ФЕДЕРАЛЬНОЕ АГЕНТСТВО

ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

ГОСТ Р ИСО/МЭК 42001— 2024

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

Система менеджмента

(ISO/IEC 42001:2023, Information technology — Artificial intelligence — Management system, IDT)

Издание официальное

Москва
Российский институт стандартизации
2024

gost.e-ecolog.ru

Предисловие

1 ПОДГОТОВЛЕН Обществом с ограниченной ответственностью «Институт развития информационного общества» (ООО «ИРИО») на основе собственного перевода на русский язык англоязычной версии стандарта, указанного в пункте 4

2 ВНЕСЕН Техническим комитетом по стандартизации ТК 164 «Искусственный интеллект»

3 УТВЕРЖДЕН И ВВЕДЕН В ДЕЙСТВИЕ Приказом Федерального агентства по техническому регулированию и метрологии от 28 октября 2024 г. № 1549-ст

4 Настоящий стандарт идентичен международному стандарту ИСО/МЭК 42001:2023 «Информационные технологии. Искусственный интеллект. Система менеджмента» (ISO/IEC 42001:2023 «Information technology — Artificial intelligence — Management system», IDT).

Наименование настоящего стандарта изменено относительно наименования указанного международного стандарта для приведения в соответствие с ГОСТ Р 1.5—2012 (пункт 3.5).

При применении настоящего стандарта рекомендуется использовать вместо ссылочных международных стандартов соответствующие им национальные стандарты, сведения о которых приведены в дополнительном приложении ДА

5 ВВЕДЕН ВПЕРВЫЕ

*Правила применения настоящего стандарта установлены в статье 26 Федерального закона от 29 июня 2015 г. № 162-ФЗ «О стандартизации в Российской Федерации». Информация об изменениях к настоящему стандарту публикуется в ежегодном (по состоянию на 1 января текущего года) информационном указателе «Национальные стандарты», а официальный текст изменений и поправок — в ежемесячном информационном указателе «Национальные стандарты». В случае пересмотра (замены) или отмены настоящего стандарта соответствующее уведомление будет опубликовано в ближайшем выпуске ежемесячного информационного указателя «Национальные стандарты». Соответствующая информация, уведомление и тексты размещаются также в информационной системе общего пользования — на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет (*gost.e-ecolog.ru

© ISO, 2023
© IEC, 2023
© Оформление. ФГБУ «Институт стандартизации», 2024

Настоящий стандарт не может быть полностью или частично воспроизведен, тиражирован и распространен в качестве официального издания без разрешения Федерального агентства по техническому регулированию и метрологии

Содержание

1 Область применения 1
2 Нормативные ссылки 1
3 Термины и определения 1
4 Среда организации 4
4.1 Понимание организации и ее среды 4
4.2 Понимание потребностей и ожиданий заинтересованных сторон 5
4.3 Определение области применения системы менеджмента искусственного интеллекта 5
4.4 Система менеджмента искусственного интеллекта 5
5 Лидерство 5
5.1 Лидерство и приверженность 5
5.2 Политика в области искусственного интеллекта 6
5.3 Функции, ответственность и полномочия 6
6 Планирование 6
6.1 Действия в отношении рисков и возможностей 6
6.2 Цели искусственного интеллекта и планирование их достижения 8
6.3 Планирование изменений 9
7 Средства обеспечения 9
7.1 Ресурсы 9
7.2 Компетентность 9
7.3 Осведомленность 9
7.4 Обмен информацией 9
7.5 Документированная информация 10
8 Деятельность 10
8.1 Планирование и управление 10
8.2 Оценка рисков искусственного интеллекта 11
8.3 Обработка рисков искусственного интеллекта 11
8.4 Оценка воздействия системы искусственного интеллекта 11
9 Оценка результатов деятельности 11
9.1 Мониторинг, измерение, анализ и оценка 11
9.2 Внутренний аудит 11
9.3 Анализ со стороны руководства 12
10 Улучшения 12
10.1 Постоянное улучшение 12
10.2 Несоответствия и корректирующие действия 12
Приложение А (обязательное) Меры и цели управления 14
Приложение В (обязательное) Руководство по внедрению мер управления по обработке рисков искусственного интеллекта 18
Приложение С (справочное) Потенциальные организационные цели и источники рисков, связанные с применением искусственного интеллекта 37
Приложение D (справочное) Использование системы менеджмента искусственного интеллекта в разных областях и сферах деятельности 39
Приложение ДА (справочное) Сведения о соответствии ссылочных международных стандартов национальным стандартам 41
Библиография 42

Введение

Технологии искусственного интеллекта (ИИ) все чаще применяются во всех сферах деятельности, в которых используются информационные технологии, и, как ожидается, станут одним из основных факторов, влияющих на экономическое развитие. Вследствие этой тенденции, некоторые приложения могут привести к возникновению социальных проблем в ближайшие годы.

Цель настоящего стандарта — помочь организациям ответственно выполнять свою роль в отношении систем ИИ (например, использовать, разрабатывать, осуществлять мониторинг работы или предоставлять продукты или услуги, использующие ИИ). Применение ИИ освещает такие вопросы, как:

- автоматическое принятие решений с использованием ИИ непрозрачным и необъяснимым способом может потребовать специального управления, выходящего за рамки управления классическими информационными системами;
- использование анализа данных, инженерии знаний и машинного обучения, а не предписанной человеком логики проектирования систем, с одной стороны, расширяет возможности применения систем ИИ, а с другой — изменяет способ разработки, обоснования и развертывания таких систем;
- при функционировании систем ИИ с непрерывным обучением меняется их поведение, поэтому к ним требуется особое внимание.

Настоящий стандарт содержит требования к созданию, внедрению, поддержанию в рабочем состоянии и постоянному улучшению системы менеджмента ИИ в среде организации. Организациям следует обращать основное внимание на применении требований к характеристикам, специфическим для ИИ. Ввиду определенных особенностей ИИ, таких как способность к постоянному обучению и улучшению либо отсутствие прозрачности или объяснимости, может потребоваться использование различных мер предосторожности в случае, если при выполнении задачи с помощью ИИ возникают дополнительные опасения по сравнению с выполнением задачи традиционным способом. Внедрение системы менеджмента ИИ для расширения существующих структур управления является стратегическим решением для организации.

На создание и внедрение системы менеджмента ИИ оказывают влияние следующие факторы: потребности и цели организации, процессы, размер и структура, а также ожидания различных заинтересованных сторон. Другими факторами, влияющими на создание и внедрение системы менеджмента ИИ, являются многочисленные варианты использования ИИ и необходимость соблюдения надлежащего баланса между механизмами стратегического управления и инновациями. Организации могут предпочесть применять эти требования, используя подход, основанный на оценке рисков, чтобы гарантировать, что соответствующий уровень контроля применяется только в отношении конкретных вариантов использования, услуг или продуктов ИИ в пределах сферы деятельности организации. Ожидается, что все эти факторы влияния со временем будут меняться, поэтому следует время от времени проводить их ревизию.

Система менеджмента ИИ должна быть интегрирована с процессами организации и общей структурой управления. При проектировании процессов, информационных систем и разработке мер управления должны быть учтены конкретные факторы, связанные с ИИ. Критически важными примерами таких процессов управления являются:

- определение организационных целей, вовлечение заинтересованных сторон и формирование организационной политики;
- управление рисками и возможностями;
- управление факторами, связанными с надежностью систем ИИ, такими как защита, безопасность, справедливость, прозрачность, качество данных и качество систем ИИ на протяжении всего их жизненного цикла;
- управление взаимоотношениями с поставщиками, партнерами и третьими сторонами, которые предоставляют или разрабатывают системы ИИ для организации.

В настоящем стандарте содержатся рекомендации по развертыванию применимых мер управления для поддержки таких процессов и отсутствуют конкретные указания по процессам управления. Для внедрения таких важнейших процессов как управление рисками, жизненным циклом и качеством данных, которые подходят для конкретных случаев использования ИИ, продуктов или услуг, организация может сочетать общепринятые концепции, другие международные стандарты и свой собственный опыт.

Организация, соответствующая требованиям настоящего стандарта, может создать свидетельство своей ответственности и подотчетности касательно своей роли в отношении систем ИИ.

