



Анна Абрамова
Анастасия Рыжкова
Юлия Церех

**ПОКАЗАТЕЛЬ ОЦЕНКИ ДЕЯТЕЛЬНОСТИ
КЛЮЧЕВЫХ АКТОРОВ /СТЕЙКХОЛДЕРОВ
В ОБЛАСТИ РАЗВИТИЯ ГРАМОТНОСТИ
ПО ВОПРОСАМ ЭТИЧЕСКИХ АСПЕКТОВ
ИСКУССТВЕННОГО ИНТЕЛЛЕКТА**

коллекция исследований ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ
ДЛЯ РАЗВИТИЯ

приоритет2030[^]
Лидерами становятся

**МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ИНСТИТУТ
МЕЖДУНАРОДНЫХ ОТНОШЕНИЙ
МИНИСТЕРСТВА ИНОСТРАННЫХ ДЕЛ
РОССИЙСКОЙ ФЕДЕРАЦИИ**

Центр искусственного интеллекта МГИМО

Анна Абрамова, Анастасия Рыжкова, Юлия Церех

**«ЭТИЧЕСКАЯ ОЦЕНКА ИИ НА НАЦИОНАЛЬНОМ И
МЕЖДУНАРОДНОМ УРОВНЯХ. КЛЮЧЕВЫЕ СУБИНДЕКСЫ:
ПОКАЗАТЕЛЬ ОЦЕНКИ ДЕЯТЕЛЬНОСТИ КЛЮЧЕВЫХ АКТОРОВ
/СТЕЙКХОЛДЕРОВ В ОБЛАСТИ РАЗВИТИЯ ГРАМОТНОСТИ ПО
ВОПРОСАМ ЭТИЧЕСКИХ АСПЕКТОВ ИСКУССТВЕННОГО
ИНТЕЛЛЕКТА»**

Москва, 2022

Авторы:

Анна Абрамова, кандидат экономических наук, директор Центра искусственного интеллекта МГИМО, заведующая кафедрой Цифровой экономики и искусственного интеллекта группы АДВ МГИМО-Университета

Анастасия Рыжкова, кандидат технических наук, научный сотрудник Центра искусственного интеллекта МГИМО

Юлия Церех, младший научный сотрудник Центра искусственного интеллекта МГИМО

Этические аспекты искусственного интеллекта (ИИ) становится одним из основных элементов мягкого права в регулировании национального и международного рынков. В декабре 2021 г. ЮНЕСКО приняла Рекомендацию об этических аспектах искусственного интеллекта, в которой излагаются подходы к международному мягкому регулированию, уделяющему особое внимание этике. Международная система оценки этических аспектов ИИ может стать основой для оценки этического воздействия в сочетании с Рекомендацией и Принципами ИИ ОЭСР, а также подходами к классификации ИИ, разработанной экспертами ОЭСР. Грамотность в области ИИ является частью группы субиндексов в рамках методологии, которые влияют на всех участников в рамках жизненного цикла системы ИИ. Являясь важным элементом цифровой грамотности, она сталкивается с почти такими же проблемами в плане развития. На грамотность в области ИИ влияет цифровой разрыв и гибкость подходов в национальной системе образования в отношении данных и экономики ИИ.

«Оценка этических аспектов ИИ на национальном и международном уровнях. Ключевые субиндексы: грамотность в области ИИ.»

Работа является частью коллекции научных исследований «ИИ для развития» Центра ИИ МГИМО.

JEL F01, F20, F42, F53, F55, F60, F63, F68

Ключевые слова: искусственный интеллект, этика, индекс, государственная политика, правительство.

© 2022 МГИМО. Все права защищены. Короткие фрагменты текста, не превышающие двух абзацев, могут цитироваться без явного разрешения при условии полной ссылки на источник, включая примечание ©.

