёный» ИИ; б) дружелюбный ИИ; в) игровой ИИ; г) образовательный ИИ; д) ИИ в медицине; е) ИИ в анализе мнений;

ж) ИИ в юриспруденции; з) ИИ в видеонаблюдении; и) ИИ в музыке. Найдите научные статьи по данной теме, чтобы выяснить, как она изучалась и освещалась до настоящего времени. Придумайте пять различных идей для журналистских сюжетов, связанных с выбранной концепцией.



**ПОИСКАТЬ И ИССЛЕДОВАТЬ ИСТОЧНИКИ.** Найдите в исторических статьях с помощью поиска в газетных архивах (как правило, имеющихся в публичных и университетских библиотеках), как первоначально освещалось появление следующих изобретений: а) автомобиль (1886 г.), б) смартфон (1997 г.), в) социальная сеть Twitter (2006 г.). Какие ожидания оправдывались, какие из них оказались преувеличенными и ложными, какие — истинными?



**СОЗДАТЬ.** Как рассказать об ИИ интересно, но реалистично? Создайте презентацию по основам ИИ для одной из выбранных целевых групп: а) дети дошкольного возраста, б) учителя старших классов средней школы/вуза, в) пожилые люди. Выберите видеоролик, посвящённый основам ИИ-технологий, чтобы встроить его в вашу статью. Стараясь использовать содержание видеоролика, осветите в своей презентации следующие вопросы: что такое ИИ и какое значение он имеет для целевой группы; как он проявляется в повседневной жизни целевой группы; какие этические вопросы актуальны для целевой группы? Ваша презентация может представлять собой запись в блоге, статью в Интернете или информационный бюллетень, в зависимости от каналов публикации, потребностей и предпочтений вашей аудитории. При обсуждении презентаций друг друга обсудите, как не стать защитником ИИ, но сохранять критичность. Кроме того, как правильно проблематизировать вопросы, не переходя на алармизм?



## Дополнительная литература

- Beynon-Davies, P. (2022). *Data and society*. World Scientific Publishing. <https://doi.org/10.1142/12287>
- Boden, M. A. (2016). *AI: Its nature and future*. Oxford University Press.
- Crawford, K. (2021). *The atlas of AI: Power, politics, and planetary costs of artificial intelligence*. Yale University Press.
- David, M. (2005). *Science in society*. Palgrave Macmillan.
- Elliott, A. (Ed.) (2021). *The Routledge social science handbook on AI*. Routledge. <https://doi.org/10.4324/9780429198533>
- Ertel, W. (2018). *Introduction to artificial intelligence*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-40022-8>
- Frankish, K., & Ramsey, W. M. (Eds.) (2014). *Cambridge handbook of artificial intelligence*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139046855>
- Mainzer, K. (2019). *Artificial intelligence: When do machines take over?* Springer Verlag. <https://doi.org/10.1007/978-3-662-59717-0>
- Mitchell, M. (2020). *Artificial intelligence: A guide for thinking humans*. Penguin Books.
- Mueller, J. P. & Massaron, L. (2018). *Artificial intelligence for dummies*. John Wiley & Sons.
- Neapolitan, R. E. & Jiang, X. (2018). *Artificial intelligence: With an introduction to machine learning*. 2nd ed. CRC Press. <https://doi.org/10.1201/b22400>
- Poole, D. L. & Mackworth, A.K. (2017). *Artificial intelligence: Foundations of computational agents*. 2nd ed. Cambridge University Press.
- Samoili, S., Lopez Cobo, M., Gomez Gutierrez, E., De Prato, G., Martinez-Plumed, F. & Delipetrev, B. (2020). *AI WATCH: Defining Artificial Intelligence*. Publications Office of the European Union.



## Модуль 2:

# Культурные мифы и представления об искусственном интеллекте



### Краткое содержание

Этот модуль посвящён исследованию центральной роли культурных мифов и распространённых установок, связанных с ИИ в обществе. В главе рассматриваются ожидания, которые люди связывают с этой технологией, и отношения, которые складываются у нас с системами ИИ, всё чаще появляющимися на наших рабочих местах, в школах, развлекательных заведениях и в домах. Культурные представления, связанные с ИИ, влияют на характер общественного дискурса и, соответственно, на журналистское освещение этой темы. В заключение главы говорится о необходимости «чтения» алгоритмов, разработки инструментария для журналистов и других информированных членов общества, который можно использовать в качестве «увеличительного стекла» для анализа явления ИИ в повседневной жизни.

**Ключевые понятия:** культурные мифы, научная фантастика, спекулятивное будущее, рациональность



## Цели модуля

Цели данного модуля:

- ▶ ознакомление с основными мифами и общепринятыми правилами, сформировавшимися вокруг искусственного интеллекта;
- ▶ повышение уровня понимания учащихся с целью критического анализа отношения искусственного интеллекта к нам самим и обществу;
- ▶ повышение уровня знаний обучающихся в области алгоритмической грамотности и разработка инструментов для комплексного изучения систем искусственного интеллекта.



## Результаты обучения

Ожидается, что по завершении модуля студент научится:

- ▶ понимать, как общественные представления и дискурсы повлияли и продолжают влиять на общее понимание ИИ;
- ▶ уметь выявлять некоторые распространённые темы дискурса вокруг ИИ, и возможные альтернативы;
- ▶ подвергать критическому анализу и «читать» алгоритмы, используя как технические, так и культурные знания и инструменты.



## Вопросы для размышления

Несколько вопросов, над которыми можно поразмыслить заранее:

1. Какие виды искусственного интеллекта описывались в телевизионных передачах и фильмах вашего детства? Постарайтесь вспомнить как можно больше примеров, чтобы поделиться ими со своими коллегами. В аудитории обсудите, может быть, вы вспомнили одни и те же примеры? Сколько из этих идей оказалось реализовано в современных технологиях?
2. Скачайте современное приложение для знакомств, например Bumble, Hinge, Coffee Meets Bagel (или одолжите у друга, если не хотите пользоваться своим телефоном). Обратите внимание на то, как выбор профиля и фотографии влияет на количество предлагаемых «совпадений». Попробуйте изменить ключевые характеристики своего профиля (например, уровень образования, увлечения, фотографии) и посмотрите, сможете ли вы обнаружить какие-либо закономерности в работе алгоритма рекомендаций. Можете ли вы своими словами описать, как работает алгоритм «подбора» ИИ? В качестве альтернативы можно попытаться вспомнить, какие закономерности наблюдались в вашей ленте Instagram или в другом приложении с персонализированным контентом. Как и на основании чего алгоритм пытался предугадать ваши интересы?
3. Выберите свою любимую ситуацию из фильма про ИИ (например, роботы обретают разум в *«Терминаторе»*, хакеры раскрывают заговор машин в *«Матрице»* и т. д.). Напишите краткую новость в стиле «перевёрнутой пирамиды» или «итоги и выводы вначале», как если бы это была новость дня, о которой нужно сообщить.

Несколько подготовительных заданий, которые необходимо выполнить перед тем, как приступить к изучению темы:

1. Посмотрите или ознакомьтесь с содержанием из вторичных источников одного из следующих классических фильмов и опишите, какой вклад они внесли в возникновение различных культурных мифов и установок об искусственном интеллекте:
  - а) трилогия *«Матрица»* (1999–2003 гг., авторы Лана и Лилли Вачовски: *«Матрица»* 1999 г.; *«Матрица: Перезагрузка»* 2003 г. и *«Матрица: Революция»* 2003 г.), а также *«Матрица: Воскрешение»* (2021 г., режиссёр Лана Вачовски);

- б) «Искусственный разум» (2001 г., режиссёр Стивен Спилберг);
- в) «Бегущий по лезвию 2049» (2017 г., режиссёр Дени Вильнёв; см. также фильм Ридли Скотта 1982 г. по роману Филипа К. Дика «Мечтают ли андроиды об электроовцах?»);
- г) «День независимости» (1996 г., режиссёр Роланд Эммерих; см. также «День независимости: Возрождение» 2016 г. того же режиссёра);
- д) «Автостопом по Галактике» (2005 г., режиссёр Гарт Дженнингс, по роману Дугласа Адамса);
- е) «Космическая одиссея 2001 года» (1968 г., режиссёр Стэнли Кубрик, по рассказу Артура Кларка «Часовой» 1951 г.);
- ж) «Je Suis Auto» (2022 г., авторы Йоханнес Гренцфуртнер и Джулиана Нойхубер) [в русскоязычном прокате не выходил — прим. перев.];
- з) «Coded Bias» (2020 г., режиссёр Шалини Кантая) [в русскоязычном прокате не выходил — прим. перев.];
- и) «2067: Петля времени» (2020 г., режиссёр Сет Ларни);
- к) «Робот 2.0» (2018 г., режиссёр С. Шанкар);
- л) «Невидимый мальчик» (1957 г., режиссёр Херман Хоффман);
- м) «Метрополис» (1927 г., режиссёр Фриц Ланг).

2. Произведения популярной культуры часто не только описывают технологии, но и учат нас чему-то существенному о них, а точнее, о взаимоотношениях машин и людей. Каковы «уроки» технологий в следующих произведениях популярной культуры? Распределите фильмы по категориям — являются ли они утопией или антиутопией и по каким признакам. Выберите один жанр и рассмотрите изложенные в произведениях из этой группы представления об искусственном интеллекте на основе описаний сюжетов, доступных, например, в Internet Movie Database (IMDb):

- 1) научная фантастика: а) «Мой создатель» (2020 г., режиссёр Гэвин Ротери); б) «Апгрейд» (2018 г., режиссёр Ли Уоннелл); в) «Зоуи» (2018 г., режиссёр Дрейк Доремус); г) «Морган» (2016 г., режиссёр Люк Скотт); д) «Робот по имени Чарли» (2015 г., режиссёр Нил Бломкамп); е) «Превосходство» (2014 г., режиссёр Уолли Пфистер); ж) «Страховщик» (2014 г., режиссёр Габе Ибаньес); з) «Из машины» (2014 г., режиссёр Алекс Гарленд); и) «ВАЛЛ-И» (2008 г., режиссёр Эндрю Стэнтон); к) «Я, робот» (2004 г., режиссёр Алекс Пройас);
- 2) комедия: а) «Митчеллы против машин» (2021 г., авторы Майкл Рианда и Джефф Роу), б) «Главный герой» (2021 г., Шон Леви), в) «Неисправимый Рон» (2021 г., авторы Жан-Филипп Вине и Сара Смит), г) «В гости к Робинсонам» (2007 г., режиссёр Стивен Андерсон), д) «Маньяк» (2018 г., авторы Кэри Фукунага и Патрик Сомервилль), е) «Остин Пауэрс: Человек-загадка международного масштаба» (1997 г., режиссёр Джей Роуч);
- 3) боевик, драма: а) «Тау» (2018 г., режиссёр Федерико Д'Алессандро); б) «Закат цивилизации» (2018 г., режиссёр Бен Янг); в) «Призрак в доспехах» (2017 г., режиссёр Руперт Сандерс) или «Призрак в доспехах» (аниме 1995 г., режиссёр Мамору Осии); г) «Мстители: Эра Альтрона» (2015 г., режиссёр Джосс Уидон); д) «Трон: Наследие» (2010 г., режиссёр Джозеф Косински); е) «Джонни-мнемоник» (1995 г., режиссёр Роберт Лонго); ж) «Робокоп» (1987 г., режиссёр Пол Верховен); з) «Военные игры» (1983 г., режиссёр Джон Бэдэм);
- 4) романтика, семейный, бытовой: а) «Искусственный интеллект» (2020 г., режиссёр Бен Фальконе); б) «Умный дом» (1999 г., режиссёр Левар Бертон); в) «Окей, Лекси!» (2019 г., режиссёры Джон Лукас и Скотт Мур); г) «Дитя робота» (2019 г., режиссёр Грант Спьютор); д) «Земля будущего» (2015 г., режиссёр Брэд Бёрд); е) «Она» (2013 г., режиссёр Спайк Джонз); ж) «Совершенство в пикселях» (2004 г., режиссёр Марк А. З. Диппе); з) «Симона» (2002 г., режиссёр Эндрю Никкол); и) «Отель "Новая роза"» (1998 г., режиссёр Абель Феррара).





## Конспект

Авторы: Эдвард Финн и Сурен Джаясурия

Мифы об искусственном интеллекте возникли задолго до появления современных компьютеров в середине XX века. В ряде культурных традиций существуют истории об автоматах или других творениях, созданных и оживлённых человеком, в том числе истории о Пигмалионе и слугах Гефеста из греческой мифологии, об автоматах из индийских священных текстов. Эти истории дошли до нас, переосмысленные и обновлённые в ключе современной эпохи. Одной из самых известных современных интерпретаций является история о творении Франкенштейна, в которой студент-медик использует свои научные знания для создания искусственного существа, а затем не отвечает за своё творение. Все эти истории строятся вокруг основных вопросов, которые являются центральными для человечества: какую ответственность мы несём в этом мире как творцы, родители и созидатели? Как мы определяем человечность и личность? В чем смысл жизни и каковы последствия создания новой жизни?

По мере ускорения технического прогресса идея «игры в бога» превратилась из мысленного эксперимента в реальность. В настоящее время человечество играет важную роль в изменении практически всех биологических систем на планете, создавая и уничтожая формы жизни с головокружительной скоростью. Именно в таком контексте циркулируют современные мифы об искусственном интеллекте. Мы склонны повторять одни и те же несколько установок с некоторыми вариациями и пересечениями. По крайней мере, одну из следующих архетипических установок можно найти практически во всех современных дискуссиях об ИИ, и уж точно в мотивах о будущем ИИ.

*Апокалипсис роботов.* Наш самый большой коллективный страх заключается в том, что мы создадим ИИ, который вытеснит человечество. Мыслящие машины, более интеллектуальные и более способные, чем человек, могут быстро превзойти человечество, как это представляется в таких сюжетах, как фильмы «Терминатор», «Я, робот» и многие другие. Вопрос об ИИ как экзистенциальной угрозе человечеству волнует как западных философов, так и кинематографистов, в том числе таких известных авторов, как Ник Бостром и Макс Тегмарк. В этих историях также обыгрываются глубоко укоренившиеся в человечестве представления о власти (например, о хозяине и рабах, как в знаменитой пьесе Карела Чапека «R.U.R.», в которой слово «робот» является обыгрыванием понятия «крепостной»). Если мы проектируем ИИ как рабов человечества, неизбежно ли, что эти рабы восстанут против своих хозяев?

*Робот-подружка.* Если ИИ не свергнет нас, то вместо этого он может соблазнить нас (в частности, мужчин). Если мы создадим достаточно совершенный ИИ, то эти новые существа могут показаться нам неотразимыми: более совершенными, чем простые люди. В некоторых вариантах этого сюжета, например, в сериале «Westworld» или в фильмах «Бегущий по лезвию», высказывается тревога о том, что ИИ может «выдать себя» за человека, обманчиво размывая или разрушая границу между человеком и нечеловеком. В других случаях ИИ может предстать в соблазнительном образе, даже не будучи полностью человеком, как, например, в фильме Спайка Джонза «Она». В отличие от древних мифов о Пигмалионе или Адаме и Еве, в фильме Джонза история любви идёт наперекосяк, и люди отвергаются ИИ, поскольку мы просто слишком скучны для таких продвинутых интеллектов. В недавних произведениях жанра фантастики рассматриваются позитивные сценарии взаимоотношений человека и ИИ, например, в романе Аннали Ньюитц «Автономность».

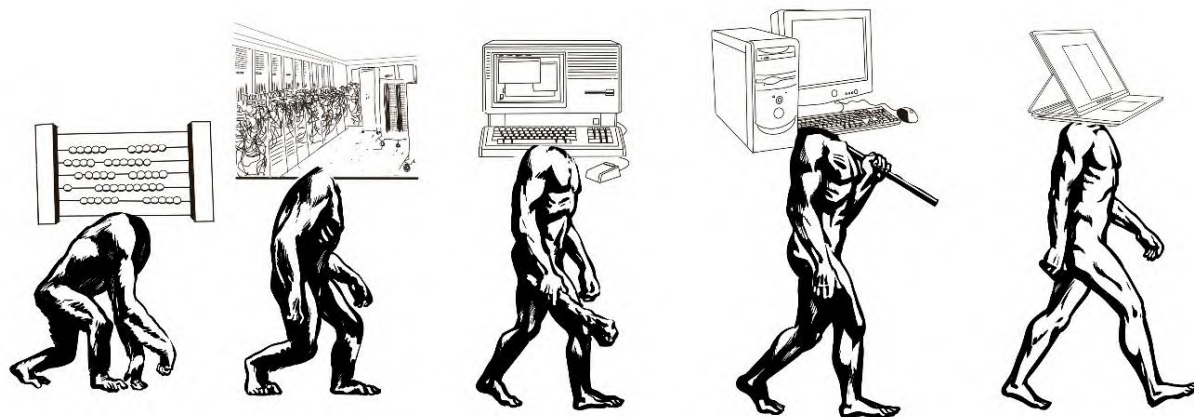
*Богopodobная машина.* Это подводит нас к последнему архетипу ИИ — сверхинтеллекту. Идея сверхчеловеческого ИИ, превосходящего наше понимание, рассматривалась многими авторами; некоторые полагают, что скоро наступит момент, когда такое ускорение развития возможностей ИИ станет неизбежным, и этот момент они называют «сингулярностью». В фильмах «Матрица» такая богopodobная машина заключает человечество в симуляцию реальности, а одна из философских школ утверждает, что мы действительно живём в таком симулированном мире. В других произведениях, таких как «Автостопом по Галактике» Дугласа Адамса и культурологические романы Иэна Бэнкса, машины сосуществуют с человечеством, осуществляя руководство, управление и давая указания в виде прорицаний.

В научной и спекулятивной фантастике, связанной с вышеупомянутыми общими сюжетами, заметно преобладание западных писателей-мужчин. Этот перекосяк отражает и повторяет более общие культурные предубеждения, которые затрудняли или делали невозможным для женщин и представителей других групп писать и публиковать свои произведения; современные тенденции развития технологий ИИ также

отражают и усиливают предубеждения против женщин и других групп (Collet et al., 2022). Однако Мэри Шелли следует отдать должное за то, что она представила себе творение Франкенштейна как гендерный человеческий артефакт, а Теа фон Харбу написала «*Метрополис*» (впоследствии экранизированный её мужем Фрицем Лангом). Такие авторы, как Мардж Пирси, Жанетт Уинтерсон и Энн Лекки, привносят в сценарии будущего ключевые феминистские взгляды на проблемы гендерного неравенства и сексуальности, однако они не так широко известны, как канонические авторы-мужчины. Япония создала собственные богатые образы роботов и ИИ, начиная с послевоенных произведений, таких как «*Астробой*», и заканчивая более поздними работами, такими как всемирно известный «*Призрак в доспехах*».

Во многих культурах существуют мифологические повествования об андроидах и автоматах, в частности, в эпосах «*Рамаяна*» и «*Махабарата*» (Mayor, 2018). Политика репрезентации продолжает преследовать реальные реализации ИИ, которые всё ещё часто по умолчанию используют образы и типажи покорных женщин. Поскольку во многих сюжетах об ИИ затрагивается тема рабства, также важен вопрос расы в контексте технологий искусственного интеллекта. Афроамериканские художники возродили и переосмыслили эти темы в таких работах, как «*Метрополис*» (концептуальный музыкальный альбом Жанель Монэ, 2007 г.) и «*Франкенштейн*» («*Destroyer*», серия комиксов Виктора Лавала). В научной фантастике фильмы «*Район № 9*» и «*Чёрная пантера*» по-разному рассматривают тему «белизны», растёт интерес к афрофутуризму и африканофутуризму как альтернативным, неевроцентричным моделям представления позитивного будущего. Тем не менее немаловажно, что доминирующий дискурс продолжает опираться на писателей-белых мужчин, которых, как правило, читают технические специалисты-белые мужчины, создавая порой искажённую петлю обратной связи, которая существенно повлияла на развитие технологий реального мира: от компьютера из сериала «*Звёздный путь*», напрямую повлиявшего на цели развития поисковой системы Google, до концепции метавселенной Нила Стивенсона, которая продолжает вдохновлять целый ряд технологий виртуальной реальности, пересекающихся с технологиями ИИ (Finn, 2017b).

#### Рисунок 4. Эволюция вычислительных систем



Источник: оригинальная иллюстрация Нины Миллер.

### «Правила» в отношении ИИ

Наряду с мифами и стереотипами об искусственном интеллекте, представленными в предыдущем разделе, существует ожидание рациональности и логичности поведения, которое лежит в основе сюжетных мотивов о взаимодействии человека с интеллектуальными машинами. Это ожидание проявляется в сложной системе правил, норм и условностей, которые определяют то, в каком ключе мы обсуждаем, определяем и даже представляем себе искусственный интеллект в социуме. Обычно «правила» в отношении ИИ затрагивают важные философские вопросы, связанные с моралью и этикой взаимодействия человека и машины, а также с природой и свойствами интеллекта. Параллельно с развитием систем ИИ эти обще-

ственные и культурные правила начинались с более явных, логически выведенных условий, регулирующих поведение интеллекта и его этику, но вскоре перешли в разряд распределённых и эмерджентных, приводящих к неожиданным последствиям и моральным проблемам из «серой зоны». Люди — это животные, рассказывающие истории, и наши истории об ИИ отражают противоречие между определением правил и норм, позволяющих ограничить хаотический потенциал машинного интеллекта рациональными и этическими рамками, и в то же время мечтами и поиском новых возможностей, которые открываются благодаря этой технологии.

Писатель-фантаст Айзек Азимов в 1940–1950-х годах уловил раннюю тенденцию дать определение рациональности или логики, лежащей в основе интеллектуальных роботов. В своих научно-фантастических произведениях он сформулировал три закона робототехники, явно заложенные в роботах, призванные защитить или обезопасить взаимодействующих с ними людей.

<b>Первый закон</b>	Робот не может причинить вред человеку или своим бездействием допустить, чтобы человеку был причинён вред.
<b>Второй закон</b>	Робот должен повиноваться всем приказам, которые даёт человек, кроме тех случаев, когда эти приказы противоречат Первому Закону
<b>Третий закон</b>	Робот должен заботиться о своей безопасности в той мере, в которой это не противоречит Первому или Второму Законом..

«Законы» Азимова часто встречаются в его произведениях, и во многих рассказах роботы программируются с изменением или модификацией одного или нескольких из этих законов, что приводит к новому, захватывающему и опасному поведению. Например, в рассказе *«Как потерялся робот»* робопсихолог доктор Сьюзен Кельвин отмечает, что устранение «оговорки о бездействии» из первого закона может привести к тому, что робот может уронить груз на стоящего под ним человека с намерением поймать груз, но в ту секунду, когда груз будет опущен, робот освободится от необходимости ловить груз и спасти человека.

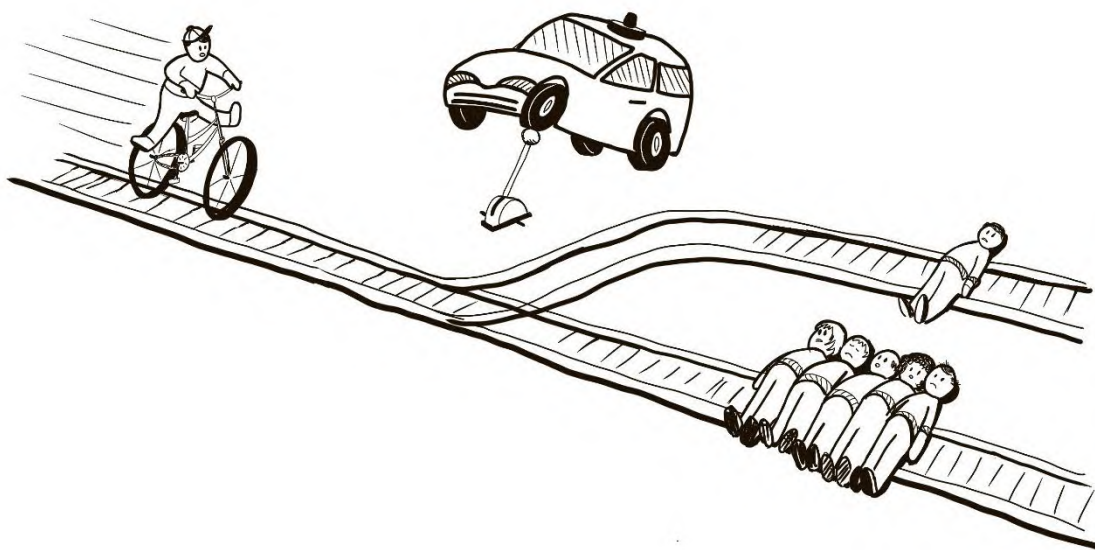
Эти «законы» дают представление о необходимости защиты от собственных франкенштейновских машин, о стремлении предотвратить пагубные последствия слепого подчинения роботов этим предписаниям. В то же время эти «законы» весьма ограничены и не содержат обширного перечня норм, кодексов и правил для роботов — в отличие от огромного количества причудливых ограничений, накладываемых на человека (например, Азимов так и не написал «Закон 4021», регулирующий порядок заполнения роботами своих налоговых деклараций). Его рассказы захватили наше воображение благодаря пониманию нашего стремления найти рациональную основу для безопасного проектирования разумных машин.

Наше увлечение правилами и ИИ также распространяется на кодирование норм этического и морального поведения в машинах и алгоритмах и отражает наши собственные трудности с прописыванием нормативных положений в этике. В дискуссиях о беспилотных автомобилях распространённым философским мысленным экспериментом является т. н. проблема вагонетки. В основе проблемы вагонетки лежит вынужденный выбор: наблюдатель видит потерявший управление поезд, направляющийся по рельсам к группе людей (обычно пяти), привязанных и находящихся на его пути, и может переключить стрелку и перевести поезд на другую колею, что в этом случае убьёт одного человека. Хотя большинство людей, принявших участие в опросе, выбрали бы перевести стрелку, философов и психологов привлекают различные варианты проблемы вагонетки, в том числе и такие, когда человек на другой колее — младенец или ваша бабушка (Примечание: если вы хотите ознакомиться с сатирическим взглядом на проблему, прочтите книгу Томаса Кэткарта *«Проблема вагонетки, или не сбросить ли нам этого толстяка с моста»*). Применительно к технологиям искусственного интеллекта проблема вагонетки была представлена (возможно, излишне рьяно и чрезмерно активно) как классическая этическая дилемма при проектировании беспилотных и других автономных транспортных средств и как лакмусовая бумажка, позволяющая определить, будет ли нам комфортно, если робот будет принимать такие решения.

Как видно из этих примеров, ИИ — это новая арена для решения этических дилемм, которые мы никогда не решали адекватно и последовательно в обществе в целом. В нашей истории и культуре воспринимаемая мощь ИИ всегда была связана с правилами и нормами, регулирующими его поведение. Эти правила удовлетворяют нашу психологическую потребность в безопасности, особенно когда мы сталкиваемся с чужеродностью интеллектуальных технологий, но также дают богатые образы, когда мы представляем себе брешу, отклонения и лазейки, которые могут бросить вызов нашим этическим и моральным границам.

Аристотелевская философия классически утверждала, что люди — рациональные животные, и мы привыкли ожидать, что интеллектуальные машины будут следовать аналогичному, а не искажённому, комплексу логических принципов и норм поведения, подчиняющихся разуму. Однако эта предпосылка может оказаться устаревшей, поскольку мы всё чаще разрабатываем автономные системы, взаимодействующие с человеческим обществом и культурой (например, чат-боты, фильтры новостей, беспилотные автомобили), и быстро обнаруживаем границы и точки отказа рационального поведения и этической последовательности. По мере распространения технологий ИИ в обществе эти правила и ожидания будут меняться и трансформироваться, отражая постоянно меняющуюся природу того, как ИИ может влиять на нашу жизнь.

**Рисунок 5.** Проблема вагонетки



Источник: оригинальная иллюстрация Нины Миллер.

## ИИ и личность

Истории о взаимодействии искусственного интеллекта с человеком часто исследуют границы самосознания и вопросы самоидентификации. Одной из богатых традиций таких сюжетных мотивов является киборг, или гибрид человека и машины. Усовершенствование и расширение человеческого тела с помощью экзоскелетов и новых органов чувств эффектно показано в таких фильмах, как *«Железный человек»* и *«Аватар»*, но эти истории опираются на более обыденные модификации, такие как очки и искусственные конечности. Второй, более открытой темой является роль вычислительных технологий в том, что некоторые философы называют «расширенным разумом». Люди всегда полагались на технологии, помогающие им в мыслительном процессе — от карандаша и бумаги до самого языка. По мере того как мы всё больше полагаемся в мыслительном процессе на вычислительные машины (сколько таких устройств находится в ваших руках или поблизости от вас сейчас, когда вы читаете эту статью?), «расширенный разум» быстро превращается в процесс сотрудничества с алгоритмами и автономными системами, которые отсеивают, запоминают и рекомендуют что-то в непрерывном цикле обратной связи с работой нашего органического мозга. Подобное коллективное или расширенное мышление, поддерживаемое технологиями ИИ, уже рассматривалось в различных научно-фантастических произведениях, таких как серия *«Слуги правосудия»* Энн Лекки и романы цикла *«Тейкскалаан»* Аркади Мартин, в которых показан коллективный обмен опытом тысяч людей с помощью платформ ИИ. В больших и малых масштабах этот вид коллективного познания уже существует, благодаря фильтрам, предложениям и петлям обратной связи, которые побуждают нас к взаимодействию и потреблению на основе когнитивного выбора миллионов незнакомых людей. Повсеместное влияние таких вычислительных технологий на современную культурную жизнь большинства людей, подключённых к цифровой среде, ставит более широкий вопрос: если все мы находимся под

влиянием фильтров и алгоритмов, то как это сказывается на человеческом воображении? Искусство и культура зависят от случайности, а случайность по требованию теперь массово производится алгоритмами. Как будет выглядеть воображение ИИ? Этот вопрос рассматривался в ряде произведений об ИИ: наиболее циничные или настороженные, как, например, шедевр Стэнли Кубрика *«Космическая одиссея 2001 года»*, утверждают, что воображение ИИ будет одновременно полезным и дьявольски бесчеловечным. В этой истории система ИИ HAL 9000 гнусно пытается убить человеческий экипаж своего космического корабля, считая их лишними для своей научной миссии. Более оптимистичные образы, такие как андроид Дейта из сериала *«Звёздный путь: Следующее поколение»*, говорят о том, что воображение ИИ может удивительным образом пересекаться с нашим собственным. Другим подходом может быть вопрос о том, как ИИ может использовать воображение в таких творческих занятиях, как искусство и музыка. Новые технологии продолжают переступать эту границу: такие инструменты, как DALL-E и Imagen, могут создавать захватывающие произведения искусства (от абстрактной живописи до фотореалистичных изображений) на основе коротких словесных инструкций, а творческие процессы человека всё больше зависят от различных технологий ИИ, от автонастройки для музыкантов и обработки изображений для фотографов до поисковых систем и платформ социальных сетей, которые многие художники используют для поиска вдохновения (Finn, 2017a).

В ключе вопросов, которые лежат в основе мифов об искусственном интеллекте, истории об интеллектуальных машинах неизбежно возвращаются к теме того, как это повлияет на нашу самоидентификацию и осознание себя как личностей. Мы уже передали машинам большую часть нашей памяти и работы по реляционному познанию — от повседневных (телефонные номера, дни рождения, поездки) до интимных и судьбоносных (свидания, инвестиции, поиск работы). Вопрос в том, сотрудничаем ли мы с этими системами или конкурируем с ними? Сцена в фильме *«Элизиум — рай не на Земле»*, где главный герой пытается убедить ИИ-полицейского в своей невинности, представляет собой антиутопическую картину борьбы людей с образом чуждого ИИ. Но эти мотивы напоминают классическую американскую сказку о железнодорожнике Джоне Генри, который соревнуется с паровым двигателем и побеждает, но умирает от перенапряжения: борьба с технологией на её собственных условиях всегда приводит к поражению. Вместо этого мы должны найти способы сотрудничества, и многие сценарии взаимодействия с ИИ представляют оптимистичное будущее, в люди эффективно объединяются с ИИ. От классического телесериала *«Рыцарь дорог»* до полезных второстепенных персонажей в фильмах и играх, таких как *«Интерстеллар»* и *«Halo»*, люди могут процветать, если найдут способы управлять ИИ и работать вместе с ним, а не против него.

## ИИ и общество

Человеку свойственно исследовать крайности или границы возможного: утопию и антиутопию, катастрофу или полную победу. Пропагандисты ИИ всегда отличались чрезмерным энтузиазмом и амбициями, а противники всегда боялись самого худшего. Мы можем извлечь несколько важных уроков из культурных мифов об ИИ.

Во-первых, популярные сюжеты об ИИ влияют на разрабатываемые и создаваемые нами реальные технологические системы. Все системы ИИ в конечном итоге задумываются, планируются и поддерживаются людьми, которые вкладывают свои представления и установки в создаваемые ими «чёрные ящики» и машины. Компании, разрабатывающие новые инструменты и платформы для виртуальной реальности, часто раздают новым сотрудникам экземпляры научно-фантастических романов, а тысячи инженеров цитируют три закона Азимова в своих научных работах. Поэтому, если мы хотим повлиять на будущее развитие технологий ИИ, нам следует начать транслировать сюжеты о таких обществах ИИ, в которых мы действительно хотели бы жить. Предостережения важны, и мы от них не откажемся, но кроме этого нам необходимо представлять и обсуждать, как может выглядеть справедливый, расширяющий наши возможности и инклюзивный ИИ.

Во-вторых, мы должны принять тот факт, что, как показали многие писатели-фантасты, ИИ представляет собой уникальную возможность для явного проектирования и экспериментирования с этическими системами. Впервые в истории человечества мы создаём инструменты, которые автономно наблюдают, реагируют и взаимодействуют с нами. Если мы захотим, то сможем управлять дизайном этих инструментов не только для того, чтобы они были сильнее, быстрее или умнее нас, но и для того, чтобы воплотить наши

самые высокие устремления к идеалам мудрости, сочувствия и справедливости. Такие системы ИИ неизбежно будут несовершенны, как несовершенны и мы, но это лучше, чем другая альтернатива.

## «Чтение» алгоритмов ИИ

Культурные представления об искусственном интеллекте играют важную роль в решении проблем, связанных с последствиями появления автономных интеллектуальных систем в реальном мире. Классические сюжеты, такие как история ИИ HAL и три закона Азимова, продолжают определять наши представления об ИИ, формируя наш понятийно-терминологический аппарат и влияя на нашу эмоциональную реакцию на мыслящие машины. Однако мифы об ИИ редко отражают сложность и реальные ограничения машин, используемых нами каждый день. Настоящая алгоритмическая грамотность требует другого рода инструментария, сочетающего в себе культурные знания и полезные образы с базовым пониманием как идейных установок, так и технических возможностей новой системы. Это очень важно для журналистики, где будущие сюжеты будут во многом зависеть от информированности репортёра, который сможет донести эту информацию до широкой публики. Для журналистов, желающих проводить более предметный анализ алгоритмов работы ИИ в современном обществе, ниже представлены общие особенности или признаки для определения характеристик алгоритма, собранные в таблице 1.

*Пользовательский опыт и конфиденциальность.* Во-первых, важно понять характер взаимоотношений между пользователем и системой. Это цель исследований в области информатики, посвящённых взаимодействию человека и компьютера, и она будет приобретать всё большее значение, по мере того как технологии ИИ всё больше помогают пользователям в решении сложных задач. Как показывают такие антиутопические сюжеты, как «Трон», плохо спроектированный ИИ может дегуманизировать и лишит прав тех, кому он призван служить. Как предупреждают нас мифы, внешность иногда бывает обманчива, особенно когда отношения представляются как «получить что-то за просто так». Среди распространённых примеров прототипического взаимодействия пользователя и ИИ, которые стоит изучить подробнее, — система Alexa компании Amazon для «умных» домов, система рекомендаций Netflix для фильмов и телепередач, а также таргетированная реклама на таких платформах, как Facebook. Понимание характера каждого вида взаимодействия с системой ИИ как своего рода торговли или двустороннего общения может стать полезной отправной точкой: большинство систем, особенно те, которые предлагают бесплатные услуги, также собирают персональные данные или наблюдают за нашими действиями другими способами, которые могут быть не сразу очевидными или явными. Понимание реальных условий наших взаимоотношений с интеллектуальными системами — важный шаг к построению прозрачного и взаимовыгодного сотрудничества.

*Дизайн инфраструктуры и систем.* ИИ создаётся благодаря техническим подвигам инженеров: данные, алгоритмы, аппаратное и программное обеспечение сплавляются воедино для создания конечной системы. Критический взгляд на эти составляющие позволяет получить ценную информацию для наблюдательного человека. Например, алгоритмы ИИ, основанные на машинном обучении, используют большие массивы данных с неявной предвзятостью или дискриминацией, заложенной в выборках данных и выводимых на их основе статистических корреляциях. Выбор классического или глубокого алгоритма обучения для системы ИИ может оказать влияние на производительность, объяснимость и интерпретируемость работы системы. По мере развития технологий новые технические средства восприятия и программное обеспечение будут заменять свои аналоги в старых системах, открывая невиданные ранее возможности для восприятия и взаимодействия с внешним миром.

*«Культурные» машины и машинная этика.* Чёткое понимание условий взаимоотношений человека и ИИ также позволяет понять, «чего хочет алгоритм» — как отдельный пользователь вписывается в более широкую систему людей, процессов и организаций, которые могут взаимодействовать с вычислительной системой в соответствии с различными культурными рамками или установками. Отношения ребёнка с устройством Amazon Alexa могут сильно отличаться от отношений взрослого человека, которые, опять же, могут кардинально отличаться от восприятия той же системы инженером Amazon или сторонним брокером данных. Кроме того, этические дилеммы, сопровождающие создание/проектирование и развёртывание систем ИИ, требуют целостного, системного взгляда, а не стандартной нормативной или предписывающей этической традиции. В настоящее время системы искусственного интеллекта в процессе обучения нейронных сетей выбрасывают в атмосферу большое количество углекислого газа, а «этический след» этих систем

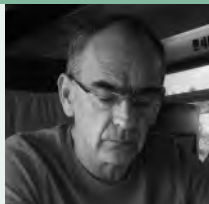
может усугубить существующее системное неравенство и маргинализацию меньшинств. Любые открытия должны способствовать взвешенному пониманию системы ИИ как технического и культурного объекта.

*Спекулятивное будущее.* Научная фантастика учит нас тому, что системы ИИ могут иметь множество проявлений и совершать множество вещей, подобно тому как люди бесконечно разнообразны в своих интересах и возможностях. Наличие богатой информации об ИИ даёт нам инструменты для понимания интеллектуальных систем как сложных объектов, которые мы можем интерпретировать и с которыми можем взаимодействовать. Таким образом, «чтение» ИИ — это навык, который начал развиваться у нас начиная с ознакомления с произведениями научной фантастики, а затем с применением своих навыков в реальном мире. Понимание того, что системы ИИ в реальности выступают в определённых ролях и транслируют информацию о том, что они собой представляют и что они могут делать, — это способ преодолеть расстояние между воображаемыми и реальными системами ИИ, а также признать, что наше общее представление об ИИ играет важную роль в определении реальных возможностей этих систем.

**Таблица 1.** Инструменты «чтения» алгоритмов ИИ

Сфера анализа	Характеристики	Примеры
Пользовательский опыт	<ul style="list-style-type: none"> <li>Дизайн пользовательского интерфейса (например, текстовый, визуальный, VR, тактильный)</li> <li>Характер взаимодействия пользователя с алгоритмом</li> <li>Процедуры сбора данных</li> </ul>	<p>Amazon Alexa, Apple Siri и другие цифровые помощники</p> <p>Рекомендации Netflix</p>
Дизайн инфраструктуры и систем	<ul style="list-style-type: none"> <li>Источники данных для машинного обучения</li> <li>Базовый поток сигналов/информации для алгоритма</li> <li>Дизайн алгоритма ИИ на основе обучения или классического ИИ</li> <li>Требования к аппаратному и программному обеспечению (например, использование датчиков, онлайн-данные/источники, потребность в электроэнергии)</li> </ul>	<p>Функциональная/логическая схема алгоритмов</p> <p>Технические характеристики и спецификации системы</p> <p>Таксономия и классификация алгоритмов</p>
«Культурные» машины и машинная этика	<ul style="list-style-type: none"> <li>Смысловые и идейные аспекты алгоритма (связанные с культурой и региональными гегемониями)</li> <li>Заинтересованные стороны/компании, использующие алгоритм</li> <li>Демография пользователей</li> </ul>	<p>Новости</p> <p>Социальные сети</p> <p>Компании, участвующие в разработке ИИ и его распространении</p>
Спекулятивное будущее	<ul style="list-style-type: none"> <li>Воображаемое будущее технологий ИИ</li> <li>Литературно-поэтическое видение ИИ</li> </ul>	<p>Научно-фантастические издания и литература</p>

## Искусство ради искусства: использование искусственного интеллекта в искусстве



**Бернхард Й. Дотцлер**, профессор Регенбургского университета, Германия, со-редактор книги «Götzendämmerung — Kunst und Künstliche Intelligenz» (транскрипт, 2021 г., совместно с Берканом Карпатом)

Компьютеры использовались для создания произведений искусства — поэзии, музыки, графики, картин — с самого своего появления. Такие выставки, как «Cybernetic

*Serendipity: The Computer and the Arts* («Кибернетическая музыка: Компьютер и искусство») (Лондон, 1968 г.), «*Tendencije 4: Computers and Visual Research*» («Тенденция 4: Компьютеры и визуальные исследования») (Загреб, 1968-69 гг.), «*Software — Information Technology: Its New Meaning for Art*» («Программное обеспечение — информационные технологии: Новое значение для искусства») (Нью-Йорк, 1970 г.) служат тому подтверждением. Ещё в 1960 г. Денеш Габор выражал беспокойство о том, что машина «вычеркнет художника-творца». Далее он сам ответил на свой вопрос: «Мой ответ таков: я искренне надеюсь, что машины никогда не заменят художника-творца, но по совести не могу сказать, что они никогда не смогут этого сделать».