Порядок, в котором в настоящем стандарте представлены требования, не отражает их важности и не подразумевает порядок, в котором они должны быть реализованы. Нумерация элементов списка носит исключительно справочный характер.

Совместимость с другими стандартами систем менеджмента

В настоящем стандарте применяется гармонизированная структура (идентичные номера разделов, идентичные названия разделов, идентичный текст, общие термины и основные определения), разработанная для улучшения согласованности между стандартами ИСО на системы менеджмента (ССМ).

Система менеджмента ИИ предъявляет специфические требования к управлению проблемами и рисками, возникающими в результате использования ИИ в организации. Этот общий подход облегчает реализацию и согласованность с другими стандартами ИСО на системы менеджмента, например со стандартами, связанными с качеством, безопасностью, защитой и неприкосновенностью частной жизни.

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

Система менеджмента

Artificial intelligence.
Management system

Дата введения — 2025—01—01

1 Область применения

Настоящий стандарт определяет требования и рекомендации по созданию, внедрению, поддержанию в рабочем состоянии и постоянному улучшению системы менеджмента искусственного интеллекта (ИИ) в среде организации.

Настоящий стандарт предназначен для использования организациями, предоставляющими или использующими продукты или услуги, применяющие системы ИИ. Настоящий стандарт призван помочь организациям ответственно разрабатывать или использовать системы ИИ для достижения своих целей и соответствовать применимым требованиям, обязательствам, связанным с заинтересованными сторонами, и ожиданиями от них.

Настоящий стандарт применим к любой организации независимо от размера, типа и рода деятельности, которая предоставляет или использует продукты или услуги, применяющие системы ИИ.

2 Нормативные ссылки

В настоящем стандарте использована нормативная ссылка на следующий стандарт [для датированной ссылки применяют только указанное издание ссылочного стандарта, для недатированной — последнее издание (включая все изменения)]:

ISO/IEC 22989:2022, Information technology — Artificial intelligence — Artificial intelligence concepts and terminology (Информационные технологии. Искусственный интеллект. Термины и определения, связанные с искусственным интеллектом)

3 Термины и определения

В настоящем стандарте применены термины по ИСО/МЭК 22989, а также следующие термины с соответствующими определениями.

ИСО и МЭК поддерживают терминологические базы данных для применения в сфере стандартизации по следующим адресам:

- платформа онлайн-просмотра ИСО: доступна по адресу: gost.e-ecolog.ru ;
- Электропедия МЭК: доступна по адресу: gost.e-ecolog.ru .

3.1 **организация** (organization): Лицо или группа людей, связанные определенными отношениями, имеющие ответственность, полномочия и выполняющие свои функции для достижения их целей (3.6).

Примечание 1 — Понятие организации включает в себя, но не ограничивается следующими примерами: индивидуальный предприниматель, компания, корпорация, фирма, предприятие, орган власти, товарищество, ассоциация, благотворительные учреждения, а также их часть или их объединение, являющиеся юридическим лицом или нет, государственные или частные.

Издание официальное

Примечание 2 — Если организация является частью более крупного предприятия, термин «организация» относится только к той части более крупного предприятия, которая входит в область применения системы менеджмента ИИ (3.4).

3.2 **заинтересованная сторона** (interested party, stakeholder): Лицо или организация (3.1), которые могут воздействовать на осуществление деятельности или принятие решения, быть подверженными их воздействию или воспринимать себя в качестве последних.

Примечание — В ИСО/МЭК 22989:2022, 5.19, представлен обзор ролей заинтересованных сторон в области ИИ.

3.3 **высшее руководство** (top management): Лицо или группа людей, осуществляющих руководство и управление организацией (3.1) на высшем уровне.

Примечание 1 — Высшее руководство имеет право делегировать полномочия и предоставлять ресурсы в рамках организации.

Примечание 2 — Если область применения системы менеджмента (3.4) охватывает только часть организации, под высшим руководством подразумевают тех, кто осуществляет руководство и управляет этой частью организации.

3.4 **система менеджмента** (management system): Совокупность взаимосвязанных или взаимодействующих элементов организации (3.1) для разработки политик (3.5), целей (3.6) и процессов (3.8) для достижения этих целей.

Примечание 1 — Система менеджмента может относиться к одному или нескольким аспектам деятельности.

Примечание 2 — Элементы системы менеджмента определяют структуру организации, роли и ответственность, планирование и функционирование.

3.5 **политика** (policy): Намерения и направления деятельности организации (3.1), официально сформулированные ее высшим руководством (3.3).

3.6 **цель** (objective): Результат, который должен быть достигнут.

Примечание 1 — Цель может быть стратегической, тактической или оперативной.

Примечание 2 — Цели могут относиться к разным аспектам [например, финансовые цели, цели в области здоровья и безопасности, экологии), а также применяться на разных уровнях (например, организации в целом, проекта, продукции или процесса (3.8)].

Примечание 3 — Цель может быть выражена разными способами, например, в виде намеченного результата, намерения, критерия работы, цели в области ИИ или другими словами со схожими значениями (например, целевая установка, заданная величина или задача).

Примечание 4 — В контексте системы менеджмента ИИ (3.4) цели в области ИИ, устанавливаемые организацией (3.1), согласуют с политикой в области ИИ (3.5) для достижения определенных результатов.

3.7 **риск** (risk): Влияние неопределенности.

Примечание 1 — Влияние выражается в отклонении от ожидаемого результата — позитивном или негативном.

Примечание 2 — Неопределенность является состоянием, связанным с недостатком (даже частично) информации, понимания или знания о событии, его последствиях или вероятности.

Примечание 3 — Риск часто определяют по отношениям к потенциальным событиям (как определено в Руководстве ИСО 73) и их последствиям (как определено в Руководстве ИСО 73) либо к их комбинации.

Примечание 4 — Риск часто выражается в терминах комбинации последствий события (включая изменения в обстоятельствах) и связанных с ними вероятностей (как определено в Руководстве ИСО 73) возникновения.

3.8 **процесс** (process): Совокупность взаимосвязанных и (или) взаимодействующих видов деятельности, использующих или преобразующих входы для получения результата.

Примечание — В зависимости от контекста «результат» называется выходом, продукцией или услугой.

3.9 **компетентность** (competence): Способность применять знания и навыки для достижения намеченных результатов.

3.10 **документированная информация; документ** (documented information): Информация, которая должна управляться и поддерживаться организацией (3.1), и носитель, который ее содержит.

Примечание 1 — Документированная информация может быть любого формата, на любом носителе и происходить из любого источника.

Примечание 2 — Документированная информация может относиться:
- к системе менеджмента (3.4), включая соответствующие процессы (3.8);
- информации, созданной для функционирования организации (документация);
- свидетельствам достигнутых результатов (записи).

3.11 **результаты деятельности** (performance): Измеримый итог.

Примечание 1 — Результаты деятельности могут относиться к количественным и качественным полученным выводам.

Примечание 2 — Результаты деятельности могут относиться к менеджменту действий, процессам (3.8), продукции, услугам, системам или организациям (3.1).

Примечание 3 — В контексте настоящего стандарта термин «результаты деятельности» относится как к результатам, достигнутым с помощью систем ИИ, так и к результатам, связанным с системой менеджмента ИИ. Правильное толкование этого термина становится понятным из контекста его употребления.

3.12 **постоянное улучшение** (continual improvement): Повторяющаяся деятельность по улучшению результатов деятельности (3.11).

3.13 **результативность** (effectiveness): Степень реализации запланированной деятельности и достижения запланированных результатов.

3.14 **требование** (requirement): Установленная потребность или ожидание, которое обычно предполагается или является обязательным.

Примечание 1 — Слова «обычно предполагается» означают, что это общепринятая практика организации (3.1) и заинтересованных сторон (3.2), что рассматриваемые потребности или ожидания предполагаются.

Примечание 2 — Установленным является такое требование, которое определено, например, в документированной информации (3.10).

3.15 **соответствие** (conformity): Выполнение требования (3.14).

3.16 **несоответствие** (nonconformity): Невыполнение требования (3.14).