Фото на обложке: canva.com

Оглавление

Список сокращений.....	5
Введение	6
Непростой путь от цифровой грамотности к грамотности в области ИИ.....	7
Методология разработки субиндекса в рамках Международной системы оценки этических аспектов ИИ: грамотность в области ИИ	8
Выводы	13
Использованная литература.....	14

Список сокращений

ИИ	Искусственный интеллект
ИКТ	Информационно-коммуникационные технологии
ОЭСР	Организация экономического сотрудничества и развития
ООН	Организация Объединенных Наций
ЮНКТАД	Конференция Организации Объединенных Наций по торговле и развитию
ЮНЕСКО	Организация Объединенных Наций по вопросам образования, науки и культуры

Введение

Этот документ является следующим шагом в продолжающемся исследовании Центра ИИ МГИМО по разработке международной системы оценки этических аспектов искусственного интеллекта. В отдельных статьях подробно обсуждаются все ключевые субиндексы, включая основные индикаторы и сложные вопросы.

Международная система оценки этических аспектов искусственного интеллекта — это комплексный подход, в котором участвуют все ключевые участники жизненного цикла ИИ — государство, бизнес, гражданское общество, исследовательские центры/мозговые центры. Более того, мы предлагаем оценку по трем субиндексам областей, которые способствуют устойчивому развитию ИИ на всех этапах и влияют на все группы участников – грамотность в области ИИ, инвестиции в НИОКР и развитие инфраструктуры ИКТ.

Опубликованная в феврале 2022 г. Общая структура международной системы оценки этических аспектов искусственного интеллекта (Абрамова, Рыжкова и Церех, 2022) охватывает все основные элементы индекса, представляющего группы ключевых показателей. Этот исследовательский документ сосредоточен на подробном освещении вклада субиндекса «Грамотность в области ИИ», обозначены актуальные проблемы и вовлеченность основных участников.

Структура исследования следующая: первый раздел введение, второй посвящен актуальным вопросам цифровой грамотности и грамотности в области ИИ, третья часть раскрывает методологию субиндекса «Грамотность в области ИИ» в рамках Международной системы оценки этических аспектов ИИ. Последний раздел посвящен обсуждению наиболее сложных вопросов практической реализации субиндекса в контексте концепции жизненного цикла ИИ и цифрового неравенства.

Авторы благодарны руководителям и координаторам проекта «Приоритет 2030» за возможность проведения исследования.

Непростой путь от цифровой грамотности к грамотности в области ИИ

Основные элементы цифровой грамотности охватывают компетенции, позволяющие каждому человеку быть конкурентоспособным в эпоху цифровой экономики¹. Для оценки уровня профессионализма в сфере ИКТ 6 основных навыков могут быть рассмотрены – грамотность в области информации, данных и медиа, способность к цифровому обучению и развитию, цифровые коммуникации, digital communication, участие и сотрудничество, создание цифровых благ, решение проблем и инновации².

В целях проведения исследований термин цифровой грамотности получил развитие в контексте развития секторов – финансового и здравоохранения. Цифровая финансовая грамотность была рассмотрена в качестве ключевой при принятии решений в финансовом секторе. Это охватило “знание цифровых финансовых продуктов и услуг, осведомленность о цифровых финансовых рисках, знание цифрового контроля финансовых рисков, знаний о правах пользователей, процедуры возмещения.” (Kumar et al. 2022) Цифровая грамотность в здравоохранении или «электронная грамотность в здравоохранении» фокусируется на индивидуальных способностях к получению доступа, понимания и возможности воспользоваться материалами и технологиями в сфере цифрового здравоохранения для улучшения качества жизни» (Griebel, et al, 2018).

Грамотность в области ИИ основана на подходах к цифровой грамотности с акцентом на развитие ИИ. Грамотность в области ИИ является отдельным направлением в рамках цифровой грамотности (Yang 2022). Согласно исследованию (Long and Magerko, 2020), грамотность в области ИИ является «набором компетенций, которые позволяют индивидууму критически оценивать технологии ИИ, коммуницировать и сотрудничать с использованием ИИ и использовать ИИ в качестве инструмента онлайн, дома и на рабочем месте». Грамотность в области ИИ имеет четыре основные направления – знать и понимать ИИ, использовать и оценивать ИИ, этические аспекты ИИ (Ng et al 2021). Так, грамотность в сфере ИИ, будучи менее масштабной в сравнении с цифровой грамотностью, располагает потенциалом стать одним из основных навыков в среднесрочной перспективе с ростом использования приложений с использованием ИИ в широком перечне социально-экономических направлений.