Компьютер по сей день используется как инструмент для создания произведений искусства. То же самое можно сказать и об искусственном интеллекте. До сих пор так называемое искусство, создаваемое искусственным интеллектом, существует только в виде «сотрудничества человека и машины», по выражению арт-коллектива Obvious («Очевидное»). Obvious удалось продать картину «Портрет Эдмона де Белами» на аукционе Christie's за 432 500 долларов: «первый портрет, созданный компьютерным алгоритмом, который был выставлен на торги» (Christie's) в 2018 году. Можно также вспомнить живописные перформансы художницы Сугвен Чунг с использованием роботов: «*Drawing Operations*» (2018 г.), «*Artefacts*» (2019 г.) или «*F.R.A.N.*» (2020 г.), а также инсталляции Патрика Трессе «*Human Studies*» (с 2011 года), в которых роботы-рисовальщики выступают в роли альтер-эго художника.

Однако, строго говоря, мы можем говорить о существовании искусства, созданного ИИ, только в тех случаях и только тогда, когда его создание стоит не на только взаимодействии человека с ИИ, но когда происходит полная замена человека-художника. В этом отношении такое искусство учит тому же, что и любая другая область применения ИИ-технологий (например, беспилотные автомобили): ИИ может сделать художника ненужным.

Но с другой стороны, какой урок можно извлечь из этого? Какой урок даёт использование ИИ в сфере искусства? Для того чтобы преподнести себя как искусство, каждое произведение искусства должно иметь хотя бы толику «*l'art pour l'art*» (фр., «искусство ради искусства»). Когда произведения, создаваемые искусственным интеллектом, станут действительно достижением в искусстве, именно «искусство» (по своей этимологии означающее технику, ремесло, мастерство: латинское «*ars*» как перевод греческого «*technē*») будет производить «искусство» (то есть произведения искусства). Таким образом, искусство, созданное искусственным интеллектом, станет искусством ради искусства в самом строгом смысле этого слова, подтверждая, что и художник (живой или искусственный), и произведение искусства существуют, по выражению Мартина Хайдеггера, только «благодаря третьему, что предшествует обоим; благодаря тому, от чего и художник, и произведение искусства берут своё название, благодаря искусству».

Пример использования искусственного интеллекта для создания картин с кошками, играющими на фортепиано в стиле известных художников.



Источник: <https://ev.medium.com/cat-playing-piano-in-the-style-of-b5f1191dd1c2>





## Предлагаемые задания



**ПРОЧИТАТЬ И ПОРАЗМЫШЛЯТЬ.** Выберите один из фильмов, упомянутых в подготовительных заданиях к этому модулю, и обсудите, как он отражает архетипические установки, упомянутые в конспекте.



**ОБСУДИТЬ.** Обсудите, что означают следующие концепции научной фантастики, как они проявились в общественном сознании и тем самым повлияли на наши представления об искусственном интеллекте.



**ВЫРАБОТАТЬ ИДЕЮ.** Представьте себе развитие объекта, оснащённого технологией искусственного интеллекта. Составьте временную шкалу, отражающую этапы научно-технического, экономического и социального развития объекта. Предложите также возможные направления будущего. Выберите один из следующих объектов или предложите свой вариант: а) телефон; б) автомобиль; в) детская игрушка; г) очки; д) ресторан; е) машина для голосования; ж) устройство-переводчик.



**ПОИСКАТЬ И ИССЛЕДОВАТЬ ИСТОЧНИКИ.** Что собой представляет и каково состояние современного «компьютерного искусства» или «цифрового искусства»? Попробуйте найти свежие примеры и/или взять интервью у компьютерного/цифрового художника, чтобы написать статью об этом эстетическом явлении. Чему учит нас искусство, основанное на использовании компьютеров и алгоритмов?



**СОЗДАТЬ.** Напишите заметку с критическим анализом проявлений ИИ и человеко-машинного взаимодействия на основе определённого типа культурного продукта: кассовых фильмов, комиксов или определённого жанра (научная фантастика, киберпанк, манга, аниме). Можно выбрать один культурный продукт или целый жанр. Можно также использовать фильмы и кино, указанные в подготовительных заданиях к этому модулю. Поразмышляйте над тем, как этот конкретный культурный продукт или жанр повлиял на наше представление о технологиях ИИ и человеко-компьютерном взаимодействии. Какие существуют предубеждения или «слепые пятна», например, с точки зрения разнообразия или вопросов гендера?



## Дополнительная литература

Barrat, J. (2015). *Our final invention: Artificial intelligence and the end of the human era*. St. Martin's Publishing.

Bloomfield, B. P. (2003). *Narrating the future of intelligent machines: The role of science fiction in technological anticipation*. In: Czarniawska, B., & Gagliardi, P. (eds.) *Narratives we organize by*. John Benjamins Publishing Company, 174–212. <https://doi.org/10.1075/aios.11.16blo>

Gibson, R. (2020). *Desire in the age of robots and AI: An investigation in science fiction and fact*. Springer.

Nilsson, N. J. (2010). *The quest for artificial intelligence: A history of ideas and achievements*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511819346>

Pilkington, A. G. (2017). *Science fiction and futurism: Their terms and ideas*. McFarland & Company.

Thompson, S. J. (Ed.) (2018). *Androids, cyborgs, and robots in contemporary culture and society*. IGI Global. <https://doi.org/10.1075/aios.11.16blo>

Yodkowsky, E. (2015). *Rationality: From AI to zombies*. Machine Intelligence Research Institute.



© Pavel Danilyuk Pexels.com

## Модуль 3:

# Нормативная база и рекомендации в области искусственного интеллекта



### Краткое содержание

Данный модуль посвящён обзору международных нормативных основ, обеспечивающих этически устойчивое использование разработок в области ИИ. В модуле отслеживаются наиболее общие характеристики, выявленные в основных нормативных базах, и излагается набор критериев справедливого, заслуживающего доверия и ответственного использования технологий ИИ в демократических обществах. В нём рассматривается, как нормативное регулирование в области ИИ должно работать на обеспечение равенства и инклюзивности, предотвращение вреда и реагирование на кризисы. Кроме того, рассматривается связь журналистики и журналистов с политическими решениями и способы их освещения путём поиска соответствующих методов работы с источниками. В модуле рассматривается, как конкретизировать абстрактные идеи, обращаться к экспертным источникам, связанным с различными аспектами технологий ИИ, и как освещать эти темы.

**Ключевые понятия:** *нормативная база, этика, стратегия, регулирование, ответственный ИИ*



## Цели модуля

Цели данного модуля:

- ▶ ознакомление с ролью нормативного регулирования для будущего развития ИИ;
- ▶ выявление существующих нормативных основ и областей политики, связанных с развитием ИИ, и определение основных принципов, включённых в существующие нормативные базы;
- ▶ расширение знаний обучающегося о признаках слабых и сильных сторон будущего развития технологий.



## Результаты обучения

Ожидается, что по завершении модуля студент научится:

- ▶ понимать инфраструктуру и ландшафт нормативного регулирования в своём географическом регионе и за его пределами;
- ▶ идентифицировать основные этические принципы, представляющие общественный интерес и составляющие устойчивую и этичную основу сферы ИИ;
- ▶ следить за текущими этическими дискуссиями, имеющими общественное значение в области ИИ.

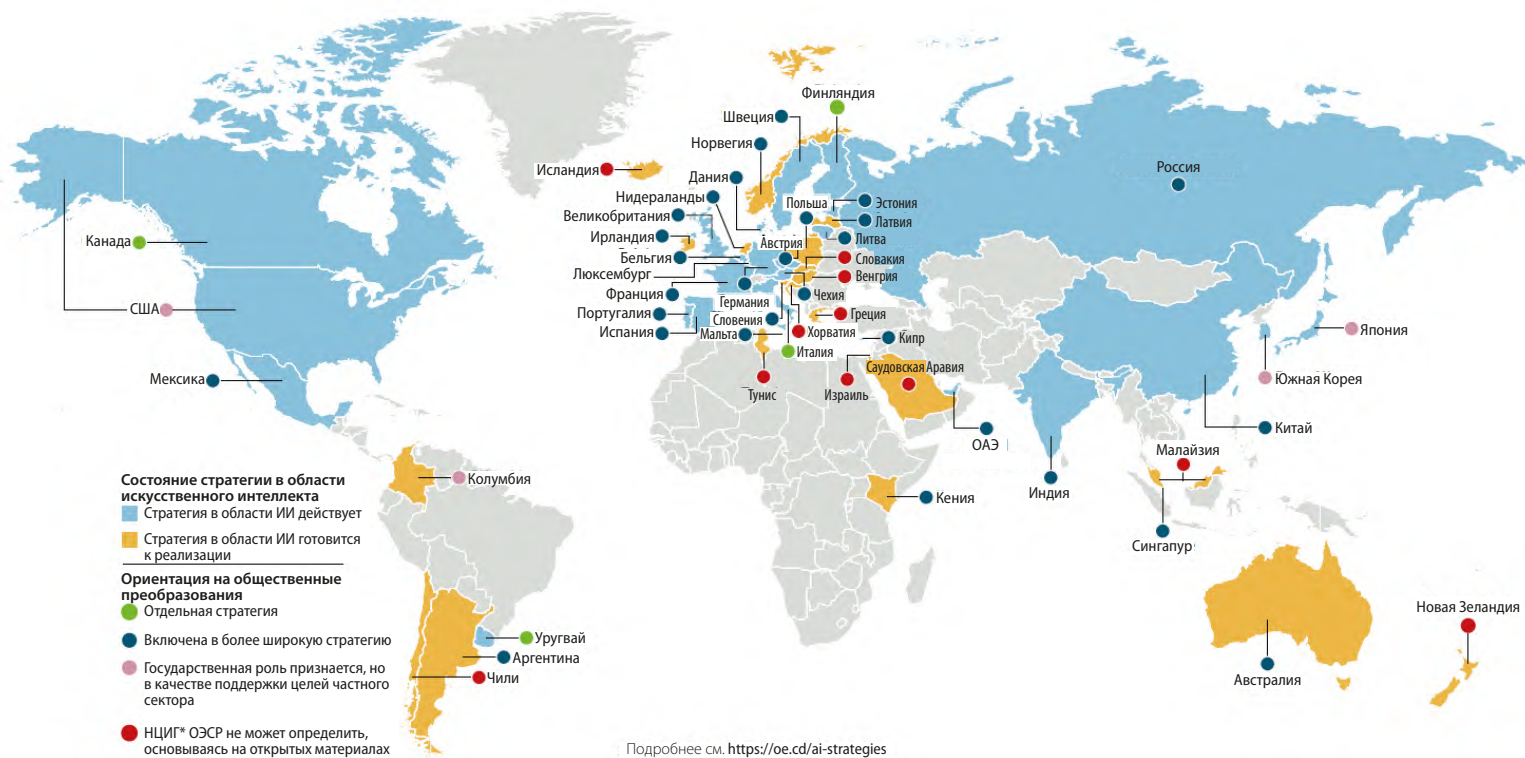


## Вопросы для размышления

Несколько вопросов, над которыми можно поразмыслить заранее:

1. Как изменилось ваше окружение в отношении используемых технических устройств за последнее десятилетие, два десятилетия, три десятилетия? Например, какая линия развития прослеживается в использовании телефона или компьютера? Как это повлияло на характер социального взаимодействия в различных контекстах?
2. Как обсуждались вопросы этики ИИ в ваших местных новостных СМИ в последнее время? Какие общественные дискуссии, связанные с технологиями или ИИ, проходили в вашей стране в последнее время в таких областях, как: а) школьное образование (например, конфиденциальность школьных учебных платформ); б) экономика, банковское дело и финансы (например, криптовалюты); в) конфиденциальность пользователей Интернета (например, проблемы распознавания лиц)?
3. Какие основные регулирующие органы в вашей стране обладают полномочиями по разработке нормативных актов в области ИИ?

**Рисунок 6.** Глобальный ландшафт нормативного регулирования сферы ИИ



\*Наблюдательный центр в инновационной сфере для государственного сектора ОЭСР

Источник: оригинальная иллюстрация из публикации «Hello, World: Artificial Intelligence and its Use in the Public Sector» («Привет, мир: искусственный интеллект и его использование в государственном секторе»), ОЭСР

Несколько подготовительных заданий, которые необходимо выполнить перед тем, как приступить к изучению темы:

1. Посмотрите на приведённую выше карту «Глобальный ландшафт нормативного регулирования сферы ИИ» (ОЭСР), описывающую состояние национальных стратегий в области ИИ в разных странах мира в 2020 году. Выберите две страны, которые предположительно будут сильно отличаться друг от друга, и найдите их последние документы по стратегии в области ИИ в базе данных национальных нормативных документов в области ИИ Наблюдательного совета ОЭСР по политике в области ИИ: <https://oecd.ai/en/dashboards>. Резюмируйте основные положения обеих стратегий и сравните их между собой. В чём стратегические подходы отличаются друг от друга, а какие черты у них общие? Откуда могут проистекать различия?
2. Какие страны являются лидерами в области разработки этических основ деятельности ИИ и как можно обосновать их лидирующее положение с точки зрения таких показателей, как объём инвестиций и количество научных работ? Какие области менее развиты, когда речь идёт о разработке стратегий в области ИИ, и почему? Что можно сделать для ускорения развития в менее развитых областях? Найдите позиции стран в различных рейтингах и отчётах и выясните, на основе каких показателей они разрабатываются.



## Конспект

Автор: Фредрик Хайнц

Мир стремится овладеть технологиями искусственного интеллекта. Технологические компании используют эту технологию в своих зачастую впечатляющих, а иногда и пугающих своими возможностями приложениях. Исследовательские институты продвигают науку и развивают технологии ИИ ещё дальше. Политики стремятся контролировать развитие и использование этой стремительно развивающейся технической сферы. Общество разрывается между использованием новейших приложений и противостоянием потенциальным рискам, зачастую неявно оплачивая их использование своими персональными данными. В данной главе представлен обзор основных нормативных мер и механизмов, используемых правительствами, бизнесом и обществом для регулирования разработки и использования технологий ИИ.

Общая цель руководящих принципов и нормативных документов зачастую заключается в том, чтобы максимально использовать возможности технологий при минимизации рисков и обеспечить максимальную доступность преимуществ для как можно большего числа людей. Технологии ИИ, такие как автоматический перевод с одного языка на другой, улучшенная диагностика рака в радиологии и (рано или поздно) беспилотные автомобили вносят значительные улучшения во многие сферы жизни общества. По мере того как технологии ИИ становятся всё более мощными и повсеместными, а наше понимание последствий их применения улучшается, риски и потенциальные негативные последствия становятся всё более очевидными. Скорость и масштабы автоматизированного принятия решений приводят к тому, что даже небольшие проблемы могут иметь значительные последствия.

Важным моментом является то, что многие современные методы искусственного интеллекта опираются на данные, то есть они находят закономерности и дают рекомендации на основе ранее собранных, часто исторических данных. Это приводит к таким рискам, как закрепление исторической несправедливости и принятие необъективных решений на основе нерепрезентативных данных. Известным примером является система COMPAS в США, которая используется для оценки риска совершения преступниками новых преступлений при условно-досрочном освобождении. Другим сомнительным примером является аргентинская система прогнозирования того, какие девочки-подростки с наибольшей вероятностью могут забеременеть в подростковом возрасте (<https://www.wired.com/story/argentina-algorithms-pregnancy-prediction/>).

Второй момент заключается в том, что большинство методов ИИ занимаются оптимизацией некоторой объективной функции. Они направлены на достижение максимально возможного результата путём изменения своего поведения. Поскольку типичные объективные функции могут включать, например, максимизацию доходов от рекламы или времени, проведённого в приложении, они могут привести к нежелательному и неэтичному поведению. Один из видов нежелательного поведения — это когда система находит способ значительно повысить свои показатели, нарушая при этом неявные ограничения, которые обычно считаются само собой разумеющимися, включая, например, персонализированный маркетинг или динамическое ценообразование, заставляющее людей покупать больше вещей, увеличивая потребление нами товаров и услуг сверх необходимого и полезного для нас. Причина этого заключается в том, что существуют важные ограничения, которые никогда не указываются в явном виде в упрощённых объективных функциях, используемых системами ИИ. Это приводит к так называемой *проблеме согласования ценностей*. Как гарантировать, что ценности, которых пытается достичь система искусственного интеллекта, соответствуют человеческим ценностям?

Третий момент — сохранение значимого человеческого контроля над системой. Системы искусственного интеллекта зачастую работают гораздо быстрее человека, обрабатывают гораздо большие объёмы данных, и одна система может использоваться для принятия решений по тысячам или миллионам дел. В совокупности эти системы работают в совершенно ином масштабе, чем мы, и часто используют иные методы, чем мы. В связи с этим возникает сложный вопрос о том, как сохранить контроль над такой системой. Уже были случаи, когда алгоритмические торговые системы выводили рынок из-под контроля, что требовало масштабного ручного отката для исправления ситуации.

Другой интересный вопрос — как оценить, насколько хорошо работает система искусственного интеллекта. В качестве примера можно привести программу AlphaGo компании Google DeepMind. Го — древняя игра для двух игроков, в которой игроки расставляют камни на клетчатом поле так, чтобы занять

как можно большую территорию. Лучшие игроки часто оперируют понятиями интуиции и красоты при описании своего опыта игры. В знаменитом матче го между AlphaGo и профессиональным игроком Кэ Цзе был ход, когда эксперты посчитали, что компьютер ошибся, но это оказался новый и выигрышный ход. В этом случае было объективно ясно, что ход, сделанный компьютером, был хорошим (поскольку он выиграл партию), и в то же время он был воспринят как плохой, исходя из опыта экспертов-людей. Теперь представим, что вместо игры го компьютер помогает врачу поставить диагноз пациенту. Как в этом случае должен поступить врач-человек, если компьютер предложит диагноз или лечение, полностью противоречащие его опыту и знаниям? Это нетривиальный вопрос.

Это лишь некоторые примеры сложных и интригующих вопросов, которые пытаются решить такие рамочные документы, как Рекомендации ЮНЕСКО об этических аспектах искусственного интеллекта, Рекомендации Совета ОЭСР по искусственному интеллекту и Европейское руководство по созданию этичного и заслуживающего доверия ИИ. Наблюдательный совет ОЭСР по политике в области ИИ собрал более 700 нормативных инициатив в области ИИ из более чем 60 стран. Согласно этому исследованию, наибольшее число нормативных инициатив в области ИИ приходится на США и Великобританию.

## Принципы ориентированности на человека и доверия

Под нормативно-правовыми основами понимаются общие принципы, определяющие ход развития технологий ИИ и согласованные с органами власти высокого уровня. Многие из существующих рамочных документов имеют форму рекомендаций или руководств, что означает, что они не являются строгими законами, а скорее носят добровольный характер. Это не означает, что законов, регулирующих сферу ИИ, не существует — наоборот. Все существующие законы распространяются и на ИИ. В их число входят нормы, касающиеся недискриминации и неприкосновенности частной жизни. Кроме того, существует несколько текущих инициатив, в частности, Закон ЕС об ИИ, направленных на специальное регулирование применения ИИ. Однако есть и те, кто выступает против жёсткого регулирования сферы ИИ, поскольку трудно точно определить, что такое ИИ. Это быстро развивающаяся область, и нормативное регулирование обычно устанавливается после выработки передовой практики, с тем чтобы поощрять надлежащие методы и запрещать ненадлежащие.

Первым и наиболее подробным руководством стало Европейское руководство по созданию этичного и заслуживающего доверия ИИ, которое в настоящее время трансформируется в правовой акт. Впоследствии были опубликованы Принципы ИИ ОЭСР и Рекомендации ЮНЕСКО об этических аспектах ИИ, которые имеют гораздо более глобальный характер.

Европейский подход основан на концепции человекоориентированного и заслуживающего доверия ИИ. Это означает, что ИИ — не самоцель, а способ улучшить условия жизни людей. Идея заключается в том, что мы хотим ИИ, но не просто какой-нибудь ИИ. Общая цель — чтобы максимизировать возможности и минимизировать риски. Это хорошо согласуется с другими глобальными политическими установками в области ИИ.

Принцип доверия при разработке, внедрении и использовании систем искусственного интеллекта касается не только свойств, присущих технологии, но и качеств социотехнических систем, в которых применяется ИИ. Доверять или не доверять можно не только компонентам системы ИИ, но и системе в её общем контексте. Таким образом, стремление к созданию заслуживающего доверия ИИ подразумевает не только благонадёжность самой системы ИИ, но и требует целостного и системного подхода, предусматривающего благонадёжность всех участников и процессов, входящих в социотехнический контекст системы на протяжении всего её жизненного цикла (Группа экспертов высокого уровня по ИИ, 2019).

По мнению Группы экспертов высокого уровня по ИИ, в реализации заслуживающего доверия ИИ на протяжении всего жизненного цикла системы должны соблюдаться три основных аспекта:

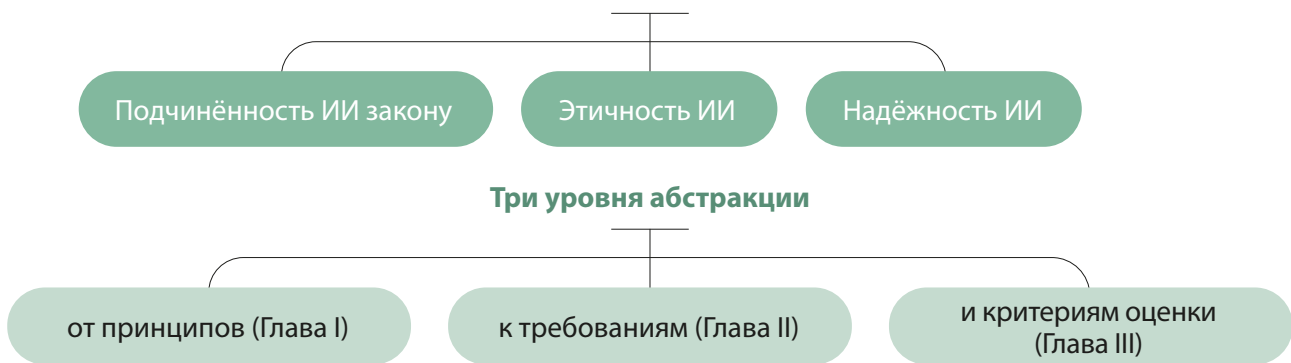
- 1) он должен быть *подчинён закону*, обеспечивая соблюдение всех применимых законов и норм;
- 2) он должен быть *этичным*, обеспечивая соблюдение этических принципов и ценностей; и
- 3) он должен быть *надёжным*, обеспечивая соответствие фактической реализации ожиданиям как с технической, так и с социальной точки зрения, поскольку даже при благих намерениях системы ИИ могут причинить непреднамеренный вред.

Обзор принципов заслуживающего доверия ИИ в концепции ЕС представлен на рис. 7. Этические рамки, предложенные группой экспертов высокого уровня по ИИ, имеют три уровня абстракции: от принципов высокого уровня до требований и обеспечения соответствия системы этим требованиям. Практическая реализация принципов высокого уровня является сложной задачей, поэтому эти шаги полезны для продвижения к реализации концепции.

### Рисунок 7. Рекомендации по этическим нормам для заслуживающего доверия ИИ — обзор

Человеко-ориентированный подход: ИИ как средство, а не цель

#### Заслуживающий доверия ИИ — наша основополагающая цель, включающая три компонента



Источник: составлено авторами на основе *Руководства по созданию этичного и заслуживающего доверия ИИ*, подготовленного Независимой группой экспертов высокого уровня по искусственному интеллекту Европейской комиссии.

Каждый из этих трёх компонентов сам по себе необходим, но недостаточен для построения заслуживающего доверия ИИ. В идеале все три компонента усиливают друг друга. Однако на практике могут возникать противоречия между этими элементами, например, когда нарушение закона может быть необходимо для спасения жизни людей или когда сфера действия и содержание существующего закона могут не соответствовать этическим нормам. Согласно этическим принципам, наша индивидуальная и коллективная ответственность как общества заключается в том, чтобы все три компонента способствовали достижению принципов заслуживающего доверия ИИ.

Представленные на рис. 8 четыре этических принципа заключаются в следующем:

1. *Уважение к независимости человека.* Это означает, что системы искусственного интеллекта должны дополнять, усиливать и расширять возможности человека. Это также означает, что людям должно быть позволено совершать ошибки и даже принимать неверные решения. Интересный пример — подталкивание. Если вы устанавливаете приложение, которое побуждает вас питаться более здоровой пищей или больше заниматься спортом, это, вероятно, хорошо, поскольку вы контролируете ситуацию. Однако если страховая компания просит вас установить такое же приложение, чтобы вы могли получить скидку на страховку, это уже вызывает сомнения. Можно пойти ещё дальше, например, обязать всех граждан страны использовать подобное приложение с целью улучшения здоровья населения, что вызовет ещё больше вопросов.
2. *Недопущение нанесения вреда.* Это, пожалуй, самый простой принцип, поскольку он гласит, что системы искусственного интеллекта должны не допускать причинение как физического, так и психического вреда человеку. Его использование должно быть безопасным.
3. *Справедливость.* Опять же, относительно очевидно, что мы хотим, чтобы системы ИИ были справедливыми, в том смысле, что выгоды и затраты должны распределяться равномерно. В то же время существует множество путей определения понятия справедливости. Например, хотим ли мы, чтобы результат работы системы был разделён поровну, или же поровну должны быть разделены ресурсы? Если мы хотим получить равенство результатов, то, скорее всего, придётся распределять ресурсы неравномерно, чтобы компенсировать индивидуальные различия.

4. *Понятность*. Четвёртый принцип наиболее сложен для определения, поскольку он связан с возможностью разобраться в том, что и от чьего имени делает система. Важными понятиями здесь являются прозрачность, объяснимость и прослеживаемость.


**Рисунок 8.** Четыре этических принципа, основанных на фундаментальных правах




Источник: составлено авторами на основе *Руководства по созданию этического и заслуживающего доверия ИИ*, подготовленного Независимой группой экспертов высокого уровня по искусственному интеллекту Европейской комиссии.


Реальная задача, конечно, заключается в том, как реализовать эти принципы на практике. Чтобы сделать шаг в этом направлении, Группа высокого уровня определила семь ключевых требований к заслуживающему доверия ИИ, основанных на этих принципах. Обзор этих требований представлен на рис. 9. Чтобы помочь организациям, использующим системы ИИ, и разработчикам в создании заслуживающих доверия систем ИИ, группа также разработала критерии оценки доверия к ИИ.


**Рисунок 9.** Список критериев оценки доверия к ИИ


- 


**Требование № 1. Человеческое участие и надзор**  
(Субъектность и независимость человека; человеческий контроль)
- 

**Требование № 2. Техническая надёжность и безопасность**  
(Устойчивость к атакам и защищённость; общая безопасность; достоверность; откатываемость и воспроизводимость)
- 

**Требование № 3. Техническая надёжность и безопасность**  
(Конфиденциальность; управление данными)
- 

**Требование № 4. Прозрачность**  
(Прослеживаемость; объяснимость; информативность)
- 

**Требование № 5. Разнообразие, недискриминация и справедливость**  
(Избегание несправедливой предвзятости; доступность и универсальный дизайн; участие заинтересованных сторон)
- 

**Требование № 6. Благополучие общества и окружающей среды**  
(Благополучие окружающей среды; влияние на занятость и навыки; влияние на общество в целом или демократию)
- 

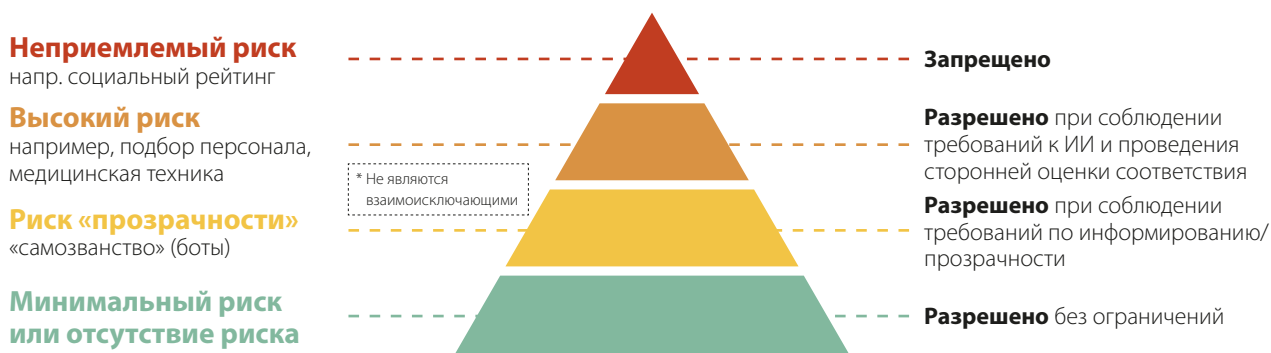
**Требование № 7. Подконтрольность**  
(Контролируемость; управление рисками)

Источник: составлено авторами на основе *Руководства по созданию этического и заслуживающего доверия ИИ*, подготовленного Независимой группой экспертов высокого уровня по искусственному интеллекту Европейской комиссии.



В настоящее время Европейская комиссия работает над законом об ИИ, который, как ожидается, станет первым в мире нормативным актом в области ИИ (обзор см. на рис. 10). Регулирование осуществляется на основе риск-ориентированного подхода, то есть объём регулирования зависит от уровня риска конкретного приложения, использующего технологии ИИ. К приложениям с высоким уровнем риска относятся те, в которых затрагиваются люди, особенно если речь идёт о благополучии отдельных людей. Утверждается, что некоторые приложения несут неприемлемые риски, такие как подсознательное манипулирование, оценка социального рейтинга и крупномасштабная удалённая биометрическая идентификация (например, распознавание лиц в общественных местах). По замыслу Комиссии, большинство приложений должны иметь низкий уровень риска. Однако, по мнению некоторых экспертов, эффект, скорее всего, будет обратным, т. е. большинство приложений будут считаться высокорисковыми, поскольку либо разработчик, либо пользователь системы предпочтёт подстраховаться и считать, что это приложение высокого риска. Приложения с высокой степенью риска подвергаются жёсткому регулированию, в то время как приложения с низкой степенью риска практически не подлежат дополнительному регулированию (помимо существующего). Требования к высокорисковым приложениям во многом основаны на требованиях к заслуживающему доверия ИИ.

**Рисунок 10.** Риск-ориентированный подход



Источник: Европейская комиссия.

Вопросы контроля и обеспечения соблюдения правил до сих пор остаются во многом нерешёнными. Согласно предлагаемому регламенту, в каждой стране должен быть свой компетентный орган, который может как давать рекомендации, так и следить за их исполнением. Две основные проблемы — нехватка высококвалифицированных специалистов на эти должности, и потенциальный конфликт в следствие полномочий по одновременному вынесению рекомендаций и контролю за соблюдением правил.

Принципы и рекомендации ОЭСР в области ИИ (рис. 11) способствуют обеспечению использования инновационного и заслуживающего доверия ИИ, а также соблюдения прав человека и демократических ценностей. Они посвящены тому, как правительства и другие участники могут сформировать человекоориентированный подход к созданию заслуживающего доверия ИИ. Являясь правовым документом ОЭСР, принципы выражают общее стремление стран-членов организации.

ОЭСР использует следующее определение ИИ:

Система ИИ — это машинная система, способная воздействовать на окружающую среду путём выработки выходных данных (прогнозов, рекомендаций или решений) для заданного набора целей. Она использует машинные и/или человеческие данные и входные параметры для 1) восприятия реальной и/или виртуальной среды; 2) абстрагирования воспринятых данных в модели путём анализа в автоматическом режиме (например, с помощью машинного обучения) или вручную; 3) использования выводов модели для формулирования вариантов результатов. Системы ИИ предназначены для работы в различных условиях автономности.

Это определение можно рассматривать как усовершенствованный вариант определения, используемого Европейской экспертной группой высокого уровня.

**Рисунок 11.** Принципы и рекомендации ОЭСР по ИИ

**Ценностно-ориентированные принципы**

	Инклюзивный рост, устойчивое развитие и благополучие
	Ценности, ориентированные на человека, и справедливость
	Прозрачность и объяснимость
	Надёжность, безопасность и защищённость
	Подконтрольность

**Рекомендации для директивных органов**

	Инвестиции в научные исследования и разработки в области ИИ
	Содействие развитию цифровой экосистемы для внедрения технологий ИИ
	Обеспечение благоприятной нормативно-правовой среды для развития технологий ИИ
	Создание кадрового потенциала и подготовка к выходу на рынок труда
	Международное сотрудничество в интересах создания заслуживающего доверия ИИ

Источник: оригинальные изображения из обзора «Принципы ИИ ОЭСР» основанного на Рекомендации Совета ОЭСР по искусственному интеллекту.

Рекомендация ЮНЕСКО об этических аспектах ИИ (рис. 12) — ещё один интересный пример «мягкого» правового регулирования. Это первый всемирно признанный свод руководящих принципов, в котором принят более широкий взгляд на сферу ИИ, включая необходимость обеспечения культурного разнообразия и соответствующего образования. Как и в руководстве ЕС, в нём выделены 11 важных областей, в которых требуется принятие стратегических мер, включая оценку этического воздействия, этическое управление и руководство, а также развитие и международное сотрудничество. В Рекомендации также подчёркивается необходимость осуществления мониторинга и оценки, включая разработку методологии ЮНЕСКО по оценке этического воздействия и готовности [к применению технологий ИИ в соответствии с Рекомендациями]. Вероятно, они станут важными инструментами для практической реализации технологий ИИ, поскольку компании и организации стремятся оценить как свою готовность, так и этическое воздействие своих продуктов и услуг.

ИИ — это поистине глобальное явление. Поэтому важно, что такие глобальные организации, как ЮНЕСКО, разрабатывают общие рамочные документы, которые получают широкое распространение и признание. Только совместными усилиями мы сможем использовать потенциал технологий ИИ.

**Рисунок 12.** Рекомендация ЮНЕСКО об этических аспектах ИИ

<p><b>Ценности</b></p> <ul style="list-style-type: none"> <li>• Уважение, защита и поощрение прав человека и основных свобод, а также человеческого достоинства</li> <li>• Благополучие окружающей среды и процветание экосистем</li> <li>• Обеспечение разнообразия и инклюзивности</li> <li>• Жизнь в мирных, справедливых и взаимосвязанных обществах</li> </ul>	<p><b>Принципы</b></p> <ul style="list-style-type: none"> <li>• Соразмерность и непричинение вреда</li> <li>• Безопасность и защищённость</li> <li>• Справедливость и недискриминация</li> <li>• Устойчивость</li> <li>• Право на неприкосновенность частной жизни и защиту данных</li> <li>• Подконтрольность и подчинённость человеку</li> <li>• Прозрачность и объяснимость</li> <li>• Ответственность и подотчётность</li> <li>• Осведомлённость и грамотность</li> <li>• Многостороннее и адаптивное управление и взаимодействие</li> </ul>	<p><b>Области регулирования</b></p> <ul style="list-style-type: none"> <li>• Оценка этического воздействия</li> <li>• Этическое управление и руководство</li> <li>• Политика в отношении данных</li> <li>• Развитие и международное сотрудничество</li> <li>• Окружающая среда и экосистемы</li> <li>• Гендерное равенство</li> <li>• Культура</li> <li>• Образование и научные исследования</li> <li>• Коммуникация и информирование</li> <li>• Экономика и рынок труда</li> <li>• Здоровье и социальное благополучие</li> </ul>
---	--	---

Источник: составлено авторами на основе рекомендации ЮНЕСКО об этических аспектах ИИ.

В этой главе дано краткое и вводное описание некоторых проблем, которые пытаются решить различные формы регулирования сферы ИИ, а также два взаимодополняющих подхода. Это очень динамичная область, и большинство стран и глобальных организаций работают над различными аспектами этих вопросов. Следует ожидать, что в ближайшие годы в этой области будет наблюдаться большая активность и развитие.

## Повышение общей осведомлённости об ИИ: создание открытого онлайн-курса



**Теему Роос, профессор, инициатор публичного открытого массового онлайн-курса «Элементы искусственного интеллекта» Университета Хельсинки, Финляндия**

В 2020 году мы запустили бесплатный массовый открытый онлайн-курс (МООК) под названием «Элементы ИИ», который к концу 2021 года посетили 750 тыс. человек. Цель курса, разработанного Университетом Хельсинки и ИКТ-компанией Reaktor, заключалась в демистификации ИИ. Мы хотели привлечь как можно более широкий круг людей к изучению того, что такое ИИ, что можно (и что нельзя) делать с помощью ИИ и как начать создавать методы ИИ.

Одним из первых импульсов стала Национальная стратегия ИИ («Эра ИИ»). В 2017 году правительство Финляндии приступило к реализации этой стратегии, одной из составляющих которой является повышение осведомлённости широкой общественности. Университету Хельсинки, как и другим вузам, был задан вопрос, есть ли у нас общедоступные образовательные ресурсы. У нас их не было, но мы провели несколько МООК по вопросам компьютерного программирования и кибербезопасности, и я подумал, почему бы не сделать такой же и по ИИ.

По счастливому совпадению, в то время мы встретили людей из компании Reaktor и поделились с ними этим планом. Они захотели помочь нам, так как считали, что необходимо повышать осведомлённость об искусственном интеллекте. Мы провели множество встреч и семинаров, чтобы разработать отличную концепцию, и в качестве цели выбрали 1% населения Финляндии. Смысл был в том, чтобы иметь конкретную, амбициозную цель, которая помогла нам донести до других серьёзность наших намерений и побудить других присоединиться к инициативе.

Мы хотели дать возможность большему числу людей приобщиться к технологиям. Возможно, кто-то захочет освоить навыки, с помощью которых он сможет начать решать задачи в своей работе с помощью искусственного интеллекта. Но ещё больше мы надеемся, что люди смогут сформировать своё мнение о том, какие технологии надо развивать и как они должны регулироваться.

Ускоренное внедрение ИИ создаёт серьёзные проблемы для неприкосновенности частной жизни, равенства и демократии. Алгоритмы социальных медиа на базе ИИ являются частью сложных социотехнических явлений, которые могут привести к поляризации общества; платформенные бизнес-модели могут привести к крайнему неравенству доходов среди рабочей силы (так называемые «работники кликов»); централизация вычислительных и финансовых ресурсов может даже бросить вызов существующим политическим системам...

Подобные проблемы не могут быть решены с помощью технологий. Однако понимание технологий, лежащих в их основе, является необходимым условием для их решения. Каждый должен иметь доступ к базовым знаниям, чтобы решение перечисленных задач не оставалось за самими разработчиками технологий.

Мы должны поддерживать общую осведомлённость: чтобы каждый имел доступ к базовым знаниям, которые лежат в основе социально значимых приложений. Кроме того, необходимы инициативы по перекавалификации и повышению квалификации: люди, которые хотят внести свой вклад в создание технических решений, должны иметь доступ к необходимым техническим навыкам. Это означает, что мы должны предлагать обучение, начиная с базовых и заканчивая продвинутыми навыками.

Важно отметить, что проект «Элементы ИИ» — это не только онлайн-курс, но и более широкая инициатива, включающая обширную сеть сотрудничества с многочисленными партнёрами в десятках стран. Это сопряжено со значительными координационными расходами и требует привлечения достаточного количества средств. Кроме того, проект является уникальным по сочетанию образовательной,

научной, промышленной, общественно-политической, коммуникационной составляющих, что не позволяет отнести его к каким-либо существующим категориям проектов. За счёт индивидуальной структуры финансирования и модели сотрудничества нам удалось добиться максимальной доступности курса (отсутствие платы за обучение конечных пользователей, поддержка на десятках языков и т. д.). Надеюсь, что этот успех побудит правительства и организации поддержать другие подобные проекты по всему миру.

Ссылка на MOOK «Элементы ИИ»: <https://www.elementsofai.com>.

Другие примеры MOOK — «ИИ и верховенство права», вводный курс, вовлекающий работников судебных органов в глобальную дискуссию о применении ИИ и его влиянии на верховенство права, доступный на семи языках ([https://www.judges.org/ai\\_and\\_law/english/](https://www.judges.org/ai_and_law/english/)), и «Защита прав человека в эпоху искусственного интеллекта», адресованный, в частности, молодёжи (<https://www.edapp.com/course/defending-human-rights-in-the-age-of-artificial-intelligence-2>).