3.17 **корректирующее действие** (corrective action): Действие, предпринятое для устранения причины несоответствия (3.16) и предупреждения его повторного возникновения.

3.18 **аудит** (audit): Систематический и независимый процесс (3.8) получения свидетельств и их объективного оценивания для установления степени соответствия критериям аудита.

Примечание 1 — Аудит может быть внутренним (аудит, проводимый первой стороной), внешним (аудит, проводимый второй или третьей сторонами) либо совместным (аудит, проводимый для двух или более систем менеджмента одновременно).

Примечание 2 — Внутренний аудит проводится самой организацией (3.1) или от ее имени внешней стороной.

Примечание 3 — «Свидетельства аудита» и «критерии аудита» определены в ИСО 19011.

3.19 **измерение** (measurement): Процесс (3.8) определения значения.

3.20 **мониторинг** (monitoring): Определение статуса системы, процесса (3.8) или действия.

Примечание — Для определения статуса может возникнуть необходимость проверить, проконтролировать или отследить.

3.21 **управление (риском)** (control): Меры, направленные на сохранение и изменение риска (3.7).

Примечание 1 — Управление риском охватывает, но не ограничивается этим: процессы, политику, устройства, методы и другие средства, используемые для сохранения и изменения риска.

Примечание 2 — Управление риском не всегда может привести к запланированным или ожидаемым результатам изменения риска.

[ИСО 31000:2018, 3.8]

3.22 **руководящий орган** (governing body): Лицо или группа лиц, которые отвечают за работу организации и ее соответствие требованиям.

Примечание 1 — Некоторые организации, особенно небольшие, могут не иметь управляющего органа, отдельного от высшего руководства.

Примечание 2 — В состав руководящего органа могут входить, но не ограничиваются нижеследующим: совет директоров, комитеты правления, наблюдательный совет, совет попечителей или надзирателей.

[ИСО/МЭК 38500:2015, 2.9]

3.23 **информационная безопасность** (information security): Сохранение конфиденциальности, целостности и доступности информации.

Примечание — Этот термин может включать в себя и другие дополнительные свойства, такие как подлинность, подотчетность, неотказуемость и достоверность.

[ИСО/МЭК 27000:2018, 3.28]

3.24 **оценка воздействия системы ИИ** (AI system impact assessment): Формализованный документированный процесс, посредством которого организация, разрабатывающая, предоставляющая или использующая продукты или услуги с применением ИИ, выявляет, оценивает воздействие на отдельных лиц, группы лиц, социальные группы и принимает соответствующие меры.

3.25 **качество данных** (data quality): Характеристика данных, показывающая степень их соответствия требованиям организации к данным для решения соответствующей задачи.

[ИСО/МЭК 5259-1:2024, 3.4]

3.26 **заявление о применимости** (statement of applicability): Документирование всех необходимых мер управления (3.21) и обоснование для их включения или исключения.

Примечание 1 — Организация может выбрать необходимые ей меры управления, перечисленные в приложении А, или рассмотреть необходимость дополнительных мер управления, установленных самой организацией.

Примечание 2 — В соответствии с требованиями настоящего стандарта все выявленные риски должны быть задокументированы организацией. Все выявленные риски и меры по управлению рисками (меры управления), установленные для их устранения, должны быть отражены в заявлении о применимости.

4 Среда организации

4.1 Понимание организации и ее среды

Организация должна определить связанные с ее целями внешние и внутренние факторы, которые влияют на способность достичь намеченные результаты для ее системы менеджмента ИИ.

Организация должна определить, уместно ли учитывать при этом изменение климата. Организация должна учитывать предполагаемое назначение систем ИИ, которые разрабатываются, предоставляются или используются организацией. Организация должна определить свои роли в отношении систем ИИ.

Примечание 1 — Чтобы понять организацию и ее контекст, для организации может быть полезно определить свою роль в отношении систем ИИ. Эти роли могут включать, но не ограничиваются нижеследующими:

- поставщики решений по ИИ, включая поставщиков платформ ИИ, поставщиков продуктов или услуг ИИ;
- производители ИИ, включая разработчиков ИИ, проектировщиков ИИ, операторов ИИ, тестировщиков и оценщиков ИИ, специалистов по развертыванию ИИ, специалистов по человеческому фактору ИИ, экспертов в предметной области, специалистов по оценке воздействия ИИ, поставщиков, специалистов по управлению и надзору за ИИ;
- потребители ИИ, включая пользователей ИИ;
- партнеры по ИИ, включая системного интегратора ИИ и поставщика данных;
- субъекты ИИ, включая субъекты данных и другие субъекты;
- соответствующие органы власти, включая директивные и регулирующие органы.

Подробное описание этих ролей приведено в ИСО/МЭК 22989. Кроме того, типы ролей и их взаимосвязь с жизненным циклом системы ИИ также описаны в системе управления рисками ИИ [29]. Роли организации определяют применимость и степень применимости требований и мер управления, изложенных в настоящем стандарте.

Примечание 2 — Внешние и внутренние факторы, подлежащие решению в соответствии с настоящим пунктом, могут варьироваться в зависимости от ролей и юрисдикции организации и их влияния на ее способность достигать предполагаемых результатов при помощи системы менеджмента ИИ. Они могут включать в себя, но не ограничиваются, нижеследующим:

а) рассмотрением вопросов, связанных с внешней средой организации, таких как:
   1) применимые юридические требования, включая запрещенное использование ИИ;
   2) политики, руководящие принципы и решения регулирующих органов, оказывающие влияние на толкование или обеспечение соблюдения юридических требований при разработке и использовании систем ИИ;
   3) стимулы или последствия, связанные с предполагаемой целью и использованием систем ИИ;
   4) культура, традиции, ценности, нормы и этика в отношении разработки и использования ИИ;
   5) конкурентная среда и тенденции для новых продуктов и услуг, использующих системы ИИ;

b) рассмотрением вопросов, связанных с внутренней средой организации, таких как:
1) организационная среда, управление, цели (см. 6.2), политики и процедуры;
2) договорные обязательства;
3) предполагаемое назначение системы ИИ, которая будет разработана или использована.

Примечание 3 — Роль организации может быть определена обязательствами, связанными с категориями данных, которые обрабатывает организация [например, обработчик персональных данных (ПДн) или оператор ПДн при обработке ПДн]. Информация о ПДн и о ролях в обработке ПДн приведена в [23]. Роли также могут определяться юридическими требованиями, характерными для систем ИИ.

4.2 Понимание потребностей и ожиданий заинтересованных сторон

Организация должна определить:
- заинтересованные стороны, имеющие отношение к системе менеджмента ИИ;
- соответствующие требования этих заинтересованных сторон;
- какие из этих требований будут выполнены с помощью системы менеджмента ИИ.

Примечание — Соответствующие заинтересованные стороны могут иметь требования, связанные с изменением климата.

4.3 Определение области применения системы менеджмента искусственного интеллекта

Организация должна определить границы системы менеджмента ИИ и охватываемую ею деятельность, чтобы установить область ее применения.

При определении области применения организация должна рассматривать:
- внешние и внутренние факторы (см. 4.1);
- требования (см. 4.2).

Область применения должна быть задокументирована.

Область применения системы менеджмента ИИ должна определять деятельность организации в отношении требований настоящего стандарта к системе менеджмента ИИ, руководству, планированию, поддержке, эксплуатации, производительности, оценке, улучшению, мерам управления и целям.

4.4 Система менеджмента искусственного интеллекта

Организация должна разработать, внедрить, поддерживать в рабочем состоянии, постоянно улучшать и документировать систему менеджмента ИИ, включая необходимые процессы и их взаимодействие в соответствии с требованиями настоящего стандарта.

5 Лидерство

5.1 Лидерство и приверженность

Высшее руководство должно демонстрировать свое лидерство и приверженность в отношении системы менеджмента ИИ посредством:
- обеспечения разработки политики (см. 5.2) и целей (см. 6.2) в области ИИ, которые согласуются со стратегическим направлением деятельности организации;
- обеспечения интеграции требований системы менеджмента ИИ в деловые процессы организации;
- обеспечения доступности ресурсов, необходимых для системы менеджмента ИИ;
- распространения в организации понимания важности результативного управления ИИ и соответствия требованиям системы менеджмента ИИ;
- обеспечения достижения системой менеджмента ИИ намеченных результатов;
- вовлечения, руководства и оказания поддержки участия работников в обеспечении результативности системы менеджмента ИИ;
- поддержки постоянного улучшения;
- поддержки других соответствующих руководителей в демонстрации ими лидерства в сфере их ответственности.