¹ What is digital literacy?

https://www.westernsydney.edu.au/studysmart/home/study_skills_guides/digital_literacy/what_is_digital_literacy

² Building Digital Capability. <https://www.jisc.ac.uk/rd/projects/building-digital-capability>

Методология разработки субиндекса в рамках Международной системы оценки этических аспектов ИИ: грамотность в области ИИ

Субиндекс Грамотность в области ИИ относится к кросс-секторальной части в рамках Международной системы оценки этических аспектов ИИ. Авторы предлагают ту же методологию, что и для групп основных участников. Такой подход помогает объединить в одну сложную структуру наиболее важные части аспектов внедрения этических аспектов ИИ.

Во-первых, авторы рассмотрели и сравнили пять наиболее распространенных и проверенных методов оценки сложных и динамических систем, таких как:

- мозговой штурм;
- анализ слабых и сильных сторон;
- метод построения карты;
- метод Дельфи;
- экспертная оценка.

Каждый из рассмотренных методов имеет свои особенности и ограничения в применении.

Для расчета субиндекса по грамотности в области ИИ авторы рекомендуют использовать метод экспертных оценок, поскольку это ценный компромисс между высокой стоимостью, длительными расчетами и глубиной исследования.

1) Мозговой штурм

Для уровня грамотности в области ИИ этот метод нереалистичен ввиду того, что очень сложно собрать всех лиц, принимающих решения, в одном месте и в одно время. Одним из вариантов решения могло бы стать создание соответствующего подразделения в рамках Министерств образования с последующим созданием расширенных рабочих групп.

Как правило, мозговой штурм проводится внутри проектной команды с возможностью привлечения к работе стороннего эксперта. Эксперт может обладать широкими или, наоборот, узкоспециальными знаниями, что, по мнению руководителя проектной группы, важно при реализации проекта.

Алгоритм метода достаточно прост и состоит из нескольких шагов:

1. Участники составляют максимально подробный список параметров, актуальных для проекта. В данном случае это может быть детализировано до отдельных направлений в деятельности министерств и ведомств и проблем связанных с этическими аспектами ИИ.

2. Параметры с наименьшей вероятностью реализации удаляются из лонг-листа большинством участников. При анализе этических аспектов ИИ, часть из них может быть признана общей для большинства ведомств.

Достоинства метода: скорость получения результата, простота реализации метода.

Недостатки метода: качество анализа напрямую зависит от опыта и кругозора лиц, участвующих в мозговом штурме.

Возможность применения метода оценки этических аспектов внедрения технологий ИИ:

- требует привлечения опытной проектной команды для внедрения аналогичных продуктов,
- высокая стоимость,
- сложность привлечения профильных специалистов.

2) Анализ слабых и сильных сторон

Для уровня грамотности в области ИИ этот метод нереалистичен ввиду того, что очень сложно проводить регулярное исследование, основанное на анализе слабых и сильных сторон деятельности организаций в контексте этических аспектов ИИ, с большим списком задействованных чиновников.

Метод аналогичен методу анализа предположений, однако команда проекта составляет список потенциальных параметров, выявляя и впоследствии анализируя их слабые/сильные стороны.

Преимущества: детальное рассмотрение параметров Индекса.

Недостатки:

- длительная реализация метода;
- чрезмерная детализация метода;
- качество анализа напрямую зависит от опыта и кругозора привлекаемых специалистов.

Возможность применения метода оценки этических аспектов внедрения технологий ИИ: команда проекта при недостаточном опыте может упустить существенные параметры и аспекты.

3) Метод составления диаграмм

Для грамотности в области ИИ этот метод нереалистичен ввиду того, что он крайне сложен для нескольких специалистов, которые могут описать сложную ситуацию в области этических аспектов ИИ, связанных с развитием образования на протяжении жизни человека.

Метод осуществляется внутри проектной команды с возможностью приглашения внешнего эксперта. Анализ проходит в три этапа:

- составление причинно-следственных связей,
- создание блок-схемы реализуемых процессов,
- составление диаграмм воздействия.

Преимущества: качественный учет потенциальных рисков проектов, связанных с этическими аспектами применения ИИ.

Недостатки: реализация метода построения диаграмм требует от проектной команды навыков работы с данным методом и значительных временных затрат.

Возможность применения метода оценки этических аспектов использования технологий ИИ: применение этого навыка требует специальных компетенций и опыта.