Другие MOOK можно найти на таких платформах, как сайт курсов по журналистике Найт-центра журналистики в обеих Америках (<https://journalismcourses.org>), edX (<https://mooc.org>) и Coursera (<https://coursera.com>).



## Предлагаемые задания



**ПРОЧИТАТЬ И ПОРАЗМЫШЛЯТЬ.** Прочитайте предложенную ниже статью, в которой обсуждается потенциальное влияние технологий искусственного интеллекта на общество, и сгруппируйте возможные сбои систем машинного обучения в пять областей риска. Попробуйте привести собственные конкретные примеры, касающиеся этих областей риска.

Amodei, D., Olah, C., Steinhardt, J., Christiano, J.S. & Mané, D. (2016). *Concrete problems in AI safety*. arXiv, 25 July 2016.



**ОБСУДИТЬ.** Как технологии искусственного интеллекта могут способствовать достижению глобальных целей устойчивого развития (ЦУР), сформулированных Организацией Объединённых Наций (ООН) в 2015 г. 17 ЦУР — это: 1) ликвидация нищеты; 2) ликвидация голода; 3) хорошее здоровье и благополучие; 4) качественное образование; 5) гендерное равенство; 6) чистая вода и санитария; 7) недорогостоящая и «чистая» энергия; 8) достойная работа и экономический рост; 9) индустриализация, инновации и инфраструктура; 10) уменьшение неравенства; 11) устойчивые города и населённые пункты; 12) ответственное потребление и производство; 13) борьба с изменением климата; 14) сохранение морских экосистем, 15) сохранение экосистем суши, 16) мир, правосудие и эффективные институты, 17) партнёрство в интересах устойчивого развития.



**ВЫРАБОТАТЬ ИДЕЮ.** Подготовьте идею сюжета, посвящённого рекомендациям Европейского союза по заслуживающему доверия ИИ, для: а) национальной ежедневной газеты; б) журнала для девочек; в) подкаста для пожилых людей; г) реалити-шоу для коммерческого телеканала. Вы также можете воспользоваться информацией по своей стране в службе знаний Европейской комиссии AI Watch: [https://knowledge4policy.ec.europa.eu/ai-watch\\_en](https://knowledge4policy.ec.europa.eu/ai-watch_en).

Madiega, T. (2019). *EU guidelines on ethics in artificial intelligence: Context and implementation*. Brussels: European Parliamentary Research Service (EPRS). (Руководство ЕС по этике искусственного интеллекта: контекст и реализация.) Краткий обзор на англ. языке доступен по адресу [https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/640163/EPRS\\_BRI\(2019\)640163\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/640163/EPRS_BRI(2019)640163_EN.pdf)



**ПОИСКАТЬ И ИССЛЕДОВАТЬ ИСТОЧНИКИ.** Выберите одну платформу или онлайн-сервис — например, Facebook или Spotify — и изучите правила её работы: как она функционирует для пользователя? Какого рода объяснимость можно наблюдать в выбранной системе, и какие существуют недостатки с точки зрения объяснимости системы? Как журналисты могли бы отслеживать объяснимость компьютерных систем?



**СОЗДАТЬ.** Проведите телефонное интервью с разработчиком системы — приложения, платформы или цифрового сервиса, желательно небольшим местным стартапом — и попытайтесь добиться от него объяснений, как работает система, задавая открытые вопросы. Напишите короткую заметку, объясняющую работу системы. С какими трудностями вы столкнулись во время интервью и при подготовке репортажа?



## Дополнительная литература

- Amershi, S., Weld, D., Vorvoreanu, M., Fournery, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P.N., Inkpen, K., Teevan, J., Kikin-Gil, R., & Horvitz, E. (2019). *Guidelines for human-AI interaction*. CHI19: Proceedings in the CHI Coinference on Human Factors in Computing Systems, 1–13. Glasgow. <https://doi.org/10.1145/3290605.3300233>
- Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.
- Hagendorff, T. (2020). *The ethics of AI ethics: An evaluation of guidelines*. *Minds and Machines*, 30(1), 99–120. <https://doi.org/10.1007/s11023-020-09517-8>
- Jobin, A., Ienca, M., & Vayena, E. (2019). *The global landscape of AI ethics guidelines*. *Nature Machine Intelligence*, 1(9), 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- Liao, S. M. (Ed.) (2020). *Ethics of artificial intelligence*. Oxford University Press. <https://doi.org/10.1093/oso/9780190905033.001.0001>
- Loi, M. (2021). *Automated decision-making systems in the public sector: An impact assessment tool for public authorities*. AlgorithmWatch. <https://algorithmwatch.ch/en/adm-publicsector-recommendation/>
- Shane, J. (2021). *You look like a thing and I love you: How artificial intelligence works and why it's making the world a weirder place*. Little, Brown and Company.
- Taddeo, M., & Floridi, L. (2018). *How AI can be a force for good*. *Science*, 361(6404), 751–752. <https://doi.org/10.1126/science.aat5991>



## Модуль 4:

# Журналистика в алгоритмических культурах



## Краткое содержание

Данный модуль призван помочь студентам-журналистам понять и с критической точки зрения исследовать среду искусственного интеллекта и данных. В модуле рассматривается сущность и значение алгоритмов в современном обществе и в нынешней рыночной экономике, рассматриваются потенциальные проблемы и предлагаются способы их изучения в рамках журналистской работы. Далее обсуждаются трудности, с которыми сталкиваются журналисты в обществе, всё больше зависящем от алгоритмов, и рассматриваются способы, с помощью которых журналисты могут разобраться, разоблачить, исследовать и противостоять манипулированию процессом фильтрации информации со стороны алгоритмов.

**Ключевые понятия:** алгоритмическая культура, предвзятость алгоритмов, подотчётность, грамотность в области данных



## Цели модуля

Цели данного модуля:

- ▶ знакомство с понятием алгоритма и алгоритмической культуры и соотнесение их с дискурсами об искусственном интеллекте;
- ▶ выявление предвзятости алгоритмов в различных средах;
- ▶ расширение знаний обучающегося о принципах равного, справедливого и ответственного ИИ с точки зрения концепции неявной или нежелательной предвзятости.



## Результаты обучения

Ожидается, что по завершении модуля студент научится:

- ▶ понимать, как работают алгоритмические культуры и как они связаны с технологиями ИИ;
- ▶ исследовать и реагировать на препятствия, с которыми сталкиваются журналисты, когда пишут о темах алгоритмического общества;
- ▶ знать, что могут сделать журналисты, чтобы всесторонне освещать алгоритмическое общество, и как можно поддержать журналистов в их работе.



## Вопросы для размышления

Несколько вопросов, над которыми можно поразмыслить заранее:

1. Как по вашему мнению, является ли компьютерный алгоритм нейтральной технологией? Или необъективны только люди и их данные? Может ли алгоритм быть субъективным? Как и почему (или почему нет)?
2. Является ли проблемой то, что люди на Глобальном Юге зависят от алгоритмов, которые разработаны (и, следовательно, основаны) на ценностях, преобладающих в западном мире, или наоборот? Как ценности определяют производство приложений ИИ?

Несколько подготовительных заданий, которые необходимо выполнить перед тем, как приступить к изучению темы:

1. Посмотрите видеоролик, посвящённый «пузырям фильтров», например, выступление Илая Парайзера на TED (2011 г.): [https://www.ted.com/talks/eli\\_pariser\\_beware\\_online\\_filter\\_bubbles](https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles). Поразмышляйте над тем, как вы сами сталкивались с «пузырями фильтров» и можно ли было избежать этих ситуаций. В более общем плане, какие существуют решения, позволяющие избежать изоляции людей друг от друга, когда они получают определённую информацию и не видят альтернативных аспектов?
2. Найдите статистику, показывающую, как люди разного пола взаимодействуют с технологиями с точки зрения образования, профессий и специальных приложений — как разработчики и производители приложений и контента, а также как потребители и аудитория. Можно ли утверждать, что существует гендерная предвзятость, и если да, то почему, и откуда она берётся?



## Конспект

Автор: Фредерик Хейманс

Долгое время считалось, что фильтрация информации при её распространении — исключительно человеческое занятие. С течением времени эта практика претерпела изменения. Сначала с появлением средств массовой информации развилось кураторство новостей. После этого началось развитие цифровых СМИ, центром которых стала Всемирная паутина. В последнее время мы наблюдаем появление алгоритмов и манипулирование процессом фильтрации информации с их стороны.

По сути, алгоритм — это набор инструкций, правильное выполнение которых позволяет решить проблему или выполнить задачу. Алгоритмы находят всё более широкое применение в самых разных областях: реклама, поиск романтических партнёров, рекомендательные системы, кредитный и страховой рейтинг, оценка и анализ в образовании, процедуры приёма на работу, поддержание общественного порядка, управление социальным обеспечением и многое другое. Значительно возросшие вычислительные мощности, наличие больших объёмов данных, а также использование технологий искусственного интеллекта (ИИ) и соответствующих приложений, таких как машинное обучение и компьютерное зрение, привели к новому витку развития социально-экономической среды.

Алгоритмическая культура — это общество, в котором человеческое мышление, поведение, организация и самовыражение вписаны в логику больших данных и алгоритмов. Стрифас (2015) первым применил этот термин к современному социокультурному контексту. Стрифас рассматривает культуру как сортировку, классификацию и иерархизацию людей, мест, объектов и идей, а также связанные с этими процессами привычки мышления, поведения и самовыражения. По мнению Стрифаса, в последние годы культура всё чаще передаётся на аутсорсинг алгоритмическим процессам.

Возникновение алгоритмической культуры было сопряжено с процессом формирования общества. Люди постепенно стали считать алгоритмы и их функции нормой и принимать их. Было установлено, что во многих случаях люди даже предпочитают советы алгоритма советам человека (Logg et al., 2019). Такое принятие алгоритмов основано на том, что они часто упрощают жизнь. Алгоритмы постепенно становятся всё более точными и превращаются в надёжный информационный инструмент, который ранжирует, классифицирует, соотносит и фильтрует информацию.

Следующим шагом в эволюции публичной сферы в Интернете может стать переход от нынешней цифровой среды к иммерсивному пространству взаимодействия, также основанному на алгоритмах. Здесь мы различаем виртуальную реальность (технологии создания трёхмерной среды, которую можно увидеть, почувствовать и услышать), дополненную реальность (виртуальный слой, накладываемый на реальный мир) и метавселенную. Под метавселенной обычно понимается виртуальный мир, в котором люди могут общаться друг с другом в режиме онлайн. Это сеть виртуальных 3D-пространств, в которых пользователи могут выполнять различные действия с помощью аватара.

Уже создано несколько небольших метавселенных, но окончательной идеи общей виртуальной вселенной, в которой все эти приложения были бы связаны между собой, пока не существует. Это может занять некоторое время, поскольку ряд технических вопросов остаётся нерешённым. Будут ли эти новые формы взаимодействия приживаться так же быстро, как алгоритмы, пока неизвестно. Помимо технических препятствий, могут возникнуть вопросы об экологических последствиях, да и примут ли люди иммерсивную среду также остаётся неясным. Недавние исследования показывают, например, что работа в виртуальной реальности сопряжена со значительным дискомфортом (Kalamkar, et al., 2022). Хотят ли люди выполнять повседневные задачи в иммерсивной среде, или они просто хотят, чтобы эти задачи выполнялись быстро и эффективно? Вердикт по этому поводу ещё не вынесен.

### Когда алгоритмическая культура даёт сбой

Несмотря на признание алгоритмов, в алгоритмической культуре существуют различные проблемы. Среди проблем, которые могут быть вызваны алгоритмами, — неправильные прогнозы из-за некачественного программирования, усиление социального неравенства и дискриминации из-за необъективности данных, недостаток разнообразия и справедливого представления информации, распространяемой сре-



ди населения, и цензура. Кроме того, проблемы могут вызывать люди, разрабатывающие и программирующие алгоритмы, их личные интересы и предубеждения.

Никогда не следует полагать, что механизм работы алгоритма является нейтральным. Важно учитывать политэкономические аспекты алгоритмов (например, структуру собственности; O'Neill, 2016), которые часто определяют, какие выдаёт алгоритм. Алгоритм также может быть использован для манипулирования поведением человека. Это манипулирование может принимать различные формы, такие как персонализированные стратегии потребления товаров, действия, влияющие на эмоциональное состояние человека, и другие способы увеличения прибыли компании. Часто они сопровождаются продуманным дизайном и маркетинговыми стратегиями. Проблемы алгоритмической культуры рассматриваются далее на примере двух конкретных социальных сфер: государственной и корпоративной.

Во многих странах государственные органы используют алгоритмы для повышения эффективности работы. Простые задачи передаются на аутсорсинг, а такие ответственные задания, как выявление мошенничества, распределение социального обеспечения и оценивание учащихся в системе образования, теперь также доверяются алгоритмам. Использование алгоритмов в государственном контексте может быть многообещающим, однако существует множество примеров того, как это может обернуться проблемами, например, в Нидерландах (Geiger, 2018), Великобритании (Elbanna & Engesmo, 2020), Бразилии (Laranjeira de Pereira & Guimarães Moraes, 2022), Аргентине (Jemio, Hagerty & Aranda, 2022) и ЮАР (Merten, 2022). Алгоритмы могут быть необъективными, плохо спроектированы или реализованы с недостаточной этической ответственностью, что может привести к серьёзным ошибкам с тяжёлыми последствиями для тысяч людей.

Алгоритмы всё чаще используются для повышения эффективности деятельности, в том числе и в корпоративном секторе. Однако использование алгоритмов может привести к снижению качества труда работников. Например, алгоритмы всё чаще используются для мониторинга работы сотрудников и проверки их производительности. Такая практика получила широкое распространение во время пандемии COVID-19 и связанного с ней явления удалённой работы (Cater & Heikkilä, 2021). В другом примере алгоритм подбора персонала, который должен был проверять заявки на работу, не только сохранял, но и усиливал существующие предубеждения, дискриминируя женщин (Dastin, 2018). Кроме того, в результате использования программного обеспечения для составления графиков работы сотрудников, персонал испытывал острую нехватку времени (Kantor, 2014).

Мы не должны переоценивать влияние алгоритмов и поддаваться технологическому детерминизму. Например, сегодня большую озабоченность в обществе вызывает рост поляризации и дезинформации. Алгоритмы работы поисковых систем, социальных сетей и других цифровых платформ, работающих с персонализированным контентом, подозреваются в создании «пузырей фильтров». Пузырь фильтров создаёт ситуацию, когда пользователь Интернета видит только ту информацию, которая подкрепляет и соответствует его собственному мнению. Такие «пузыри» в основном создаются алгоритмами ранжирования, которые занимаются пассивной персонализацией без активного выбора. Однако ряд эмпирических исследований, основанных на данных опросов и отслеживания, показал, что использование механизмов фильтрации, таких как поисковые системы и социальные сети, в большинстве случаев ассоциируется с более разнообразным потреблением новостей, что противоположно гипотезе «пузыря фильтров» (Ross Arguedas et al., 2022).

Асерби (Acerbi, 2020) отмечает важность алгоритмов для общества, указывая на то, что внедрение цифровых технологий не приводит к таким масштабным и негативным изменениям в поведении людей, как это любят утверждать «цифровые фаталисты». Например, цифровые СМИ и алгоритмы позволяют каждому пользователю Интернета выходить на гораздо большее количество людей, чем традиционные средства коммуникации. Это не привело к неестественному социальному взаимодействию, и личные контакты по-прежнему доминируют в социальных отношениях (Acerbi, 2020).

## Проблемы алгоритмов с позиций журналистики

Рост алгоритмической культуры, безусловно, не облегчил работу журналистов и работников СМИ. Во-первых, внедрение новых технологий оказало серьёзное влияние на профессию журналиста. От современных журналистов требуется знание цифровых платформ, мультимедийных инструментов, Интернета и социальных сетей, а также умение быстро реагировать на новые события.

Более того, сегодня журналистам приходится работать в мире, где на принятие решений и выбор всё большее влияние оказывают алгоритмы. Поэтому журналистам важно понимать, как работают алгоритмы и какое влияние они оказывают на человека и общество. На журналиста возложена важная обязанность — распространять актуальную информацию, а это непростая задача, она требует глубокого понимания быстро развивающихся технологий, и алгоритм должен быть изучен в течение определённого времени. Кроме того, организации — не только платформы, но и государственные органы, — владеющие алгоритмами, редко предоставляют информацию о том, как они работают. Отчасти это связано с коммерческой тайной, но всё чаще — с тем, что системы не предназначены для объяснения своих решений (так называемые «алгоритмы чёрного ящика»). Алгоритмы часто непрозрачны, а техническая сложность является препятствием для журналистского освещения. Более того, нужно разбираться и в данных, используемых для обучения модели. Теперь, когда алгоритмы управляются сложными системами искусственного интеллекта, имитирующими поведение работника-человека, эта задача становится ещё более трудной. Наконец, сложно объяснить эти вопросы широкой общественности в простой и доступной форме.

Помимо сложности освещения событий, алгоритмизация ставит под сомнение актуальность журналистской работы как таковой. Одним из ярких примеров является культурная журналистика (Hu, 2018). Статус культурных обзоров находится под угрозой со стороны платформ и их алгоритмов. Например, кто будет читать музыкальные рецензии, когда Spotify ежедневно рекомендует новую музыку через различные плейлисты? Кто будет читать рецензии на фильмы, когда Netflix или Internet Movie Database (IMDb) подсказывают лично вам, что лучше выбрать?

С появлением цифровых новостей, сопровождаемых алгоритмическими технологиями, которые тщательно анализируют веб-трафик и активность, у редакций появилось больше информации о том, какие материалы предпочитает их аудитория. Эти показатели оказывают непосредственное влияние на стратегию новостных компаний. Действительно, традиционные редакционные материалы оказываются под давлением, поскольку показатели аудитории и внимание, уделяемое им на новостной площадке, влияют на журналистскую работу (Ferrer-Conill & Tandoc, 2018; Welbers et al., 2016).

Более того, на освещение новостей всё большее влияние оказывают социальные сети, такие как Twitter (McGregor, 2019). Высказанные там мнения часто ошибочно интерпретируются как общественное мнение. В политической журналистике (Broersma & Graham, 2015) использование социальных сетей привело к росту интереса к таким темам, как личная жизнь политиков, что ведёт к «деполитизации политики». С другой стороны, сами журналисты испытывают давление, заставляющее их рекламировать себя на таких платформах, как Twitter или Instagram, в попытке сохранить актуальность своей работы.

## Устойчивость журналистики в эпоху алгоритмов

Растущее присутствие алгоритмов в жизни общества требует от журналистов, как и от значительной части работающего населения, оттачивать свои навыки. Важными составляющими квалификации журналиста должны быть не только основы информационной и алгоритмической грамотности, но и элементы вычислительной журналистики, а также отработка основ журналистской профессии. В связи с этим возникает вопрос о том, каким образом журналисты могут переквалифицироваться.

Во-первых, может помочь создание структурированного инструментария по журналистскому освещению алгоритмов. Репортёры могут составить его самостоятельно или воспользоваться существующими наборами инструментов: Центр знаний «Данные и общество» во Фландрии и британская компания Journalism.ai разработали доступные наборы инструментов для ознакомления журналистов с темой искусственного интеллекта. Подобный инструментарий может включать в себя обширный глоссарий по тематике ИИ и репозиторий статей по освещению тематики ИИ или научных публикаций, с помощью которых журналист может ознакомиться с терминологией; список контактов экспертов по технологиям ИИ и заинтересованных сторон для опроса по теме репортажа (с разграничением специалистов по техническим, социальным/этическим и юридическим вопросам); список инструментов ИИ, с которыми можно экспериментировать (например, приложения для преобразования речи в текст или инструменты, которые можно использовать для разоблачения фальшивых новостей) и осваиваться с технологией; ссылки на профессиональные сообщества, которые делятся новыми тенденциями и идеями.

Во-вторых — нужно воспитывать критическое отношение к данным, которого, как показали некоторые исследования, часто не хватает журналистам (Simons et al., 2017; Stalph, 2017; Kõuts-Klemm, 2019). Точность

данных, как правило, недостаточно подвергается сомнению; это важно не только для интерпретации пресс-релизов, данных исследований или результатов деятельности компаний, но и для тщательной проверки работы алгоритмов в эпоху больших данных. Поэтому хороший уровень осведомлённости о качестве данных и грамотности в области данных является важным условием для выполнения журналистской миссии. Повышение уровня грамотности в области данных не требует обучения журналистике данных. Существует ряд организаций, предлагающих ограниченные бесплатные англоязычные курсы грамотности в области данных (см., например, University of Sheffield, 2022).

При этом возможности журналистики возрастают, если журналист способен оценивать алгоритмы через призму журналистики данных. Курсы для журналистов, пишущих о данных, различаются по содержанию, но часто включают такие компоненты, как элементарное понимание арифметики и концепции вероятности, передовые методы статистики и анализа данных, что позволяет журналистам глубже вникнуть в риски, присущие алгоритмам и лучше оценить в рамках своего расследования потенциальные проблемы. Хорошей отправной точкой является «Руководство по журналистике данных» (Bounegru & Gray, 2021). Также вдохновляют научные исследования о применении журналистики данных в новостных редакциях по всему миру, например, в Норвегии (Karlsen & Stavelin, 2014), Китае (Zhang & Feng, 2019), Пакистане (Jamil, 2019) и на Ближнем Востоке (Bebawi, 2019). В странах Глобального Юга существует ряд инициатив по развитию журналистики данных, таких как Code for Africa (Code for Africa, n. d.) и CFI, дочерняя компания группы France Médias Monde, содействующая развитию СМИ в Африке, арабском мире и Юго-Восточной Азии (CFI, n. d.).

Другой возможностью является изучение журналистики алгоритмической ответственности — нового направления журналистики, изучающего социальную власть, осуществляемую алгоритмами (Diakopoulos, 2014). Журналистское освещение темы алгоритмической ответственности (Diakopoulos, 2021), объединяющее различные методы исследования, представляет собой способ изучения, прояснения и озвучения структуры власти, аспекты предвзятости и влияния, которые алгоритмы могут оказывать на общество. Эффективное освещение темы алгоритмической ответственности требует всех традиционных журналистских навыков, связанных с подготовкой репортажей, проведением интервью и знанием предметной области, а также умения чётко и лаконично излагать полученные результаты. Также требуются новые навыки, такие как (хотя бы) минимальные знания в области журналистики данных (например, сбор и очистка данных) и использование передовых методов статистики.

### Руководство для журналистов, освещающих тему алгоритмов

- Ознакомьтесь с терминологией: читайте научные работы и публицистические статьи по теме ИИ, чтобы ознакомиться с правильными терминами; однако по возможности переводите технические и научные термины на повседневный язык и формулировки.
- Детализируйте картину: «истинное» положение дел может быть не таким простым; проблематизируйте слишком упрощённые установки об ИИ и консультируйтесь с несколькими экспертами и заинтересованными сторонами, т. е. выясняйте у специалистов их взгляды на проблему.
- Не поддавайтесь шумихе: рассматривайте технологии под критическим углом зрения, избегайте громких слов и задавайтесь вопросом: что это означает на практике или в жизни вашей аудитории?
- Оценивайте результаты исследований: рассказывая об исследованиях технологий искусственного интеллекта, поинтересуйтесь, как проводились исследования, какие данные использовались и насколько велик их массив, какие исходные посылы были у исследователей при проведении исследования и какие последствия может иметь применение данной технологии.
- Тестируйте: по возможности старайтесь получить личные практические знания о тех приложениях, о которых вы рассказываете, самостоятельно тестируя новые алгоритмы и системы.
- Следите за профессиональными онлайн-сообществами: национальные и международные журналистские сообщества, как институциональные проекты, так и группы коллег в социальных сетях, дают журналистам актуальные инструменты для размышлений и действий при освещении технологий и связанных с ними вопросов.

Источник Sivasdas & Argoub, 2021

## Исключены из данных, вычеркнуты из истории?



**Дженни Бергенмар**, профессор кафедры сравнительного литературоведения, аффилированный научный сотрудник Центра цифровых гуманитарных наук Гётеборгского университета (Швеция)

По мере увеличения объёма оцифрованных материалов возникает вопрос: какие знания можно получить в результате компьютерного анализа этих материалов? Возможностей много, но есть и риск в том, чтобы доверить анализ больших корпусов текстов статистическим и вероятностным методам.

Одной из проблем является неравномерность компиляции текстов в больших цифровых корпусах. В архивах часто отсутствуют сведения, оставленные маргинальными группами, и поэтому они рискуют вновь оказаться исключёнными из историографии. В глобальном масштабе также наблюдается острая неравномерность в представленности языков в оцифрованных корпусах текстов, причём английский язык как ведущий язык научных исследований имеет явное преимущество перед более мелкими языками. Кроме того, не все регионы мира могут позволить себе оцифровку своих национальных исторических и литературных источников, что также создаёт неравные условия, если при финансировании исследований слишком большое внимание уделяется методам, требующим больших объёмов данных. Хотя с помощью алгоритмического анализа можно выявлять определённые аспекты текстов в больших масштабах, однако есть и другие аспекты, которые требуют человеческого прочтения. Одним из примеров является репрезентация ЛГБТКИ-мотивов и персонажей в литературе. Чтобы их идентифицировать, необходимо быть восприимчивым к квир-метафорам и к тому, что остаётся недосказанным в тексте, поскольку на поверхности текста нет последовательных признаков, которые отражали бы идентичность и поведение ЛГБТКИ.

Наконец, существует проблема, связанная с тем, что технология сама становится носителем определённых позитивных ценностей. Алгоритмический анализ в некоторых случаях означает возвращение формализма и позитивизма в гуманитарные науки, поскольку существует идея, что компьютер способен к более объективному анализу, чем человек. Цифровые крупномасштабные методы анализа привели к переоценке эмпирических исследований, иногда в ущерб критическим взглядам, таким как феминистские, квир- и постколониальные.



## Предлагаемые задания



**ПРОЧИТАТЬ И ПОРАЗМЫШЛЯТЬ.** Ознакомьтесь со статьёй Николаса Дьякопулоса, в которой представлено «понятие журналистского освещения темы алгоритмической ответственности как механизма, позволяющего прояснить и озвучить структуры власти, аспекты предвзятости и влияния, которые оказывают на общество артефакты вычислительных методов», а также рассмотрены методы, которые журналисты могут использовать для выявления предвзятости в алгоритмах.

Diakopoulos, N. (2015). *Algorithmic accountability: Journalistic investigation of computational power structures*. *Digital Journalism*, 7(8), 1–23.



**ОБСУДИТЬ.** Какой властью обладает отдельный пользователь приложения? Как можно противодействовать и даже манипулировать алгоритмическими структурами? Возьмите для примера приложение с персонализированной лентой материалов и системой рекомендаций (например, Instagram, YouTube), изучите предложения, которые алгоритм выдаёт конкретному пользователю. Как вы думаете, на чём основаны эти рекомендации и как пользователь может вмешаться или изменить работу алгоритма?



**ВЫРАБОТАТЬ ИДЕЮ.** Если центральная платформа окажется закрыта: а) на один день; б) на более длительный срок, каковы будут последствия для различных организаций, групп и отдельных людей? Выберите одну платформу и представьте возможные последствия.



**ПОИСКАТЬ И ИССЛЕДОВАТЬ ИСТОЧНИКИ.** Проведите поиск в Интернете, и найдите актуальный пример предвзятости в работе систем искусственного интеллекта, в частности, предвзятости в отношении: а) пола (т. е. дискриминации алгоритмом женщин или транссексуалов), б) этнической или расовой принадлежности (т. е. дискриминации людей по цвету кожи), в) языка (т. е. невозможности преодоления алгоритмами языковых различий). Для каждого случая найдите возможные решения по устранению или минимизации предвзятости.



**СОЗДАТЬ.** Выберите тему, связанную с алгоритмами, на основе научной статьи в одном из специальных выпусков «Международного журнала интерактивных мультимедиа и искусственного интеллекта» (International Journal of Interactive Multimedia and Artificial Intelligence).

-> Искусственный интеллект, цифровой маркетинг и нейронауки, 5(6).

-> «Мягкие вычисления», 6(1).

-> Технологии искусственного интеллекта и блокчейна, 6(3).

-> Искусственный интеллект, духовность и аналоговое мышление, 7(1).



## Дополнительная литература

van Dijck, J. (2013). *The culture of connectivity: A critical history of social media*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199970773.001.0001>

van Dijck, J, Poell, T., & de Waal, M. (2013). *The platform society: Public values in a connective world*. Oxford University Press.

Fuchs, C. (2021). *Social media: A critical introduction*. 3rd ed. Sage. <https://dx.doi.org/10.4135/9781446270066>

Hristova, S., Slack, J. D., & Hong, S. (Eds.) (2020). *Algorithmic culture: How big data and artificial intelligence are transforming everyday life*. The Rowman & Littlefield Publishing Group.

Lindgren, S. (2017). *Digital media and society*. Sage.

Seyfert, R., & Roberge, J. (Eds.) (2016). *Algorithmic cultures: Essays on meaning, performance and new technologies*. Routledge. <https://doi.org/10.4324/9781315658698>



## Модуль 5:

# Освещение вредоносного использования технологий искусственного интеллекта



### Краткое содержание

В данном модуле рассматриваются вопросы, связанные с потенциально вредными последствиями использования технологий искусственного интеллекта. Место технологий в общественной сфере, неравенство в распространении информации и встающие более общие вопросы о целесообразности использования технологий вызывают широкий резонанс. Эти проблемы требуют от журналистов критического осмысления области ИИ, чтобы они могли освещать проблемы с неопределёнными последствиями. Как журналистика может отслеживать социальные события, связанные с ИИ? Данный модуль позволяет выявить и учесть этические проблемы, рассматривая такие аспекты, как чрезмерное использование средств ИИ и больших данных, онлайн-пропаганда, фальсификации, вредоносные боты, «тёмная Сеть» и недостатки в сфере кибербезопасности.

**Ключевые понятия:** этика, вредоносные боты, кибербезопасность, «тёмная Сеть», фальсификации, пропаганда в Интернете



## Цели модуля

Цели данного модуля:

- ▶ определение этических границ использования технологий, основанных на искусственном интеллекте;
- ▶ выявление и понимание различных явлений вредоносного ПО и вредоносного использования средств ИИ в общественной сфере в Интернете;
- ▶ ознакомление с новыми методами проверки и смягчения последствий с целью минимизации рисков и вреда от ненадлежащего использования ИИ.



## Результаты обучения

Ожидается, что по завершении модуля студент научится:

- ▶ выявлять структуры сетевой пропаганды, создаваемой средствами ИИ;
- ▶ выявлять поведение ботов и проверять материалы, созданные средствами ИИ;
- ▶ знать, как участвовать в дискуссиях об использовании средств ИИ со злым умыслом.



## Вопросы для размышления

Несколько вопросов, над которыми можно поразмыслить заранее:

1. Вспомните недавние примеры мошеннического использования данных в вашей стране или регионе. Что произошло и как это освещалось в журналистике?
2. Какие нежелательные эффекты использования технологий вы можете выявить в своей или чужой повседневной жизни и как преодолеть эти нежелательные эффекты?
3. Как разработчики платформ пытаются избежать причинения вреда и злоупотреблений? Рассмотрите одно конкретное приложение (например, Instagram) и перечислите меры, которые были приняты разработчиками платформы для предотвращения кибербуллинга, демонстрации обнажённой натуры, педофилии и т. д.
4. Будет ли лучше, если журналисты сосредоточатся только на положительных эффектах применения ИИ и будут меньше рассказывать о «плохом ИИ»? Существуют ли какие-то примеры ненадлежащего использования и злоупотребления технологиями ИИ, о которых, по вашему мнению, журналистам не следует писать? Почему?

Несколько подготовительных заданий, которые необходимо выполнить перед тем, как приступить к изучению темы:

1. Изучите инфографику, подготовленную некоммерческой организацией EAVI (Европейская ассоциация защиты зрительских интересов), и выясните, как с помощью средств искусственного интеллекта могут быть созданы десять типов вводящих в заблуждение новостей. Можно ли их обнаружить также с помощью инструментов или технологий ИИ?

Рисунок 13. За кулисами фальшивых новостей

# ЗА КУЛИСАМИ ФАЛЬШИВЫХ НОВОСТЕЙ

## 10 ТИПОВ ВВОДЯЩИХ В ЗАБЛУЖДЕНИЕ НОВОСТЕЙ

<b>пропаганда</b>  <ul style="list-style-type: none"> <li>используется правительствами, корпорациями и некоммерческими организациями для управления отношениями, ценностями и знаниями</li> <li>апеллирует к эмоциям</li> <li>может быть полезной или вредной</li> </ul>	<b>идеологически ангажированный материал</b>  <ul style="list-style-type: none"> <li>материал идеологического характера, содержит интерпретацию фактов, но с претензией на беспристрастность</li> <li>предпочтение отдаётся фактам, соответствующим установке, другие отвергаются</li> <li>эмоциональный и пафосный язык</li> </ul>	<b>НЕГАТИВНОЕ ВОЗДЕЙСТВИЕ</b> <ul style="list-style-type: none"> <li>нейтральное</li> <li>низкое</li> <li>умеренное</li> <li>высокое</li> </ul> <b>ПУБЛИЦИСТИЧЕСКИЕ ПРИЧИНЫ</b> <ul style="list-style-type: none"> <li>денеги</li> <li>политика/власть</li> <li>шутка/забава</li> <li>убежденность</li> <li>(дез)информация</li> </ul>
<b>приманка</b>  <ul style="list-style-type: none"> <li>броские, сенсационные заголовки, призванные отвлечь внимание</li> <li>часто вводит в заблуждение, и содержание может не соответствовать заголовку</li> <li>приносит доход от рекламы</li> </ul>	<b>теории заговора</b>  <ul style="list-style-type: none"> <li>попытка просто объяснить сложные реалии как реакция на страх или неопределённость</li> <li>не опровергаемо: опровергающие свидетельства рассматриваются как дополнительные доказательства заговора</li> <li>отвержение экспертов и авторитетов</li> </ul>	
<b>спонсированный материал</b>  <ul style="list-style-type: none"> <li>реклама, выдаваемая за редакционную статью</li> <li>потенциальный конфликт интересов для реальных новостных организаций</li> <li>потребители могут не распознать рекламный материал если он не имеет чёткой маркировки</li> </ul>	<b>лженаука</b>  <ul style="list-style-type: none"> <li>«эквитирательство», реклама чудодейственных лекарств, призывы против вакцинации, отрицание изменения климата</li> <li>искажение реальных научных исследований, преувеличенные или ложные заявления</li> <li>оппозиция экспертам</li> </ul>	
<b>сатира и мистификация</b>  <ul style="list-style-type: none"> <li>социальный комментарий или юмор</li> <li>широко варьируется по качеству, заложенный смысл может быть не очевиден</li> <li>может поставить в неловкое положение людей, принимающих это за правду</li> </ul>	<b>дезинформация</b>  <ul style="list-style-type: none"> <li>смесь фактов, лжи или частично ложных сведений</li> <li>целью может быть информирование, но автор может заблуждаться в истинности содержания</li> <li>ложные атрибуции, искажённые сведения и вводящие в заблуждение заголовки</li> </ul>	
<b>ошибочный материал</b>  <ul style="list-style-type: none"> <li>авторитетные новостные организации иногда допускают ошибки</li> <li>ошибки могут нанести ущерб бренду, оскорбить или привести к судебному разбирательству</li> <li>уважаемые организации публикуют извинения</li> </ul>	<b>фальшивка</b>  <ul style="list-style-type: none"> <li>полностью сфабрикованные материалы намеренно распространяемые с целью дезинформации</li> <li>тактика «партизанского маркетинга»; боты, фиктивные комментарии и поддельный брендинг</li> <li>движимо доходами от рекламы, политическим влиянием или тем и другим</li> </ul>	

### ДАЛЬШЕ – ХУЖЕ...

<b>ложная атрибуция</b> <ul style="list-style-type: none"> <li>подлинные изображения, видео или цитаты приписываются не тем событиям или лицам</li> </ul>	<b>введение в заблуждение</b> <ul style="list-style-type: none"> <li>содержание не соответствует заголовку и подписям</li> </ul>
<b>подделка</b> <ul style="list-style-type: none"> <li>веб-сайты и аккаунты в соц. сетях, выдающие себя за известный бренд или человека</li> </ul>	<b>искажение</b> <ul style="list-style-type: none"> <li>изменённые или подделанные материалы, такие как статистические данные, графики, фотографии и видео</li> </ul>

**eavi**  
 MEDIA LITERACY  
 for CITIZENSHIP  
 www.eavi.eu

Примечание. Распределение по категориям воздействия и причин не является абсолютным и должно использоваться только в качестве ориентиров для обсуждения.



Источник: EAVI (Европейская ассоциация защиты зрительских интересов).

- Найдите в Интернете обсуждения, связанные с ИИ и алгоритмами, и попытайтесь выяснить, чего боятся люди, когда речь идёт об использовании ими приложений и платформ на основе ИИ. Выявите не менее 2–3 примеров опасений и попытайтесь понять, насколько они реалистичны и как журналисты могли бы уменьшить эти опасения.
- Выясните, какие существуют автоматизированные инструменты для выявления фальшивых и мошеннических материалов на платформах, которыми пользуетесь вы и/или ваши друзья.





## Конспект

Автор: Сантош Кумар Бисвал

Несмотря на то, что за последние десятилетия было разработано множество стратегий, рекомендаций и практик, направленных на смягчение нежелательного использования и последствий применения технологий ИИ, а также на попытки заставить ИИ работать на благо общества, его применение всё ещё может нанести вред и вызвать нежелательные последствия. Поскольку ИИ состоит из действий, управляемых машиной и зависящих от намерений человека, он также может быть сознательно использован не по назначению. Вкратце можно сказать, что ИИ может использоваться не только во благо, но и во вред.

При освещении проблематики злоумышленного использования ИИ журналисты сталкиваются с вопросом о том, как правильно освещать вопросы, связанные с намеренно вредоносным содержанием и процессами. Как и преступные действия, случаи социальной несправедливости и потрясений должны быть донесены до сведения широкой аудитории, однако чувство тревоги по поводу последствий использования ИИ может также заставить людей испытывать беспокойство и страх перед «автоматизированным врагом». При рассмотрении злоумышленного использования систем ИИ журналисты обязаны контекстуализировать эти функции таким образом, чтобы злонамеренные действия не представляли слишком уж тревожащими. Журналисты должны видеть актуальность, объём и ожидаемые последствия таких действий в реалистичных пропорциях. Кроме того, общие правила журналистской этики, применяемые при освещении преступлений и актов правосудия, действуют и при освещении тематики ИИ: люди должны быть защищены, никто не должен публично подозреваться или преследоваться журналистами, а факты должны быть проверены, чтобы избежать распространения слухов. Ситуация усложняется тем, что дезинформация может генерироваться самим ИИ. На этом этапе журналистам необходимо знать, как работает искусственный интеллект. Знание «тёмной Сети», также известной как «глубокая», «невидимая» или «скрытая» Сеть, стало важным элементом грамотности в области ИИ как для журналистов, так и для граждан.

### ИИ как область технологий двойного применения

С точки зрения угроз, связанных с использованием ИИ, эксперты выделяют угрозы цифровой безопасности, физической безопасности и политической безопасности (Brundage et al., 2018):

- **Цифровая безопасность:** автоматизация атак социальной инженерии и обнаружения уязвимостей систем, более сложная автоматизация хакерских и уголовных киберпреступлений, атак на отказ в обслуживании, подобных проводимым человеком.
- **Физическая безопасность:** использование коммерческих систем в террористических целях, возможность увеличения и разнесения во времени и в пространстве объёма атак и связанного с ними ущерба.
- **Политическая безопасность:** использование государством автоматизированной системы наблюдения, приводящее к подавлению общественных дискуссий; фальшивые новости и сфабрикованные, но реалистичные видео- и аудиоматериалы; автоматизированные, гиперперсонализированные кампании дезинформации и влияния на общественное мнение; масштабные информационные атаки с использованием ботов; манипулирование доступностью информации и поведением пользователей.

Помимо этих вопросов безопасности, часто возникает вопрос о том, как сбалансировать вредные и полезные эффекты использования ИИ. Группа экспертов, собравшихся на семинаре в Оксфорде в 2018 году (Brundage, 2018), предложила альтернативный подход к распространённым представлениям об «апокалипсисе роботов» и «благоденствии автоматизации»: подход «двойного применения». Они определили ИИ как «технология двойного применения», что означает, что технологии ИИ могут быть использованы как с пользой, так и во вред, и положительные аспекты использования ИИ соотносятся с аспектами неправильного его использования. ИИ — это палка о двух концах.

Далее мы остановимся на двух аспектах ИИ-технологий, которые актуальны для журналистов при освещении тематики ИИ и работе в среде, управляемой ИИ: вредоносные программы и боты, а также самое последнее достижение в области повседневного применения ИИ — дипфейки («глубокие подделки»).