Примечание 1 — Употребление слова «бизнес» в настоящем стандарте можно толковать в широком смысле как означающее те виды деятельности, которые являются ключевыми для целей существования организации.

Примечание 2 — Создание, поощрение и моделирование внутри организации культуры ответственного подхода к использованию, разработке и управлению системами ИИ может стать важной демонстрацией приверженности и лидерства со стороны высшего руководства. Обеспечение осведомленности и соблюдения такого ответственного подхода, а также поддержка системы менеджмента ИИ посредством лидерства могут способствовать успеху системы менеджмента ИИ.

5.2 Политика в области искусственного интеллекта

Высшее руководство должно разработать политику в области ИИ, которая:

a) соответствует намерениям организации;

b) создает основу для установления целей в области ИИ (см. 6.2);

c) включает в себя обязательство соответствовать применимым требованиям;

d) включает в себя обязательство постоянно улучшать систему менеджмента ИИ.

Политика в области ИИ должна:

- быть доступной и документированной;
- ссылаться на другие политики организации, где это применимо;
- доводиться до сведения работников организации;
- быть доступной для заинтересованных сторон, где это необходимо.

Меры и цели управления для разработки политики в области ИИ перечислены в таблице А.1 (пункт А.2). Руководство по внедрению этих мер управления приведено в В.2.

Примечание — Рекомендации для организаций при разработке политик в области ИИ приведены в [27].

5.3 Функции, ответственность и полномочия

Высшее руководство должно обеспечить в организации распределение соответствующих обязанностей, ответственности и полномочий.

Высшее руководство должно распределить обязанности, ответственность и полномочия:

a) для обеспечения соответствия системы менеджмента ИИ требованиям настоящего стандарта;

b) отчетности высшему руководству о результатах функционирования системы менеджмента ИИ.

Примечание — Меры управления для определения и распределения ролей и обязанностей приведены в таблице А.1 (пункт А.3.2). Руководство по внедрению этих мер управления приведено в В.3.2.

6 Планирование

6.1 Действия в отношении рисков и возможностей

6.1.1 Общие положения

При планировании в системе менеджмента ИИ организация должна учесть факторы (см. 4.1) и требования (см. 4.2), а также определить риски и возможности, подлежащие рассмотрению, в целях:

- обеспечения уверенности в том, что система менеджмента ИИ может достичь намеченных результатов;
- предотвращения или уменьшения их нежелательного воздействия;
- достижения постоянного улучшения.

Организация должна установить и поддерживать в актуальном состоянии критерии рисков ИИ, позволяющие:

- отличать приемлемые риски от неприемлемых;
- проводить оценку рисков ИИ;
- проводить обработку рисков ИИ;
- проводить оценку воздействия рисков ИИ.

Примечание 1 — Рекомендации по определению степени и типа рисков, которые организация готова принять или сохранять, приведены в [27] и [14].

Организация должна определять риски и возможности в соответствии со следующими факторами:

- предметная область и среда применения системы ИИ;
- предполагаемое использование;
- внешняя и внутренняя среда (см. 4.1).

Примечание 2 — В рамках системы менеджмента ИИ можно рассматривать более одной системы ИИ. В этом случае возможности и варианты использования следует определять для каждой системы ИИ или группы систем ИИ.

Организация должна планировать:

a) действия по рассмотрению этих рисков и возможностей;

b) то, каким образом:

1) интегрировать и внедрять эти действия в процессы системы менеджмента ИИ;

2) оценивать результативность этих действий.

Организация должна сохранять документированную информацию о действиях, предпринятых для выявления и рассмотрения рисков и возможностей ИИ.

Примечание 3 — Рекомендации по организации управления рисками для организаций, предоставляющих или использующих продукты, системы и услуги ИИ, приведены в [14].

Примечание 4 — Среда организации и ее деятельность могут оказывать влияние на деятельность организации по управлению рисками.

Примечание 5 — Способ определения риска и, следовательно, представления об управлении рисками может варьироваться в зависимости от различных сфер деятельности. Определение риска, изложенное в 3.7, позволяет получить широкое представление о риске, применимое к любой сфере деятельности, например, к упомянутым в D.1. Роль организации в рамках оценки рисков заключается в том, чтобы сначала принять видение риска, адаптированное к ее среде. Это может включать подход к рискам с помощью определений, используемых в различных сферах деятельности, для которых разрабатывается и используется система ИИ, таких как определение из [7].

6.1.2 Оценка рисков искусственного интеллекта

Организация должна определить и внедрить процесс оценки рисков ИИ, который должен:

a) основываться на политике в области ИИ (см. 5.2) и целях ИИ (см. 6.2) и согласоваться с ними.

Примечание — При оценке последствий в рамках 6.1.2 d) 1) организация может использовать оценку воздействия системы ИИ, как указано в 6.1.4;

b) быть разработан таким образом, чтобы повторные оценки рисков ИИ давали непротиворечивые, достоверные и сопоставимые результаты;

c) определять риски, которые помогают или препятствуют достижению целей ИИ;

d) проводить анализ рисков ИИ:

1) оценивать потенциальные последствия для организации, отдельных лиц и социальных групп, которые могут произойти в результате наступления выявленных рисков;

2) оценивать реальную вероятность наступления выявленных рисков;

3) определять уровни рисков;

e) оценивать риски ИИ, т. е.:

1) сравнивать результаты анализа рисков ИИ с критериями рисков, установленными в соответствии с 6.1.1;

2) определять приоритетность обработки проанализированных рисков ИИ.

Организация должна хранить документированную информацию о процессе оценки рисков ИИ.

6.1.3 Обработка рисков искусственного интеллекта

Опираясь на результаты оценки рисков, организация должна определить процесс обработки рисков ИИ:

a) для выбора подходящих вариантов обработки рисков ИИ;

b) определения всех мер управления, необходимых для реализации выбранного(ых) варианта(ов) обработки рисков ИИ; сравнения мер управления, определенных в соответствии с указанными в приложении А для проверки того, что никакие необходимые меры управления не были упущены.

Примечание 1 — В приложении А приведен перечень основных мер управления для достижения организационных целей и рассмотрения рисков, связанных с проектированием и использованием систем ИИ;

c) определения мер управления из приложения А, необходимых для реализации вариантов обработки рисков ИИ;

d) определения необходимости дополнительных мер управления (помимо указанных в приложении А) для реализации всех вариантов обработки рисков;

e) обзора приведенных в приложении В руководящих указаний по внедрению мер управления, определенных в b) и c).

Примечание 2 — Цели управления неявным образом включены в выбранные меры управления. Организация может выбрать необходимые ей меры и цели управления, перечисленные в приложении А. Приведенные в приложении А меры и цели управления не являются исчерпывающими, и организация может рассмотреть необходимость дополнительных мер управления и целей их применения. При необходимости организация может разрабатывать меры управления или брать их из существующих источников. Управление рисками ИИ может быть интегрировано в другие системы менеджмента, если это применимо;

f) подготовки заявления о применимости мер управления, которое содержит: необходимые меры управления [см. b), c) и d)]; обоснование их применения; информацию о том, реализованы или нет необходимые меры управления; обоснование неприменения мер управления, представленных в приложении А. Обоснование для неприменения мер управления может включать случаи, когда меры управления не считаются необходимыми в результате оценки риска и когда они не предусматриваются применимыми внешними требованиями (или подпадают под исключения в соответствии с ними).

Примечание 3 — Организация может предоставить документированные обоснования для исключения любых целей управления в целом или для конкретных систем ИИ, будь то перечисленные в приложении А или установленные самой организацией;

g) разработки плана обработки рисков ИИ.