4) метод Дельфи

Для уровня грамотности в сфере ИИ этот метод нереалистичен, так как имеет слишком большой срок реализации.

Достоинства: Качественная проработка этических аспектов применения ИИ в рамках решения проблем повышения грамотности в сфере ИИ.

Недостатки: метод требует длительной реализации и финансовых ресурсов для реализации.

Возможность применения метода оценки этических аспектов внедрения технологий ИИ: метод требует больших затрат времени и средств.

5) Метод экспертных оценок

Метод экспертных оценок аналогичен методу Дельфи, однако предполагает открытый опрос экспертов.

Преимущества: качественная проработка выявления потенциальных рисков.

Недостатки: требуется создание базы специалистов, готовых участвовать в большом опросе для исследования

Возможность применения метода оценки этических аспектов внедрения технологий ИИ: метод требует много времени.

Формула расчета

Авторы основывали оценку групп показателей на индексе значимости, который рассчитывается по формуле:

$$r_{ij}^k = \alpha_{ij} \beta_{ij}^k, \quad (1)$$

Где

r_{ij}^k - значимость i -го показателя, оцененного j -м респондентом, с точки зрения влияния на k -фактор,

$i = (1 \dots N)$, где N – количество рассматриваемых в исследовании параметров,

$j = (1 \dots n)$, где n — количество полученных ответов,

$k = (1 \dots 5)$, где $1 \dots 5$ – номера групп влияния соответственно (соответственно стоимость, время выполнения ИТ-проекта, качество продукта, среда, безопасность),

α_{ij} - вес значимости показателя i , оцененный j -м респондентом,

β_{ij}^k - величина «эффекта» влияния показателя на рассматриваемого заинтересованного лица и/или преследуемые им цели.

Для оценки среднего значения показателей рассчитывается Индекс значимости показателя по формуле:

$$R_i^k = \frac{\sum_{j=1}^n r_{ij}^k}{n} = \frac{1}{n} \sum_{j=1}^n \alpha_{ij} \beta_{ij}^k \quad (2)$$

Предлагаемые параметры для расчета

УЧАСТИЕ БИЗНЕСА

- Размер инвестиций в проекты связанные с грамотностью в сфере ИИ
- Оценка уровня грамотности сотрудников в сфере ИИ
- Динамика изменения грамотности сотрудников сфере ИИ

УЧАСТИЕ ГОСУДАРСТВА

- Количество программ в области грамотности в сфере ИИ
- Средняя численность участников в государственных программах в области грамотности в сфере ИИ: местный и национальный уровни
- Соблюдение этических требований ЮНЕСКО
- Масштаб государственных инвестиций в проекты, связанные с грамотностью в сфере ИИ

ОБРАЗОВАТЕЛЬНАЯ СИСТЕМА:

- Число образовательных проектов для молодежи
 - школы (К, К+1 уровень)
 - университеты
- число проектов для людей пожилого возраста
- число программ
 - частные
 - государственные
 - частные-государственные
- средняя численность участников программ. Уровень местный и национальный
- Соблюдение этических требований ЮНЕСКО
- Масштаб инвестиций в соответствующие проекты

Выводы

Грамотность в сфере ИИ — это новое быстрорастущее направление цифровой грамотности, поскольку количество областей применения ИИ увеличивается. Грамотность в сфере ИИ основана на общей грамотности и базовых навыках в цифровой грамотности. Но проблемы с базовой грамотностью не решены, что ограничивает достижение целевых показателей по цифровой грамотности, указанных в перечне ЦУР. Это означает, что риск расширения разрыва между регионами сохраняется и угрожает социально-экономической стабильности в глобальном измерении.

Так, в зависимости от уровня общей цифровой грамотности навыки в области ИИ будут варьироваться от ограниченного понимания возможностей и вызовов технологии, включая этические аспекты, до создания приложений с использованием ИИ для личных и профессиональных нужд. В Международной системе оценки этических аспектов ИИ мы рассматриваем субиндекс Грамотность с области ИИ как оказывающий воздействие на всех участников в рамках всего жизненного цикла систем ИИ. Более того индикаторы в субиндексе подобраны таким образом, чтобы отражать подход к обучению в течение всей жизни и охватывают все основные уровни и всех участников.