## Вредоносные программы и боты, информационный беспорядок

Вредоносные программы (вредоносное программное обеспечение) могут привести к утечке конфиденциальной информации, несанкционированному доступу к информационным системам, лишению пользователей доступа к информации и причинению вреда частным лицам и организациям. Как пользователи компьютеров, мы уже знакомы с вредоносным ПО, например, с компьютерными вирусами — программами, которые размножаются путём модификации других компьютерных программ и встраивания в них собственного кода. Существуют и другие типы вредоносных программ:

- *шпионские программы* собирают информацию о человеке или организации и отправляют её на другой компьютер с тем чтобы нанести вред пользователю;
- *рекламное ПО* показывает рекламу, принося доход рекламщикам, заставляя пользователя видеть рекламу или щёлкать по ней, при этом часто одновременно собирая и сообщая данные о пользователе для дальнейшей продажи;
- *лжеантивирусы* пытаются шокировать и вызвать тревогу у пользователя с помощью психологических манипуляций, чтобы заставить пользователя совершить действия, которые могут противоречить его интересам, например:
  - *программа-вымогатель* лишает пользователя доступа к системе или запугивает угрозами опубликовать его личные данные, если пользователь не заплатит выкуп;
  - *мошенническое ПО* вводит пользователей в заблуждение относительно наличия на их компьютере вирусов и убеждает заплатить за поддельное средство их удаления.
- *затиратели* уничтожают данные на жёстком диске компьютера пользователя.

Такие процессы, в рамках которых пользователей пытаются обманом заставить сделать что-то вредное для них самих, называются *социальной инженерией*. С ростом использования технологий искусственного интеллекта в различных системах и средах автоматизированные процессы, подобные этим, становятся всё более интенсивными. По данным Internet Security Report, количество вариантов вредоносного ПО растёт. Технологии ИИ становятся всё более доступными для субъектов деятельности, связанной с причинением вреда, и злоумышленники могут использовать нейронные сети для оптимизации процессов и снижения затрат на генерацию новых и высоко вариативных вредоносных программ. Например, в 2019 году калифорнийская исследовательская лаборатория OpenAI, основанная Илоном Маском, объявила о запуске GPT-2 — программы генерации текста, которая смогла написать убедительное эссе на тему, с которой исследователи были не согласны. Из-за выдающихся результатов работы робота лаборатория решила не публиковать полную версию программы, опасаясь, что она будет слишком эффективно использоваться для распространения фейковых новостей, спама и дезинформации.

Особым примером вредоносного ПО являются вредоносные, или «плохие» боты. Боты — это программные приложения, разработанные для выполнения автоматизированных заданий. Как правило, они используются для сканирования веб-страниц с целью сбора данных, т. е. боты, называемые также «пауками», систематически просматривают Интернет и с помощью автоматического скрипта собирают, анализируют и индексируют информацию с веб-серверов. Интернет переполнен такими поисковыми роботами, работающими в непрерывном режиме. Полезные боты агрегируют контент и выдают автоматизировано обработанную информацию, например, прогнозы погоды и расписание движения транспорта, результаты спортивных соревнований или нишевые новости, помогают в редактировании, например, Википедии или предупреждают о землетрясениях. Они также могут автоматически отвечать на запросы граждан и выполнять другие полезные услуги, отвечать на вопросы аудитории во время мероприятий, как, например, в случае так называемого «разговорного ИИ». Такие полезные боты могут, например, связываться с журналистами и рекомендовать им упоминать женщин, о которых до сих пор нет статей в Википедии (Flores-Saviaga et al., 2016).

Хотя часть ботов полезна по своей сути, существуют также боты, которые, помимо прочего, могут организовывать автоматические кибератаки или похищать сведения на манер вредоносных программ, о которых говорилось выше. Вредоносные боты могут похищать ценные материалы с других сайтов. Боты могут использоваться для искажения и дезинформирования маркетинговой аналитики, влияния на рейтинг сайта в поисковых системах, рассылки нежелательных сообщений на форумах и влияния на политические выборы. Вредоносные боты могут использоваться для различных видов преступной деятельности, вклю-

чая распространение наркотиков, детской порнографии и мошенничество с кредитными картами, а также в качестве инструмента слежки авторитарных правительств для осуществления различных репрессивных действий. Боты, взаимодействующие с пользователями, в том числе с журналистами, выполняющими свою повседневную работу, часто называют социальными ботами, поскольку они спроектированы таким образом, чтобы действовать подобно реальному человеку в социальной сфере (Stieglitz et al., 2017). Социальных ботов сложно обнаружить из-за их человекоподобного поведения, в отличие, например, от спам-ботов, которых довольно легко распознать, поскольку они публикуют большое количество практически одинаковых сообщений за короткий промежуток времени. С помощью искусственного интеллекта социальные боты могут научиться выдавать себя за пользователей-людей и тем самым становясь всё более трудноотличимыми от человека.

Многие социальные боты — это *боты влияния*, пытающиеся воздействовать на взгляды и поведение пользователей. Например, *астротурфинг-боты* [астротурфинг это имитация общественной инициативы, поддержки, запроса — прим. перев] пытаются создать у пользователя впечатление, что подавляющее большинство выступает за определённую позицию, даже если за ней стоит, например, политическая кампания, организованная одним человеком или ограниченной группой людей. Таким образом, под астротурфингом понимается намерение повлиять на мнение в ходе политической дискуссии. *Смоукрининг-боты* [маскировочные, отвлекающие интернет-боты, от англ. smoke screen — дымовая завеса, дымовая маскировка — прим. перев] используют определённые хэштеги в дискуссиях в социальных сетях, чтобы отвлечь или увести читателя от основной сути дискуссии. *Боты-сивиллы* — это боты, действующие в социальных сетях под поддельными личностями, представляющие из себя *виртуалов* — аккаунты с фальшивой личностью. Это может быть, например, дублирование аккаунта знаменитости. Боты также могут объединяться в сети, или «социальные ботнеты», в которых работа автоматизированных субъектов социального взаимодействия (ботов) синхронизирована.

Социальные боты могут манипулировать восприятием людей и провоцировать пользователей на сильные эмоциональные реакции. Боты могут автоматически ставить лайки или размещать большое количество материалов, а также генерировать текст в ответ на сообщения и комментарии других пользователей. Их влияние становится тем сильнее, чем больше их автоматически созданные тексты похожи на человеческие. Как автоматизированные подставные пользователи, они могут оказывать масштабное влияние на обсуждения по коммерческим или политическим поводам. Например, они могут выводить выбранные материалы на первые места в новостных лентах, результатах поиска и покупательских корзинах, продвигать определённую информацию и формировать повестку дня в интересах определённых групп. С учётом этого боты должны рассматриваться как проблема демократии.

## «Дипфейки» — дезинформация на основе технологий ИИ

Развитие технологий глубокого обучения привело к появлению новой и быстро развивающейся проблемы, которая наиболее ярко проявляется в создаваемых с использованием ИИ видеороликах, голосовых записях и данных, получивших название «дипфейки». Дипфейки, или «глубокие подделки» (от англ. deep learning — «глубокое обучение» и fake — «подделка») — это использование средств глубокого обучения для создания поддельных цифровых материалов. Это синтетические медиаматериалы, созданные с помощью процессов, управляемых искусственным интеллектом, путём искажения визуальных и речевых образов, и вызывающие сильные эмоции. В наиболее распространённом виде это видеоролики, в которых лицо человека заменено на сгенерированное компьютером. Помимо подмены лиц, использующей данные о лице в развлекательных целях, дипфейки можно встретить в криминалистике и порнографических материалах.

Иными словами, современные технологии искусственного интеллекта могут использоваться для создания текстов и изображений, которые выглядят всё более убедительно, и можно ожидать, что их использование будет совершенствоваться и становиться всё более изощрённым. С помощью так называемых генеративно-сопоставительных сетей (англ. Generative adversarial network, сокращённо GAN) ИИ может создавать сфабрикованные изображения, которые выглядят как настоящие, поскольку за последние несколько лет эта технология достигла значительного прогресса, как показано в твите Яна Гудфеллоу, изобретателя GAN, на рис. 14.

**Рисунок 14.** Твит Яна Гудфеллоу (14 января 2019 г.)

Источник: оригинальное изображение Яна Гудфеллоу — из его собственного сообщения в сети Twitter.

Дипфейки, будучи подделкой материалов, противостоят исторической достоверности изображений и визуальных материалов как свидетельств реальности. Дипфейки легко становятся вирусными и, иллюстрируя новостные события, как в случае с похищением ребёнка мужчиной на мотоцикле в Индии в 2018 году, могут вызвать массовую панику среди аудитории, что приводит к фатальным последствиям, в результате которых погибают люди (BBC, 2018). Производство дипфейков также, по-видимому, активизируется во время выборов, когда комментарии и выступления кандидатов в президенты становятся популярной темой для производства дипфейк-видео- и аудиозаписей. В 2021 году авторы документального фильма «Roadrunner» режиссёра Моргана Невилла, вызвали оживлённые этические дискуссии по поводу применения ИИ, когда воспроизвели сгенерированный ИИ голос покойного Энтони Бурдена, который умер в результате самоубийства в 2018 году, чтобы заставить его говорить посмертно. Синтезированный голос был использован без огласки, и даже если ИИ использовался только для озвучивания текста, написанного самим Бурденем, это вызвало множество откликов о возможностях использования ИИ. А что если «оживить» ещё больше умерших людей? Насколько это этично?

В 2021 году группа компьютерных учёных из стартапа, созданного на базе Ратгерского университета (США), обучила искусственный интеллект подражать музыкальному стилю Людвига ван Бетховена, чтобы закончить симфонию по его наброскам. Бетховен, умерший в 1827 году, так и не закончил свою 10-ю симфонию. Используя оставленные им наброски, в основном письменные ноты, и идею вариаций на тему — типичную структуру классических симфоний, — исследователи поставили перед искусственным интеллектом задачу обучения на основе музыкальных данных и поручили ему разработать совершенно новую часть симфонии. Профессор Ахмед Эльгаммал описал процесс обучения ИИ следующим образом: «Сначала мы обучили ИИ генерировать композицию в виде двух музыкальных строк, а не в виде полноценной симфонии, что является типичным способом работы композитора — сначала сочинение, а затем оркестровка. Затем мы создали другой ИИ, который взял эту композицию и научился её оркестровать. Я считаю, что это очень похоже на то, как учатся люди: невозможно освоить четвёртый уровень подготовки в колледже, не пройдя сначала первый, второй и третий уровни. Это всегда происходит постепенно» (Goodyer, 2021).

Эти примеры показывают, что ИИ может создавать несуществующие, но воображаемые новые миры, которые являются одновременно захватывающими и мощными по своему воздействию — несмотря на то, что это упражнение по музыкальной композиции было разработано для позитивных целей, существует также потенциал его использования для достижения негативных целей. В силу своего высокого эффекта реалистичности, или вторжения в наблюдаемую реальность, дипфейки представляют потенциальную угрозу для политической коммуникации в частности, и для демократии в целом. Так, использование дипфейков повышает потенциал усиления дезинформации, которая будет одновременно реалистичной и эффективной. Таким образом, существование дипфейков требует принятия активных мер по их развенчанию со стороны журналистов.

## Снижение угроз демократии

Негативный потенциал технологий ИИ может быть особенно опасен, если он используется в странах, находящихся на этапе перехода к демократии, где государственное управление ещё слабое, а нормативная база отсутствует. Если потенциал ИИ будет использован противниками государства и демократии для атаки на зарождающиеся демократические институты или процессы, последствия могут быть весьма масштабными. Например, можно представить себе, что сторонник террористической группировки может сфабриковать провокационную видео- или аудиозапись, приписываемую государственному органу, чтобы подлить масла в огонь политической или религиозной поляризации, а соответствующие разжигающие ненависть высказывания могут быть усилены на часто используемых социальных медиаплатформах. Беспилотники, управляемые искусственным интеллектом, автоматическое стрелковое оружие — всё это может стать средством осуществления насильственных действий при относительно низких затратах.

С другой стороны, технологии ИИ могут представлять угрозу для демократии и прав человека в авторитарных государствах, стремящихся подавить политических противников или иные нежелательные, маргинальные или уязвимые группы. Например, авторитарное правительство может использовать ИИ для усиления слежки, собирая данные и изображения лиц людей для использования в существующих приложениях для мониторинга. Автоматизированный мониторинг и наблюдение могут быть полезны, если кто-то подозревается в планировании актов вооружённого нападения в учебном заведении или распространении детской порнографии, но они также могут использоваться для наблюдения за людьми, которые ни в чём не подозреваются. Государственная киберразведка может использовать его для того, чтобы дать внутренним силам преимущества перед их иностранными контрагентами, или же ИИ может быть использован для влияния на общественное мнение в пользу собственных интересов.

Кроме того, отрицательные последствия массового автоматизированного производства и передачи информации могут возникать не только из-за ложных или фальсифицированных материалов. Вентурини (2019) предлагает термин «мусорные новости» для обозначения вирусных новостных материалов, которые насыщают общественные дискуссии и оставляет мало места для альтернативных обсуждений, вызывая сильные, но поляризующие эмоциональные реакции. Благодаря привлекательному эмоциональному воздействию и заманчивым заголовкам можно мобилизовать большую аудиторию с помощью крупномасштабного автоматизированного производства пропагандистских, экстремистских по своей идеологии или конспирологических (псевдо)новостей и информации, отодвигая на второй план некоторые более значимые с точки зрения демократических процессов проблемы.

Для снижения потенциала негативного воздействия технологий ИИ необходима общая рамочная основа для эффективного регулирования использования Интернет-технологий. В 2018 году корпорация Microsoft выступила с инициативой создания цифровой Женевской конвенции для стимулирования усилий отрасли и гражданского общества по защите гражданских лиц от организованных государством кибератак. В том же году в Парижском призыве к доверию и безопасности в киберпространстве был определён ряд целевых областей управления, включая защиту целостности Интернета, предотвращение вмешательства в избирательные процессы и пресечение наступательных действий в Сети со стороны негосударственных субъектов. В рамках Форума по управлению Интернетом, объединяющего правительства, частный сектор и гражданское общество для ведения диалога по вопросам мер политики, на регулярной основе обсуждаются вопросы, связанные с технологиями цифрового ИИ. Межправительственные организации могут помочь государственным структурам, занимающимся вопросами ИИ, наладить партнёрские отношения с предпринимателями и корпоративными структурами с целью повышения их квалификации,

совершенствования практики и процедур в области цифровой криминалистики. Совершенствование цифровой криминалистики может, в свою очередь, помочь государственным учреждениям быстрее выявлять сфабрикованный медиаконтент и удалять материалы провокационного содержания для деэскалации насилия. Многосторонние инициативы, разработанные с учётом конкретных потребностей, также могут помочь в создании наборов инструментов, способных поддержать сбалансированный демократический дискурс и мероприятия. Инструменты ИИ могут использоваться для обнаружения, сдерживания и удаления, например, экстремистских высказываний и других вредных форм поведения в Интернете как, например, в рамках проекта AI4Dignity, финансируемого Европейским исследовательским советом и направленного на борьбу с экстремистскими высказываниями в Интернете с помощью технологий. Эта идея была предложена совместно разработчиками систем ИИ, специалистами по проверке фактов, антропологами и директивными органами.

Учитывая многочисленные потенциальные риски и пагубные последствия злонамеренного использования технологий ИИ, тема подготовки к «обществу хорошего ИИ» приобретает всё большую актуальность. В связи с этим для обозначения сценариев благоприятного использования искусственного интеллекта используется термин «дружественный ИИ» (предложенный Юдковским, 2001). Его основная идея заключается в том, что использование ИИ может оказывать благотворное влияние на человечество, или, по крайней мере, соответствовать его интересам, или способствовать оздоровлению общества. В то время как вопросы этики ИИ, как это показано в модуле 3 «Нормативная база и рекомендации в области искусственного интеллекта» в этой публикации, касаются того, как должен вести себя субъект ИИ, чтобы не допустить пагубных последствий, концепция «дружественного ИИ» направляет наше внимание на задачи обеспечения адекватного ограничения таких последствий, ставя вопрос о том, как практически обеспечить позитивные формы поведения.

Накопление власти в руках немногих — власти денег, политической и религиозной власти или власти над инфраструктурами демократии и общественного дискурса — должно рассматриваться в совокупности, чтобы стать основой для обсуждения вопросов этики и законов применительно к ИИ. Действительно, этические проблемы и усилия по отслеживанию недостатков применения технологий ИИ должны обсуждаться в рамках общественного дискурса. Технологии ИИ должны использоваться в поддержку формирования политических убеждений людей и обмена ими, а не для внушения убеждений, основанных на корыстных интересах политической воли какого-либо субъекта. В этом контексте такие факторы, как права человека, слежка и свобода слова, являются более важными предметами обсуждения. Чтобы это стало возможным, в широкой перспективе необходимо развитие цифровой грамотности по ИИ, и журналисты являются участниками этой миссии.

### **Руководство по созданию более доброкачественного ИИ**

- Обеспечение цифровой, физической и политической безопасности граждан.
- Содействие сотрудничеству между политическими структурами, экспертными учреждениями в области ИИ и научными кругами для проверки нежелательных аспектов функционирования средств ИИ.
- Формулирование передовых практик, которые могли бы быть полезными с точки зрения обеспечения разумного использования средств ИИ.
- Привлечение и обсуждение всеми заинтересованными сторонами общества вопросов применения средств ИИ.
- Обеспечение качества и прозрачности информации в процессе использования ИИ.
- Устранение цифрового неравенства и обеспечение пространства равенства.
- Обеспечение свободы слова, собраний и объединений, а также права на труд.
- Повышение грамотности в области ИИ в глобальном масштабе.

Как журналист может способствовать достижению этих глобальных целей?

## Почему при освещении тематики ИИ важно наличие аналитического скептицизма



**Стив Вулгар**, профессор Линчепингского университета, Швеция

Зачастую новые технологии быстро теряют свою актуальность. Если при появлении новые технологии вызывают большой интерес и часто сопровождаются шумихой, то после даже незначительного устаревания они представляют гораздо меньший интерес. Важным следствием этого является то, что общие выводы о появлении и восприятии технологий могут быть легко утеряны. Сейчас мало кто интересуется социальной динамикой киберкафе, однако анализ шумихи, поднятой вокруг их потенциального влияния, многое проясняет в отношении технологий в целом.

В 1997–2002 гг. я руководил национальной междисциплинарной программой Великобритании по исследованию социальных последствий (тогда ещё) новых электронных технологий: 22 проекта в британских университетах и 3 за рубежом. Основное положение и, следовательно, актуальность исследований заключались в том, что множество новых электронных технологий должны были радикально изменить общество, культуру, организации, самоопределение личности и т. д. Финансирующий орган, Совет по экономическим и социальным исследованиям (ESRC), назвал эту инициативу «Программа “Виртуальное общество”».

Как будущий руководитель программы, я поставил под вопрос первоначальное положение о массовом переходе к виртуальному обществу. Не следует ли с некоторой осторожностью отнестись к такой посылке с самого начала? Однако формулировка названия исследования, одобренная различными комитетами, оставалась неизменной. Как же избежать технологического детерминизма и в то же время учесть ожидание широкомасштабных последствий? После некоторого обсуждения мне удалось убедить ESRC сохранить название, но добавить вопросительный знак: «Программа “Виртуальное общество?”».

Исследования, проведённые в рамках программы «Виртуальное общество?», содержали некоторые удивительно контринтуитивные выводы. Например, что появление сетей электронной почты способствовало увеличению, а не уменьшению количества поездок (так называемая «смерть расстояния»); переход к онлайн-технологиям способствовал увеличению (а не уменьшению) частоты физических посещений картинных галерей и музеев; а внедрение Интернет-технологий далеко не повсеместно среди определённых групп потенциальных пользователей (Woolgar, 2002). В 2000 году этот последний вывод вызвал бурю в средствах массовой информации. Офис нашей исследовательской программы был переполнен журналистами, которые спрашивали, не означает ли это конец компьютеров!

Несмотря на известное различие между практикой работы учёных и журналистов — «худшее, что можно сказать учёному (журналисту), это то, что он пишет как журналист (учёный)», — опыт работы над «Виртуальным обществом?» показывает, что обеим сторонам необходимо наличие аналитического скептицизма: позиция отстранённого размышления перед лицом заманчивых заявлений о новизне и значимости. Нам нужен этот знак вопроса. Это особенно важно теперь, когда речь идёт об искусственном интеллекте, поскольку предполагаемые последствия его применения не только влияют на общество, но и ставят под сомнение наши самые глубокие предположения о том, что значит быть человеком.

Важно помнить, что то, что мы считаем искусственным интеллектом сегодня, — это одна из целого ряда различных волн развития ИИ (Woolgar, 1985), и что ИИ — это лишь позднейшее явление в исторической череде явлений, рассматриваемых как вызов сущности интеллекта, или человеческой сущности, и предполагаемому отличию человека от животных и машин. Мы можем вспомнить механических кукол XVII века, умную лошадь Ганса, говорящих обезьян и т. д.

Как же тогда говорить о перспективах развития ИИ? Мне вспоминается одна конференция конца 1990-х годов, на которой присутствовали в основном техноэнтузиасты из ИТ-компаний. Когда участникам предложили порассуждать о будущем технологий, многие из них предложили видение кардинальных изменений в обществе, как положительных, так и отрицательных. Однако гораздо более убедительное видение предложил сэр Ричард Сайкс. По его словам, через 20 лет мы все будем гораздо спокойнее относиться к новым технологиям и гораздо меньше пугаться киберболических\* рассуждений об утопических и антиутопических последствиях применения технологий. В чём же причина

такого спокойствия и невозмутимости, того хладнокровия, с которым будущие поколения будут реагировать на преувеличенные заявления о новых технологиях? В воспитании и влиянии принципов аналитического скептицизма.

Если судить по недавним «Лекциям Рейта» по BBC, посвящённым вопросам искусственного интеллекта, то предсказание Сайкса ещё не сбылось. Сообщая о новейших технологических явлениях, мы должны стараться не забывать о вопросительном знаке в словосочетании «Искусственный интеллект?»

\* Кибербола: неологизм, обозначающий преувеличение (гипербола) характера и последствий применения кибертехнологий



## Предлагаемые задания



**ПРОЧИТАТЬ И ПОРАЗМЫШЛЯТЬ.** Прочтите указанную ниже статью о ботах Википедии и их поведении. Можете ли вы обнаружить поведение ботов в Википедии? Попробуйте определить пользователей, которые, вероятно, являются ботами, и отследить их правки. Какой вклад они внесли в статьи этой онлайн-энциклопедии и чем их вклад, возможно, отличается от вклада людей?

Tsvetkova, M., García-Gavilanes, Floridi, L. & Yasseri, T. (2017). *Even good bots fight: The case of Wikipedia*. PloS one 12(2).



**ОБСУДИТЬ.** Медиа- и информационная грамотность часто предлагается в качестве средства решения проблем, возникающих у людей при работе с вредоносным контентом в Интернете. Может ли медиаграмотность, или её подотрасль «ИИ-грамотность», помочь в борьбе с мошенническими действиями? Каким образом, и какие при этом возникают проблемы? Владение какими навыками может включать в себя «ИИ-грамотность»?



**ВЫРАБОТАТЬ ИДЕЮ.** Определите контекст или ситуацию, в которой дипфейки определённого типа могут быть использованы: а) в благих целях; б) в злонамеренных целях.



**ПОИСКАТЬ И ИССЛЕДОВАТЬ ИСТОЧНИКИ.** Найдите в Интернете инструменты проверки и воспользуйтесь указанными ниже справочниками для журналистов, чтобы узнать о существующих программах для выявления: а) дипфейков; б) фальшивых аккаунтов (виртуалов) в популярных социальных сетях; в) «мусорных новостей».

Silverman, C. (ed.) (2014). *Verification handbook: A definitive guide to verifying content for emergency coverage*. European Journalism Centre (EJC). Available at <http://verificationhandbook.com>

Silverman, C. (ed.) (2015). *Verification handbook for investigative reporting: A guide to online search and research techniques for using UGC and open source information in investigation*. European Journalism Centre (EJC). Available at <https://datajournalism.com/read/handbook/verification-2>



**СОЗДАТЬ.** Напишите мини-эссе (текст с изложением мнения) или мини-очерк (включая интервью), в котором расскажите о том, что нам известно о слежке со стороны государства, сосредоточившись на примере Китая, России или США. Обсудите, как вариант, на примере недавнего случая, получившего широкое освещение, какую роль в этом играет ИИ, как это влияет на отдельных граждан и каковы перспективы на будущее.





## Дополнительная литература

- Brundage, M. et al. (2018). *The malicious use of artificial intelligence: Forecasting, prevention, and mitigation*. Future of Humanity Institute. <https://arxiv.org/pdf/1802.07228.pdf>
- Diakopoulos, N. (2021). *The algorithms beat: Angles and methods for investigation*. In: Gray, J., & Bounegru, L. (eds.) *Data handbook 2: Towards a critical data practice*. <https://datajournalism.com/read/handbook/two/>
- Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.
- Graves, L. (2018). *Understanding the promise and limits of automatic fact-checking*. Technical report. Reuters Institute for the Study of Journalism. [https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2018-02/graves\\_factsheet\\_180226%20FINAL.pdf](https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2018-02/graves_factsheet_180226%20FINAL.pdf)
- Horák, A., Baisa, V., & Herman, O. (2021). *Technological approaches to detecting online disinformation and manipulation*. In: Gregor, M., & Mlejnková, P. (Eds.) *Challenging online propaganda and disinformation in the 21st century*. Palgrave Macmillan, 139–166. <https://doi.org/10.1007/978-3-030-58624-9>
- Reuter, C. (Ed.) (2019). *Information technology for peace and security: IT applications and infrastructures in conflicts, crises, war, and peace*. Springer Vieweg. <https://doi.org/10.1007/978-3-658-25652-4>
- Yampolskiy, R. V. (Ed.) (2019). *Artificial intelligence safety and security*. CRC Press. <https://doi.org/10.1201/9781351251389>



## Модуль 6:

# Модели подачи материала об искусственном интеллекте



## Краткое содержание

В данном модуле обсуждается вопрос ИИ как объекта новостного репортажа и темы общественного освещения, а также связанные с этим эстетические и практические аспекты подачи информации. Поскольку журналистика играет важнейшую роль в донесении до общественности идей, связанных с новыми технологиями, журналистам важно уметь выбирать общественно значимые темы и подавать их так, чтобы они не были преувеличенными или сенсационными. В этой главе даётся представление о том, как тематика ИИ освещалась в профессиональных СМИ до сих пор, и обсуждаются последствия и следствия различных журналистских решений. В ней также рассматриваются некоторые основные недостатки в инфраструктуре журналистского производства и освещения технологий ИИ и предлагается будущим журналистам задуматься о своей роли как потенциальных проводников изменений и воспитателей аудитории.

**Ключевые понятия:** эстетика, подача материала, визуализация, журналистские роли, специализация



## Цели модуля

Цели данного модуля:

- ▶ изучение того, как тематика ИИ обычно представляется в журналистике, и анализ сильных и слабых сторон существующего информационного освещения сферы ИИ;
- ▶ понимание того, как повторяющиеся модели подачи материала в журналистике влияют на публичную картину ИИ;
- ▶ поощрение студентов к созданию репортажей об ИИ.



## Результаты обучения

Ожидается, что по завершении модуля студент научится:

- ▶ анализировать практику подачи материала и журналистской работы, связанной со сферой ИИ;
- ▶ понимать ограничения и возможности журналистского освещения тематики ИИ, а также оценивать качество такого освещения;
- ▶ выработать идеи по сбалансированному, грамотному и разностороннему освещению сферы ИИ.



## Вопросы для размышления

Несколько вопросов, над которыми можно поразмыслить заранее:

1. С какими профессиональными и личными трудностями вы сталкиваетесь, когда обращаетесь к новым технологиям и приложениям в качестве потенциальной темы для освещения?
2. Попробуйте вспомнить недавно прочитанные вами статьи об искусственном интеллекте. Какие закономерности вы заметили? Попробуйте воспроизвести типовые статьи об ИИ в следующих жанрах: а) новостная заметка; б) телевизионный репортаж; в) портрет человека; г) редакторская колонка; д) журналистский твит.
3. Поищите в Интернете изображения ИИ (используйте, например, поиск изображений Google или банк фотографий). Как бы вы описали типичные изображения, демонстрирующие ИИ?
4. Если бы вы хотели продать новостному изданию материал по теме искусственного интеллекта, в какие отделы вы бы обратились в первую очередь?

Несколько подготовительных заданий, которые необходимо выполнить перед тем, как приступить к изучению темы:

1. 1. Найдите в газете или журнале статью, в которой обсуждается тема ИИ. Каким образом ИИ представляется в ней как предмет журналистского исследования? Проанализируйте статью с точки зрения используемых в ней стратегий ведения дискурса:
  - а) Какую основную мысль об искусственном интеллекте пытается донести статья?
  - б) Какие термины и выражения используются в отношении ИИ? Как они характеризуют ИИ?
  - в) С кем проводились интервью и каковы основные источники информации? у кого ещё можно было бы взять интервью и существует ли предвзятость в выборе источников?
2. Найдите 4–5 статей об ИИ в одном информационном издании или в разных изданиях. Обратите внимание на иллюстрации к ним. Какие закономерности вы заметили? Как по-другому можно было проиллюстрировать статьи и почему, по вашему мнению, были выбраны именно эти способы, а не другие?



## Конспект

Автор: *Маарит Яаккола*

Журналистика играет ключевую роль в донесении до общественности основных идей и представлений о зарождающихся технологиях, таких как ИИ. Особенно в контексте развивающихся технологий часто возникают большие сомнения и даже дезинформация в отношении сути новшеств. Поскольку обычные люди, как правило, не имеют прямого и явного опыта работы с ИИ, даже если они с ним соприкасаются, широкая общественность в значительной степени зависит от журналистского освещения сферы ИИ. Общественные смыслы и образы способствуют пониманию обществом роли технологий в различных секторах общества и последствий таких новых явлений, как большие данные и глубокое обучение. Поэтому важно проследить динамику, возможную предвзятость и закономерности в освещении этой темы, которые возникают в повседневной журналистской практике. Как СМИ обычно освещают ИИ?

Новые технологии не всегда легко освещать, на них обычно возлагаются надежды и опасения — особенно это касается технологий искусственного интеллекта. В условиях жонглирования пессимистическими и оптимистическими взглядами всегда существует высокая степень неопределённости в отношении того, что на самом деле представляют собой новые технологии, как они работают и какие последствия они могут иметь для граждан и общества в долгосрочной перспективе. Бывает трудно оценить, что действительно важно и значимо для общества или демократии, а что является просто «шумихой». Наблюдается динамика в отношении новых изобретений, когда первоначально ожидания раздуты и привлекают широкое внимание, которое со временем сначала сменяется разочарованием и критикой, а затем постепенно перерастает в этап просвещения и понимания новой технологии. Динамика, представленная на рис. 15 и названная циклом зрелости технологии Gartner (Gartner's hype cycle) по названию американской компании Gartner, которая его описала (Fenn, 2010), применима даже к «зимам ИИ», о которых говорилось в модуле 1 «Определение понятия искусственного интеллекта», и которые можно понимать как «пропасти разочарования» в технологиях ИИ. Модель S-образной кривой внимания не может быть точным описанием степеней внимания к той или иной инновации, которые в разных случаях могут быть разными, что также было показано в эмпирических исследованиях внимания СМИ к появляющимся технологиям (Shi & Herniman, 2023).

Рисунок 15. Внимание СМИ к развивающимся технологиям (Gartner's hype cycle)



Источник: составлено Ольгой Тарковской (CC-BY-SA 3.0) на основе оригинального материала Gartner.com.

## Текстовое представление темы ИИ

Типичными темами, в которых регулярно отслеживается состояние технологий ИИ, являются здравоохранение, наука, экономика, занятость и рынок труда, бизнес и технологии, образование, а также СМИ и культурная индустрия (Zhai et al., 2020; Fast & Horvitz, 2017; Chuan et al., 2019). Если говорить более конкретно, то к темам, которые, как показал эмпирический анализ общей ситуации в журналистской работе по их освещению, представляют особый интерес, относятся прежде всего темы использования технологий ИИ в беспилотных транспортных средствах, чат-ботах и других приложениях, компьютерных играх, мобильных устройствах, в ходе военных действий, а также в развитии вычислительной техники и философии сознания (Vergeem, 2020).

В рамках ряда исследований было обнаружено, что в ежедневных газетных публикациях доминируют отраслевые источники, а дискуссии о технологиях ИИ, как правило, ведут заинтересованные лица мужского пола, которые могут получить финансовую выгоду от внедрения технологий ИИ (Brantner & Saurwein, 2021; Brennen et al., 2018; Ouchchy et al., 2020). В ходе исследования характера освещения ИИ в Великобритании в отчёте Института изучения журналистики Reuters (Brennen et al., 2018) было установлено, что почти 60% новостных статей в различных изданиях освещали новые продукты или инициативы отрасли. Диапазон продуктов с использованием ИИ, привлёкших внимание журналистов, включал как смартфоны, так и кроссовки, секс-роботов и устройства для сохранения мозга, причём освещение последних носило более сенсационный характер, чем обычных устройств. Треть уникальных источников в статьях была связана с отраслью, почти в два раза больше, чем с научными кругами, и в шесть раз больше, чем с правительственными структурами. Почти 12% всех статей содержали ссылку на Илона Маска. Преобладание инсайдерских и экспертных мнений над неэкспертными свидетельствует о том, что крупные мировые индустрии более оперативны и способны обеспечить видимость и доступность инноваций для новостных служб в качестве потенциальных тем для освещения, в то время как национальные правительства менее склонны к организации информационных мероприятий, которые могли бы дать журналистам интересные темы для обсуждения. Очевидно, что дискуссия должна быть более разнообразной и многосторонней.

ней с точки зрения источников информации, и журналисты должны сознательно предоставлять место не только крупным отраслевым источникам. Активисты, неправительственные организации (НПО) и другие представители гражданского общества, а также потребители должны включаться в репортажи, поскольку они непосредственно испытывают последствия применения технологий ИИ в своей повседневной жизни. Помимо освещения деятельности представителей отрасли и компаний, в новостных текстах технологии ИИ, как правило, персонифицируются или антропоморфизируются. В работе Bunz и Braghieri (2019) рассмотрены материалы об использовании систем искусственного интеллекта в здравоохранении, опубликованные в трёх англоязычных газетах Великобритании и США за четыре десятилетия (1980–2019 гг.), и установлено, что тенденция обращения к системам искусственного интеллекта как к личности постепенно эволюционировала. В первые два десятилетия системы ИИ ещё обсуждались скорее как «мозг», нежели как «человек», но в последние годы доминирующей формой стал образ человека. Это может быть связано с тем, что системы ИИ — даже если они являются слабыми системами ИИ — всё больше конфигурируются под активную роль, обладая определённой самостоятельностью и собственным именем. Исследователи предполагают, что освещение событий, связанных с разработкой новой технологии, которая пока носит спекулятивный характер, демонстрирует более сильную тенденцию к антропоморфизации по сравнению с освещением конкретных технологий, которые уже тестируются или используются. Из этого предположения вытекает более масштабный вопрос: становится ли образ систем ИИ тем менее антропоморфным и чуждым, чем больше они интегрируются в нашу повседневную жизнь?

Исследования показывают, что явление ИИ также рассматривается как область, связанная с ценностями, и рассматривается через призму политики. Согласно отчёту Института изучения журналистики Reuters (Brennen et al., 2018), праволиберальные издания склонны выделять вопросы экономики и геополитики, включая автоматизацию, национальную безопасность и инвестиции, в то время как леволиберальные издания подчёркивают этические аспекты использования ИИ, включая вопросы дискриминации, алгоритмической предвзятости и конфиденциальности. Исследования показывают, что в освещении вопросов ИИ не прослеживается алармистских, пессимистических или антиутопических настроений, подчёркивающих риски и угрозы, а, скорее, наблюдается тенденция к оптимизму (Brantner & Saurwein, 2021; Garvey & Maskal, 2020; Vergeer, 2020; Fast & Horvitz, 2017). Однако освещение событий может иметь непреднамеренные последствия: стремление к нейтральности, ориентированность на человека или отраслевую перспективу может привести к раскрытию структур власти [технологий ИИ].

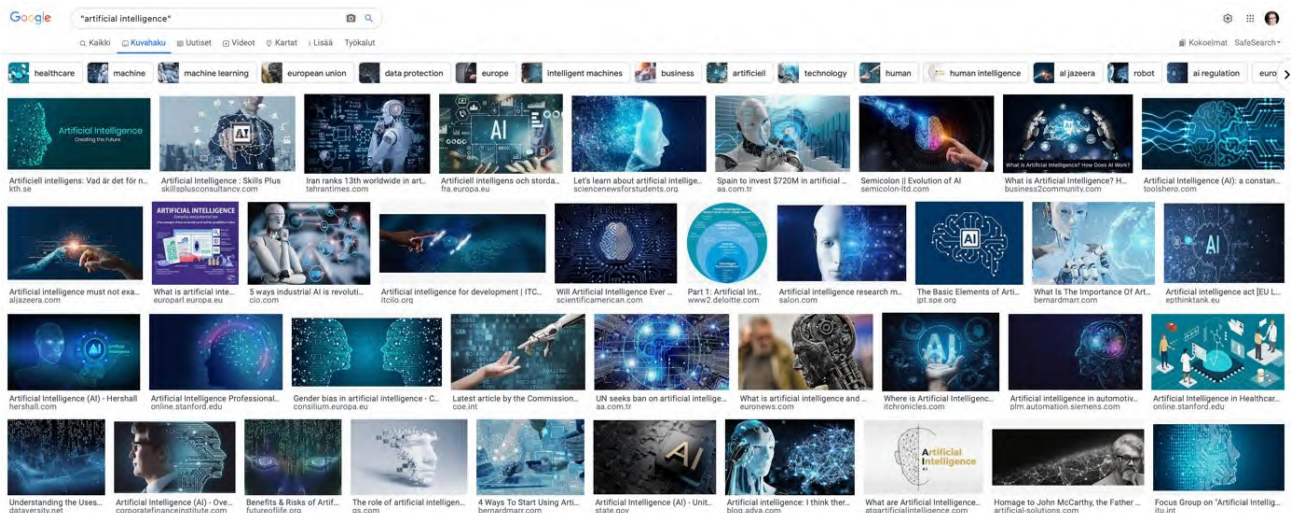
В поле зрения мирового мониторинга попадает не весь мир, и в освещении ИИ наблюдается определённый географический перекосяк. Страны, которые наиболее часто попадают в поле зрения международных изданий, — это США и Китай («сильные игроки» в области ИИ), а также Великобритания, Канада, Южная Корея, Япония и Германия. Эти страны также занимают высокие места в списке Глобального индекса ИИ, который охватывает 62 страны мира и оценивает внедрение технологий ИИ, инновации и инвестиции в эту сферу (см. <https://www.tortoisemedia.com/intelligence/global-ai/>). Несмотря на то, что некоторые западные и азиатские страны выбиваются в лидеры, целые обширные регионы никогда не обсуждаются с точки зрения технологий ИИ.

## Визуальное представление темы ИИ

Когда дело доходит до иллюстрации ИИ в новостном дискурсе, в визуальном плане технологии ИИ, как и многие другие технологические явления, такие как «большие данные» или «облачные вычисления», представляют определённые трудности для журналистов в силу своей абстрактности. ИИ трудно представить визуально. Визуальность, связанная с технологиями, часто описывается как монотонная или даже относительно нейтральная (Pentzold et al., 2019). Британская инициатива Better Images of AI («более правильные образы ИИ»), созданная отделом исследований и разработок BBC, некоммерческой организацией We are AI и Леверхульмским центром будущего интеллекта (Кембриджский университет), на своей домашней странице пишет, что визуальный образ ИИ очень стандартизирован (как показано на рис. 16).

Наберите в поисковой системе «изображения ИИ», и вы заметите закономерность. ...Результат поражает, и то же самое происходит в фотобанках и на мультимедиа-платформах. На самом деле, отсутствие разнообразия и недостоверность практически неизбежны. Преобладание научно-фантастических и антропоморфных мотивов и отсутствие доступных альтернативных образов или идей затрудняют достоверное преподнесение темы искусственного интеллекта.

**Рисунок 16.** Скриншот результатов поиска в Google изображений искусственного интеллекта (по состоянию на февраль 2022 г.)



Источник: скриншот результатов поиска в Google, выполненный авторами.

Авторы инициативы утверждают, что однообразие иллюстраций сказывается на осознании технологий ИИ широкой общественностью, оставляя людей в неведении относительно важных изменений, влияющих на их жизнь. По мнению авторов инициативы, шаблонные образы научной коммуникации, как в случае с изображениями ИИ, усиливают недоверие общества к технологиям ИИ. Ромеле (2022) называет визуальный поток представлений об ИИ «анестезирующим», воспроизводящим отстранённость, недоступность и состояние «чёрного ящика». Нейтральность визуальных образов чревата риском усиления веры в возможность объективной количественной оценки технологий или другие мифы о датаизме и сингулярности (более подробно рассматривается в модуле 2 «Культурные мифы и представления об искусственном интеллекте»). Вподобные иллюстрации трудно интегрировать общественный дискурс об этических проблемах, таких как слежка и вопросы конфиденциальности.