Организация должна получить от назначенного руководства согласование плана обработки рисков ИИ и принятие остаточных рисков ИИ владельцами рисков. Необходимые меры управления должны быть:

- согласованы с целями, указанными в 6.2;
- доступны в виде документированной информации;
- доведены до сведения внутри организации;
- доступны заинтересованным сторонам в случае необходимости.

Организация должна хранить документированную информацию о процессе обработки рисков ИИ.

6.1.4 Оценка воздействия системы искусственного интеллекта

Организация должна определить процесс оценки потенциальных последствий для отдельных лиц или групп лиц, для тех и других либо для социальных групп, которые могут возникнуть в результате разработки, предоставления или использования систем ИИ.

Оценка воздействия системы ИИ должна определять потенциальные последствия развертывания, предполагаемого использования и прогнозируемого неправильного использования системы ИИ для отдельных лиц или групп лиц, для тех и других либо для социальных групп.

Оценка воздействия системы ИИ должна учитывать конкретную техническую и социальную среду, в которой развернута система ИИ, и применимые юрисдикции.

Результат оценки воздействия на систему должен быть задокументирован. При необходимости результаты оценки воздействия на систему могут быть доступны соответствующим заинтересованным сторонам способом, определенным организацией.

Организация должна учитывать результаты оценки воздействия системы ИИ в своей оценке рисков (см. 6.1.2). В таблице А.1 (пункт А.5) приведены меры управления для оценки воздействия систем ИИ.

Примечание — В некоторых случаях (например, в системах ИИ, критически важных для обеспечения безопасности или конфиденциальности) организация может потребовать, чтобы оценка воздействия предметно-ориентированной системы ИИ (например, влияющей на безопасность, конфиденциальность или на систему охраны), проводилась в рамках общей деятельности организации по управлению рисками.

6.2 Цели искусственного интеллекта и планирование их достижения

В организации должны быть установлены цели ИИ применительно к соответствующим функциям и уровням управления организацией.

Цели ИИ должны:

a) быть согласованными с политикой в области ИИ (см. 5.2);
b) быть измеримыми (если это практически осуществимо);
c) учитывать применимые требования ИИ;
d) подлежать мониторингу с точки зрения достижения;
e) быть доведены до сведения всех заинтересованных сторон;
f) при необходимости актуализироваться;
g) быть доступными и применяться как документированная информация.

При планировании способов достижения целей в области ИИ организация должна определить:
- что должно быть сделано;
- какие для этого требуются ресурсы;
- кто будет нести за это ответственность;
- сроки достижения целей;
- каким образом будут оцениваться полученные результаты.

Примечание — Неисключительный перечень целей ИИ, связанных с управлением рисками, приведен в приложении С. Цели и меры управления для определения целей ответственной разработки и использования систем ИИ, а также меры по их достижению представлены в таблице А.1 (пункты А.6.1 и А.9.3). Руководство по внедрению указанных мер управления представлено в В.6.1 и В.9.3.

6.3 Планирование изменений

В тех случаях, когда организация выявляет необходимость в изменениях системы менеджмента ИИ, эти изменения должны осуществляться на плановой основе системным образом.

7 Средства обеспечения

7.1 Ресурсы

Организация должна определить и обеспечить наличие ресурсов, необходимых для создания, внедрения, поддержки и постоянного улучшения системы менеджмента ИИ.

Примечание — Цели и меры управления ресурсами ИИ приведены в таблице А.1 (пункт А.4). Руководство по внедрению указанных мер управления приведено в В.4.

7.2 Компетентность

Организация должна:
- определить необходимую компетентность лиц(а), выполняющих(его) работу под ее управлением, которая оказывает влияние на результаты деятельности и результативность системы ИИ;
- обеспечивать компетентность этих лиц на основе соответствующего образования, подготовки и (или) опыта;
- там, где это применимо, предпринимать действия, направленные на получение требуемой компетентности, и оценивать результативность предпринятых действий.

Сохранять соответствующую документированную информацию как свидетельство компетентности.

Примечание 1 — Руководство по развитию человеческих ресурсов, включая рассмотрение необходимого опыта, приведено в В.4.6.

Примечание 2 — Применимые действия могут включать, например проведение обучения, наставничество или перераспределение обязанностей среди имеющихся работников; или же наем лиц, обладающих требуемым уровнем компетентности.

7.3 Осведомленность

Лица, выполняющие работу под управлением организации, должны быть осведомлены:
- о политике в области ИИ (5.2);
- своем вкладе в обеспечение результативности системы менеджмента ИИ, включая пользу от улучшения результатов деятельности ИИ;
- последствиях несоответствия требованиям системы менеджмента ИИ.

7.4 Обмен информацией

Организация должна определить порядок внутреннего и внешнего обмена информацией, относящейся к системе менеджмента ИИ, включая следующее:
- какая информация будет передаваться;
- когда будет передаваться информация;
- кому будет передаваться информация;
- каким образом будет передаваться информация.

7.5 Документированная информация

7.5.1 Общие положения

Система менеджмента ИИ организации должна включать:

a) документированную информацию, требуемую в соответствии с настоящим стандартом;

b) документированную информацию, определенную организацией как необходимую для обеспечения результативности системы менеджмента ИИ.

Примечание — Объем документированной информации системы менеджмента ИИ одной организации может отличаться от другой в зависимости:

- от размера организации и вида ее деятельности, процессов, продукции и услуг;
- сложности процессов и их взаимодействия;
- компетентности работников.

7.5.2 Создание и актуализация документированной информации

При создании и актуализации документированной информации организация должна соответствующим образом обеспечить:

- идентификацию и описание (например, наименование, дата, автор или ссылочный номер);
- формат (например, язык, версия программного обеспечения, графические средства) и носитель информации (например, бумажный или электронный);
- анализ, пересмотр и одобрение с точки зрения пригодности и адекватности.

7.5.3 Управление документированной информацией

Документированная информация, требуемая системой менеджмента ИИ и настоящим стандартом, должна находиться под управлением в целях обеспечения:

a) ее доступности и пригодности для использования, где и когда это необходимо;

b) надлежащей защиты (например, от несоблюдения конфиденциальности, от ненадлежащего использования или потери целостности).

Для управления документированной информацией организация должна предусматривать следующие действия в той степени, в какой это применимо:

- распространение, обеспечение доступа, поиск и использование;
- хранение и сохранность документов, включая их читаемость;
- управление изменениями (например, управление версиями/редакциями);
- соблюдение сроков хранения и порядок уничтожения или передачи на архивное хранение.

Документированная информация внешнего происхождения, определенная организацией как необходимая для планирования и системы менеджмента ИИ, должна быть соответствующим образом идентифицирована и находиться под управлением.

Примечание — Доступ к документированной информации подразумевает разрешение только просмотра документированной информации или разрешение просмотра с полномочиями по внесению изменений в документированную информацию.

8 Деятельность

8.1 Планирование и управление

Организация должна планировать, внедрять процессы, необходимые для выполнения требований и для выполнения действий, определенных в разделе 6, и осуществлять управление этими процессами посредством:

- установления критериев для процессов;
- осуществления управления процессами на основе установленных критериев.

Организация должна внедрять меры управления, определенные в соответствии с 6.1.3, которые связаны с работой системы управления ИИ (например, меры управления, связанные с разработкой системы ИИ и жизненным циклом).

Организация должна управлять эффективностью представленных мер управления и в случае, если запланированные результаты не достигнуты, должна рассмотреть вопрос применения корректирующих действий. В приложении А перечислены эталонные средства контроля, а в приложении В приведены рекомендации по их внедрению.

Организация должна хранить документированную информацию в объеме, необходимом для обеспечения уверенности в том, что процессы выполнялись так, как это было запланировано.

Организация должна управлять запланированными изменениями и анализировать последствия непредусмотренных изменений, предпринимая, при необходимости, меры по смягчению любых негативных воздействий.

Организация должна контролировать поступающие извне процессы, а также товары и услуги, которые имеют отношение к системе менеджмента ИИ.

8.2 Оценка рисков искусственного интеллекта

Организация должна проводить оценку рисков ИИ в соответствии с 6.1.2 через запланированные периоды времени или в случае предполагаемых или произошедших значительных изменений.

Организация должна хранить документированную информацию о результатах проведенных оценок рисков ИИ.