Конечно, есть ряд общих факторов, которые влияют на качество оценки развития грамотности в области ИИ — национальные статистические подходы, возможности по сбору и оценке соответствующих данных, развитие ИКТ, гибкость национальной образовательной системы и открытость к новым цифровым технологиям.

Цифровой разрыв является одной из наиболее сложных проблем для развития системы образования. Это также отражено в субиндексе Подключение к ИКТ и инвестиций в НИОКР. Такой подход позволяет отразить потребности развития в краткосрочной и долгосрочной перспективе.

Использованная литература

AI Ethics: Another Step Closer to the Adoption of UNESCO's Recommendation. [E-resource]. Available at: URL <https://en.unesco.org/news/ai-ethics-another-step-closeradoption-unescos-recommendation-0>

L. Griebel, H. Enwald, H. Gilstad, A. Pohl, J. Moreland & M. Sedlmayr (2018) eHealth literacy research—Quo vadis?, *Informatics for Health and Social Care*, 43:4, 427-442, DOI: 10.1080/17538157.2017.1364247

P. Kumar, R. Pillai, N. Kumar et al., The interplay of skills, digital financial literacy, capability, and autonomy in financial decision making and well-being, *Borsa Istanbul Review*, <https://doi.org/10.1016/j.bir.2022.09.012>

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Conference on human factors in computing systems (CHI)*. <https://doi.org/10.1145/3313831.3376727>. Honolulu, HI, USA

Morgan, P. J., Trinh, L. Q., Huang, B., & Trinh, L. Q. (2019). The need to promote digital financial literacy for the digital age. [E-resource]. Available at: URL <https://t20japan.org/policy-brief-need-promote-digital-financial-literacy/>

Ng D.T.K., Leung J.K.L., Chu S.K.W. & Shen M.Q., Conceptualizing AI literacy: An exploratory review, *Computers and Education: Artificial Intelligence* (2021), doi: <https://doi.org/10.1016/j.caeai.2021.100041>.

UNESCO 2019. Beijing Consensus on Artificial Intelligence and Education. [E-resource]. Available at: URL <https://unesdoc.unesco.org/ark:/48223/pf0000368303>

UNESCO 2021. Forum on AI and Education engages international partners to ensure AI as a common good for education. [E-resource]. Available at: URL <https://en.unesco.org/news/unesco-forum-ai-and-education-engages-international-partners-ensure-ai-common-good-education>

UNESCO 2021. Forum on AI and Education engages international partners to ensure AI as a common good for education. [E-resource]. Available at: URL <https://en.unesco.org/news/unesco-forum-ai-and-education-engages-international-partners-ensure-ai-common-good-education>

UNESCO 2021. Recommendation on the ethics of artificial intelligence. [E-resource]. Available at: URL <https://unesdoc.unesco.org/ark:/48223/pf000038>

Weipeng Yang, Artificial Intelligence education for young children: Why, what, and how in curriculum design and implementation. Computers and Education: Artificial Intelligence, Volume 3, 2022, 100061, <https://doi.org/10.1016/j.caeai.2022.100061>.



Центр искусственного интеллекта МГИМО создан для расширения международного сотрудничества и взаимодействия со всеми субъектами цифровой экономики как на национальном, так и на международном уровне. Наше междисциплинарный научный подход сосредоточен на повестке дня международного сотрудничества, национальной политике в области ИИ и возможностях для бизнеса. Международная торговля и торговая политика (приоритет цифровой торговли), устойчивое развитие, этика ИИ — ключевые направления нашей деятельности.

На базе Университета МГИМО мы развиваем международную экспертную площадку по искусственному интеллекту с регулярными конференциями и круглыми столами, научными статьями и исследовательскими работами. Наша расширяющаяся сеть стратегических партнерств позволяет предоставлять консультационные и иные решения в области ИИ как для бизнеса, так и для государственных учреждений.

Центр основан в октябре 2021 года

Наши контакты



143007, Одинцово, Московская область,
Ново-Спортивная 3
<https://aicentre.mgimo.ru>
aicentre@inno.mgimo.ru
+7 903 623-95-15
<https://t.me/aicentremgimo>



приоритет2030[^]
Лидерами становятся