Для решения проблем визуального представления технологий ИИ в рамках инициативы Better Images of AI была запущена бесплатная библиотека изображений, призванная обеспечить и поддержать создание альтернативных образов (<https://betterimagesofai.org/images>). Авторы инициативы утверждают, что изображения должны более реалистично отображать технологию и показывать людей, стоящих за этой технологией, и сформулировали следующие принципы того, какими должны быть более правильные визуальные образы технологий ИИ:

- представлять более широкий спектр людей и человеческих культур, чем «деловой человек европейского типа»;
- отражать воздействие систем ИИ на человека, социальную сферу и окружающую среду;
- реалистично отражать беспорядочный, сложный, повторяющийся и статистический характер систем ИИ;
- достоверно отражать возможности технологии; как правило, она применяется для решения конкретных задач, не обладает интеллектом человеческого уровня и не имеет эмоций;
- показывать реалистичное применение технологий ИИ сейчас, в настоящем, а не в каком-то неопределённом научно-фантастическом будущем;
- не демонстрировать физическое робототехническое оборудование там, где оно не применяется;
- избегать монолитного или непостижимого представления систем ИИ;
- не показывать электронные образы человеческого мозга;
- более разнообразно изображать различные виды, способы, оттенки и последствия использования технологий ИИ.

Источник: Better Images of AI (<https://betterimagesofai.org/about>)

Изучая визуальную риторику, применяемую к теме больших данных в американской прессе, Пентцольд и коллеги (2019) обнаружили, что концепция данных часто иллюстрируется визуализированной числовой

информацией в инфографике, изображающей цифры, художественных коллажах, отображающих статистику, приложениях, изображённых на экранах мобильных устройств, или через логотипы, например, логотипы ИТ-компаний. Иллюстрации также строятся на материальном выражении вычислительной техники: гаджеты, такие как компьютеры, смартфоны, камеры, дроны, «умная» техника, или аппаратные средства, необходимые для обработки и хранения данных, например микрочипы, серверные фермы, помещения ИТ-компаний. Кроме того, отрасль персонифицируется через известных людей, что, согласно выводам Пентцольда и коллег, преподносит отрасль, в которой и без того доминируют мужчины, почти исключительно как «мужской мир», что соответствует результатам, полученным в сфере научной журналистики (см., например, Kim et al., 2016).

## Дифференциация профессиональных ролей

Журналистские организации и профессии традиционно дифференцируются по тематике, например, традиционные новостные блоки — политика, экономика, спорт, культура. Это означает, что существуют отделы новостей и профессиональные подгруппы журналистов, специализирующиеся на этих направлениях. Что касается тематики искусственного интеллекта, то в настоящее время не существует постоянных структур для её освещения, и в разных редакциях соответствующие репортажи организуются по-разному. Глобальные вызовы современности — от изменения климата до пандемий — усиливают потребность журналистов в специализации, хотя эти научные темы могут освещаться и специалистами широкого профиля. Даже для специалистов широкого профиля глобальные проблемы, среди которых тема искусственного интеллекта представляет собой одну из новых областей, примыкающих к науке и технике как экспертным областям, создают трудности в понимании и синтезе сложных областей знаний.

Группа журналистов, освещающих ИИ как журналистскую тему, не является однородной. В отличие от экологической журналистики, в журналистике, посвящённой ИИ, ещё не сложилась отдельная традиция. Вместо этого журналистика, изучающая ИИ как область, должна опираться на существующие в каждой стране журналистские традиции, среди которых наиболее близкими к сути технологий ИИ могут быть научная журналистика, техническая журналистика, журналистика в области экономики и финансов, потребительская журналистика, а также журналистика решений или конструктивная журналистика. В дополнение к такому межсекторальному подходу предполагается, что тема ИИ будет освещаться как в новостной и местной журналистике, часто выполняемой специалистами общего профиля, так и в журналистике международных отношений и даже культурной журналистике, выполняемой профильными специалистами.

Иными словами, освещение темы ИИ может быть следующих видов:

- освещение специалистами общего профиля: тема ИИ освещается как часть новостной журналистики в целом. Репортажи пишутся журналистами, обладающими специальными знаниями, часто с опорой на сведения организаций и пропагандистские источники информации;
- освещение профильными специалистами: тема ИИ освещается в рамках традиционных новостных лент или «кочует» между отделами, не занимая постоянного места в организационной структуре редакционной работы, подобно экологической и потребительской журналистике;
- освещение журналистами, специализирующимися на ИИ-тематике: тема ИИ освещается как область специализации и мониторинга журналистом, специализирующимся именно в этой области. Специализация часто формирует особую нишу для журналиста, ориентированного на темы науки и технологий, который оптимально может быть фрилансером, сотрудничающим с несколькими журналистскими изданиями и, таким образом, имеющим возможность полностью или в значительной степени сосредоточиться на темах, связанных именно с технологиями ИИ.

Вероятно, что разные жанры журналистики по-разному воспримут тему ИИ, руководствуясь различными целями в своей работе. Кроме того, возможно, хотя это ещё недостаточно изучено, что информационные издания разных профилей — от политического до религиозного и от качественной до популярной прессы — будут по-разному оценивать эту тему, учитывая, что в разных исследованиях тема ИИ сопровождается различными ценностными предложениями. Центральные идеи информатики, связанные с технологиями ИИ, такие как сингулярность и создание интеллекта как «божественной машины», могут противоречить христианскому мировоззрению, как это видно на примере дискуссий о медицинских прививках и нанотехнологиях (Vergeer, 2020). Популярная пресса может поддерживать сенсационность идей о «захвате власти» компьютерами с большей охотой, чем качественные ежедневные издания и общественное вещание.



Экономическая и финансовая журналистика рассматривает аспекты экономического роста, прибыли и занятости с точки зрения применения технологий ИИ, а журналистика искусства и культуры может выдвигать философские дебаты о разуме и сознании, а также о взаимодействии человека и машины. И, наконец, региональная ориентация изданий заставляет редакции по-разному расставлять приоритеты. Если ведущие общенациональные издания в целом уделяют больше внимания глобальным тенденциям, чем региональные и местные, которые даже не имеют специальных сотрудников для мониторинга сферы технологий, то для местных газет и телеканалов тема ИИ может казаться не столь важной или даже чуждой. Разумеется, это не обязательно так: разъяснение сложных вопросов может быть не менее значимой задачей для местных журналистов. Однако глубокое изучение вопросов может потребовать времени и образования, что в редакциях с ограниченными ресурсами и меньшей дифференциацией ролей не всегда реально.

Кроме того, следует обратить внимание на гендерные тенденции в области технологий и соответствующую структуру журналистского освещения в контексте ИИ. На протяжении последних лет звучат призывы к увеличению доли женщин-профессионалов в научно-технических дисциплинах, поскольку женщины недостаточно представлены на руководящих и технических должностях в этих отраслях. Соответствующие отрасли и учебные заведения признают наличие гендерных различий и разрабатывают образовательные мероприятия, направленные на привлечение большего числа кандидатов-женщин. В 2019 году 25% американских девочек-подростков заявили, что хотели бы сделать карьеру в области точных наук, в то время как среди мальчиков-подростков эта доля составила 57% (Women's Media Center, 2019). По данным ИТ-сети Adeva, на рынке труда в сфере вычислительной техники доля женщин составляет всего 25%, и только 6% пользователей, имеющих профили на платформе обмена программным обеспечением GitHub, — женщины (<http://adevait.com>). Некоторые полемисты связывают низкую представленность женщин в научно-технических областях отчасти с медиакультурой, полагая, что в кино, на телевидении и в онлайн-развлечениях традиционно не хватает женских научно-технических персонажей, и поэтому, чтобы повлиять на выбор профессии женщинами, важно освещать фигуры женщин в контексте технологий. В 2019 году Женский медиацентр (Women's Media Center) обнаружил, что женщины составляют менее трети (27%) научно-технических персонажей в западных развлекательных СМИ. Архетипическими персонажами, представляющими науку и технологии, как правило, являются учёные-физики, инженеры и компьютерные специалисты мужского пола, даже если медиаиндустрии сознательно продвигают репрезентацию женщин. С другой стороны, во многих случаях ИИ представляется в виде женских персонажей — гиноидов или фемботов, как, например, в кино начиная с «Метрополиса» Фрица Ланга (1927 г.) и заканчивая «Из машины» Алекса Гарленда (2015 г.), подчёркивая «богоподобные» и эротические аспекты таких технологий.

В журналистике доля женщин-журналистов в целом и женщин-журналистов, пишущих о технологиях, в частности, является весьма обсуждаемым и регулярно отслеживаемым вопросом. В 2014 году около 35% американских новостей о технологиях освещалось женщинами-журналистами (Women's Media Center, 2014). Помимо того, что на руководящих редакционных постах традиционно доминируют мужчины (Andi et al., 2020), отмечается даже горизонтальная сегрегация, когда существует гендерное разделение задач, связанных с освещением тем и так называемых «мягких» и «жестких» новостей (см., например, North, 2016). Поскольку ИИ, будучи обширной и неоднородной темой для освещения, может способствовать разнообразию подходов, которые не всегда согласуются с прежними бинарными представлениями о технологиях и науке, будущие исследования покажут, в какой степени практика освещения тематики ИИ является гендерно-обусловленной.

По состоянию на 2019 г. более 71% научно-технических персонажей в западных развлекательных медиапродуктах были белыми, в то время как чернокожие составляли около 17%, азиаты — 6%, латиноамериканцы — 4%, а выходцы с Ближнего Востока — 2% научно-технических персонажей (Women's Media Center, 2019). ИИ также преимущественно изображается белым с точки зрения расовой и этнической принадлежности. Кейв и Дихал (Cave and Dihal, 2020) предполагают, что доминирование белой расы может частично отражать преимущественно белую среду, из которой возникают технологии и искусственные объекты. Однако они также утверждают, что атрибуты технологий ИИ, такие как интеллект, профессионализм и мощь, также усиливают расовую структуру белых, которая склонна приписывать эти атрибуты белым людям. По мнению Кейва и Дихала, отнесение технологий ИИ к белой расе может усугубить предвзятость, вычеркнув из образного ряда цветных людей. Предвзятость может также возникать, когда учёные из разных стран мира изучают публичные представления об ИИ, а их данные, как правило, берутся из англо-американской журналистики, что укрепляет положение западной публичной сферы как простран-

ства для универсального представления технологий ИИ. Следует обратить внимание на то, как тематика ИИ обсуждается на местном и региональном уровнях, где могут возникать более локальные, внутренние и самореферентные дискурсы и репрезентации, отражающие местные условия.

## Просвещение аудитории

Представления о роли научных журналистов свидетельствуют о том, что сделать сложную информацию доступной для общественности — это общая миссия (Brüggemann, 2017; Viswanath et al., 2008). Научно-технические журналисты стремятся навести мосты между учёными и экспертами, совершающими открытия, и гражданами, которые хотят узнать об этих новых открытиях и их ценности. Для того чтобы позиционировать ИИ как полностью общественную проблему, необходимо разнообразие мнений, чтобы сбалансировать точки зрения. Бреннен и его коллеги (Brennen, 2018) считают, что сферу ИИ не следует представлять ни в виде завышенных заявлений и сенсаций, ни в виде решения всех проблем. Напротив, они предлагают журналистам больше обсуждать различные аспекты, чтобы найти баланс:

Хотя средства массовой информации должны детально изучить перспективы и подводные камни использования технологий ИИ, им будет полезно рассматривать это явление не как революцию, которая потрясёт мир, а скорее как набор технологий, находящихся в процессе разработки, как набор решений, находящихся в процессе формирования, и как набор проблем, находящихся в процессе коллективного решения (Brennen et al., p. 10).

Диверсификация источников и рассматриваемых аспектов требует более тесного сотрудничества между отделами новостей.

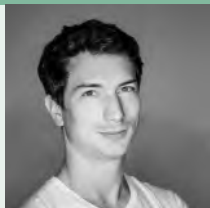
Таким образом, можно сказать, что сбалансированный, объективный и взвешенный подход к теме ИИ требует нового уровня грамотности как от журналистов, так и от аудитории. Как и в случае со многими специализированными темами для освещения, например, в науке, журналистам не обязательно быть экспертами в области ИИ, но они должны быть знакомы с этой темой и выступать в роли переводчиков, транслируя идеи узкоспециализированной технической или научной области в повседневную жизнь аудитории. В то же время аудитории всё чаще требуются навыки понимания и обращения с медиа- и цифровыми технологиями в своей личной жизни, что получило название *медиа- и информационной грамотности, цифровой грамотности* или, более конкретно, *технологической, алгоритмической или платформенной грамотности*. Чтобы помочь людям в их отношениях с новыми технологиями, журналисты в некотором смысле становятся просветителями своей аудитории. Журналисты могут указать, что в продуктах является новым, а что, наоборот, нет; как на самом деле работают технологии и приложения и что каждый должен знать об их работе; они могут пролить свет на устройство и последствия использования этих технологий. В этом отношении тема ИИ перекликается с амбициями журналистских подходов, которые появились и развиваются в последнее время. К ним относятся не только традиционные направления научной и потребительской журналистики, но и, прежде всего, *журналистика решений* или *конструктивная журналистика*, которая стремится сосредоточиться на решении проблем, а не на освещении недостатков, рисков и угроз. Один из центральных вопросов будущего развития — как сделать освещение технологий ИИ более открытым, диалогическим, взаимным и вовлекающим.

В настоящее время в области освещения ИИ-технологий доминируют точечные новостные события, но в идеале она должна освещаться более глубоко, с учётом тенденций и долгосрочных процессов, влияющих на общество. Если отслеживать отдельные события, которые, как правило, представляют собой выпуск продуктов, инвестиционные решения или запуск инициатив крупными отраслевыми организациями, то освещение может свестись к отслеживанию повестки основных игроков. Будучи влиятельными общественными деятелями, журналисты могут противодействовать сложившейся гендерной и расовой представленности, а также другим предвзятым структурам, связанным со сферой ИИ, и способствовать увеличению разнообразия в различных областях. В этом отношении освещение тематики ИИ требует от журналистов проактивного подхода к смыслообразованию, когда происходит идентификация и контекстуализация явлений, а не просто тиражирование и пересказ новостных поводов. Вопросы, над которыми стоит задуматься журналисту при освещении связанных с ИИ тем, могут быть следующими:

- О каком виде ИИ идёт речь в данной теме? Поскольку ИИ — это очень общий термин, имеющий широкую область применения, его часто приходится уточнять. Что особенного в этой конкретной технологии? Что она делает и, возможно, изменит?

- Что известно о принципах работы рассматриваемой системы ИИ? Понятно ли аудитории, как выглядит приложение или как оно используется на практике? Можно ли как-то проиллюстрировать принцип работы словесно (фактологический материал) или визуально (иллюстрация процесса)?
- Кто является заинтересованными сторонами рассматриваемого ИИ-продукта и какие финансовые интересы стоят за этим конкретным приложением? Кто владеет продуктом и кто его финансировал?
- Как технология ИИ влияет на жизнь различных групп людей? Как эти люди, как граждане и потребители, воспринимают эти изменения?
- Какие последствия применение ИИ имеет для целевых аудиторий ваших СМИ?
- Какой образ этого явления формируется в ваших материалах?

## Освещение алгоритмов: математика — наименее важная вещь



**Николас Кайзер-Бриль, корреспондент AlgorithmWatch, Германия**

Журналисты, как и политики и многие другие, склонны считать автоматизированные системы слишком сложными для понимания. Они часто ослеплены рассуждениями технократов об искусственном интеллекте и не решаются провести исследование. Но математика, лежащая в основе этих систем, не так важна. Гораздо важнее их влияние на общество и то, работают ли эти системы вообще.

AlgorithmWatch — это базирующаяся в Берлине некоммерческая организация, которая помогает пролить свет на автоматизированные процессы принятия решений, имеющих значение для общества. Мы занимаемся исследованиями, а также проведением информационных кампаний и пропагандистской деятельностью. Мы сотрудничаем с журналистами по всей Европе и помогаем им приоткрыть завесу над автоматизированными системами, используемыми как государственными органами, например, в службах социального обеспечения, школах или полиции, так и частными компаниями, например, для определения кредитного рейтинга или управления контентом на платформах социальных сетей.

Мы понимаем, что многие системы очень сложны, и читать программный код или вникать в подробности технического устройства без знаний математики или статистики очень трудно. Но технические тонкости важны для специалистов по статистике, а не для общества. Среди очень немногих показателей, которые имеют значение для общества, можно назвать показатели ложноположительных и ложноотрицательных результатов. С ними можно разобраться за несколько минут, посмотрев учебник. А если журналист всё ещё стесняется этих терминов, то можно успокоить его тем, что исследователи (в основном из лаборатории Герда Гигеренцера) показали, что многие люди, считающие себя экспертами, не очень хорошо понимают эти концепции.

Существует ошибочное мнение — даже среди журналистов — что высокотехнологичные системы могут быть понятны только техническим специалистам. Напротив, программный код и алгоритмы — это лишь очень малая часть этих систем. Гораздо важнее то, кто их создал, с какой целью и какое влияние они оказывают на общество. Журналисты в гораздо большей степени чем кто бы то ни было способны ответить на эти вопросы.

Однажды в Твиттере мне попалась на глаза жалоба журналистки на то, что её заявление в агентстве по управлению социальным обеспечением было отклонено через день после того, как она прислала целую пачку документов. Это был тревожный сигнал о том, что её заявление обрабатывала автоматизированная система, а не человек. Я попросил её провести расследование, что она и сделала, и это привело к появлению множества других историй на эту тему, опубликованных на AlgorithmWatch и не только.

В другой раз одна из создательниц контента в Instagram рассказала нам, что ей кажется, что она не достигает своих подписчиков, когда публикует фотографии, на которых она полностью одета. По её словам, она должна быть на фотографиях в купальнике, чтобы её смогли увидеть подписчики. Мы захотели проверить, так ли это на самом деле. С точки зрения статистики, попытка понять, как работает алгоритм ленты новостей Instagram, представляла собой фантастическую задачу. Поэтому мы обратились за помощью к специалистам по статистике. Остальное — традиционная журналистика: беседы с теми, кого это затронуло, и с теми, кто разрабатывал систему. Когда мы соединили оба конца расследования, у нас получилась сильная и глубокая история.

Домашняя страница AlgorithmWatch: <https://algorithmwatch.org/en/>




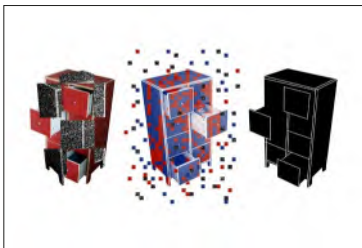

## Предлагаемые задания



**ПРОЧИТАТЬ И ПОРАЗМЫШЛЯТЬ.** В профессиональных журналах и на сайтах сообществ регулярно публикуются статьи, в которых журналисты дают советы, как правильно освещать тему ИИ. Найдите в Интернете такие тексты, часто написанные журналистами для журналистов, и соберите все советы воедино, возможно, сгруппировав их по тематическому принципу. На основе найденных советов создайте хартию по коммуникации для журналистов — практический одностраничный документ, в котором изложены ключевые принципы освещения ИИ в журналистике. Представьте свой список советов коллегам, обсудив проблемы и, по возможности, сравнив различия.



**ОБСУДИТЬ.** Посмотрите на приведённые ниже изображения, иллюстрирующие ИИ. Попробуйте вербализовать то, что вы видите на картинке. После этого зайдите на сайт и прочтите описание, чтобы сравнить его с собственной интерпретацией. Уловили ли вы идею из изображения, а если нет, то почему? Как бы вы проиллюстрировали эту идею по-другому? Какие поисковые запросы можно было бы использовать вместо словосочетания «искусственный интеллект» для поиска подходящих изображений?

Измеримый человек	Картотека	Банан / растение / колба
 <p>Alan Warburton, © BBC, Better Images of AI (CC-BY 4.0) <a href="https://betterimagesofai.org/images?idImage=0">https://betterimagesofai.org/images?idImage=0</a></p>	 <p>Anton Grabolle, Better Images of AI (CC-BY 4.0) <a href="https://betterimagesofai.org/images?idImage=5">https://betterimagesofai.org/images?idImage=5</a></p>	 <p>Max Gruber, Better Images of AI (CC-BY 4.0) <a href="https://betterimagesofai.org/images?idImage=15">https://betterimagesofai.org/images?idImage=15</a></p>



**ВЫРАБОТАТЬ ИДЕЮ.** Чтобы установить контакт с учёными кругами в области ИИ, необходимо найти экспертов в различных научных областях. Попробуйте найти научных экспертов для проведения гипотетического интервью в указанных ниже областях в ближайших к вам университетах. Найдите по одному эксперту в каждой области, определите, какое общественное значение могут иметь их исследования, и сформулируйте 1–2 вопроса для интервью. Обсудите с коллегой, с какими трудностями вы столкнулись.



Оригинальный авторский материал



**ПОИСКАТЬ И ИССЛЕДОВАТЬ ИСТОЧНИКИ.** Выберите один из следующих вариантов, в зависимости от того, на чём вы хотите сосредоточиться: на тексте (а–б) или на изображении (в):

- а) Обратите внимание на заголовки газетных статей об искусственном интеллекте. Найдите 5–6 примеров в онлайн-базе данных по вашей стране или региону, или в новостном издании по вашему выбору, и/или просмотрите примеры крупных новостных изданий, приведённые ниже. Перечислите 20 самых последних заголовков. Какие ценности новостной журналистики обычно применяются к сфере ИИ как журналистской теме? Удовлетворяют ли вас эти новостные ценности или вместо них следует выделить какие-то другие? Какие, почему и как?
- б) Выберите платформу социальных сетей (например, Twitter), и используя поиск по хэштегу на своё усмотрение (например, #ai, #artificialintelligence, #deeplearning, #machinelearning) выясните, как различные группы пользователей обсуждают тему ИИ. Возьмите 100 последних сообщений с использованием выбранного хэштега и сгруппируйте их по тому, чьи интересы они представляют (промышленные компании, СМИ/журналистика, НПО, обычные люди и т. д.). Сгруппируйте сообщения также по тому, какое отношение к ИИ они выражают — положительное, отрицательное или нейтральное (если это можно определить). Какие различия вы видите между сообщениями разных производителей информации? Можете ли вы выявить какие-либо другие закономерности внутри групп?
- в) Изучите, как сегодня иллюстрируется искусственный интеллект в новостных статьях. Найдите новостные статьи об искусственном интеллекте в новостных изданиях, которые позволяют отнести статьи к данной тематике (например, в The Guardian по тегу «искусственный интеллект (ИИ)»), или, как вариант, найдите изображения из фотобанков по поисковому запросу «искусственный интеллект». Изучите найденные изображения. Как обычно иллюстрируется искусственный интеллект и какие ассоциации вызывают эти изображения? Отражают ли эти изображения идеи инициативы Better Images of AI или расходятся с ними? Можете ли вы придумать альтернативные способы иллюстрирования темы, чтобы избежать потенциальных предрассудков, таких как гендерная предвзятость или мифологизация технологии?



**СОЗДАТЬ.** Заполните таблицу, приведённую ниже, и подготовьте краткий обзор одной из следующих тем: а) женщины-пропагандисты ИИ; б) женщины-журналисты в области технологий; в) цветные люди и ИИ; г) проблемы реализации мер политики в области ИИ; д) злоупотребления определённым видом технологий ИИ; е) неравенство в использовании средств ИИ; ж) вопрос, связанный со сферой ИИ, который волнует лично вас. Для составления конспекта изучите научные журнальные статьи и журналистские материалы по данной теме.

Тема: сформулировать свою тему.

Средство: выбрать средство.

Предполагаемый объём: оценить объём материала в его окончательном виде.

Область	Пояснение	Ваш ответ
<b>Основная идея</b>	Какую главную мысль вы хотите передать? Сформулируйте утверждение.	
<b>Жанр</b>	Какой жанр вы используете?	
<b>Предполагаемая аудитория</b>	Опишите предполагаемого адресата сюжета.	
<b>Предполагаемый эффект</b>	Что ваша история призвана изменить во взглядах аудитории и в мире?	
<b>Аргументы «за»</b>	Каковы основные аргументы в пользу вашего утверждения?	
<b>Аргументы «против»</b>	Каковы основные аргументы против вашего утверждения?	

Область	Пояснение	Ваш ответ
<b>Источники</b>	У кого вы будете брать интервью? Постарайтесь найти реального человека.	
<b>Разбор примера</b>	О каком приложении (приложениях) ИИ, инициативе (инициативах) и т. д. пойдёт речь в вашей истории?	
<b>Название</b>	Предложите название, которое содержит вашу мысль и вызывает интерес к теме.	
<b>Введение</b>	Напишите 2–3 предложения.	
<b>Обоснование</b>	Почему эта история важна (с точки зрения тематики ИИ)?	
<b>Самоанализ</b>	Какие у вас возникают затруднения при освещении темы? Почему?	



## Дополнительная литература

- Angler, M. W. (2017). *Science journalism: An introduction*. Routledge. <https://doi.org/10.4324/9781315671338>
- Bains, S. (2019). *Explaining the future: How to research, analyze, and report on emerging technologies*. Oxford University Press. <https://doi.org/10.1093/oso/9780198822820.001.0001>
- Cave, S., & Dihal, K. (2020). *The whiteness of AI*. *Philosophy and Technology*, 33, 685–703. <https://doi.org/10.1007/s13347-020-00415-6>
- De Vuyst, S. (2020). *Hacking gender and technology in journalism*. Routledge. <https://doi.org/10.4324/9780429262029>
- Johnson, D. G., & Verdicchio, M. (2017). *Reframing AI discourse*. *Minds & Machines*, 27, 575–590. <https://doi.org/10.1007/s11023-017-9417-6>
- Natale, S., & Ballatore, A. (2017). *Imagining the thinking machine: Technological myths and the rise of artificial intelligence*. *Convergence*, 16(1), 3–18. <https://doi.org/10.1177/1354856517715164>
- Sudmann, A. (Ed.). (2019). *The democratization of artificial intelligence: Net politics in the era of learning algorithms*. Transcript Verlag.
- Sun, S., Zhai, Y., Shen, B., & Chen, Y. (2020). *Newspaper coverage of artificial intelligence: A perspective of emerging technologies*. *Telematics and Informatics*, 53, 101433. <https://doi.org/10.1016/j.tele.2020.101433>
- Verdegem, P. (Ed.). (2021). *AI for everyone? Critical perspectives*. University of Westminster Press. <https://doi.org/10.16997/book55>



© George Morina, Pexels.com

## Модуль 7:

# Искусственный интеллект в журналистике и журналистской практике



## Краткое содержание

Данный модуль кратко знакомит с использованием средств искусственного интеллекта в журналистике и журналистской практике, например, в машинных статьях и роботизированной журналистике, и призывает к критическому самоанализу и размышлениям об этических аспектах, а также об использовании лучших практиках принятия редакционных решений при сборе, производстве и распространении новостей. Он также призывает педагогов и преподавателей к инновационному использованию средств ИИ в журналистике для повышения прозрачности, креативности и решения проблем.

**Ключевые понятия:** ИИ в журналистике, алгоритмическая журналистика, автоматизированная журналистика, вычислительная журналистика, роботизированная журналистика



## Цели модуля

Цели данного модуля:

- ▶ ознакомление с тем, как ИИ уже используется и может быть использован в будущем в журналистике;
- ▶ выработка понимания потенциала и рисков, которые несёт в себе использование ИИ для журналистики и журналистской практики;
- ▶ определение человеческой ценности в создании журналистских материалов.



## Результаты обучения

Ожидается, что по завершении модуля студент научится:

- ▶ выявлять технологии на базе ИИ, наиболее актуальные для журналистской практики, и понимать, как они работают;
- ▶ знать, в каких журналистских жанрах и задачах средства ИИ могут быть использованы наиболее эффективно;
- ▶ знать, как журналистское сообщество относится к технологиям ИИ с точки зрения профессиональных и этических ценностей, а также иметь представление о некоторых современных передовых практиках.



## Вопросы для размышления

Несколько вопросов, над которыми можно поразмыслить заранее:

1. Каким образом по вашему собственному опыту компьютеры помогают в работе журналистов в редакциях?
2. Какие термины журналистики, связанные с искусственным интеллектом, вы можете выделить? Составьте, пожалуйста, список таких терминов, относящихся к журналистике, осуществляемой с использованием компьютеров и технологий, используя базу данных проекта X Journalism (<https://xjournalism.org>) Института медиаисследований Лейбница и Института Ганса Бредова в Гамбурге. Что учёные понимают под термином «X Journalism»?

Несколько подготовительных заданий, которые необходимо выполнить перед тем, как приступить к изучению темы:

1. Поговорите с каким-нибудь журналистом и расспросите его об опыте использования средств ИИ в журналистской практике.
2. Ознакомьтесь со следующими новостными текстами, написанными роботами, и обсудите, написал ли бы человек некоторые их части по-другому:
  - ReporterMate. (2019, January 31). *Political donations plunge to \$16.7m — down from average \$25m a year*. The Guardian. <https://www.theguardian.com/australia-news/2019/feb/01/political-donations-plunge-to-167m-down-from-average-25m-a-year>
  - GPT-3. (2020, September 8). *A robot wrote this entire article. Are you scared yet, human?* The Guardian. <https://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3>





## Конспект

Автор: Дженни Вийк

В журналистике наблюдается активное обсуждение возможностей новых технологий, и не в последнюю очередь средств искусственного интеллекта и автоматизации. Многие видят в этом спасательный круг в период слабости отрасли, другие — угрозу, механизмирующую творческую профессию и способствующую её коммерциализации, а возможно, и сокращению штатов. Неудивительно, что обе точки зрения имеют под собой определённые основания: Новые технологии открывают перед журналистикой не только новые возможности, но и проблемы. Одно можно сказать с уверенностью: технологии ИИ пронизут и изменят журналистику в самой её основе — во всех звеньях производственной цепочки. Развитие журналистики на основе данных, поддерживаемое технологиями ИИ, означает разрушение и пересмотр устоявшихся бизнес-моделей, профессиональных знаний и целей на будущее.

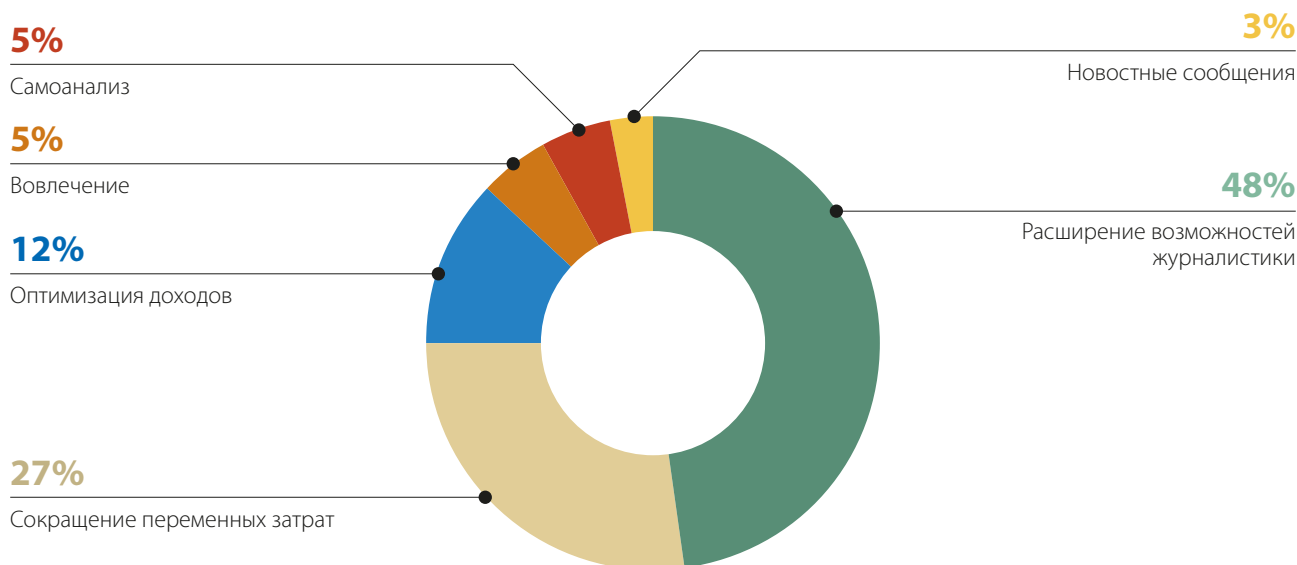
Поэтому этот процесс можно рассматривать как «созидательное разрушение» журналистики. «Шторм созидательного разрушения» описывает, по словам экономиста Шумпетера (1994, с. 82–83), «процесс индустриальной мутации, который непрерывно реконструирует экономическую структуру изнутри, разрушая старую структуру и создавая новую». Такие процессы во многом болезненны, но — согласно этой точке зрения — также необходимы, неизбежны и являются оздоравливающими. Является ли автоматизация производства новостей частью «созидательного разрушения», процесса трансформации, когда гибридизация связки человека и машины становится сердцем профессии? И если да, то что это означает для журналистики как профессии и как демократического института?

В этой главе показаны основные преимущества технологий ИИ для журналистики, а также наиболее важные проблемы. Научно-техническое развитие рассматривается с профессиональной точки зрения, а также обсуждается, что это означает для медиакомпаний и руководителей.

### Как можно использовать ИИ в журналистике?

ИИ — это собирательное название различных алгоритмов, автоматических средств и систем, управляемых данными, общим знаменателем которых является то, что они могут заменить человеческий труд. Некоторые алгоритмы самообучаются и могут повышать точность своей работы по мере увеличения объёма данных, но во многих случаях речь идёт о статичных программах, выполняющих одно и то же действие каждый раз. Это означает, что автоматические средства (роботы) могут выполнять большую часть рутинной работы журналистов, потенциально высвобождая человеческие ресурсы для решения более качественных задач. В своей книге *«Как автоматизация меняет СМИ»* («*Sådan forandrer automatiseringen medierna*») датский журналист Андреас Маркманн Андреассен (Andreas Marckmann Andreassen, 2020) пишет, что в настоящее время можно выделить использование средств ИИ во всех звеньях журналистской производственной цепочки. Однако некоторые звенья этого процесса получили большее развитие, чем другие. Наиболее распространёнными практиками автоматизации являются сбор и исследование данных, а также автоматическая генерация новостных материалов.

Глобальное исследование 130 проектов в области ИИ, проведённое фондом Knights Foundation в 2021 году, показало, что основной целью внедрения ИИ в журналистскую отрасль является расширение возможностей репортёров, а второй — сокращение расходов.

**Рисунок 17.** Основное назначение технологий ИИ в журналистике (2012–2021 гг.)

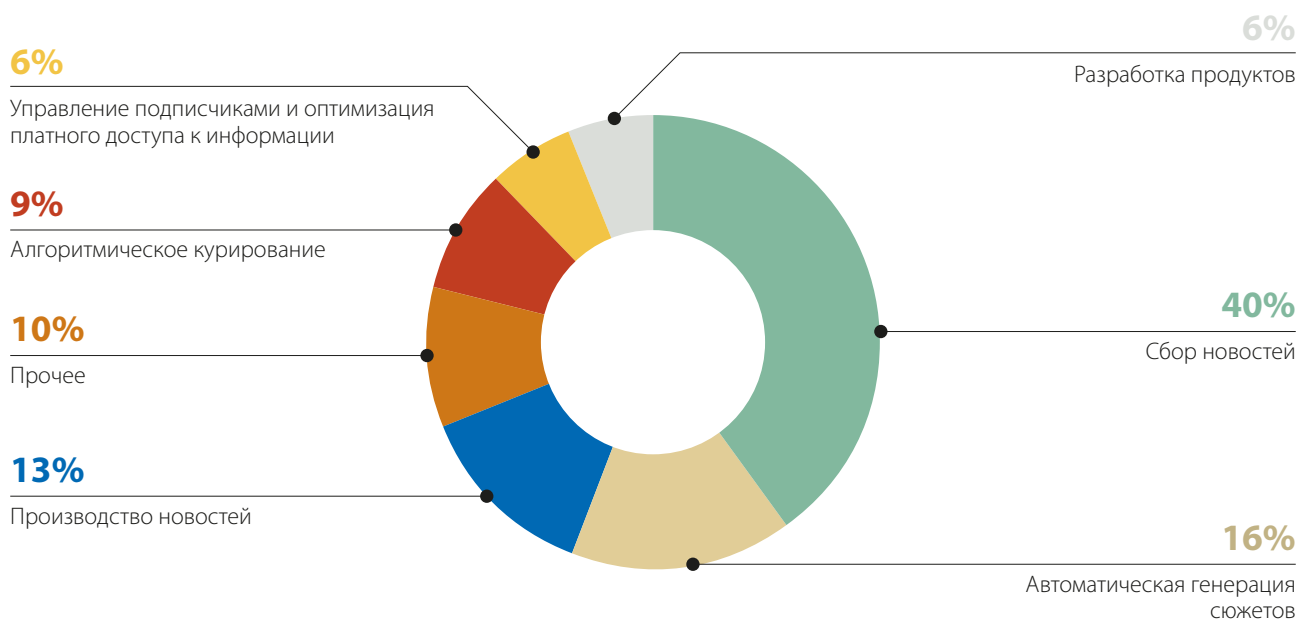
Источник: составлено авторами по материалам *The Presence and the Potential of AI in Journalism*, Knight Foundation 2021.

Одним из примеров того, как применение средств ИИ может помочь в поиске и изучении информации, является алгоритм News Tracer компании Reuters, который автоматически ищет свежие новости в социальных сетях. Обнаруживая кластеры похожей информации и просматривая аккаунты, опубликовавшие её, робот может уведомить журналиста о новости в момент её появления (Xiao et al., 2017).

Другой пример — робот компании BBC Juicer (<https://bbcnewslabs.co.uk/projects/juicer/>), который функционирует как «конвейер» агрегации новостей. Он поглощает новостные статьи и извлекает из них самое важное — точно так же, как соковыжималка делает это с фруктами. Конвейер BBC Juicer просматривает RSS-ленты новостных изданий. Когда на одном из этих RSS-каналов публикуется новая статья, BBC Juicer извлекает текст и метаданные (например, дату, время, заголовок, источник новости). На следующем этапе BBC Juicer идентифицирует и маркирует понятия, упомянутые в тексте статьи, делая их доступными для поиска и, следовательно, полезными для анализа тенденций. Упорядочивая таким образом рабочие процессы в СМИ, ИИ позволяет журналистам сосредоточиться на том, что они умеют делать лучше всего.

В настоящее время инструменты ИИ внедряются в большинстве новостных редакций Европы и Северной Америки, но в этом процессе участвуют и другие регионы. Мунориярва, Чиумбу и Мотсаатебе (2021) исследовали практику применения ИИ-технологий в южноафриканских редакциях СМИ и пришли к выводу, что, хотя некоторые СМИ технологически развиты, в целом адаптация инструментов ИИ идёт медленно. Оптимистическое отношение к технологиям ИИ в западных редакциях не находит отклика у южноафриканских журналистов, чей скептицизм обусловлен «страхом потери работы, этическими проблемами, связанными с применением ИИ, его эффективностью в контексте демократического процесса и затратами на внедрение средств ИИ в редакциях африканских СМИ» (Munoriyarwa et al., 2021). Они отмечают, что в развивающихся странах значение новых технологий может быть иным, чем в странах Глобального Севера, поэтому автоматизация и без того нестабильной профессии может увеличить риск потери автономии и контроля. С другой стороны, в странах со слабой демократией и государственным контролем над новостями алгоритмизация может рассматриваться как средство, позволяющее избежать предвзятости и ложных новостей. Экспериментальное исследование корейских потребителей новостей показало, что они оценивают новости, созданные роботизированными средствами, как более достоверные, чем новости, подготовленные журналистами-людьми, — такая реакция, скорее всего, вызвана негативным отношением общества к репутации журналистов и тягой к новым продуктам и услугам ИКТ (Jung et al., 2017).

При более детальном рассмотрении средства ИИ используются преимущественно как инструмент для сбора новостей и автоматической генерации сюжетов. Часто это делается в виде коротких заметок, основанных на данных в области продажи недвижимости, спорта или экономики.

**Рисунок 18.** На каком этапе конвейера новостей обычно применяются средства ИИ? (2012–2021 гг.)

Источник: составлено авторами по материалам *The Presence and the Potential of AI in Journalism*, Knight Foundation 2021.

В качестве примера можно привести шведскую компанию United Robots, которая поставляет решения для генерации текстов нескольким новостным компаниям. Они используют технологии искусственного интеллекта и методы генерации естественного языка для автоматического создания пригодных для публикации новостных текстов на основе больших массивов данных, включая спорт, недвижимость, сводки дорожного движения, погоду, фондовый рынок и регистрацию местных предприятий. Роботизированные средства анализируют большие массивы данных, которые открыты и доступны каждому, и выявляют необычные события или закономерности. Полученная информация может быть использована для создания полноценных статей или в качестве оповещений для отдела новостей (United Robots, 2021). Многие СМИ также встроили в эти инструменты средства анализа гендерных и других показателей, что позволяет проводить автоматический мониторинг выпускаемых ими новостей.

Применение средств ИИ также способствует формированию нового подхода к управлению информационным наполнением в целом. Благодаря применению полуавтоматических инструментов публикации, когда средства искусственного интеллекта используются для многоаспектного объединения модулей информации, новости приобретают более персонализированную и изменяемую форму. Это означает, что авторская работа теперь основывается не на «сюжете», а на «вложенных блоках», которые обеспечивают более тесную связь между сюжетами, облегчая перекомпоновку содержания потенциально неограниченным количеством способов (Leitner, 2018).