8.3 Обработка рисков искусственного интеллекта

Организация должна реализовывать план обработки рисков ИИ в соответствии с 6.1.3 и осуществить проверку его эффективности.

Для новых рисков, выявленных в процессе проведения оценки рисков и требующих обработки, должен быть выполнен процесс обработки рисков.

В случае выявления неэффективности способов обработки рисков, определенных планом обработки рисков, они должны быть пересмотрены и повторно подтверждены в соответствии с 6.1.3, и план обработки рисков должен быть актуализирован.

Организация должна хранить документированную информацию о результатах проведенных мероприятий по устранению рисков ИИ.

8.4 Оценка воздействия системы искусственного интеллекта

Организация должна проводить оценку воздействия системы ИИ в соответствии с 6.1.4 через запланированные промежутки времени или в случае предполагаемых значительных изменений.

Организация должна хранить документированную информацию о результатах проведенных оценок воздействия системы ИИ.

9 Оценка результатов деятельности

9.1 Мониторинг, измерение, анализ и оценка

Организация должна определить:

- объекты мониторинга и измерения;
- методы проведения мониторинга, измерения, анализа и оценки, необходимые для обеспечения достоверных результатов;
- периодичность проведения мониторинга и измерения;
- когда результаты мониторинга и измерения должны быть проанализированы и оценены.

Необходимо обеспечить наличие документированной информации, свидетельствующей о полученных результатах.

Организация должна оценить результаты деятельности и результативность системы менеджмента ИИ.

9.2 Внутренний аудит

9.2.1 Общие положения

Организация должна проводить внутренние аудиты через запланированные периоды времени для получения информации, что система менеджмента ИИ:

a) соответствует:
   1) собственным требованиям организации к ее системе менеджмента ИИ;
   2) требованиям настоящего стандарта;
b) эффективно внедрена и функционирует.

9.2.2 Программа внутреннего аудита

Организация должна планировать, разрабатывать, реализовывать и поддерживать в актуальном состоянии программу(ы) аудитов, включая периодичность и методы проведения аудитов, а также ответственность, планируемые для проверки требования и предоставление отчетности.

Программа(ы) аудитов должна(ы) разрабатываться с учетом важности проверяемых процессов и результатов предыдущих аудитов.

Организация должна:

a) определить цели, критерии и область проверки для каждого аудита;

b) отбирать аудиторов и проводить аудит таким образом, чтобы обеспечить объективность и беспристрастность процесса аудита;

c) обеспечивать передачу информации о результатах аудитов соответствующим руководителям.

Организация должна сохранять соответствующую документированную информацию как свидетельство реализации программы аудита и полученных результатов аудита.

9.3 Анализ со стороны руководства

9.3.1 Общие положения

Высшее руководство должно анализировать через запланированные периоды времени систему менеджмента ИИ в целях обеспечения ее постоянной пригодности, адекватности и результативности.

9.3.2 Входные данные анализа со стороны руководства

Анализ со стороны руководства должен включать в себя рассмотрение:

a) степени реализации решений, осуществляемых по результатам предыдущих анализов со стороны руководства;

b) изменений во внешних и внутренних факторах, касающихся системы менеджмента ИИ;

c) изменений в потребностях и ожиданиях заинтересованных сторон, касающихся системы менеджмента ИИ;

d) информации о результатах деятельности в системе менеджмента ИИ, включая тенденции, относящиеся:

1) к выявлению несоответствий и применению корректирующих действий;
2) результатам мониторинга и измерений;
3) результатам аудитов;

e) потенциальных возможностей для постоянного улучшения системы.

9.3.3 Выходные данные анализа со стороны руководства

Итоги анализа со стороны руководства должны включать в себя решения, относящиеся к возможностям постоянного улучшения и необходимости внесения изменений в систему менеджмента ИИ.

Организация должна хранить документированную информацию как свидетельство результатов анализа со стороны руководства.

10 Улучшения

10.1 Постоянное улучшение

Организация должна постоянно улучшать пригодность, адекватность и результативность системы менеджмента ИИ.

10.2 Несоответствия и корректирующие действия

При выявлении несоответствий организация должна:

a) реагировать на данное несоответствие и, насколько это применимо:

1) предпринимать действия по управлению и коррекции выявленного несоответствия;
2) предпринимать действия в отношении последствий данного несоответствия;

b) оценивать необходимость действий по устранению причин данного несоответствия с тем, чтобы избежать его повторного появления или появления в другом месте посредством:

1) анализа несоответствия;
2) определения причин, вызвавших появление несоответствия;
3) определения наличия аналогичного несоответствия или возможности его возникновения где-либо еще;

c) осуществлять необходимые корректирующие действия;
d) анализировать результативность каждого предпринятого корректирующего действия;
e) вносить при необходимости изменения в систему менеджмента ИИ.

Корректирующие действия должны соответствовать последствиям выявленных несоответствий.

Необходимо обеспечить наличие документированной информации, свидетельствующей:

- о характере несоответствий и любых последующих предпринятых действий;
- результатах всех корректирующих действий.

Приложение А
(обязательное)

Меры и цели управления

А.1 Общие положения

Перечисленные в таблице А.1 цели, а также меры управления служат организации ориентиром для достижения организационных целей и устранения рисков, связанных с проектированием и эксплуатацией систем ИИ. Перечень мер управления, содержащийся в данной таблице, не является исчерпывающим, и организация может разрабатывать и внедрять собственные меры управления (см. 6.1.3).

В приложении В приведены рекомендации по внедрению мер управления, перечисленных в таблице А.1.

Таблица А.1 — Меры управления и цели их применения

А.2 Политики в области ИИ		
А.2.1 Цель: обеспечить получение от руководства руководящих указаний и поддержки систем ИИ в соответствии с деловыми требованиями		
	Тема	Меры управления
А.2.2	Политика в области ИИ	Организация должна задокументировать политику развития и использования систем ИИ
А.2.3	Согласование с другими организационными политиками	Организация должна определить, каким образом цели организации в отношении систем ИИ могут повлиять на другие ее политики и каким образом другие политики могут оказаться применимыми в отношении данных целей
А.2.4	Пересмотр политики в области ИИ	Политика в области ИИ должна пересматриваться в плановом порядке (и дополнительно по мере необходимости) для обеспечения ее постоянной уместности, адекватности и эффективности
А.3 Организация внутренней деятельности		
А.3.1 Цель: обеспечить в организации подотчетность, поддерживающую ее ответственный подход к внедрению, эксплуатации и управлению системами ИИ		
	Тема	Меры управления
А.3.2	Роли и обязанности в сфере ИИ	Роли и ответственность в сфере ИИ должны быть определены и распределены в соответствии с потребностями организации
А.3.3	Информирование о проблемах	Организация должна определять и внедрять процессы, позволяющие сообщать о проблемах, связанных с ролью организации в отношении системы ИИ на протяжении всего ее жизненного цикла
А.4 Ресурсы систем ИИ		
А.4.1 Цель: обеспечить учет организацией ресурсов системы ИИ (включая ее компоненты и активы) для полного понимания и рассмотрения связанных рисков и последствий		
	Тема	Меры управления
А.4.2	Документирование ресурсов	Организация должна выявить и задокументировать соответствующие ресурсы, необходимые для выполнения операций на конкретных стадиях жизненного цикла системы ИИ, а также для иных связанных с ИИ видов деятельности, актуальных для организации
А.4.3	Ресурсы данных	В рамках идентификации ресурсов организация должна документировать информацию о ресурсах данных, используемых для системы ИИ