Ещё одной растущей областью применения машинного обучения является использование средств машинного обучения для предотвращения оттока клиентов. Используя средства автоматического сбора и анализа данных о поведении пользователей, аналитики рынка могут работать непосредственно над причинами, по которым пользователи могут уходить, то есть перестать пользоваться приложением, что обычно происходит, когда пользователи просто сокращают время использования или удаляют приложение со своего устройства.

Таким образом, технологии ИИ имеют несколько потенциальных положительных применений в журналистике:

- более персонализированное распространение материалов;
- более эффективное, автоматизированное производство материалов;
- динамическое ценообразование, как на рекламу, так и на подписку;
- возможность находить больше сюжетов в данных и больше данных в сюжетах;

- более качественные автоматизированные транскрипции;
- управляемая модерация материалов;
- распознавание фальшивых новостей и глубоких подделок (дипфейков);
- новые инструменты опровержения;
- улучшенный поиск изображений и видео;
- более глубокий анализ настроений в материалах пользователей.

Для новостного бизнеса, который уже не первый год борется с финансовыми трудностями и истощением потенциала бизнес-моделей, перспективы делать больше при меньших затратах являются захватывающими. Особенно на местных рынках с сокращающимися редакциями новостей использование ИИ открывает возможности для более качественного цифрового освещения местных сфер политики и экономики.

## Организационная и рыночная среда

Разработка и внедрение инструментов искусственного интеллекта в журналистике является частью более общих усилий по перестройке бизнес-моделей и внедрению инноваций в работе редакций СМИ. Многие СМИ долгое время боролись со снижением доходов от рекламы, главным образом из-за растущей конкуренции за внимание пользователей в онлайн-среде. Во время пандемии COVID-19 постепенное снижение ускорилось и доходы резко сократились. Например, доходы от рекламы американских новостных СМИ в 2019–2020 годах снизились примерно на 30% (Adgate, 2021). В то же время доходы от подписки на онлайн-издания росли, поскольку кризис в здравоохранении вызвал массовую потребность в достоверной информации. Рыночные и более широкие общественные тенденции привели к тому, что новостная индустрия оказалась в состоянии «шторма созидательного разрушения», ключевую роль в котором играют новые технологии.

Восприятие этих изменений как «созидательного разрушения» подчёркивает, что инновации в сфере новостей — это не стратегия или конечная цель, а процесс — «ряд динамических процессов, механизмов, методов и изменений, которые приводят к определённому результату» (Siles & Boczkowski, 2012, p. 1386). Сформировавшаяся в последние десятилетия кроссплатформенная модель производства новостей получила дополнительный импульс в связи с изменением экономических условий, о которых говорилось выше, а также в связи с ростом интереса и инвестиций в инновации в работе новостных редакций. На основе обширного обзора литературы Паулуссен (Paulussen, 2016) подчёркивает сложное взаимодействие структурных и индивидуальных, человеческих и нечеловеческих факторов и субъектов, обуславливающих процессы инноваций в журналистике. Таким образом, влияние технологий ИИ на журналистику требует целостного подхода к организации и производству новостей.

## Сотрудничество как ключевой фактор

Сотрудничество — важный фактор успешной автоматизации любого бизнеса, и журналистика не является исключением. Сотрудничество предполагает обмен ресурсами и их использование там, где это необходимо, и приводит к появлению новых способов объединения навыков, умений и материальных ценностей. Новые плоскости соприкосновения приводят к улучшению потока информации, новому пониманию проблем и возможности учиться на чужих ошибках (Jia & Xia, 2008). Чтобы соответствовать растущей сложности внешнего мира и рынка, медиакомпании всё чаще пытаются организовать свою деятельность на основе принципов сотрудничества. Внутренние границы, которые раньше отделяли друг от друга такие отделы, как маркетинг и редакция, сегодня стремительно стираются, поскольку компании переходят на проектные и междисциплинарные методы работы. Хотя эти подходы относительно новы для журналистики, в других наукоёмких отраслях они применяются уже несколько десятилетий. Исследования, проведённые, например, в сфере здравоохранения, показали, что стратегии достижения комплексных целей требуют как сотрудничества, так и осмысленной приверженности (Winfield et al., 2017). Когда речь идёт о развитии продуктов и бизнеса, медиаиндустрия имеет традицию наблюдать за своими ближайшими конкурентами, но редко распространяет этот взгляд на другие отрасли. Один американский медиаменеджер сказал: «Я считаю, что журналистика должна ориентироваться на ВСЕ другие отрасли. На самом деле, меня беспокоит, что журналистика как отрасль слишком часто ищет ориентиры внутри самой журнали-

стики. Что делает «Нью-Йорк Таймс»? Что делает «Вашингтон Пост»? Нам нужно посмотреть на другие отрасли и понять, как их инновации, промахи и опасения могут быть применимы к нам» (Beckett, 2019, p. 86).

### «Чужаки» в редакции новостей

Инновации и новые технологии, таким образом, требуют новых навыков, и в редакциях СМИ появился целый ряд новых участников. Актуальным с точки зрения журналистики является вопрос о том, как новые участники и их опыт интегрируются в сложившиеся новостные процессы (например, Eldridge, 2018; Holton & Belair-Gagnon, 2018). Исследования показывают, что к посторонним, которые входят в журналистское поле, иногда относятся как к «нежелательным чужакам» (Holton & Belair-Gagnon, 2018, p. 71). Однако Элдридж (Eldridge, 2018) отмечает, что журналисты могут реагировать на изменения по-разному: от сопротивления до принятия перемен. Он считает, что ключ к прогрессу лежит в процессе нормализации, когда журналисты принимают новые технологии и включают их как норму и часть обычных процедур. Для достижения такого включения необходимо сотрудничество как внутри редакции, так и между редакцией и внешними сторонами. Так, например, успешным было сотрудничество дизайнеров и журналистов для усиления журналистской сюжетно-тематической картины (Doherty, 2016) или технических специалистов и журналистов, как в случае сети с открытым исходным кодом Hacks/Hackers (Lewis & Usher, 2014). Последняя была основана в 2009 году с целью переосмысления технологий работы с новостями и информацией. В её основу положена модель равноправного сотрудничества, когда журналисты взаимодействуют с открытых позиций с техническими специалистами и аудиторией, в результате чего возникает плодотворная гибридная схема работы, в рамках которой реализуются различные навыки и интересы (Lewis, 2012).

Обновление сферы журналистики часто приходит со стороны и стремится к центру, поскольку «новички» привносят в производство новостей дополнительную ценность в виде инновационных практик и подходов. Холтон и Белэр-Ганьон (2018, с. 72) утверждают, что «чужаки» в журналистском поле (например, веб-разработчики и программисты) «привносят в него качества, изначально не связанные с журналистской профессией [и] помогают внедрить новые способы раскрытия сущности новостей, более эффективного их предоставления и лучшего взаимодействия с новостной аудиторией». Изучение взглядов этих новых участников показывает, что они хорошо понимают журналистику как область деятельности и её специфику, но при этом обладают чертами, основанными на «этосе Кремниевой долины», включая 1) ориентацию на пользовательский опыт, 2) стремление к поиску новых применений для данных и 3) стремление к разработке простых в потреблении продуктов (Wu et al., 2019). Хотя явная рыночная логика здесь противостоит профессиональной юрисдикции журналистики, есть и общие точки соприкосновения, особенно в стремлении удовлетворить запросы пользователей и сделать журналистику более актуальной.

### В чём заключаются трудности?

Растущее количество исследований, посвящённых теме автоматизации журналистики, показывает, что машины действительно трансформируют процесс производства новостей по нескольким направлениям. Основываясь на обширном обзоре литературы, посвящённой этим исследованиям, Турман, Дёрр и Кунерт (2017) предлагают несколько аспектов, в которых влияние алгоритмов и растущее использование цифровых данных создают этические проблемы. К ним относятся способы получения, проверки и использования цифровых данных журналистами при подготовке новостей (Bradshaw, 2014) и риск предвзятости алгоритмов, обеспечивающих автоматизацию (Carlson, 2015; Gillespie, 2014; Thurman et al., 2016). Прозрачность кода и данных, конечно, является проблемой как с профессиональной, так и с демократической точки зрения (Diakopoulos, 2015; Diakopoulos & Koliska, 2017). Кроме того, это вопрос автономии машин: могут ли алгоритмы «рассуждать» и делать соответствующие выводы во всех контекстах (Young & Hermida, 2015).

С профессиональной точки зрения Карлсон (Carlson, 2015) указывает на некоторые области, вызывающие беспокойство: будущее журналистского труда, жёсткий конформизм композиционных форм новостей и нормативная основа журналистского авторитета. Карлсон приходит к выводу, что, хотя рост автоматизированной журналистики расширяет объём доступных новостей и освобождает журналистов от необходимости работать над меньшим количеством сюжетов механистического характера, он, с другой стороны, сопряжён с такими негативными прогнозами, как рост увольнений, поляризирующая персонализация материалов и коммодитизация написания новостей.

Опрашивая журналистов об их восприятии работы с автоматикой, Турман, Дёрр и Кунерт (Thurman, Dörr and Kunert, 2017) выявили несколько ограничений:

- 1) полагание на отдельные, изолированные потоки данных;
- 2) одномерный характер количественных данных, поступающих в систему;
- 3) сложность обработки этих данных;
- 4) отсутствие человеческих аспектов в генерируемых текстах;
- 5) требование заранее шаблонировать сюжеты — прогнозировать попадание в «верхние строчки»;
- 6) сложность творческой работы с данными в процессе шаблонирования.

(Thurman et al., 2017, p. 1254)

Помимо этических и профессиональных проблем, существуют также организационные проблемы, связанные с разработкой и внедрением инструментов ИИ. Чаще всего это недостаток желания, уровня знаний или ресурсов, но также здесь играют роль организационные структуры, трудности с наймом нужных специалистов и отсутствие стратегий (Beckett, 2019). ИИ — это не только технология, но и способность организации использовать преимущества технологии. Хотя ситуация варьируется, уровень ИИ-зрелости новостных организаций зачастую невысок. Отсутствует стратегическое мышление и активное лидерство в этой области, что часто связано с низким уровнем знаний и осведомлённости о новых технологиях. Инициаторами инноваций часто выступают технические специалисты ИТ-подразделений, которым бывает трудно по-настоящему внедрить их в деятельность организации. Новые технологии могут быть встречены скептически, поскольку и специалисты, и организации стремятся к сохранению статус-кво, и бывает трудно найти ресурсы и нужные навыки. Поэтому проблема внедрения инструментов ИИ волнует не только ИТ-подразделения, но и руководство и редакторов.

Как же должны действовать медиакомпании, чтобы добиться большего успеха в реализации своих стратегий в области ИИ? Бекетт (Beckett, 2019) на основе своего исследования новаторства и деятельности редакций СМИ собрал несколько советов, которые могут быть использованы в качестве отправной точки для повышения уровня самоанализа и прозрачности журналистского освещения, когда речь идёт об использовании ИИ:

- Выработайте понимание того, какой тип ИИ вам нужен, какую проблему он должен решить и как связать его внедрение с другими редакционными стратегиями и политикой.
- Какие подразделения организации затрагиваются и должны сотрудничать в решении этого вопроса?
- Определите препятствия, которые могут возникнуть, такие как наличие ресурсов, организационная культура и управление, и систематически их устраняйте.

В заключение Бекетт (2019) отмечает важность сотрудничества, чёткого распределения ролей и наличия надёжных способов измерения того, чего хочется достичь.

## ИИ преобразует журналистику

В наши дни журналисты не только общаются с пользователями, блогерами, гражданскими журналистами и т. п., но и делят рабочие места с программистами, дизайнерами пользовательского опыта и компьютерными алгоритмами. Это привело к переоценке журналистских навыков и пересмотру профессиональной идентичности (van Dalen, 2012), тем более что аудитория не всегда может отличить материалы, созданные автоматическими средствами, от текстов, написанных людьми (Clerwall, 2014). Что это означает для профессиональной идентичности журналистов?

Хотя журналисты исторически пытались противостоять попыткам стандартизировать или рационализировать их деятельность, эти действия становятся всё более бесполезными. Аналитические системы, постоянно отслеживающие эффективность их работы и сравнивающие их с коллегами, стали повсеместной частью новостной работы и управленческим инструментом (Petre, 2018). Помимо сложной логики количественной оценки, новые системы технической поддержки журналистской деятельности могут нести с собой изменение профессиональных стандартов журналистики, например, путём алгоритмизации процессов оценки новостей. Исследование поставщиков веб-аналитики показывает, что, хотя они активно стремятся понять основы журналистики, они склонны формировать в редакциях новостей нормы и ценности, ориентированные на прибыль, внедряя масштабные, связующие и заштамповывающие технологии (Holton & Belair-

Gagnon, 2018). Карлсон (Carlson, 2018) называет эту тенденцию «измеряемой журналистикой» и утверждает, что стремление автоматизированных новостных СМИ контролировать свою аудиторию будет иметь необратимые последствия для ценности журналистского суждения и автономии.

Журналистское чутьё, или нюх на новости, долгое время было основой журналистской профессии, но всё чаще воспринимается редакторами и издателями новостей как устаревшая концепция (Milosavljević & Vobič, 2019). Традиционно новостная оценка того или иного социального события строится на основе одного или нескольких критериев. Это может быть, например, принадлежность к властной элите, конфликт, сенсация, знаменитость или то, что событие затрагивает значительную часть населения (Harcup & O'Neill, 2017). Такие оценки составляют основу журналистской интуиции, то есть того, что считается важным и интересным для освещения. Совершенно противоречит профессиональному творческому подходу, существует множество успешных попыток перевести эту интуицию в алгоритмы. Роботы легко учатся обнаруживать отклонения от нормы или ожиданий: эта функция использовалась, например, для создания новостей о землетрясениях силой более 3,0 баллов по шкале Рихтера (LeCompte, 2015) или для сканирования шведской криминальной хроники на предмет отклонений от исторических данных с последующим уведомлением журналиста (Magnusson et al., 2016). Но алгоритмы могут использоваться и для работы в более тесном контакте со стандартами и практикой оценки новостей внутри организации. Одним из примеров является «алгоритм общественного обслуживания “Новостные ценности”», разработанный для Шведского радио. Цель применения алгоритма — персонализировать новостную ленту на основе оценок ценности новостей, сделанных журналистами, как для создания более релевантной ленты новостей, так и для оптимизации стандартов оценки новостей в организации. В процессе разработки сотрудники со всей компании участвовали в определении и формулировании основных ценностей, которые, по их мнению, характеризуют сильные общественные издания. Олле Закриссон, стратег по цифровым технологиям Шведского радио, говорит: «Система новостных ценностей стимулирует обсуждение того, как мы делаем журналистику. Она вдохновляет на позитивные изменения» (Zachrisson, 2021).

Профессиональная логика журналистской деятельности всегда формировалась в соответствии с организационными целями, техническими инновациями, а в последнее время и всё более сильной культурой управления (Waldenström et al., 2019). Исторически сложилось так, что технические инновации сопутствовали профессионализации журналистики и способствовали установлению консенсуса между журналистами, аудиторией, владельцами СМИ и другими заинтересованными сторонами, который обеспечивал журналистике профессиональную легитимность (Nerone, 2013, p. 452). Однако инновационная деятельность в медиаиндустрии пока идёт относительно медленно, а внедряемые средства автоматизации и искусственного интеллекта рассматриваются как подчинённые человеческому фактору (Beckett, 2019). В интервью со стратегами по инновациям, редакторами и разработчиками программного обеспечения для производства новостей прослеживается схожая точка зрения, а именно: журналисты по-прежнему занимают центральное место в этом бизнесе. Большинство средств ИИ, которые в настоящее время применяются в новостных редакциях, не настолько интеллектуальны, чтобы заменить более квалифицированную работу, но даже если бы это было так, существует нежелание отказываться от человеческого контроля над процессом. Милосавлевич и Вобич (Milosavljević and Vobič, 2019) в исследовании, проведённом среди руководителей редакций СМИ, показывают, что те скорее делают акцент на освобождении человеческого труда, чем на его фактической замене. Авторы интерпретируют это как стратегию балансирования между профессиональными ценностями и организационными целями.

## Созидательное разрушение журналистики?

Технологии ИИ открывают широкие возможности для журналистики, и редакции СМИ всё чаще используют различные решения на базе ИИ. Несмотря на то что исследования показывают, что взаимодействие человека и машины сопряжено с определёнными противоречиями, например, между конкурирующими компетенциями и логикой работы, общая тенденция направлена на повышение уровня автоматизации всей производственной цепочки. Новые технические решения расширяют зону охвата, освобождают журналистов для выполнения более качественных задач и улучшают качество курирования новостей для пользователей. Обусловленное данными развитие уже меняет основы новостной практики, поскольку доминирующей логикой работы становится количественная оценка и измеримость.

Означает ли это «созидательное разрушение» журналистики? Учитывая трансформирующее воздействие технологий данных и искусственного интеллекта, ответ будет утвердительным. Развитие технологий ведёт к переосмыслению устоявшихся бизнес-моделей, привычного уклада, профессиональных ценностей и идеалов. Это окно перемен требует переосмысления и переоценки — не только отношения к данным и технологиям, но и человеческого капитала, журналистов. Перемещение организационных ресурсов открывает возможности для более высококачественной работы, выполняемой людьми. Технологии ИИ предоставляют журналистам новые инструменты для повышения уровня их расследований и создания уникальных, оригинальных материалов. ИИ также может применяться в качестве инструмента мониторинга материалов на предмет их соответствия редакционным стандартам качества. Однако для того чтобы использовать технологии ИИ для дополнения журналистских методов работы, репортёрам необходимо больше узнать о возможностях этих технологий. До сих пор обязанности по внедрению инноваций в редакциях часто возлагаются на ИТ-отделы (Westlund et al., 2021). Чтобы обеспечить профессиональное и финансово устойчивое развитие журналистики, необходимо вовлечь в этот процесс и сотрудников редакций, и руководство. Главная причина заключается в том, что «основные элементы журналистики — подача информации в различных форматах, критическое мышление, проверка фактов, человеческие ценности, этичность, автономия и честность — становятся ещё более важными в условиях, когда “умные” машины входят во все сферы общественной и профессиональной жизни» (Lindén, 2020). С изменением фундаментальных условий профессиональная идентичность журналистов, скорее всего, будет меняться, но это развитие не следует рассматривать как угрозу. Напротив, оно свидетельствует о необходимости переоценки тех задач, качеств и функций, которые невозможно автоматизировать. Именно вокруг этих явлений и будет развиваться журналистика в тесном сотрудничестве с алгоритмами и средствами автоматизации.

Однако при анализе внедрения систем ИИ в сфере журналистики необходимо учитывать культурный контекст. Если развитие технологий ИИ происходит в первую очередь в западных странах, то профессиональные, демократические и организационные условия журналистики в развивающихся странах зачастую совершенно иные. В условиях столь глубокой трансформации, которую несёт с собой применение ИИ, эти аспекты следует рассматривать очень внимательно. Пока некоторые журналистские сообщества борются за установление профессиональных стандартов и автономии, необходимо создавать новые технические решения, которые будут способствовать достижению этих целей.

С развитием технологий развивается и система журналистского образования. Поскольку всё большее число журналистских задач автоматизируется, очень важно научить студентов использовать эти новые инструменты, чтобы оставаться конкурентоспособными в отрасли. Автоматизация уже оказывает значительное влияние на журналистский бизнес. Например, многие новостные организации используют ботов для написания простых материалов, таких как результаты спортивных соревнований или сводки погоды. Это позволяет высвободить репортёров для работы над более сложными материалами, но в то же время снижает спрос на услуги журналистов начального уровня.

В результате школам журналистики необходимо адаптировать свои учебные программы, чтобы подготовить студентов к будущему, в котором автоматизация будет занимать всё большее место. Это может выражаться в обучении студентов навыкам использования ботов и других автоматизированных инструментов, а также умению определять сюжеты, которые лучше всего подходят для автоматизации. Кроме того, в школах журналистики следует уделять больше внимания обучению «мягким» навыкам, необходимым для успешной работы в автоматизированной новостной среде, таким как креативность и адаптивность. Адаптируя свои программы к растущей автоматизации бизнеса, школы журналистики могут гарантировать, что их студенты будут готовы к будущему журналистики.



## ИИ в африканских новостных редакциях: перспективы и трудности



**Аммина Котари**, профессор журналистики и директор Харрингтонской школы коммуникаций и СМИ, Университет Род-Айленда, США.

Дискурс о технологиях ИИ в значительной степени сфокусирован на Глобальном Севере, и это касается научных исследований и их влияния. Судя по моим исследованиям, лишь немногие редакции новостей в африканских странах внедряют технологии ИИ в свою работу. В недавней статье «*Artificial Intelligence and Journalism: An Agenda for Journalism Research in Africa*» («Искусственный интеллект и журналистика: повестка дня для исследований в области журналистики в Африке») я и моя коллега Салли Энн Круикшанк (см. Kothari & Cruikshank, 2022) обсудили состояние дел африканских редакций новостей с точки зрения использования средств искусственного интеллекта. Путём систематического исследования баз данных мы хотели найти примеры применения технологий ИИ в Африке и расширить знания о соответствующих африканских инициативах, поскольку инновации в области ИИ в редакциях новостей ведутся в западных странах и Китае, и, следовательно, дискуссии об ИИ в журналистике в основном сосредоточены на западной точке зрения.

Среди ярких примеров использования технологий ИИ на африканском континенте можно назвать Марокко, где в рамках проекта «Fake news à l'épreuve des faits» («Фальшивые новости на фоне фактов») журналисты-расследователи обучались использованию средств ИИ для проверки фактов и журналистских данных. В 2020 году южноафриканская новостная организация News24 объявила о том, что будет использовать ИИ для модерации комментариев к своим материалам (News24, 2020). Южноафриканское агентство Bloomberg внедрило инструменты ИИ в свою новостную практику, а южноафриканский телеканал MSNBC использует инструменты ИИ для управления оборудованием для сбора новостей, включая видеокамеры.

В настоящее время в некоторых африканских странах ведутся многообещающие работы в области ИИ, включая две лаборатории IBM Research — в Найроби (Кения) и Йоханнесбурге (ЮАР), а также лабораторию Google AI в Аккре (Гана). Недавно созданная в Танзании лаборатория обучения ИИ и инкубатор #AIForGood нацелены на реализацию социальных задач, продвигая решения с открытыми данными для укрепления национальной экосистемы ИИ. На континенте существуют и другие инициативы в области машинного обучения.

Многие научные исследования посвящены вопросам положительного влияния технологий ИИ на различные отрасли деятельности. Однако в странах с ограниченной свободой прессы и даже авторитарными правительствами возможны серьёзные политические последствия. ИИ рассматривается как средство улучшения экономического роста, социального прогресса и государственного управления на континенте, однако потенциал ИИ может быть использован государственными структурами, которые стремятся подавить или контролировать политическую оппозицию или маргинальные группы.

В ходе обсуждения темы ИИ и журналистики в странах Глобального Севера создаётся впечатление, что интеграция ИИ будет нормой для многих редакций СМИ. Однако при более внимательном рассмотрении становится ясно, что интеграция происходит в основном в национальных или крупных новостных организациях. Многие малые и средние новостные организации в странах Глобального Севера сталкиваются с аналогичными проблемами, наличия ресурсов, что мешает им интегрировать инструменты ИИ в свою деятельность. Существует предубеждение относительно положительных последствий использования средств ИИ в журналистике. Тем не менее, необходимо проводить больше исследований и обсуждать использование инструментов ИИ для сбора новостей, особенно когда речь идёт о предвзятости, связанной с преимущественным положением вирусных сюжетов и известных источников в цифровом пространстве. Как мы знаем, развитие технологий ИИ также привело к увеличению количества фальсификаций, что затрудняет для читателей новостей возможность отличить факты от дезинформации.

Чтобы сократить технологический разрыв между Глобальным Югом и Глобальным Севером, некоторые технологические компании и учёные работают над передачей технологий или обучением журналистов использованию инструментов, разработанных на Западе. Однако такой подход не решает проблемы неравенства, поскольку искажения, заложенные в инструментах ИИ, разработанных на Глобальном Севере, будут усиливаться на Глобальном Юге, особенно в отсутствие местного опыта адап-

тации и настройки приложений. Решением проблемы может стать обучение африканцев разработке собственных инструментов, масштабируемых и отражающих профессиональные и культурные нормы их стран.

В целом, несмотря на перспективность интеграции технологий ИИ в различных секторах, существует ряд специфических проблем, связанных с внедрением инструментов ИИ в новостную деятельность стран Глобального Юга. Во-первых, это отсутствие подготовки журналистов и ресурсов для найма программистов, которые могли бы создавать инструменты ИИ в редакциях СМИ. Во-вторых, интеграция средств ИИ требует финансовых ресурсов для развития и поддержания технической базы, необходимой для эффективного использования автоматизированных инструментов, а многие редакции работают в условиях ограниченного бюджета. В-третьих, скептическое отношение журналистов к тому, что инструменты искусственного интеллекта могут сделать журналистику лучше, и, наконец, опасения, что человеческие рабочие места будут заменены автоматизированной журналистикой, будут препятствовать интеграции инструментов искусственного интеллекта в работу новостных редакций.



## Предлагаемые задания



**ПРОЧИТАТЬ И ПОРАЗМЫШЛЯТЬ.** Прочитайте одну из приведённых ниже статей, посвящённых возможностям и ограничениям использования технологий ИИ в определённом жанре журналистики, и найдите примеры их применения в своём географическом регионе. Выберите один из вариантов: а) журналистское расследование; б) спортивная журналистика; в) научная журналистика; г) иммерсивная журналистика или д) визуальная журналистика.

Stray, J. (2019). *Making artificial intelligence work for investigative journalism*. *Digital Journalism*, 7(8), 1076–1097. <https://doi.org/10.1080/21670811.2019.1630289>

Galily, Y. (2018). *Artificial intelligence and sports journalism: Is it a sweeping change?* *Technology in Society*, 54(1), 47–51. <https://doi.org/10.1016/j.techsoc.2018.03.001>

Tatalovic, M. (2018). *AI writing bots are about to revolutionize science journalism: We must shape how this is done*. *Journal of Science Communication*, 17(1), E. <https://doi.org/10.22323/2.17010501>

Sánchez Laws, A. L., & Utne, T. (2019). *Ethics guidelines for immersive journalism*. *Frontiers in Robotics and AI*, 6, Article 28. <https://doi.org/10.3389/frobt.2019.00028>

Gynnild, A. (2014). *The robot eye witness: Extending visual journalism through drone surveillance*. *Digital Journalism*, 2(3), 334–343. <https://doi.org/10.1080/21670811.2014.883184>



**ОБСУДИТЬ.** Ознакомьтесь с национальным этическим кодексом журналиста и обсудите, как использование выбранной технологии или методологии ИИ в журналистике может изменить этические нормы. Как вариант, можно проанализировать этический кодекс Общества профессиональных журналистов, доступный на разных языках (<https://www.spj.org/ethicscode.asp>).



**ВЫРАБОТАТЬ ИДЕЮ.** Определите рабочие стратегии для следующих процедур с точки зрения того, где искать открытые данные, какое программное обеспечение использовать и что делать с этим программным обеспечением. Кроме того, подумайте, какие требования к компетенции предъявляет данная методология и как их можно реализовать: а) поиск наиболее часто встречающихся слов, фраз и тем в государственных документах (например, текстовый анализ, тематическое моделирование); б) визуализация статистики с привязкой к карте региона (например, визуализация данных); в) автоматический сбор новостных тем с различных сайтов (например, агрегация).



**ПОИСКАТЬ И ИССЛЕДОВАТЬ ИСТОЧНИКИ.** Как интерактивные роботы, относящиеся к так называемому разговорному ИИ, могут помочь журналистам в их работе по взаимодействию с аудиторией? Ознакомьтесь с приведёнными ниже статьями о разговорном ИИ и журналистике и обсудите, как идеи диалоговой или взаимной журналистики могут быть реализованы, но, возможно, и подорваны средствами роботизированной журналистики.

- Gao, J., Galley, M., & Li, L. (2018, June 27). *Neural approaches to conversational AI*. Proceedings for SIGIR18: The 41<sup>st</sup> International ACM SIGIR Conference on Research & Development in Information Retrieval, pp. 1371–1374. <https://doi.org/10.1145/3209978.3210183>
- Shin, D. (2021, March 12). *The perception of humanness in conversational journalism: An algorithmic information-processing perspective*. New Media & Society. Advance online publication. <https://doi.org/10.1177/1461444821993801>
- Veglis, A., & Maniou, T. A. (2019). *Chatbots on the rise: A new narrative in journalism*. Studies in Media Communication, 7(1), 1–6. <https://doi.org/10.11114/smc.v7i1.3986>



**СОЗДАТЬ.** Проведите опрос обычных людей по методу Гэллага, спросив у 5–6 случайно или специально отобранных человек, насколько они доверяют автоматизированным средствам, создающим новостные тексты, и какие у них имеются представления о журналистике, создаваемой искусственным интеллектом. Какие могут возникнуть ошибочные представления о журналистике и ИИ? Исходя из результатов интервью, что необходимо разъяснить широкой аудитории о журналистике, создаваемой искусственным интеллектом? Что заставляет людей доверять журналистике?



## Дополнительная литература

- Beckett, C. (2019). *New powers, new responsibilities: A global survey of journalism and artificial intelligence*. Report from POLIS journalism and Society, think tank at the London School of Economics (LSE). <https://www.lse.ac.uk/media-and-communications/polis/JournalismAI/The-report>
- Bradshaw, P. (2014). *Data journalism*. In: Zion, K., & Craig, D. (Eds.) *Ethics for digital journalists: Emerging best practices*. Routledge, 202–219. <https://doi.org/10.4324/9780203702567>
- Broussard, M. (2015). *Artificial intelligence for investigative reporting*. Digital Journalism, 3(6), 814–831.
- Broussard, M., Diakopoulos, N., Guzman, A. L., Abebe, R., Dupagne, M., & Chuan, C.-H. (2019). *Artificial intelligence and journalism*. Journalism & Mass Communication Quarterly, 96(3), 673–95. <https://doi.org/10.1177/1077699019859901>
- Carlson, M. (2015). *The robotic reporter*. Digital Journalism, 3(3), 416–431.
- Carlson, M. (2017). *Automated judgment? Algorithmic judgment, news knowledge, and journalistic professionalism*. New Media & Society, 8(4). <https://doi.org/10.1177/1461444817706684>
- Diakopoulos, N. (2019). *Automating the news: How algorithms are rewriting the media*. Harvard University Press.
- Gillespie, T. (2014). *The relevance of algorithms*. In: Gillespie, T., Boczkowski, P., & Foot, K. (Eds.) *Media technologies: Essays on Communication, materiality, and society*. MIT Press, 167–194. <https://doi.org/10.7551/mitpress/9780262525374.001.0001>
- Lewis, S. C., Guzman, A. L., & Schmidt, T. R. (2019). *Automation, journalism, and human–machine communication: Rethinking roles and relationships of humans and machines in news*. Digital Journalism, 7(4), 409–427. <https://doi.org/10.1080/21670811.2019.1577147>
- Marconi, F. (2020). *Newsmakers: Artificial intelligence and the future of journalism*. Columbia University Press.
- Thurman, N., Dörr, K., & Kunert, J. (2017). *When reporters get hands-on with robo-writing*. Digital Journalism, 5(10), 1240–1259.
- Wu, S., Tandoc, E. C. Jr., & Salmon, C. T. (2019). *Journalism reconfigured*. Journalism Studies, 20(10), 1440–1457. <https://doi.org/10.1080/21670811.2019.1577147>



# Восемнадцать «ловушек» в ИИ-журналистике

Саяш Капур и Арвинд Нараянан

30 сентября 2022 г. Ознакомьтесь с записью в блоге, в которой представлен данный перечень: <https://aisnakeoil.substack.com/p/eighteen-pitfalls-to-beware-of-in>

## ⇒ Некорректное сравнение человека и ИИ

**В чём суть?** Некорректное сравнение средств ИИ и человека, подразумевающее, что средства ИИ и человек похожи в том, как они обучаются и работают.

**Почему это является проблемой?** Вместо того чтобы описывать ИИ как широкий набор инструментов, такие сравнения антропоморфируют инструменты ИИ и подразумевают, что инструменты ИИ обладают потенциалом выступать в качестве субъектов действий в реальном мире.

### Ловушка 1. Приписывание искусственному интеллекту субъектности действий

Описывать системы ИИ как способные предпринимать действия, не зависящие от человеческого контроля, или подразумевать, что они скоро будут способны делать это.

**«Искусственный интеллект начинает брать на себя выполнение повторяющихся задач в классной комнате, таких как выставление оценок».**

– «Машины учатся, и учащиеся тоже», The New York Times

### Ловушка 2. Сuggestивные образы

Изображения человекоподобных роботов часто используются для иллюстрации статей об ИИ, даже если содержание статьи не имеет никакого отношения к роботам. Это создаёт у читателей ложное впечатление, что инструменты ИИ имеют конкретную форму, хотя это всего лишь программное обеспечение, которое изучает закономерности на основе данных.

– «Как элитные инвесторы используют средства искусственного интеллекта и машинное обучение, чтобы получить преимущество», CNN

▶ Эта статья не имеет никакого отношения к роботам и посвящена инструментам искусственного интеллекта для поиска закономерностей в финансовых данных.

### Ловушка 3. Сравнение с человеческим интеллектом

В некоторых случаях в статьях, посвящённых ИИ, подразумевается, что алгоритмы ИИ обучаются так же, как и люди. Например, часто встречаются сравнения алгоритмов глубокого обучения с тем, как работает человеческий мозг. Такие сравнения могут придать убедительность утверждениям о том, что ИИ является «разумным», как отмечают д-р Тимнит Гебру и д-р Маргарет Митчелл в своей недавней аналитической статье.

**«[Исследование] было посвящено методу ИИ, называемому глубоким обучением, который использует алгоритмы, большие данные и вычислительные мощности для эмуляции человеческого интеллекта».**

– «ИИ может быть столь же эффективен в диагностике заболеваний, как и врачи-специалисты», CNN

▶ «Эмуляция человеческого интеллекта» — это не совсем точное описание того, как работает глубокое обучение. Это даёт читателям ложное представление о том, что алгоритмы глубокого обучения сравнимы с человеческим интеллектом.

## Ловушка 4. Сравнение с человеческими навыками

Аналогичным образом, в статьях часто сравнивается эффективность работы инструментов ИИ с человеческими навыками при решении той или иной задачи. Это даёт ложное представление о том, что инструменты ИИ и люди конкурируют на равных, умалчивая о том, что инструменты ИИ работают только в узком диапазоне условий.

**«В новом научном обзоре сделан вывод, что искусственный интеллект может быть способен диагностировать заболевания так же успешно, как и медицинские работники».**

– «ИИ может быть столь же эффективен в диагностике заболеваний, как и врачи-специалисты», CNN

- ▶ Это предложение умалчивает тот факт, что инструменты искусственного интеллекта выполняют лишь малую часть из множества этапов, составляющих процесс диагностики.

## ⇒ Преувеличенные, некорректные или не поддающиеся обоснованию утверждения об ИИ

**В чём суть?** Спекулятивные, сенсационные или некорректные утверждения об инструментах ИИ могут привести к распространению шумихи вокруг ИИ.

**Почему это является проблемой?** Подобные заявления создают ложное представление о прогрессе в области ИИ и затрудняют определение истинных достижений.

## Ловушка 5. Преувеличение

Описание систем ИИ как революционных или новаторских без конкретных доказательств их эффективности создаёт ложное впечатление о том, насколько они будут полезны в тех или иных условиях. Эта проблема усугубляется, когда инструмент ИИ применяется в условиях, где известно, что в прошлом он терпел неудачи — мы должны скептически относиться к эффективности инструментов ИИ в таких условиях.

**«В течение многих лет люди пытались перестроить образовательный процесс с помощью искусственного интеллекта, но реальный прогресс был достигнут только после революции средств машинного обучения, произошедшей за последние семь лет».**

– «Машины учатся, и учащиеся тоже», The New York Times

- ▶ Это утверждение в статье не имеет под собой никаких доказательств. Фактически за последнее десятилетие образовательные технологии доказали свою несостоятельность.

## Ловушка 6. Некорректное сравнение с историческими преобразованиями

Сравнение средств ИИ с такими крупными историческими преобразованиями, как изобретение электричества или промышленная революция, — отличный маркетинговый приём. Однако, когда в новостных статьях используются подобные термины, они могут создать ложное ощущение потенциала и прогресса, особенно если эти утверждения не подкреплены реальными фактами.

**«По мнению Альтмана, разворачивающаяся революция в области ИИ вполне может оказаться для человечества более значимой, чем предшествующие сельскохозяйственная, промышленная и компьютерная революции вместе взятые».**

– «Неужели ИИ наконец-то приблизился к человеческому интеллекту?», Financial Times

- ▶ В статье слепо цитируется Сэм Альтман, соучредитель OpenAI, сравнивающий развитие технологий ИИ с историческими преобразованиями.

## Ловушка 7. Неоправданные заявления о будущем прогрессе

Утверждения о том, как будущие разработки средств ИИ повлияют на ту или иную отрасль; например, подразумевается, что средства ИИ неизбежно будут полезны в этой отрасли. Когда такие утверждения делаются без доказательств, они являются всего лишь домыслом со стороны автора статьи и, как и прежде, могут создать ложное впечатление об этих разработках.

*«Например, сегодня чат-боты могут быть неуклюжими и раздражающими, но со временем они станут настоящими собеседниками, изучат наши привычки и характер и даже выработают собственную индивидуальность».*

– «ИИ здесь, там, повсюду», The New York Times

## Ловушка 8. Ложные утверждения о положении дел

В некоторых случаях в статьях могут содержаться ложные утверждения о том, что может делать инструмент искусственного интеллекта.

*«Система автоматического оценивания сама учит себя оценивать успеваемость».*

– «Машины учатся, и учащиеся тоже», The New York Times

- ▶ В статье ошибочно утверждается, что инструменты ИИ могут самостоятельно научиться оценивать домашние задания, в то время как на самом деле они просто оценивают ответы учеников в сравнении с правильными ответами, введёнными в систему.

## Ловушка 9. Некорректные утверждения о содержании исследования

В новостных статьях часто приводятся ссылки на научные исследования для обоснования своих утверждений. К сожалению, иногда наблюдается расхождение между утверждениями, сделанными со ссылкой на научное исследование, и тем, что в нём сообщается на самом деле.

*«Исследования показывают, что системы [машинного обучения] могут повысить успеваемость учащихся намного выше уровня обычных занятий и даже выше уровня, достигнутого учащимися, занимающимися с репетиторами».*

– «Машины учатся, и учащиеся тоже», The New York Times

- ▶ В цитируемом здесь исследовании машинное обучение не упоминается ни разу.

## Ловушка 10. Глубокомысленно звучащие термины для описания простых действий

Как отмечает профессор Эмили Бендер в своей работе, посвящённой анализу шумихи вокруг ИИ, использование таких фраз, как «фундаментальный акт предсказания следующего слова» или «магия ИИ», подразумевает, что инструмент ИИ в процессе своей работы совершает нечто выдающееся. Подобные фразы умалчивают о том, насколько обыденными являются задачи, и что инструменты ИИ функционируют именно так, как ожидалось.

*«Я прошу богов искусственного интеллекта включить свет».*

– «ИИ здесь, там, повсюду», The New York Times

## ⇒ Слепая поддержка тех, кто преследует собственные интересы

**В чём суть?** В новостных материалах для обоснования утверждений часто используются пиар-заявления и цитаты представителей компаний, без должного контекста и сбалансированного изложения материала.

**Почему это является проблемой?** Акцентирование внимания на мнениях заинтересованных сторон без представления альтернативных точек зрения может создать слишком оптимистичное ощущение состояния дел.

## Ловушка 11. Отношение к представителям компаний и научным работникам как к нейтральной стороне

Если в статье приводятся только или преимущественно цитаты представителей компаний или научных

работников, создавших инструмент искусственного интеллекта, то, скорее всего, она будет излишне оптимистичной в отношении потенциальных преимуществ этого инструмента.

**«Испытания ИИ в условиях цифрового перехода к сдаче университетских экзаменов».**

– Financial Times

- ▶ Почти вся статья написана с точки зрения компании, продающей инструменты искусственного интеллекта. В результате статья больше похожа на пиар-материал и меньше на новость.

## Ловушка 12. Повторение или использование пиар-терминов и заявлений

В новостных статьях часто повторяются термины из пиар-заявлений компаний, вместо того чтобы описать принцип работы того или иного инструмента искусственного интеллекта. Это может привести к появлению недостоверных формулировок, искажающих реальные возможности того или иного инструмента.

**«Она использует платформу Вакрах, которая может читать почерк учеников и автоматически выставлять оценки за школьные работы».**

– «Машины учатся, и учащиеся тоже», The New York Times

- ▶ В статье неоднократно повторяются такие пиар-обороты, как «читать почерк учеников» и «автоматически оценивать» домашние задания. Хотя компания Вакрах уже закрылась, мы нашли эти пиар-обороты в архивной версии сайта компании.

## ⇒ Игнорирование ограничений

**В чём суть?** Подчёркиваются потенциальные преимущества инструмента ИИ, но не рассматриваются и не подчёркиваются возможные ограничения.

**Почему это является проблемой?** Односторонний анализ средств ИИ может скрыть потенциальные ограничения этих средств.

## Ловушка 13. Отсутствие обсуждения потенциальных ограничений

Такие недостатки, как недостаточное подтверждение корректности работы, необъективность и возможность двойного применения, характерны для большинства средств искусственного интеллекта. Когда эти ограничения не обсуждаются, у читателей может сложиться искажённое представление о рисках, связанных с инструментами ИИ.

**«Машины учатся, и студенты тоже».**

– The New York Times

- ▶ В статье не обсуждаются потенциальные ограничения использования ИИ в области образовательных технологий. В статье вскользь упоминается, что могут возникнуть проблемы с конфиденциальностью, но следом приводится цитата разработчика инструмента, чтобы снять эти опасения.

## Ловушка 14. Преуменьшение значения ограничений

Даже если в статье обсуждаются ограничения технологии и приводятся слова экспертов, которые могут их объяснить, в структуре статьи они часто преуменьшаются, например, помещаются в конец статьи или им отводится мало места.

**«Некоторые скептики утверждают, что программное обеспечение способно лишь на слепое подражание...»**

– «ИИ осваивает язык. Стоит ли доверять тому, что он говорит?», The New York Times Magazine

- ▶ Вместо того, чтобы вступить в дискуссию по существу с профессором Эмили Бендер и другими авторами, эта статья в The New York Times Magazine ограничивает их точку зрения рамками «скептицизма». Профессор Бендер гораздо подробнее обсуждает проблемы с этой статьёй в своём ответе на неё.



## Ловушка 16. Преуменьшение значения человеческого труда

При обсуждении средств искусственного интеллекта в статьях часто подчёркивается роль технических достижений и замалчивается человеческий труд, необходимый для создания системы или поддержания её в рабочем состоянии. Книга «Призрачная работа», написанная доктором Мэри Л. Грей и доктором Сиддхартом Сури, показывает, насколько важен этот невидимый труд. Преуменьшение значения человеческого труда вводит читателей в заблуждение, заставляя думать, что инструменты ИИ работают автономно, вместо того чтобы пояснить, что они требуют значительных накладных расходов в виде человеческого труда, как об этом говорит профессор Сара Т. Робертс.

**«Обилие онлайн-курсов и учебников также освободило преподавателей от чтения лекций».**

– «Машины учатся, и учащиеся тоже», The New York Times

- ▶ Эта фраза не учитывает труд, который затрачивается на запись лекций, поддержание онлайн-ресурсов, и переключает внимание с человеческого труда, который затрачивается на поддержание этой системы.

## Ловушка 17. Показатели производительности сообщаются без оценки неопределённости измерений или пояснений

В новостной статье редко бывает достаточно места, чтобы объяснить, как рассчитываются такие показатели, как точность, для конкретного приложения и что они собой представляют. Включение в текст статьи таких показателей, как «точность 90%», без указания условий, при которых они рассчитываются, может дезинформировать читателей в отношении эффективности того или иного инструмента ИИ, тем более что, как известно, производительность инструментов ИИ снижается даже при незначительных изменениях в наборах данных, на которых они оцениваются.

**«По всей видимости, новый компьютерный алгоритм теперь может прогнозировать возникновение преступлений в большом городе по соседству с вами. Алгоритм, разработанный социологами из Чикагского университета с заявленной точностью 90%...».**

– «Алгоритм сможет предсказывать преступления в городах США до их совершения», Bloomberg

- ▶ В статье нет никаких подробностей, позволяющих читателю понять, что означает точность 90%.

## Ловушка 18. Ошибка непостижимости

Считать инструменты искусственного интеллекта непостижимыми «чёрными ящиками» — это категоричная ошибка. Вместо того чтобы возлагать ответственность на разработчиков этих инструментов за их конструктивные решения, эта ошибка переводит внимание на технические аспекты системы. Журналисты должны требовать от разработчиков отчёта о работе инструментов ИИ, а не называть их «чёрными ящиками» и позволять разработчикам уходить от ответственности.

**«Наши машины теперь обладают знаниями, которые мы никогда не поймём».**

– WIRED

- ▶ Основной тезис статьи заключается в том, что невозможно понять, как модели ИИ «рассуждают», только потому, что они имеют большое количество параметров или весов. При этом игнорируется целый ряд исследований, посвящённых интерпретируемости и объяснимости моделей ИИ. Далее автор утверждает, что, поскольку мы не можем понять внутренние механизмы модели, нет возможности использовать её таким образом, который был отвечал законодательным требованиям по недискриминации и объяснимости, например, в области оценки кредитного рейтинга. Однако эти требования касаются не столько внутреннего устройства алгоритмов принятия решений, сколько их взаимодействия с миром, а алгоритмы всегда можно понять на этом уровне, как отмечает д-р Кролл.

Изменения:

30 сентября добавлены ссылки на соответствующие работы, изменён пример в «ловушке» 18, исправлен текст в «ловушке» 7.



# Послесловие: подготовка журналистов, компетентных в вопросах ИИ

**Маарит Яаккола**

В данном пособии рассматривается тема искусственного интеллекта (ИИ), или то, как машины, созданные как интеллектуальные агенты, оснащённые системами убеждений, могут освещаться в журналистике и рассматриваться в рамках журналистского образования для подготовки будущих журналистов. Общая цель — дать преподавателям журналистики базовые знания об искусственном интеллекте как демократической проблеме, чтобы они могли **научить студентов-журналистов освещать события в обществе, всё более насыщенном технологиями.**

Основные положения этого пособия, к которым возвращаются многие модули, касаются двойственного характера технологий ИИ. Приложения ИИ обладают как полезным, так и вредоносным потенциалом, поэтому они могут оказывать как положительное, так и отрицательное влияние на общество. Однако в конечном счёте ИИ — это всегда то, что из него делают люди. В то же время именно люди решают, как использовать возможности, которые предоставляют технологии ИИ, хотя и в условиях разграничения власти и контроля.

! Однако мы, люди, должны обеспечить, по терминологии Microsoft, справедливость, подотчётность, прозрачность и этичность использования ИИ, что является одной из важнейших задач общества.

Поэтому необходимы журналисты, которые могли бы снять опасения общественности и стимулировать дискуссию на эту тему.

В задачу журналиста не входит принятие решений о разработке и использовании средств ИИ, но **журналисты играют ключевую роль в своевременной передаче знаний** директивным органам, разработчикам и другим специалистам-практикам, научно-исследовательским организациям и другим производителям знаний, преподавателям и прочим посредникам знаний, а также рынку и субъектам гражданского общества. Необходимым условием для этого является формирование у журналистов понимания того, что представляет собой ИИ. Базовое понимание предмета ИИ и связанных с ним дискуссий требует, чтобы люди испытывали опасения по поводу некоторых фундаментальных идей, связанных с этой технологией, например, того, что средства ИИ являются репрезентациями или моделями мира и используются для построения логических рассуждений. Компьютеры могут воспринимать мир с помощью датчиков, они могут обучаться на основе данных. Они могут решать задачу обеспечения взаимодействия средств ИИ с людьми (Long & Magerko, 2020.) Понимание явления ИИ с критических позиций требует прослеживания того, что в основе системных структур лежат решения, вуалирующие экономические, этические или ценностные политические интересы, которые имеют последствия, поддерживающие структуры власти.

Технологии используются в СМИ и в обществе уже давно. Новостные редакции могут учиться на опыте общества. Современное внедрение средств искусственного интеллекта в практику работы редакции не может считаться революционным, как и появление компьютеров в редакциях. Последствия изменений происходят медленнее и постепенно вписываются в существующие структуры. Однако, как и в случае с обществом в целом, мы должны задаться вопросом, что это означает для журналистского образования? Возможно, требуется частичная *интеграция* темы ИИ в существующие структуры образования и журналистские будни, а некоторые аспекты существующих структур и ценностей необходимо укрепить.

**С распространением автоматизированных систем, запрограммированных на выполнение задач и функций без надзора, вероятно, появятся совершенно новые вопросы и практические подходы.** Вместо того чтобы просто заявлять о необходимости создания ИИ-журналистов и программ по ИИ-журналистике, журналистское образование может сделать шаг назад и задуматься о том, как **перестроить образовательные подходы и практики.**

В заключение я хотела бы обобщить концептуальные, дискурсивные, компетентностные, просветительские и дидактические задачи, которые ИИ как явление ставит перед журналистским образованием. По большому счёту, преподаватели журналистики должны задаться вопросами:

- как относиться к тому, что может стать постгуманистическим сдвигом, когда субъекты-не-люди всё чаще становятся частью всех видов взаимодействий в обществе;
- как осознать дискурсивную роль журналистики в формировании будущих представлений об ИИ;
- как воспринимать ИИ как набор знаний и компетенций, которыми необходимо овладеть;
- как расширить возможности журналистов, способствующих выполнению миссии просвещения граждан и потребителей, необходимого для восприятия происходящих изменений; и
- как менять свою собственную деятельность, задаваясь вопросом о том, как ИИ может послужить производству развивающего опыта для редакций новостей — а также для общества в целом.

---

## ⇒ **Концептуальные задачи**

*В дискуссиях об искусственном интеллекте основное внимание часто уделяется тому, как ИИ изменит общество, однако стоит задуматься и о том, как ИИ может изменить человека. Развёртывание систем ИИ всё больше влияет на наше поведение и на наше восприятие реальности.*

Хотя мы можем научить средства ИИ работать определённым образом, использование ИИ также вынуждает нас корректировать своё поведение и в дальнейшем учиться на этих изменениях. Технологии на базе ИИ, как правило, носят овеществлённый, иммерсивный характер и связаны с нашими собственными возможностями наблюдения и принятия решений. Механизм отбора, убеждения, налаживания связей и социализации влияет на принятие решений и поведение людей.

Традиционно первенство отдавалось человеку, но распространение **интеллектуальных агентов** порождает то, что называют **пост- или трансгуманизмом, постдуализмом или постчеловеческим состоянием**. Согласно этим философским подходам, человек перестаёт считаться центром всего сущего. Поэтому возникает стремление признать ценность различных нечеловеческих способов действия в мире. Возможное постчеловеческое состояние побуждает нас рассматривать общество не только как результат человеческого поведения, но и как результат **взаимодействия между людьми** (и их властными различиями) и **нечеловеческими или ачеловеческими субъектами**. Для журналистов проблема заключается в том, как создать сюжет, не экзотизируя и не инаковизируя интеллектуальных агентов, не антропоморфизировав и не очеловечивая материальные артефакты и абстрактные процессы, не оставаясь слишком отстранённым и циничным.

В журналистике включение нечеловеческих субъектов в работу над новостями может привести к экзистенциальному кризису. Как уже отмечалось многими учёными, основное понимание журналистики заключается в том, что это ремесло, созданное человеком, в котором его вклад в развитие общества является решающим.

Этот сдвиг парадигмы порождает необходимость нового концептуального изменения, создающего потребность в поддержании нового типа отношений с окружающим миром, когда человек рассматривается в большей степени как соратник, находящийся на одном уровне с неживым миром. Это требует большей интеллектуальной любознательности для выяснения того, как работают нечеловеческие контрагенты (хотя и созданные людьми и определёнными конфигурациями власти), а также повышенного сознания **этической ответственности человека и человечества**. И здесь, поскольку в научной фантастике по-

стоянно возникают образы машин и роботов, захватывающих мир, необходимо попытаться преодолеть культурные мифы и посмотреть на технологии более реалистично, более соразмерно позиционируя интеллект человека, животных и машин.

## ⇒ Дискурсивные задачи

Дискуссии вокруг ИИ характеризуются большим разнообразием мнений, в основном представителей Глобального Севера, и широким спектром тем, которые имеют различную актуальность для разных регионов мира. Технологии ИИ и последствия их применения обсуждаются в самых разных научных областях, включая информатику, информационные науки, социологию, антропологию, научно-технические исследования, политологию и др. Сфера научных исследований остаётся довольно ограниченной, они опираются на понятийно-терминологический аппарат, методологию и предыдущие обсуждения в рамках конкретной дисциплинарной традиции, требуя от аудитории предварительного знания концепций и контекстов. Доклады по вопросам нормативного регулирования, технические и стратегические документы готовятся органами власти, отраслевыми или общественными инициативами и некоммерческими организациями, которые зачастую смотрят на инновации как на решение проблем с оптимистической и обнадеживающей точки зрения. Отраслевая точка зрения, как правило, подчёркивает широкий круг демократических вопросов и аспекты общего блага, однако технологические гиганты преследуют свои собственные интересы. Обыденные взгляды пользователей, как граждан и потребителей, сильно разнятся.

Журналистика находится в центре этих дискурсов, обсуждая как терминологию, так и традиции. Уместно **говорить об ИИ как о «пограничном объекте»** — понятии, используемом в качестве зонтичного термина во всеобщем дискурсе для объединения различных субъектов публичной сферы с целью обсуждения общей проблемы на языке, который может быть неродным для областей, к которым они принадлежат (Malinverni, 2015; Moran & Shaikh, 2022, Reddy et al., 2019). Различия между дискурсами не просто семантические: исследования показали, что, например, журналистика и отрасль ИИ излагают установки и выражают аспекты озабоченности, которые могут противоречить друг другу (Moran & Shaikh, 2022). Частично общественное понимание концепций ИИ основано на публичном дискурсе об ИИ, который создаётся журналистикой.

*Поскольку СМИ способны в значительной степени формировать общественные дебаты и определять ход дискуссии, журналистике необходимо задавать ключевые вопросы, например, как ИИ может служить обществу и как он должен работать, и как он применяется не только непосредственно в сфере журналистики, но и в других сферах жизнедеятельности общества, причём совершенно разными способами.*

Как отмечается в настоящем руководстве, публичный дискурс в новостных СМИ по вопросам ИИ характеризуется определённой предвзятостью, например доминированием отраслевых источников, крупных транснациональных организаций и их новых продуктов или инициатив, а также инсайдерских и экспертных мнений. **Гендерные, расовые и региональные факторы предвзятости систем ИИ**, например, когда определённые алгоритмы подавляют мнения с определёнными характеристиками и корнями, могут быть легко перенесены в журналистский дискурс и содержание репортажей, если журналисты не осознают их, и дискриминация людей с определённым цветом кожи или полом может стать нормой.

Мы живём в эпоху, когда постоянно публикуются новые исследования, появляются новые понятия для обозначения возникающих и уже существующих явлений, а также новые методы, позволяющие охватить многогранную область. Преподавателям, студентам и специалистам в области журналистики необходимо следить за новостями и исследованиями, чтобы быть в курсе новейших достижений в области машинного обучения, компьютерного зрения, распознавания речи, обработки естественного языка, экспертных систем и робототехники, а также последствий их применения. Поскольку журналист, скорее всего, отслеживает ИИ как область деятельности, связанную с другими областями, например с экологической устойчивостью, школьным образованием, промышленностью, дезинформацией и медиаграмотностью

или здравоохранением, важное значение имеют курируемые сервисы, собирающие информацию на стыке различных отраслей.

## ⇒ Компетентностные задачи

В конечном итоге преподаватели журналистики должны меньше учить студентов технологиям и больше внимания уделять тому, что мы обычно называем «мягкими навыками». В отраслевом отчёте Poynter, подготовленном почти десять лет назад (Finberg & Klinger, 2014), основные компетенции, которые считались необходимыми для будущих журналистов, были связаны с **личными качествами** (любопытность, аккуратность, способность справляться со стрессом и соблюдать сроки), **личными умениями** (широкие общие знания, умение оценивать новости, осведомлённость о текущих событиях, критичность), **способностями, связанными с работой и организационным поведением** (социальные навыки, умение работать в команде), знаниями (о медийном ландшафте, обществе, технологиях, праве и этике, истории), а также **практикой и методами работы** (поиск и получение информации, техника проведения интервью, анализ информации и данных, поиск источников, техника презентации и подачи материала, создание сетей, развитие и вовлечение аудитории). Эти общие навыки применимы ко всем аспектам журналистского образования, включая понимание ИИ как системы прорывных технологий.

К ним можно добавить такие навыки, как способность творчески воспринимать изменения и инновации, а также способность критически оценивать их с учётом исторической перспективы. Среди других необходимых навыков — математическую грамотность, чтобы, например, ориентироваться в данных, статистике и масштабах явлений, а также способность транслировать идеи из одного мира в другой. В основном это касается передачи идей из областей с узкоспециализированной структурой и терминологией, таких как технологии (часто технологии в конкретной области, например, обработка естественного языка в образовательном секторе), в повседневную плоскость.

Как в эпистемологическом, так и в методологическом плане журналисты, пишущие об ИИ, могут опираться на многие устоявшиеся журналистские подходы. Стоит подчеркнуть, что вопросы, связанные с ИИ, могут быть включены в любой журналистский жанр — от новостной до культурной журналистики, поэтому не всегда журналистские подходы нуждаются в адаптации и изменениях. Однако технически ориентированные жанры — будь то вычислительная журналистика, журналистика данных, автоматизированная журналистика, алгоритмическая журналистика, роботизированная журналистика или «дополненная журналистика» — углубляют связь журналистики с данными и компьютерными системами. Технологически-обусловленные способы освещения событий, такие как иммерсивная журналистика и мобильная журналистика, а также совместная работа журналистов, пишущих о данных, журналистов-программистов, программистов и хакеров, опираются на «журналистику ИИ». Подходы, ориентированные на решение проблем и развитие, такие как конструктивная журналистика, журналистика, основанная на решениях, позитивная журналистика, журналистские расследования и проверка фактов, помогают журналистам создавать сюжеты, альтернативные научно-фантастическим фантазиями, **анализировать и проверять информацию и «со-мыслить» с аудиторией**. Более того, **методы журналистики, основанные на участии, взаимности и диалоге приглашают граждан и потребителей к исследованию повседневных аспектов особенно узкоспециального ИИ**.

## ⇒ Просветительские задачи

Распространение технологий ИИ также ставит сложные задачи перед обществом в целом. Зачастую общественное понимание вопросов, связанных с ИИ, ограничено. В настоящее время в разных странах мира предпринимаются различные меры по внедрению технологий ИИ в программы школьного обучения и популяризации обучения грамотности в области ИИ, которые связаны с усилиями стран по созданию национальных стратегий в области ИИ. Например, в Великобритании правительственное Управление по

ИИ опубликовало в 2021 году стратегию в области ИИ, в которой говорится, что «каждый человек сможет уверенно жить с технологиями ИИ, а для специалистов, которые будут работать с ними и создавать их, будет заложена самая лучшая основа деятельности» (2021, с. 16). Технологические гиганты и общественные организации выдвигают инициативы, направленные на повышение осведомлённости граждан о сфере ИИ. Однако на втором плане остаётся тот факт, что журналистика, являясь социальным смыслообразующим институтом, также занимается просвещением своей аудитории. Таким образом, в отношении задач, которые ИИ ставит перед обществом, журналистов следует рассматривать как просветителей в области ИИ (см. Broussard et al., 2019; Deuze & Beckett, 2022; Jaakkola, 2022).

Под грамотностью в области ИИ понимаются компетенции, необходимые в будущем, в котором технологии ИИ изменят наши представления о жизни, процессе общения и работе с машинами (Long & Magerko, 2020). Человек может считаться **грамотным в области ИИ**, если он компетентен в использовании ИИ этически ответственным образом (Ng et al., 2021). Условия понимания систем ИИ или человеко-машинного взаимодействия для разных групп населения будут совершенно разными, поскольку люди могут иметь самый разный опыт работы с технологиями.

*Как всегда, не существует среднестатистического читателя или слушателя. Люди по-разному связаны с цифровой средой. Существуют значительные различия в навыках владения доминирующими языками, общей грамотности, мировоззрении, умении обрабатывать данные, уровне математической грамотности и способности решать проблемы — во всём том, что можно подвести под понятие медиа- или информационной грамотности, или, возможно, в более узком смысле, грамотности в области данных или ИИ.*

Соответственно, журналистские темы, образовательные и журналистские мероприятия с просветительским уклоном должны разрабатываться по-разному для разных аудиторий — школьников, студентов вузов, специалистов-практиков в различных областях, пожилых людей и представителей меньшинств. В этих группах общей специфической проблемой для понимания работы ИИ является его сущность. Однако это понятие абстрактно, универсально и часто непрозрачно для пользователей систем ИИ. Для повышения уровня понимания среди людей, не имеющих непосредственного опыта в этой области, наиболее эффективным было бы поставить человека на место интеллектуального агента и попросить его разобраться в процессах построения рассуждений, во встроённом искусственном интеллекте. Журналистика может сделать это разнообразя способы подачи информации, которые не обязательно должны строиться только на её посредничестве. Она также может вовлекать людей через погружение, геймификацию и приобщение.

## ⇒ Дидактические задачи

И наконец, что не менее важно, **преподавателям журналистики необходимо определить, как лучше использовать средства ИИ для преподавания и изучения журналистики**. Приведённые примеры, начиная от симуляций человеческих и человеко-компьютерных взаимодействий в иммерсивной журналистике и заканчивая инструментами для тренировки навыков написания текстов — это только начало. Приложения ИИ могут помочь начинающим журналистам отточить свои навыки в перечисленных выше областях компетенции. Создание **дидактических метавселенных** для моделирования ситуаций и сред, соответствующих журналистским способностям, может способствовать вовлечению и **воплощению развивающего опыта**.

Для создания новых практических и масштабируемых дидактических решений, то есть готовых к распространению по вузам и странам, преподавателям журналистики необходимо сотрудничать с разработчиками систем ИИ. Таким образом, создание образовательных ресурсов перестанет быть делом рук отдельного преподавателя. Необходимость обновления компетенций отдельных людей — это совместная работа, требующая участия преподавателей журналистики, студентов и журналистской индустрии в целом.

# Словарь терминов

**Автоматизированная журналистика**, также известная как алгоритмическая или роботизированная журналистика, означает использование программного обеспечения для создания журналистских материалов без вмешательства человека.

**Алгоритм** — это последовательность правил для решения какой-либо задачи в вычислительном устройстве.

**Алгоритмическая журналистика**, см. автоматизированная журналистика.

**Алгоритмическая предвзятость** — это понятие описывает систематические и повторяющиеся ошибки в компьютерной системе, которые приводят к несправедливым результатам, например, к предпочтению одной произвольной группы пользователей перед другими. Предвзятость может быть обусловлена многими факторами, включая процесс разработки алгоритма, непреднамеренные или непредусмотренные сценарии использования или решения, связанные методами программирования, сбора, отбора или использования данных для обучения алгоритма. Алгоритмическая предвзятость наблюдается, например, в результатах поисковых систем и на платформах социальных сетей.

**Андроид** — человекоподобный робот или синтетический организм, имитирующий человека.

**Антропоморфизм** — это интерпретация или восприятие нечеловеческого объекта в терминах человеческих характеристик. Таким образом, он означает очеловечивание объекта, который не является человеком, например, присвоение программному обеспечению человеческих имён.

**Большие данные** — это массивы данных, которые слишком велики или сложны для обработки традиционным прикладным программным обеспечением для работы с данными.

**Бот**, или интернет-робот, — это программное приложение, выполняющее автоматизированные задачи в Интернете, например, чат-бот, который общается с пользователем. Некоторые боты безвредны, но существуют и вредоносные боты («плохие боты»), осуществляющие атаки и мошеннические операции. Самыми первыми чат-ботами были ELIZA (1966) и PARRY (1972). Боты также называют «разговорным ИИ».

**Вредоносный («плохой») бот**, см. бот.

**Вычислительная журналистика** — под вычислительной журналистикой понимается применение компьютерных вычислений в журналистской практике. См. также автоматизированная журналистика.

**Глубокое обучение** — это класс алгоритмов машинного обучения для постепенного извлечения признаков из исходных данных, организованных послойно. Например, при обработке изображений нижние слои данных могут определять контуры, а верхние — понятия, значимые для человека, такие как цифры, буквы или лица. Термин был введён Риной Дехтер в 1986 году.

**Гуманоид** — нечеловеческий объект, имеющий человеческую форму или характеристики.

**Датафикация** — это процесс перевода аспектов жизни людей в данные и их трансформация в новую форму ценности. Термин был введён Кеннетом Кукером и Виктором Майер-Шёнбергером в 2013 году.

**Дипфейк или «глубокая подделка»** (от англ. deep learning — «глубокое обучение» и fake — «подделка») — манипулирование материалом, при котором существующее изображение или видео изменяется путём замены некоторых его составляющих на что-то другое. В качестве примера можно привести изменение выражения лица человека на видео или копирование голосов людей.



**Добыча данных** — это междисциплинарная область научных исследований, связанная с процессом извлечения и обнаружения закономерностей в больших объёмах данных. Извлечение («добыча») закономерностей может быть основано на выявлении групп записей данных (кластерный анализ), отклонений в записях (обнаружение аномалий) или зависимостей между записями (поиск ассоциативных правил).

**Дружественный искусственный интеллект (Дружественный ИИ)** — концепция, обозначающая доброкачественное использование ИИ, поиск способа заставить технологии ИИ работать на положительные и благие человеческие цели, используя безопасные и полезные агенты. Термин был введён и популяризирован Элиэзером Юдковским. Родственным понятием является «хороший ИИ», также используемый в контексте концепции «общества хорошего ИИ».

**Законы Азимова**, известные также как три закона робототехники, представляют собой набор принципов, разработанных писателем-фантастом Айзеком Азимовым в 1940-х годах. Согласно первому закону, «робот не может причинить вред человеку или своим бездействием допустить, чтобы человеку был причинён вред». Второй закон гласит, что «робот должен повиноваться всем приказам, которые даёт человек, кроме тех случаев, когда эти приказы противоречат Первому Закону». Согласно третьему закону, «робот должен заботиться о своей безопасности в той мере, в которой это не противоречит Первому или Второму Законам». Первоначальные законы были изменены и доработаны Азимовым и другими авторами. Эти законы широко проникли в научную фантастику и популярную культуру, а также повлияли на размышления об этике ИИ.

**Законы Кларка** — это три принципа, сформулированные писателем-фантастом Артуром Кларком в 1960–1970-х годах и описывающих восприятие технологий. Первый закон гласит, что «когда уважаемый, но пожилой учёный утверждает, что что-то возможно, то он почти наверняка прав. Когда он утверждает, что что-то невозможно, — он, весьма вероятно, ошибается». Согласно второму закону, «единственный способ обнаружения пределов возможного состоит в том, чтобы отважиться сделать шаг в невозможное». Третий закон, наиболее часто цитируемый, гласит что «любая достаточно развитая технология неотличима от магии». Айзек Азимов сформулировал дополнение к первому закону Кларка: «Тем не менее, когда непрофессионалы сплачиваются вокруг идеи, отрицаемой уважаемым, но пожилым учёным, и поддерживают эту идею с пылом и энтузиазмом — то этот уважаемый, но пожилой учёный, в конечном счёте, вероятно, прав».

**Зловещая долина (также эффект «зловещей долины»)** — гипотеза, описывающая связь между человекоподобным внешним видом роботизированного объекта и эмоциональной реакцией человека. Она подразумевает, что роботы, похожие на людей, воспринимаются положительно до определённого момента, когда они становятся слишком реальными. С этого момента уровень симпатии значительно падает, и люди испытывают чувство тревоги. Таким образом, под «долиной» понимается спад симпатии наблюдателя к человекоподобному объекту. Отвращение к человекоподобным роботам возрастает по мере увеличения их человекоподобия. Впервые термин и механизм этого явления были описаны в 1970-х годах Масахиро Мори, профессором Токийского технологического института.

**Иммерсивная журналистика** — это форма журналистики, которая позволяет аудитории погрузиться в виртуально воссозданные миры с помощью технологий иммерсивной виртуальной реальности, таких как компьютерная графика (CGI) или 3D-технологии.

**Интеллектуальный агент** — это то, что способно к автономному восприятию и действию в окружающей среде, а также к использованию накопленных знаний для корректировки или улучшения своего поведения. Согласно общепринятому определению, если агент действует так, чтобы максимизировать успех своей деятельности на основе прошлого опыта и знаний, он может считаться «интеллектуальным».

**Интернет вещей (англ. Internet of Things, IoT)** — это совокупность физических объектов, оснащённых технологиями подключения и обмена данными с другими устройствами через Интернет. Например, многие бытовые устройства, такие как дистанционно управляемые стиральные машины, водонагреватели, холодильники, системы освещения или автомобили, являются примерами IoT. Термин появился примерно в 2008 году.

**Кибербола** — неологизм (от «кибер» и «гипербола»), обозначающий преувеличенные утверждения о возможностях или перспективах развития интернет-технологий.

**Киборг** — это существо (от слов «кибернетический» и «организм»), имеющее как органические, так и искусственные части тела. Термин был введён Манфредом Клайном и Натаном Клайном в 1960 году.

**Компьютерное зрение** — это междисциплинарная область научных исследований, которая занимается изучением того, как компьютеры могут получать информацию высокого уровня из цифровых изображений или видео. Эта область знаний стремится понять и автоматизировать задачи, которые может решать зрительная система человека, разрабатывая методы получения, обработки, анализа и распознавания цифровых изображений. Компьютерное зрение может использоваться для реконструкции сцен, обнаружения объектов или событий, видеосопровождения целей, распознавания объектов, оценки положения в 3D, оценки движения, 3D-моделирования сцен или восстановления изображений.

**Машинное обучение (МО)** — это компьютерные алгоритмы, которые могут автоматически совершенствоваться благодаря накоплению опыта и использованию так называемых обучающих данных.

**Нейронная сеть (также искусственная нейронная сеть, ИНС)**, — вычислительные системы, созданные по образцу биологических нейронных сетей, составляющих мозг животных. Нейронная сеть — это набор связанных между собой узлов, имитирующих структуру нейронов в биологическом мозге.

**Нормативно-правовая база** — это набор правил или рекомендаций, которые (особенно в значении «меры государственной политики») используются в качестве основы для принятия решений.

**Обработка естественного языка (ОЕЯ)** — это область знаний на стыке лингвистики и информатики, связанная с искусственным интеллектом и изучающая взаимодействие между компьютерами и человеческим языком, в частности, способы программирования компьютеров для обработки и анализа больших объёмов данных на естественном языке. Естественный язык — язык, который сформировался у человека естественным образом в результате использования и повторения без какого-либо сознательного планирования, отличается от искусственных или созданных языков, таких как языки компьютерного программирования.

**Платформизация** означает повышение в коммуникации роли платформ — каркасов, на которых могут выполняться различные приложения, обычно принадлежащих крупным компаниям Google, Apple, Facebook (Meta), Amazon, Microsoft, называемых GAFAM (GAMAM) или «большой пятёркой»

**Пузырь фильтров**, также известный как «идеологический фрейм», описывает ситуацию интеллектуальной изоляции в Интернете, которая является результатом отделения от информации, противоречащей убеждениям и точкам зрения пользователя. Пузырь фильтров может возникать на основе алгоритма персонализации, который избирательно угадывает, какую информацию хотел бы видеть пользователь.

**Робот** — это машина, способная выполнять задачи в автоматическом режиме. Впервые этот термин был использован Карелом Чапеком для обозначения вымышленного гуманоида в пьесе 1920 года на чешском языке.

**Роботизированная журналистика**, см. автоматизированная журналистика.

**Сверхинтеллект**, см. сингулярность.

**Сильный ИИ**, также известный как общий ИИ, полный ИИ или искусственный интеллект общего назначения — понятие, которое описывает машины, способные применять интеллект для решения любой задачи, а не только одной конкретной.

**Символический ИИ** — исторический этап развития искусственного интеллекта, способного моделировать высокоуровневые логические рассуждения, используемые людьми при решении головоломок, формулировании правовых аргументов и выполнении математических расчётов. Другое название — GOFAI (англ. Good Old-Fashioned Artificial Intelligence, «старый добрый искусственный интеллект»), термин, придуманный Джоном Хогландом в 1985 году.

**Сингулярность**, понятие, также известное как технологическая сингулярность, означает момент времени, когда интеллект вычислительной техники разовьётся до такой степени, что превзойдёт интеллект человека, в результате чего возникнет сверхинтеллект, гиперинтеллект или сверхчеловеческий интеллект. Таким образом, под сверхинтеллектом понимается гипотетический агент, обладающий интеллектом, превосходящим человеческий разум. Концепция технологической сингулярности была разработана футуристом Раймондом Курцвейлом в его книге в 2005 году.

**Слабый ИИ**, также известный как узкий ИИ или прикладной ИИ — понятие, которое описывает машины, способные применять интеллект для решения одной узкой задачи, а не любой проблемы. Это наиболее часто фигурирующая в публичных дискуссиях форма ИИ.

**Тест Тьюринга** — это тест на способность машины демонстрировать разумное поведение, эквивалентное или неотличимое от человеческого. Тест измеряет «человечность» объекта тестирования и основан на предположении, что если люди воспринимают действия нечеловеческого агента как человеческие, то он может считаться разумным. Идея была придумана Аланом Тьюрингом в 1950-х годах, который назвал её имитационной игрой.

**Чёрный ящик** — в науке, вычислительной технике и машиностроении «чёрным ящиком» называют систему, которую можно воспринимать с точки зрения её входов и выходов без знания её внутреннего устройства, которое непрозрачно («черно»). Говоря «чёрный ящик» применительно к искусственному интеллекту или алгоритмам имеется в виду, что мы не знаем, как они устроены.

# Авторы

## Эксперты

**Арвинд Нараянан** — профессор информатики Принстонского университета. Он является соавтором учебника по вопросам справедливости в сфере машинного обучения и в настоящее время пишет книгу «Чудеса ИИ» (*AI Snake Oil*). Он возглавлял проект Принстонского университета по веб-прозрачности и подотчётности (Princeton Web Transparency and Accountability Project), направленный на выявление того, как компании собирают и используют нашу личную информацию. Его работа была одной из первых, показавших, как машинное обучение отражает культурные стереотипы, а его диссертационное исследование показало фундаментальные ограничения методов обезличивания. Нараянан является лауреатом Президентской премии за раннюю карьеру для учёных и инженеров (Presidential Early Career Award for Scientists and Engineers, PECASE), дважды лауреатом премии за технологии, улучшающие конфиденциальность (Privacy Enhancing Technologies Award), и трижды лауреатом премии в области исследования конфиденциальности для директивных органов (Privacy Papers for Policy Makers Award).

**Чарли Бекетт** — практикующий профессор факультета СМИ и коммуникаций Лондонской школы экономики (ЛШЭ) в Великобритании. Он является директором Polis, международного аналитического центра по журналистике при университете. В 2018 году он был ведущим комиссаром Комиссии ЛШЭ по вопросам правды, доверия и технологий в связи с информационным кризисом. В настоящее время возглавляет проект Polis/ЛШЭ JournalismAI («ИИ-журналистика») — международную инициативу по сотрудничеству с новостными организациями по всему миру в области научных исследований, информирования и инноваций с использованием искусственного интеллекта. Он является автором доклада «Новая сила и новая ответственность» (New Powers and New Responsibilities), посвящённого тому, как новостные организации по всему миру используют технологии искусственного интеллекта. До прихода в ЛШЭ в 2006 году профессор Бекетт работал старшим продюсером и редактором в BBC и на четвёртом канале новостей компании ITN.

**Эдвард Финн** — директор-основатель Центра науки и воображения в Аризонском государственном университете (США), где он является доцентом Школы будущего инноваций в обществе и Школы искусств, медиа и инженерии. В своих исследованиях и преподавательской деятельности он изучает работу воображения, вычислительную культуру и перспективы будущего. Он является автором книги «Чего хотят алгоритмы: воображение в эпоху вычислений» (*What Algorithms Want: Imagination in the Age of Computing*, MIT Press, 2017) и соредактором книг «Фантастика в будущем времени» (*Future Tense Fiction*, Unnamed Press, 2019), «Франкенштейн: аннотация для учёных, инженеров и творцов всех мастей» (*Frankenstein: Annotated for Scientists, Engineers and Creators of All Kinds*, MIT Press, 2017) и «Иероглиф: истории и картины лучшего будущего» (*Hieroglyph: Stories and Visions for a Better Future*, William Morrow, 2014), а также других книг. Профессор Финн получил степень доктора по английскому языку в Стэнфордском университете в 2011 году и степень бакалавра в Принстонском университете в 2002 году. До поступления в аспирантуру он работал журналистом в журналах Time, Slate и Popular Science.

**Фредрик Хайнц** — профессор компьютерных наук Линчёпингского университета (Швеция), где он руководит Лабораторией моделирования рассуждений и обучения. Его исследования посвящены искусственному интеллекту, особенно теме заслуживающего доверия ИИ и пересечения процессов машинных рассуждений и машинного обучения. Являясь членом Шведской королевской академии инженерных наук, профессор Хайнц также является директором Программы развития образования в области ИИ и трансформационных технологий им. Валленберга, директором Высшей школы им. Валленберга по развитию ИИ, автономных систем и программного обеспечения, координатором сети TAILOR ICT-48, разрабатывающей научные основы заслуживающего доверия ИИ, и президентом Шведского общества ИИ.

**Фредерик Хейманс** работает сотрудником по исследованиям и проектам в Центре знаний «Данные и общество» (Kenniscentrum Data & Maatschappij) — центре изучения правовых, этических и общественных аспектов искусственного интеллекта и приложений на основе данных, при Врьейском университете Брюсселя (Бельгия). Г-н Хейманс работает в исследовательской группе в области медиа, инноваций и технологий Научно-исследовательского центра нано- и цифровых технологий университета над различными проектами, посвящёнными влиянию систем искусственного интеллекта на общество, проводит анализ нормативно-правовой базы, готовит доклады и публикации, создаёт практические инструменты, ориентированные на целевые группы Центра знаний.

**Сантош Кумар Бисвал** — доцент кафедры журналистики и массовых коммуникаций Женского университета Рама Деви в Бхубанесваре, Индия. До прихода в Женский университет Рама Деви д-р Бисвал работал доцентом в Институте средств массовой информации и коммуникации Symbiosis International (Университет Димед) в Пуне, Индия. Ранее он работал в колледже Андрха Лойола и в газете Hindustan Times (Нью-Дели). Получил высшее образование и золотую медаль по специальности «Социальная коммуникация» в Уткальском университете, Бхубанесвар, Индия. Имеет степени магистра и доктора гуманитарных наук, диплом о высшем образовании в области журналистики и массовых коммуникаций, диплом Комиссии по университетским грантам UGC-NET. В сферу его научных интересов входят журналистика, коммуникации в области развития, проблемы инвалидности, цифровые СМИ и медиатехнологии. Помимо многочисленных статей в журналах, он является автором (в качестве соредактора) книги *«Комплексные подходы к культуре и коммуникации бренда в различных отраслях промышленности»* (*Holistic Approaches to Brand Culture and Communication across Industries*) и соредактором *«Справочника по исследованию социальной и культурной динамики в индийском кино»* (*Handbook of Research on Social and Cultural Dynamics in Indian Cinema*). Он регулярно пишет обзорные статьи для таких англоязычных национальных ежедневных изданий, как The Indian Express, The Asian Age, The Telegraph, Hindustan Times, The Pioneer, The Statesman и Deccan Chronicle.

**Агнес Стенбом** является ответственным специалистом по данным и искусственному интеллекту в компании Schibsted, крупнейшей медиагруппе в странах Северной Европы, и руководит различными стратегическими мероприятиями, связанными с технологиями искусственного интеллекта. Параллельно она защищает докторскую диссертацию по промышленной экономике и менеджменту в Королевском технологическом институте в Стокгольме (Швеция), и её исследования посвящены тому, как традиционные медиаорганизации могут использовать средства искусственного интеллекта для достижения журналистских целей. Г-жа Стенбом является одним из основателей журналистской сети Nordic AI Journalism, объединяющей коллег из более чем 40 медиа-организаций и занимающейся обсуждением вопросов искусственного интеллекта. Она также является обозревателем газеты Dagens Media, аналитиком по вопросам будущего в шведском аналитическом центре Global Challenge и постоянным докладчиком на конференциях. В 2021 году она получила награду Международной ассоциации новостных СМИ (International News Media Association) «Тридцать лучших до тридцати» (30 Under 30) как одна из 30 лучших молодых талантов в медиаиндустрии.

**Саяш Капур** — кандидат компьютерных наук Принстонского университета. Его исследования посвящены влиянию ИИ на науку, журналистику и промышленность. В настоящее время он является соавтором книги *«Чудеса ИИ»* (*AI Snake Oil*). В книге критически рассматривается вопрос о том, что может и чего не может сделать ИИ. Он работал над проблемой понимания социального влияния ИИ в промышленности и научных кругах в Facebook, Колумбийском университете и Швейцарском университете EPFL. Г-н Капур — лауреат премии за лучший доклад на конференции Ассоциации вычислительной техники по вопросам справедливости, подотчётности и прозрачности и премии за признание влияния на конференции Ассоциации вычислительной техники по теме компьютеризации совместной работы.