Продолжение таблицы А.1

А.4.4	Инструментальные ресурсы	В рамках идентификации ресурсов организация должна документировать информацию об инструментальных ресурсах, используемых для системы ИИ
А.4.5	Системы и вычислительные ресурсы	В рамках идентификации ресурсов организация должна документировать информацию о системе и вычислительных ресурсах, используемых для системы ИИ
А.4.6	Человеческие ресурсы	В рамках идентификации ресурсов организация должна документировать информацию о человеческих ресурсах и компетенциях, используемых для разработки, развертывания и эксплуатации, управления изменениями, обслуживания, передачи и вывода из эксплуатации, а также проверки и интеграции системы ИИ
А.5 Оценка воздействия систем ИИ		
А.5.1 Цель: провести оценку воздействия системы ИИ на отдельных лиц, группы людей, или и тех и других, а также социальные группы, на которые влияет система ИИ на протяжении всего ее жизненного цикла		
	Тема	Меры управления
А.5.2	Процесс оценки воздействия системы ИИ	Организация должна разработать процесс проведения оценки потенциальных последствий для отдельных лиц, групп людей или и тех и других, а также социальных групп, которые могут появиться в результате разработки или использования систем ИИ на протяжении ее жизненного цикла
А.5.3	Документирование оценки воздействия системы ИИ	Организация должна документировать результаты оценки воздействия системы ИИ и сохранять соответствующие документы в течение установленных сроков хранения
А.5.4	Оценка воздействия системы ИИ на отдельных лиц и группы лиц	Организация должна проводить оценку и документировать потенциальное воздействие систем ИИ на отдельных лиц, организации или сообщества на протяжении всего жизненного цикла системы
А.5.5	Оценка воздействия систем ИИ на общество	Организация должна проводить оценку и документировать потенциальное воздействие своих систем ИИ на общество на протяжении всего их жизненного цикла
А.6 Жизненный цикл системы ИИ		
А.6.1 Руководство по управлению разработкой системы ИИ		
А.6.1.1 Цель: обеспечить определение и документирование организацией целей и внедрения процессов ответственного проектирования и разработки систем ИИ		
	Тема	Меры управления
А.6.1.2	Цели ответственной разработки системы ИИ	Организация должна определить и задокументировать цели, которыми следует руководствоваться при ответственной разработке систем ИИ, а также учитывать эти цели и интегрировать меры по их достижению в жизненный цикл разработки
А.6.1.3	Процессы ответственного проектирования и разработки системы ИИ	Организация должна определить и задокументировать конкретные процессы ответственного проектирования и разработки системы ИИ

Продолжение таблицы А.1

А.6.2 Жизненный цикл системы ИИ		
А.6.2.1 Цель: определить критерии и требования для каждой стадии жизненного цикла системы ИИ		
	Тема	Меры управления
А.6.2.2	Требования и спецификация к системе ИИ	Организация должна определить и задокументировать требования к новым системам ИИ или существенным усовершенствованиям существующих систем
А.6.2.3	Документация по проектированию и разработке системы ИИ	Организация должна документировать проектирование и разработку системы ИИ на основе организационных целей, задокументированных требований и установленных в спецификациях критериев
А.6.2.4	Верификация и валидация системы ИИ	Организация должна определить и задокументировать меры верификации и валидации для системы ИИ и указать критерии для их использования
А.6.2.5	Развертывание системы ИИ	Организация должна задокументировать план развертывания и обеспечить выполнение соответствующих требований до инициирования самого развертывания
А.6.2.6	Функционирование и мониторинг системы ИИ	Организация должна определить и задокументировать необходимые элементы для непрерывной работы системы ИИ. Как минимум, это должно включать мониторинг системы и ее производительности, ремонт, обновления и поддержку
А.6.2.7	Техническая документация по системе ИИ	Организация должна определить, какая техническая документация по системе ИИ необходима для каждой соответствующей категории заинтересованных сторон, таких как пользователи, партнеры, контролирующие органы, и предоставить им техническую документацию в соответствующей форме
А.6.2.8	Ведение журналов событий системой ИИ	Организация должна определить, на каких стадиях жизненного цикла ИИ должно быть включено ведение журналов событий, но как минимум, при использовании системы ИИ
А.7 Данные для систем ИИ		
А. 7.1 Цель: обеспечить понимание организацией роли и воздействия данных в системах ИИ при применении и разработке, предоставлении или использовании систем ИИ на протяжении их жизненного цикла		
	Тема	Меры управления
А.7.2	Данные для разработки и усовершенствования системы ИИ	Организация должна определять, документировать и внедрять процессы управления данными, связанные с разработкой систем ИИ
А.7.3	Сбор данных	Организация должна определить и задокументировать подробную информацию о сборе и отборе данных, используемых в системах ИИ
А.7.4	Качество данных для систем ИИ	Организация должна определить и задокументировать требования к качеству данных и обеспечить соответствие данных, используемых для разработки и эксплуатации системы ИИ, этим требованиям
А.7.5	Происхождение данных	Организация должна определить и задокументировать процесс проверки и документирования сведений о происхождении данных, используемых в ее системах ИИ на протяжении жизненного цикла данных и системы ИИ
А.7.6	Подготовка данных	Организация должна определить и задокументировать свои критерии для отбора методов подготовки данных и сами методы подготовки данных, которые будут использоваться

Окончание таблицы А.1

А.8 Информация для заинтересованных сторон систем ИИ		
А.8.1 Цель: обеспечить предоставление соответствующим заинтересованным сторонам необходимой информации для понимания и оценки рисков и их последствий (как положительных, так и отрицательных)		
	Тема	Меры управления
А.8.2	Системная документация и информация для пользователей	Организация должна определить и предоставить необходимую информацию пользователям системы
А.8.3	Внешняя отчетность	Организация должна предоставить возможности для отчетности о неблагоприятных воздействиях системы ИИ
А.8.4	Информирование об инцидентах	Организация должна определить и задокументировать план информирования пользователей системы об инцидентах
А.8.5	Информация для заинтересованных сторон	Организация должна определить и задокументировать свои обязательства по предоставлению информации о системе ИИ заинтересованным сторонам
А.9 Использование систем ИИ		
А.9.1 Цель: обеспечить ответственное использование систем ИИ организацией в соответствии с политиками организации		
	Тема	Меры управления
А.9.2	Процессы ответственного использования систем ИИ	Организация должна определить и задокументировать процессы ответственного использования систем ИИ
А.9.3	Цели ответственного использования системы ИИ	Организация должна определить и задокументировать цели, которыми следует руководствоваться при ответственном использовании систем ИИ
А.9.4	Предполагаемое использование системы ИИ	Организация должна обеспечить, чтобы система ИИ использовалась в соответствии с предполагаемым использованием системы ИИ и сопровождающей ее документацией
А.10 Взаимоотношения с третьими сторонами и клиентами		
А.10.1 Цель: обеспечить, чтобы организация понимала свои обязанности и оставалась подотчетной за них, а риски соответствующим образом распределялись при участии третьих сторон на любой стадии жизненного цикла системы ИИ		
	Тема	Меры управления
А.10.2	Распределение обязанностей	Организация должна обеспечить распределение обязанностей в рамках жизненного цикла своей системы ИИ между организацией, ее партнерами, поставщиками, заказчиками и третьими сторонами
А.10.3	Поставщики	Организация должна разработать процесс, гарантирующий, что использование ею услуг, продуктов или материалов, предоставляемых поставщиками, соответствует подходу организации к ответственной разработке и использованию систем ИИ
А.10.4	Заказчики	Организация должна обеспечить, чтобы ее ответственный подход к разработке и использованию систем ИИ учитывал ожидания и потребности клиентов

Приложение В
(обязательное)

Руководство по внедрению мер управления по обработке рисков искусственного интеллекта

В.1 Общие положения

Руководство по внедрению, приведенное в настоящем приложении, относится к мерам управления, перечисленным в таблице А.1. В настоящем приложении содержится информация в поддержку внедрения мер управления, перечисленных в таблице А.1, и достижения цели управления, но организации не обязаны документировать или обосновывать включение или исключение руководства по внедрению в заявление о применимости (см. 6.1.3).

Руководство по внедрению не всегда является подходящим и достаточным во всех ситуациях, и не всегда соответствует специфическим требованиям организации к управлению. Организация может расширять или изменять руководство по внедрению или формировать свое собственное руководство по внедрению мер управления в соответствии со своими конкретными требованиями и потребностями в обработке рисков.

Настоящее приложение следует использовать в качестве руководства для определения и внедрения мер управления процессом обработки рисков ИИ в системе менеджмента ИИ, определенной в настоящем стандарте. Дополнительные организационные и технические меры управления, отличные от тех, которые включены в настоящее приложение, при необходимости могут быть определены путем оценки рисков (см. обработку рисков системой управления ИИ в 6.1.3).