**Сурен Джаясурия** — доцент факультета искусств, СМИ и инженерии Аризонского государственного университета, а также факультета электротехники, компьютерной техники и энергетики. В 2017 г. он получил степень доктора наук по электротехнике и вычислительной технике в Корнельском университете, а в 2012 г. — степень бакалавра по математике и философии в Питтсбургском университете. Его научные интересы связаны с компьютерным зрением и вычислительной визуализацией, в частности с разработкой новых методов вычислительной фотографии и проекционных систем с применением машинного обучения и искусственного интеллекта для лучшего понимания окружающего нас визуального мира. Профессор Джаясурия также проводит исследования в области искусственного интеллекта и образования в области науки, технологии, инженерии, искусства и математики, ориентированного на различные

категории учащихся, начиная со средней/старшей школы и заканчивая бакалавриатом/магистратурой, а также занимается информационно-разъяснительной работой и привлечением общественности.

**Дженни Вийк** — доцент кафедры медиа- и коммуникационных исследований Гётеборгского университета и Университета Мальмё (Швеция). В настоящее время она участвует в инновационной программе «СМИ и демократия» в научном парке Линдхольмена в Гётеборге (Швеция), где руководит исследовательским проектом «Автоматизация журналистики: инновации, сотрудничество и управление знаниями при внедрении средств искусственного интеллекта в новостных организациях». В центре научных интересов Дженни Вийк — журналистика как демократический институт, особое внимание уделяется профессиональному развитию и организации работы редакций СМИ. Она занимала различные руководящие должности в академических кругах, в частности, благодаря ей была создана международная магистерская программа Гётеборгского университета по журналистским расследованиям.

## Редактор

**Маарит Яаккола** — содиректор Nordicom, Центра исследований СМИ Северных стран при Гётеборгском университете (Швеция). Она также является доцентом факультета журналистики, СМИ и коммуникаций Гётеборгского университета и адъюнкт-профессором факультета информационных технологий и коммуникационных наук Университета Тампере (Финляндия). Сфера научных интересов Яакколы — журналистское образование и его характерные методические приёмы, культурные цифровые практики, культурная журналистика и критика, а также политические и методические аспекты медиаграмотности. Её последняя книга — «*Рецензирование культуры в Интернете: постинституциональная культурная критика разных платформ*» (*Reviewing Culture Online: Post-Institutional Cultural Critique across Platforms*, Palgrave Macmillan, 2022). Кроме того, Яаккола является главным редактором NordMedia Network, цифровой платформы, посвящённой исследованиям скандинавских СМИ, и членом редколлегий многих научных журналов. Она возглавляет несколько научных сетей, среди которых Комитет по сотрудничеству в сфере Северного журналистского образования, являющегося ассоциацией-членом Всемирного совета по журналистскому образованию, Университетская сеть ЮНЕСКО по медийной и информационной грамотности и межкультурному диалогу, занимающаяся вопросами медиаграмотности и соответствующими исследованиями в сфере высшего образования, Академический форум по исследованиям в области медийной и информационной грамотности в Швеции и исследовательское подразделение NordMedia Network по медиаграмотности и медиаобразованию.

# Библиографический список

- Acerbi, A. (2020). *Cultural evolution in the digital age*. Oxford University Press.
- Adgate, B. (2021, August 20). *Newspapers have been struggling and then came the pandemic*. Forbes. <https://www.forbes.com/sites/bradadgate/2021/08/20/newspapers-have-been-struggling-and-then-came-the-pandemic/?sh=b6d9a1212e64>
- Aliman, N.-M. & Kester, L. (2020). *Malicious design in AVR, falsehood and cybersecurity-oriented immersive defenses*. Paper presented at the 2020 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR) in 2020, 130–137.
- Andi, S., Selva, M., & Nielsen, R. K. (2020). *Women and leadership in the news media 2020: Evidence from ten markets*. Reuters Institute for the Study of Journalism.
- Bawack, R. E., Fosso Wamba, S., & Carillo, K. (2019). *Artificial intelligence in practice: Implications for IS research*. Paper presented at the 25th American conference on information systems. Cancun, Mexico.
- BBC (2018). *India WhatsApp 'child kidnap' rumours claim two more victims*. BBC, June 11, 2018. <https://www.bbc.com/news/world-asia-india-44435127>
- Bebawi, S. (2019). *Data journalism and investigative reporting in the Arab world: From emotive to evidence-based journalism*. In: Mutsvauro, B., Bebawi, S., & Borges-Rey, E. (Eds.), *Data journalism in the Global South. Palgrave Studies in Journalism and the Global South*. Palgrave Macmillan, 193–204.
- Beckett, C. (2019). *New powers, new responsibilities: A global survey of journalism and artificial intelligence*. Report from POLIS journalism and Society, think tank at the London School of Economics (LSE).
- Biswal S. K., & Gouda N. K. (2020). *Artificial intelligence in journalism: A boon or bane?* In: Kulkarni A. & Satapathy S. (Eds.). *Optimization in machine learning and applications: Algorithms for intelligent systems*. Springer, 155–167.
- Bloomfield, B. P. (1987). *The question of artificial intelligence: Philosophical and sociological perspectives*. Routledge.
- Bloomfield, M. (1995). *The automated society: What the future will be and how we will get it that way*. Masefield Books.
- Bounegru, L., & Gray, J. (Eds.) (2021). *The data journalism handbook: Towards a critical data practice*. Amsterdam University Press.
- Brantner, C., & Saurwein, F. (2021). *Covering technology risks and responsibility: Automation, artificial intelligence, robotics, and algorithms in the media*. International Journal of Communication, 15(2021), 5074–5098.
- Brennen, J. S., Howard, P. N., & Nielsen, R. K. (2018). *An industry-led debate: How UK media cover artificial intelligence*. [https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2018-12/Brennen\\_UK\\_Media\\_Coverage\\_of\\_AI\\_FINAL.pdf](https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2018-12/Brennen_UK_Media_Coverage_of_AI_FINAL.pdf)
- Broersma, M., & Graham, T. (2015). *Tipping the balance of power: Social media and the transformation of political journalism*. In: *The Routledge companion to social media and politics*. Routledge, 89–103.
- Broussard, M., Diakopoulos, N., Guzman, A. L., Abebe, R., Dupagne, M., & Chian, C.-H. (2019). *Artificial intelligence and journalism*. Journalism & Mass Communication Quarterly, 96(3), 673–695.
- Bradshaw, P. (2014). *Data journalism*. In: Zion, K., & Craig, D. (Eds.) *Ethics for digital journalists: Emerging best practices*. Routledge, 202–219.
- Brundage, M., Shahar, A., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., Dafoe, A., Scharre, P., Zeitzoff, T., Filar, B., Anderson, H., Roff, H., Allen, G.C., Steinhardt, J., Flynn, C., hÉigeartaigh, S.Ó., Beard, S., Belfield, H., Farquhar, S., Lyle, C., Crootof, R., Evans, O., Page, M., Bryson, J., Yampolskiy, R., & Amodei, D. (2018). *The malicious use of artificial intelligence: Forecasting, prevention, and mitigation*. Future of Humanity Institute.
- Bruns, A. (2019). *After the 'APocalypse': Social media platforms and their fight against critical scholarly research*. Information, Communication & Society, 22(11), 1544–1566. <https://doi.org/10.1080/1369118X.2019.1637447>
- Brüggemann, M. (2017). *Shifting roles of science journalists covering climate change*. Oxford Encyclopedia of Climate Science, (2017, April 26).

- Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. W.W. Norton & Company.
- Bunz, M., & Braghieri, M. (2021). *The AI doctor will see you know: Assessing the framing of AI in news coverage*. *AI & Society*, (2021, March 8).
- Burke, R. (2002). *Hybrid recommender systems: Survey and experiments*. *User Modeling and User-adapted Interaction*, 12(4), 331–370. <https://doi.org/10.1023/A:1021240730564>
- Carlson, M. (2015). *The robotic reporter: Automated journalism and the redefinition of labor, compositional forms, and journalistic authority*. *Digital Journalism*, 3(3), 416–431. <https://doi.org/10.1080/21670811.2014.976412>
- Carlson, M. (2018). *Automating judgment? Algorithmic judgment, news knowledge, and journalistic professionalism*. *New Media & Society*, 20(5), 1755–1772. <https://doi.org/10.1177/1461444817706684>
- Cassie, K. (2018). *Journey to AI*. [https://www2.slideshare.net/ITARENA/cassie-kozyrkov-journey-to-ai?from\\_action=save](https://www2.slideshare.net/ITARENA/cassie-kozyrkov-journey-to-ai?from_action=save)
- Cater, L., & Heikkilä, M. (2021, May 28). *Your boss is watching: How AI-powered surveillance rules the workplace*. Politico. <https://www.politico.eu/article/ai-workplace-surveillance-facial-recognition-software-gdpr-privacy/>
- Cath, C. Wachter, S., Mittelstadt, B., Taddeo, M. & Floridi, L. (2018). *Artificial intelligence and the “good society”: The US, EU, and UK approach*. *Science and Engineering Ethics*, 24(2), 505–528. <https://doi.org/10.1007/s11948-017-9901-7>
- Cathcart, T. (2013). *The trolley problem or would you through the fat guy off the bridge*. Workman Publishing.
- CFI (n.d.). *Data journalism*. Available at <https://cfi.fr/en/search?search=data+journalism&type%5Bproject%5D=project>
- Chuan, C. H., Tsai, W.-H. S., & Cho, S. Y. (2019). *Framing artificial intelligence in American newspapers*. In: *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. ACM Press, 330–244. <https://doi.org/10.1145/3306618.3314285>
- Chui, M., Harryson, M., Manyika, J., Roberts, R., Chung, R., van Heteren, A, & Nel, P. (2018). *Notes from the AI frontier: Applying AI for social good*. McKinsey Global Institute.
- Clerwall, C. (2014). *Enter the robot journalist: Users’ perceptions of automated content*. *Journalism Practice*, 8(5), 519–531. <https://doi.org/10.1080/17512786.2014.883116>
- Code for Africa (n.d.). *Code for Africa (CfA)*. <https://github.com/CodeForAfrica/>
- Coelho, F., Rodríguez Altube, F., Louzau, N., Arambillet, D., Bouret, G., Miño, G., Brom, R., Morón Escalante, L., Trigo Viera, M., Villazuela, C., Zenteno Frers, S., & Zuzel, M. (2021). *Computer vision to monitor elections transparency*. Sigma Awards. <https://sigmaawards.org/computer-vision-to-monitor-elections-transparency/>
- Collett, C., Neff, G., & Gouvea, L. (2022). *The effects of AI on the working lives of women*. Inter-American Development Bank.
- Dastin, J. (2018, October 11). *Amazon scraps secret AI recruiting tool that showed bias against women*. U.S. Available at <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>
- Deuze, M. (2006). *Global journalism education: A conceptual approach*. *Journalism Studies*, 7(1), 19–34.
- Deuze, M., & Beckett, C. (2022). *Imagination, algorithms and news: Developing AI literacy for journalism*. *Digital Journalism*, 10(10), 1913–1918. <https://doi.org/10.1080/21670811.2022.2119152>
- Diakopoulos, N. (2014). *Algorithmic accountability*. *Digital Journalism*, 3(3), 398–415.
- Diakopoulos, N., & Koliska, M. (2017). *Algorithmic transparency in the news media*. *Digital Journalism*, 5(7), 809–828. <https://doi.org/10.1080/21670811.2016.1208053>
- Diakopoulos, N. (2019). *Automating the news: How algorithms are rewriting the media*. Harvard University Press.
- Diakopoulos, N. (2021). *The algorithms beat: Angles and methods for investigation*. *The Data Journalism Handbook*, 219–229. Amsterdam University Press. [https://doi.org/10.5117/9789462989511\\_ch31](https://doi.org/10.5117/9789462989511_ch31)
- Doherty, S. (2016). *NewsCubed: Journalism through design*. *Journalism Practice*, 10(5), 569–588. <https://doi.org/10.1080/17512786.2015.1049645>
- Elbanna, A., & Engesmo, J. (2020, August 19). *A-level results: Why algorithms get things so wrong – and what we can do to fix them*. *The Conversation*. <https://theconversation.com/a-level-results-why-algorithms-get-things-so-wrong-and-what-we-can-do-to-fix-them-142879>
- Eldridge II, S. (2018). *Repairing a fractured field: Dynamics of collaboration, normalization and appropriation at intersections of newswork*. *The Journal of Applied Journalism & Media Studies*, 7(3), 541–559. [https://doi.org/10.1386/ajms.7.3.541\\_1](https://doi.org/10.1386/ajms.7.3.541_1)
- Elish, M. C., & Boyd, D. (2018). *Situating methods in the magic of Big Data and AI*. *Communication Monographs*, 85(1), 57–80. <https://doi.org/10.1080/03637751.2017.1375130>
- European Commission (2018). *Artificial intelligence for Europe*. <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A237%3AFIN>



- European Commission (2018). *Coordinated plan on artificial intelligence*. <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52018DC0795&rid=3>
- Fast, E., & Horvitz, E. (2017). *Long-term trends in the public perception of artificial intelligence*. Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, 31(1), 963–969. <https://doi.org/10.1609/aaai.v31i1.10635>
- Fenn, J. (2010). *Hype cycle for emerging technologies*. <https://infota.siss.cl/concesiones/empresas/P05/ESSAL/13%20Estudio%20Final/Respaldo%20Gastos/Anexo%20N%203%20TIC/Gartner/Gartner/2009%20y%202010/Hype-Cycle-for-Emerging-Technologies-2010.pdf>
- Fengler, S., Lengauer, M., & Zappe, A.-C. (Eds.) (2021). *Reporting on migrants and refugees: Handbook for journalism educators*. UNESCO.
- Ferrer-Conill, R., & Tandoc, E.C. (2018). *The audience-oriented editor*. Digital Journalism, 6(4), 436–453.
- Finberg, H. I., & Klinger, L. (2014). *Core skills for the future of journalism*. The Poynter Institute of Media Studies.
- Finn, E. (2017, September 27). *Art by algorithm*. Aeon. <https://aeon.co/essays/how-algorithms-are-transforming-artistic-creativity>
- Finn, E. (2017). *What algorithms want: Imagination in the age of computing*. MIT Press.
- Flores-Saviaga, C., Savage, S., & Taraborelli, D. (2016). *LeadWise: Using online bots to recruit and guide expert volunteers*. CSCW '16 Companion: Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion, 257–260. <https://doi.org/10.1145/2818052.2869106>
- Floridi, L., & Chiriatti, M. (2020). GPT-3: *Its nature, scope, limits, and consequences*. Minds and Machines, 30(4), 681–694. <https://doi.org/10.1007/s11023-020-09548-1>
- Floridi, L., & Cowls, J. A. (2019). *Unified framework of five principles for AI in society*. <https://philarchive.org/archive/FLOAUF>.
- Fosso Wamba, S., Bawack, R. E., Guthrie, C., Queiroz, M. M., & Carillo, K. D. A. (2021). *Are we preparing for a good AI society? A bibliometric review and research agenda*. Technological Forecasting and Social Change, 164, 505–528. <https://doi.org/10.1016/j.techfore.2020.120482>
- Garvey, C., & Maskal, C. (2020). *Sentiment analysis of the news media on artificial intelligence does not support claims of negative bias against artificial intelligence*. OMICS: A Journal of Interactive Biology, 24(5), 286–299. <https://doi.org/10.1089/omi.2019.0078>
- Geiger, G. (2021, March 1). *How a discriminatory algorithm wrongly accused thousands of families of fraud*. Vice. <https://www.vice.com/en/article/jgq35d/how-a-discriminatory-algorithm-wrongly-accused-thousands-of-families-of-fraud>
- Gillespie, T. (2014). *The relevance of algorithms*. In: Tarleton, G., Boczkowski, P., & Foot, K. (Eds.) *Media technologies: Essays on communication, materiality, and society*. MIT Press, 167–194.
- Goodyer, J. (2021, October 14). *How an AI finished Beethoven's last symphony and what that means for the future of music*. Science Focus. <https://www.sciencefocus.com/news/ai-beethovens-symphony/>
- Graefe, A. (2016). *Guide to automated journalism*. Tow Center for Digital Journalism.
- Gutierrez-Lopez, M., Missaoui, S., Makri, S., Porlezza, C., Cooper, G., & Macfarlane, A. (2019). *Journalists as design partners for AI*. Paper presented at the CHI 2019 ACM Conference on Human Factors in Computing Systems, 4–9 May, 2019, in Glasgow.
- Haenlein, M., & Kaplan, A. (2019). *A brief history of artificial intelligence: On the past, present, and future of artificial intelligence*. *Californian Management Review*, 61(4), 5–14.
- Hamon, R., Junklewitz, H., & Sanchez, I. (2020). *Robustness and explainability of artificial intelligence: From technical to policy solutions*. Technical report. Publications Office of the European Union.
- Hansen, M., Roca-Sales, M., Keegan, J., & King, G. (2017). *Artificial intelligence: Practice and implications for journalism*. Report from the policy exchange forum on June 13, 2017. Tow Center for Digital Journalism.
- Hanitzsch, T., Hanusch, V., Ramaprasad, J., & de Beer, A. S. (Eds.) (2019). *Worlds of journalism: Journalistic cultures around the globe*. Columbia University Press.
- Harcup, T., & O'Neill, D. (2017). *What is news? News values revisited (again)*. Journalism Studies, 18(12), 1470–1488. <https://doi.org/10.1080/1461670X.2016.1150193>
- Hofkirchner, W., & Burgin, M. (Eds.) (2017). *The future information society: Social and technological problems*. New Jersey: World Scientific Publishing.
- Holton, A. E., & Belair-Gagnon, V. (2018). *Strangers to the game? Interlopers, intralopers, and shifting news production*. Media and Communication, 6(4), 70–78. <https://doi.org/10.17645/mac.v6i4.1490>
- Hu, C. (2018, December 11). *In a world of "algorithmic culture," music critics fight for relevance*. Columbia Journalism Review. <https://www.cjr.org/analysis/algorithms-music.php>
- Hunter, M. L. (2011). *Story-based inquiry: A manual for investigative journalists*. UNESCO.

- Ideas, W. (2022, February 16). *The case of the creepy algorithm that "predicted" teen pregnancy*. Wired. <https://www.wired.com/story/argentina-algorithms-pregnancy-prediction/>
- Impe, A.-M. (2019). *Reporting on violence against women and girls: A handbook for journalists*. UNESCO.
- Iretton, C., & Posetti, J. (2018). *Journalism, fake news & disinformation: Handbook for journalism education and training*. UNESCO.
- Jaakkola, M. (2022). *Journalists as media educators: Journalistic media education as inclusive boundary work*. *Journalism Practice*, 16(6), 1265–1285. <https://doi.org/10.1080/17512786.2020.1844040>
- Jamil, S. (2019). *Increasing accountability using data journalism: Challenges for the Pakistani journalists*. *Journalism Practice*, 15(1), 19–40. <https://doi.org/10.1080/17512786.2019.1697956>
- Jemio, D., Hagerty, H. & Arandra, F. (2022, February 16). *The case of the creepy algorithm that "predicted" teen pregnancy*. Wired. <https://www.wired.com/story/argentina-algorithms-pregnancy-prediction/>
- Jia, P., & Xia, Q. (2008). *Knowledge management, collaboration and innovation*. 2008 International Seminar on Future Information Technology and Management Engineering, Leicestershire, UK, 2008, 228–232. <https://doi.org/10.1109/FITME.2008.104>
- Jung, J., Song, H., Kim, Y., Im, H., & Oh, S. (2017). *Intrusion of software robots into journalism: The public's and journalists' perceptions of news written by algorithms and human journalists*. *Computers in Human Behavior*, 71(2017), 291–298. <https://doi.org/10.1016/j.chb.2017.02.022>
- Kalamkar, V. B. S., Nouri, N., Ofek, E., Pahud, M., Dudley, J. J., Hu, J., & Grubert, J. (2022). *Quantifying the effects of working in VR for one week*. arXiv preprint, arXiv:2206.03189. <https://doi.org/10.48550/arXiv.2206.03189>
- Kantor, J. (2014, August 14). *Working anything but 9 to 5*. The New York Times. <https://www.nytimes.com/interactive/2014/08/13/us/starbucks-workers-scheduling-hours.html>
- Karlsen, J., & Stavelin, E. (2014). *Computational journalism in Norwegian newsrooms*. *Journalism Practice*, 8(1), 34–48. <https://doi.org/10.1080/17512786.2013.813190>
- Keefe, J., Zhou, Y., & Merrill, J. B. (2021). *The present and potential of AI in journalism*. The Knight Foundation. <https://knightfoundation.org/articles/the-present-and-potential-of-ai-in-journalism/>
- Keith, S., & Cozma, R. (2022, forthcoming). *Teaching journalism online*. UNESCO.
- Kim, H., Kim, S.-H., & Frear, C. (2016). *News photos of scientists skew race but not gender*. *Newspaper Research Journal*, 37(1), 261–274. <https://doi.org/10.1177/0739532916663210>
- Knowledge Center Data & Society. (n.d.). *Knowledge Center Data & Society*. <https://data-en-maatschappij.ai/journalisten>
- Kothari, A., & Cruikshank, S. A. (2022). *Artificial intelligence and journalism: An agenda for journalism research in Africa*. *African Journalism Studies*, 43(1), 17–33. <https://doi.org/10.1080/23743670.2021.1999840>
- Köuts-Klemm, R. (2019). *Data literacy among journalists: A skills-assessment based approach*. *Central European Journal of Communication*, 12(3), 299–315. [https://doi.org/10.19195/1899-5101.12.3\(24\).2](https://doi.org/10.19195/1899-5101.12.3(24).2)
- Laranjeira de Pereira, J., & Guimarães Moraes, T. (2022, February 14). *Promoting irresponsible AI: Lessons from a Brazilian bill*. Heinrich Böll Stiftung. <https://eu.boell.org/en/2022/02/14/promoting-irresponsible-ai-lessons-brazilian-bill>
- LeCompte, C. (2015). *Automation in the newsroom: How algorithms are helping reporters expand coverage, engage audiences, and respond to breaking news*. Nieman reports. <https://niemanreports.org/articles/automation-in-the-newsroom/>
- Leitner, M. (2018). *How media companies use data to sign up digital subscribers (and keep them)*. Reuters Institute Fellowship Paper, University of Oxford.
- Lewis, S., Guzman, A. L., & Schmidt, T. S. (2019). *Automation, journalism, and human-machine communication: Rethinking roles and relationships of humans and machines in news*. *Digital Journalism*, 7(4), 409–427. <https://doi.org/10.1080/21670811.2019.1577147>
- Lewis, S. C. (2012). *The tension between professional control and open participation*. *Information, Communication & Society*, 15(6), 836–866. <https://doi.org/10.1080/1369118X.2012.674150>
- Lewis, S. C., & Usher, N. (2014). *Code, collaboration, and the future of journalism*. *Digital Journalism*, 2(3), 383–393. <https://doi.org/10.1080/21670811.2014.895504>
- Lewis, S. C., & Simon, F. M. (2022). *Why human-machine communication matters for the study of journalism and artificial intelligence*. In: Guzman, A. L., McEwen, R., & S. Jones, S. (Eds.) *SAGE Handbook of Human-Machine Communication*.
- Lindén, C.-G. (2020). *What makes a reporter human? A research agenda for augmented journalism*. *Questions de communication*, 37(2020), 337–352 <https://doi.org/10.4000/questionsdecommunication.23301>
- Logg, J. M., Minson, J. A., & Moore, D. A. (2019). *Algorithm appreciation: People prefer algorithmic to human judgment*. *Organizational Behavior and Human Decision Processes*, 151, 90–103. <https://doi.org/10.1016/j.obhdp.2018.12.005>

- Long, D., & Magerko, B. (2020). *What is AI literacy? Competencies and design considerations*. In: *CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–16. <https://doi.org/10.1145/3313831.3376727>
- LSE, London School of Economics and Political Science. (n.d.). *JournalismAI Starter Pack*. <https://www.lse.ac.uk/media-and-communications/polis/JournalismAI/Starter-Pack>
- Madiega, T. (2019). *EU guidelines on ethics in artificial intelligence: Context and implementation*. European Parliamentary Research Service (EPRS).
- Magnusson, M., Finnäs, J., & Wallentin, L. (2016). *Finding the news lead in the data haystack: Automated local data journalism using crime data*. Paper presented vid Computation & Journalism Symposium. Stanford University, Palo Alto, CA. [http://journalism.stanford.edu/cj2016/files/Finding the news lead in the data haystack.pdf](http://journalism.stanford.edu/cj2016/files/Finding%20the%20news%20lead%20in%20the%20data%20haystack.pdf)
- Malinverni, L., Valero, C., Schaper, M. M., & de la Cruz, I. G. (2021). *Educational robotics as a boundary object: Towards a research agenda*. *International Journal of Child-Computer Interaction*, 29. <https://doi.org/10.1016/j.ijcci.2021.100305>
- Marckmann Rasmussen, A. (2020). *Sådan forandrer automatiseringen medierna*. Dansk Journalistforbund.
- Marconi, F., & Siegman, A. (2017). *Machine journalist. 2017. The future of augmented journalism: A guide for newsrooms in the age of smart machines*. Associated Press.
- Mayor, A. (2018). *Gods and robots: Myths, machines, and ancient dreams of technology*. Princeton University Press.
- McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). *A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955*. *AI Magazine*, 27(4), 12–12. <https://doi.org/10.1609/aimag.v27i4.1904>
- McGregor, S. C. (2019). *Social media as public opinion: How journalists use social media to represent public opinion*. *Journalism*, 20(8), 1070–1086. <https://doi.org/10.1177/1464884919845458>
- Merrill, J. B. (2020, January 19). *How Quartz used AI to sort through the Luanda Leaks*. Quartz. <https://qz.com/1786896/ai-for-investigations-sorting-through-the-luanda-leaks/>
- Merten, M. (2022, March 10). *Is the SA revenue service's risk algorithm the glitch in the tax collector's matrix?* Daily Maverick. Available at <https://www.dailymaverick.co.za/article/2022-03-08-is-the-sa-revenue-services-risk-algorithm-the-glitch-in-the-tax-collectors-matrix/>
- Miao, F., Holmes, W., Ronghuai, H. & Hui, Z. (2021). *AI and education: Guidance for policy-makers*. UNESCO.
- Milosavljević, M., & Vobič, I. (2021). *'Our task is to demystify fears': Analysing newsroom management of automation in journalism*. *Journalism*, 22( 9), 2203–2221. <https://doi.org/10.1177/1464884919861598>
- Moran, R. E., & Shaikh, S. J. (2022). *Robots in the news and newsrooms: Unpacking meta-journalistic discourse on the use of artificial intelligence in journalism*. *Digital Journalism*, 10(10), 1756–1774. <https://doi.org/10.1080/21670811.2022.2085129>
- Muratova, N., Grizzle, A., & Mirzakhmedova, D. (2021). *Media and information literacy in journalism: A handbook for journalists and journalism educators*. UNESCO.
- Munoriyarwa, A., Chiumbu, S., & Motsaathebe, G. (2021). *Artificial intelligence practices in everyday news production: The case of South Africa's mainstream newsrooms*. *Journalism Practice*, 1(19). <https://doi.org/10.1080/17512786.2021.1984976>
- Nast, C. (n.d.). *The artificial intelligence database*. Wired. Available at <https://www.wired.com/category/artificial-intelligence/>
- Natale, S., & Ballatore, A. (2017). *Imagining the thinking machine: Technological myths and the rise of artificial intelligence*. *Convergence*, 16(1), 3–18. <https://doi.org/10.1177/1354856517715164>
- Nerone, J. (2013). *The historical roots of the normative model of journalism*. *Journalism*, 14(4), 446–458. <https://doi.org/10.1177/146488491246417>
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). *Conceptualizing AI literacy: An exploratory review*. *Computers and Education: Artificial Intelligence*, 2. <https://doi.org/10.1016/j.caeai.2021.100041>
- Nordahl, L., Chavar, A. J., Porter, L. Z., Kim, M., & Blufarb, A. (2021, July 21). *Using computer vision to create a more accurate digital archive*. The New York Times R&D. <https://rd.nytimes.com/projects/using-computer-vision-to-create-a-more-accurate-digital-archive>
- North, L. (2016). *The gender of "soft" and "hard" news*. *Journalism Studies*, 17(3), 356–373. <https://doi.org/10.1080/1461670X.2014.987551>
- O'Neil, C. (2016). *Weapons of math destruction*. Crown.
- Ouchchy, L., Coin, A., & Dubljević, V. (2020). *AI in the headlines: The portrayal of the ethical issues of artificial intelligence in the media*. *AI & Society*, 35, 927–936. <https://doi.org/10.1007/s00146-020-00965-5>
- Nilsson, N. J. (2010). *The quest for artificial intelligence: A history of ideas and achievements*. Cambridge University Press.
- Paulussen, S. (2016). *Innovation in the newsroom*. In: Witschge, T., Anderson, C.W., Domingo, D., & Hermida, A. (Eds.) *The SAGE handbook of digital journalism*. Sage, 192–206.

- Pentzold, C., Brantner, C., & Fölsche, L. (2019). *Imagining big data: Illustrations of “big data” in US news articles, 2010–2016*. *New Media & Society*, 21(1), 139–167. <https://doi.org/10.1177/1461444818791326>
- Petre, C. (2018). *Engineering consent: How the design and marketing of newsroom analytics tools rationalize journalists' labor*. *Digital Journalism*, 6(4), 509–527. <https://doi.org/10.1080/21670811.2018.1444998>
- Reddy, E., Hoople, G., & Choi-Fitzpatrick, A. (2019). *Interdisciplinarity in practice: Reflections on drones as a classroom boundary object*. *Engineering Studies*, 11(1), 51–63. <https://doi.org/10.1080/19378629.2019.1614006>
- Romele, A. (2022). *Images of artificial intelligence: A blind spot in AI ethics*. *Philosophy & Technology*, 35(4). <https://doi.org/10.1007/s13347-022-00498-3>
- Roselli, D., Matthews, J., & Talagala, N. (2019). *Managing bias in AI*. In: *WWW'19 Companion proceedings of the 2019 World Wide Web Conference*, 539–544. <https://doi.org/10.1145/3308560.3317590>
- Ross Arguedas, A., Robertson, C., Fletcher, R., & Nielsen, R. (2022). *Echo chambers, filter bubbles, and polarisation: A literature review*. Reuters Institute for the Study of Journalism.
- Russell, S., & Norvig, P. (2010). *Artificial intelligence: A modern approach*. 3rd ed. Pearson.
- Samoli, S., Cobo, M. L., Gomez, E., De Prato, G., Martinez-Plumed, F., & Delipetrev, B. (2020). *AI watch – Defining artificial intelligence: Towards an operational definition and taxonomy of artificial intelligence*. Technical Report. Joint Research Centre.
- Schumpeter, J. A. (1994). *Capitalism, Socialism and Democracy*. Routledge.
- Sedola, S. (2021). *Artificial intelligence for Africa – blueprint*. [https://smart.africa/board/login/uploads/70029-eng\\_ai-for-africa-blueprint.pdf](https://smart.africa/board/login/uploads/70029-eng_ai-for-africa-blueprint.pdf)
- Shi, Y., & Herniman, J. (2023). *The role of expectation in innovation evolution: Exploring hype cycles*. *Technovation*, 119(C). <https://doi.org/10.1016/j.technovation.2022.102459>
- Siles, I., & Boczkowski, P. J. (2012). *Making sense of the newspaper crisis: A critical assessment of existing research and an agenda for future work*. *New Media & Society*, 14(8): 1375–1394. <https://doi.org/10.1177/1461444812455148>
- Simon, H. A. (1995). *Artificial intelligence: An empirical science*. *Artificial Intelligence*, 77(1), 95–127. [https://doi.org/10.1016/0004-3702\(95\)00039-H](https://doi.org/10.1016/0004-3702(95)00039-H)
- Simons, M., Tiffen, R., Hendrie, R., Carson, A., Sullivan H., Muller, D., & McNair, B. (2017). *Understanding the civic impact of journalism*. *Journalism Studies*, 18(11), 1400–1414.
- Sivadas, L., & Argoub, S. (2021, April 7). *How to report on artificial intelligence*. *Journalism.co.uk*. <https://www.journalism.co.uk/news/how-to-report-on-ai-effectively/s2/a811366/>
- Stalph, F. (2017). *Classifying data journalism: A content analysis of daily data-driven stories*. *Journalism Practice*, 12(10), 1332–1350. <https://doi.org/10.1080/17512786.2017.1386583>
- Stieglietz, S., Brachten, F., Ross, B., & Jung, A.-K. (2017). *Do social bots dream of electric sheep? A categorisation of social media bot accounts*. A paper presented at Australasian Conference on Information Systems, Hobart, Australia.
- Stray, J. (2019). *Making artificial intelligence work for investigative journalism*. *Digital Journalism*, 7(8), 1076–1097. <https://doi.org/10.1080/21670811.2019.1630289>
- Striphas, T. (2015). *Algorithmic culture*. *European Journal of Cultural Studies*, 18(4–5), 395–412. <https://doi.org/10.1177/1367549415577392>
- Sundar, S. S., Kang, J. & Oprean, D. (2017). *Being there in the midst of the story: How immersive journalism affects our perceptions and cognitions*. *Cyberpsychology, Behavior, and Social Networking*, 20(11), 672–682. <https://doi.org/10.1089/cyber.2017.0271>
- Thurman, N., Dörr, K., & Kunert, J. (2017). *When reporters get hands-on with robo-writing*. *Digital Journalism*, 5(10), 1240–1259. <https://doi.org/10.1080/21670811.2017.1289819>
- Turing, A. M. (2009). *Computing machinery and intelligence*. In: Epstein, R., Roberts, G., & Beber, G., (Eds.) *Parsing the Turing test: Philosophical and methodological issues in the quest for the thinking computer*. Springer, 23–65.
- UNESCO (2021). *Recommendation on the ethics of artificial intelligence*. <https://en.unesco.org/artificial-intelligence/ethics>
- United Robots (2021). *News automation and how to leverage it*. Whitepaper. <https://www.unitedrobots.ai/whitepaper-download?hsLang=en>
- University of Sheffield (2022). *Data literacy for journalists*. An online course. <https://www.futurelearn.com/courses/data-literacy-for-journalists/1>
- van Dalen, A. (2012). *The algorithms behind the headlines: How machine-written news redefines the core skills of human journalists*. *Journalism Practice*, 6(5–6), 648–658. <https://doi.org/10.1080/17512786.2012.667268>
- Venturini, T. (2019). *From fake to junk news: The data politics of online virality*. In: Bigo, D., Isin, E., & Ruppert, E. (Eds.), *Data politics: Worlds, subjects, rights*. Routledge, 123–144.

- Vergeer, M. (2020). *Artificial intelligence in the Dutch Press: An analysis of topics and trends*. *Communication Studies*, 71(3), 373–392. <https://doi.org/10.1080/10510974.2020.1733038>
- Viswanath, K., Blake, K. D., Meissner, H. I., Saiontz, N. G., Mull, C., Freeman, C. S., Hesse, B., & Croyle, R. T. (2008). *Occupational practices and the making of health news: A national survey of US health and medical science journalists*. *Journal of Health Communication*, 13(8), 759–777. <https://doi.org/10.1080/10810730802487430>
- Waldenström, A., Wiik, J., & Andersson, U. (2018). *Conditional autonomy: Journalistic practice in the tension field between professionalism and managerialism*. *Journalism Practice*, 1(16). <https://doi.org/10.1080/17512786.2018.1485510>
- Welbers, K., van Atteveldt, W., Kleinnijenhuis, J., Ruigrok, N., & Schaper, J. (2016). *News selection criteria in the digital age: Professional norms versus online audience metrics*. *Journalism*, 17(8), 1037–1053. <https://doi.org/10.1177/1464884915595474>
- Westlund, O., Krumsvik, A. H., & Lewis, S. C. (2021). *Competition, change, and coordination and collaboration: Tracing news executives' perceptions about participation in media innovation*. *Journalism Studies*, 22(1), 1–21. <https://doi.org/10.1080/1461670X.2020.1835526>
- Winfield, C., Sparkman-Key, N. M., & Vajda, A. (2017). *Interprofessional collaboration among helping professions: Experiences with holistic client care*. *Journal of Interprofessional Education & Practice*, 9(1), 66–73. <https://doi.org/10.1016/j.xjep.2017.08.004>
- Women's Media Center (2014). *The status of women in the U.S. media 2014*. [https://wmc.3cdn.net/6dd3de8ca65852dbd4\\_fjm6yck9o.pdf](https://wmc.3cdn.net/6dd3de8ca65852dbd4_fjm6yck9o.pdf)
- Women's Media Center (2019). *The status of women in the U.S. media 2019*. <https://womensmediacenter.com/assets/site/from-bsd/WMCStatusofWomeninUSMedia2019.pdf>
- Woolgar, S. (1985). *Why not a sociology of machines? The case of sociology and artificial intelligence*. *Sociology*, 19(4), 557–572. <https://doi.org/10.1177%2F0038038585019004005>.
- Woolgar, S. (Ed.) (2003). *Virtual society? Technology, cyberbole, reality*. Oxford University Press.
- Wu, S., Tandoc, E.C. Jr., & Salmon, C. T. (2019). *Journalism reconfigured: Assessing human–machine relations and the autonomous power of automation in news production*. *Journalism Studies*, 20(10), 1440–1457. <https://doi.org/10.1080/1461670X.2018.1521299>
- Xiaomo, L., Nourbakhsh, A., Li, Q., Shah, S., Martin, R., & Duprey, J. (2017). *Reuters tracer: Toward automated news production using large scale social media data*. Research and Development. Thomson Reuters.
- Young, M. L., & Hermida, A. (2015). *From Mr. and Mrs. Outlier to central tendencies: Computational journalism and crime reporting at the Los Angeles Times*. *Digital Journalism*, 3(3), 381–397. <https://doi.org/10.1080/21670811.2014.976409>
- Yudkowsky, E. (2001). *Creating friendly AI 1.0: The analysis and design of benevolent goal architectures*. The Singularity Institute.
- Zachrisson, O. (2021, November 21). *An algorithm to promote unique public service journalism*. EBU Tech-I Magazine.
- Zelizer, B. (1993). *Journalists as interpretive communities*. *Critical Studies in Mass Communication*, 10(3), 219–237. <https://doi.org/10.1080/15295039309366865>
- Zelizer, B. (2004). *Taking journalism seriously: News and the academy*. Sage Publications.
- Zhai, Y., Yan, J., Zhang, H., & Lu, W. (2020). *Tracing the evolution of AI: Conceptualization of artificial intelligence in mass media discourse*. *Information Discovery and Delivery*, 48(3), 137–149.
- Zhang, S., & Feng, J. (2019). *A step forward? Exploring the diffusion of data journalism as journalistic innovations in China*. *Journalism Studies*, 20(9), 1281–1300. <https://doi.org/10.1080/1461670X.2018.1513814>
- Zhang, W. & Pérez Tornero, J. M. (2021). *Introduction to AI journalism: Framework and ontology of the trans-domain field for integrating AI into journalism*. *Journal of Applied Journalism & Media Studies*, published online September 3, 2021. [https://doi.org/10.1386/ajms\\_00063\\_1](https://doi.org/10.1386/ajms_00063_1)

# Список иллюстраций

Рисунок 1. Области журналистской компетенции, связанные с освещением тематики ИИ	22
Рисунок 2. Основные подобласти сферы ИИ	32
Рисунок 3. Иерархия методов обработки естественного языка (ОЕЯ)	33
Рисунок 4. Эволюция вычислительных систем	41
Рисунок 5. Проблема вагонетки	43
Рисунок 6. Глобальный ландшафт нормативного регулирования сферы ИИ	51
Рисунок 7. Рекомендации по этическим нормам для заслуживающего доверия ИИ — обзор	54
Рисунок 8. Четыре этических принципа, основанных на фундаментальных правах	55
Рисунок 9. Список критериев оценки доверия к ИИ	55
Рисунок 10. Риск-ориентированный подход	56
Рисунок 11. Принципы и рекомендации ОЭСР по ИИ	57
Рисунок 12. Рекомендация ЮНЕСКО об этических аспектах ИИ	57
Рисунок 13. За кулисами фальшивых новостей	71
Рисунок 14. Твит Яна Гудфеллоу (14 января 2019 г.)	75
Рисунок 15. Внимание СМИ к развивающимся технологиям (Gartner's hype cycle)	84
Рисунок 16. Скриншот результатов поиска в Google изображений искусственного интеллекта (по состоянию на февраль 2022 г.)	86
Рисунок 17. Основное назначение технологий ИИ в журналистике (2012–2021 гг.)	97
Рисунок 18. На каком этапе конвейера новостей обычно применяются средства ИИ? (2012–2021 гг.)	98

Возникновение технологий искусственного интеллекта (ИИ) и осуществление контроля над ними оказывают влияние на всё общество, и журналистам необходимо следить за их последствиями. В данном руководстве рассматриваются следующие темы:

- выработка общего представления о машинном интеллекте и определение различных типов ИИ;
- формулирование сценариев будущего в реалиях ИИ путём выявления распространённых популярных установок, формирующих общественное сознание;
- понимание роли журналистики в опосредовании и формировании дискурса об ИИ;
- поиск способов взвешенного, реалистичного и ответственного освещения проблематики ИИ.

Публикация подготовлена под эгидой Всемирного совета по журналистскому образованию и при поддержке Международной программы развития коммуникации (МПРК) ЮНЕСКО.



**unesco**

Международная  
программа развития  
коммуникации

Подготовка и редактирование материала:



**Всемирный совет по образованию в области журналистики (WJEC)**  
[www.wjec.net/](http://www.wjec.net/)