Настоящее приложение можно рассматривать как отправную точку для разработки руководства по внедрению мер управления с учетом специфики организации.

В.2 Политики в области искусственного интеллекта

В.2.1 Цель

Обеспечивать управление и поддержку систем ИИ в соответствии с деловыми требованиями.

В.2.2 Политика в области искусственного интеллекта

Меры управления

Организация должна документировать политику разработки и использования систем ИИ.

Руководство по внедрению

Политика в области ИИ должна исходить:

- из бизнес-стратегии;
- ценностей и культуры организации, а также из степени риска, который организация готова принять или сохранить;
- уровня риска, создаваемого системами ИИ;
- юридических требований,
- совокупности рисков организации;
- воздействия на соответствующие заинтересованные стороны (см. 6.1.4).

Политика в области ИИ должна включать (в дополнение к требованиям 5.2):

- принципы, которыми руководствуется организация во всей деятельности, связанной с ИИ;
- процессы обработки отклонений и исключений из политики.

Политика в области ИИ должна, где это необходимо, принимать во внимание специфические аспекты, предоставляя дополнительные рекомендации и/или ссылаясь на другие политики, в которых эти аспекты рассматриваются. Примерами таких аспектов являются:

- ресурсы и активы ИИ;
- оценка воздействия системы ИИ (см. 6.1.4);
- разработка системы ИИ.

Разработка, закупка, эксплуатация и использование систем ИИ должны регулироваться соответствующими политиками.

В.2.3 Согласование с другими политиками организации

Меры управления

Организация должна определить, какие другие политики могут быть затронуты целями организации в отношении систем ИИ или применяться к ним.

Руководство по внедрению

Деятельность в области ИИ пересекается с деятельностью в ряде других областей, таких как качество, безопасность, защита персональных данных. Организации следует провести тщательный анализ, чтобы определить,

пересекаются ли текущие политики и где именно, и либо актуализировать эти политики, если требуются обновления, либо включить соответствующие положения в политику в области ИИ.

Дополнительная информация

Политики, установленные руководящим органом от имени организации, должны лежать в основе политики в области ИИ. Рекомендации для руководящего органа по внедрению системы ИИ и управлению ею на протяжении всего ее жизненного цикла приведены в [27].

B.2.4 Пересмотр политики в области искусственного интеллекта

Меры управления

Политики в области ИИ должны пересматриваться через запланированные периоды времени или в случае происходящих существенных изменений для обеспечения уверенности в сохранении их приемлемости, адекватности и результативности.

Руководство по внедрению

Исполнитель указанной руководством роли должен нести ответственность за разработку, анализ, пересмотр и оценку политики в области ИИ или ее составных частей. Процесс анализа и пересмотра должен включать оценку возможностей для совершенствования политик организации и ее подхода к управлению системами ИИ в случае изменений в среде организации, деловых обстоятельств, правовых условий и технического окружения.

При проведении анализа политики в области ИИ следует учитывать результаты проверок со стороны руководства.

B.3 Организация внутренней деятельности

B.3.1 Общие положения

Цель

Установить подотчетность внутри организации для поддержания ее ответственного подхода к внедрению, эксплуатации и управлению системами ИИ.

B.3.2 Роли и обязанности в области искусственного интеллекта

Меры управления

Роли и обязанности в области ИИ должны быть определены и распределены в соответствии с потребностями организации.

Руководство по внедрению

Определение ролей и обязанностей имеет решающее значение для обеспечения подотчетности всей организации с точки зрения ее роли в отношении системы ИИ на протяжении всего ее жизненного цикла. Для обеспечения охвата всех соответствующих областей организация должна учитывать политики в области ИИ, цели ИИ и выявленные риски при распределении ролей и обязанностей. Организация может расставлять приоритеты в распределении ролей и обязанностей. Примерами областей, которые могут потребовать определенных ролей и обязанностей, могут являться:

- управление рисками;
- оценка воздействия системы ИИ;
- управление активами и ресурсами;
- защита;
- безопасность;
- конфиденциальность;
- развитие;
- результаты деятельности;
- человеческий надзор;
- взаимоотношения с поставщиками;
- демонстрация способности последовательно выполнять юридические обязательства;
- управление качеством данных (на протяжении всего жизненного цикла).

Ответственность в рамках выполнения различных ролей должна быть определена на уровне, подходящем соответствующим лицам для выполнения своих обязанностей.

B.3.3 Информирование о проблемах

Меры управления

Организация должна определить и внедрить процесс, позволяющий сотрудникам организации сообщать о проблемах, связанных с ролью организации в отношении системы ИИ на протяжении всего ее жизненного цикла.

Руководство по внедрению

Механизм отчетности должен:

a) включать подбор вариантов обеспечения конфиденциальности, анонимности или и того, и другого;

b) быть доступным лицам, работающим по найму и по контракту;

c) быть укомплектованным квалифицированными специалистами;

d) устанавливать соответствующие полномочия по расследованию и разрешению споров для лиц, указанных в c);

e) предусматривать механизмы своевременного представления отчетности и доведения ее до сведения руководства;

f) обеспечивать эффективную защиту от ответных действий как для лиц, связанных с отчетностью, так и с расследованием (например, позволяя выполнять отчетность анонимно и конфиденциально);

g) обеспечивать предоставление отчетов в соответствии с 4.4 и, при необходимости, e); сохраняя конфиденциальность и анонимность a) и соблюдая общие рекомендации деловой конфиденциальности;

h) в соответствующие сроки обеспечивать механизмы реагирования.

Примечание — В рамках данного процесса организация может использовать существующие механизмы отчетности.

Дополнительная информация

В дополнение к руководству по внедрению, приведенному в этом пункте, организации следует также рассмотреть [25].

B.4 Ресурсы систем искусственного интеллекта

B.4.1 Цель

Обеспечить учет организацией ресурсов (включая компоненты и активы) системы ИИ для полного понимания рисков и воздействий, а также устранения их последствий.

B.4.2 Документация ресурсов

Меры управления

Организация должна определить и задокументировать соответствующие ресурсы, необходимые для деятельности на данных стадиях жизненного цикла системы ИИ, а также связанные с ИИ виды деятельности, имеющие значение для организации.

Руководство по внедрению

Документирование ресурсов системы ИИ имеет решающее значение для понимания рисков, а также потенциальных воздействий системы ИИ (как положительных, так и отрицательных) на отдельных лиц, группы лиц или и тех, и других, а также социальные группы. Документирование таких ресурсов (которые используют, например, диаграммы потоков данных или схемы архитектуры системы) может служить основой для оценки воздействия системы ИИ (см. B.5).

Ресурсы могут включать в себя, но не ограничиваются следующим:

- компоненты системы ИИ;
- информационные ресурсы, т. е. данные, используемые на любой стадии жизненного цикла системы ИИ;
- инструментальные средства (например, алгоритмы, модели или инструменты ИИ);
- системные и вычислительные ресурсы (например, аппаратное обеспечение для разработки и запуска моделей ИИ, хранилище данных и инструментальные средства);
- человеческие ресурсы, то есть лица, обладающие необходимым опытом (например, для разработки, продаж, обучения, эксплуатации и технического обслуживания системы ИИ) в соответствии с ролью организации на протяжении всего жизненного цикла системы ИИ.

Ресурсы могут предоставляться самой организацией, ее клиентами или третьими лицами.

Дополнительная информация

Документация о ресурсах также может помочь определить наличие ресурсов, и, в случае их отсутствия, организации следует пересмотреть спецификацию проектирования системы ИИ или требований к ее развертыванию.

B.4.3 Ресурсы данных

Меры управления

В рамках идентификации ресурсов организация должна документировать информацию о ресурсах данных, используемых для системы ИИ.

Руководство по внедрению

Документация по данным отражает, но не ограничивается этим:

- происхождение данных;

- дату последнего обновления или модификации данных (например, тег даты в метаданных);
- категории данных — для машинного обучения (например, обучающие, валидационные, тестовые и производственные данные);
- категории данных (см. [12]);
- процесс маркировки данных;
- предполагаемое использование данных;
- качество данных (например, как описано в серии стандартов ИСО/МЭК 5259[1)]);
- применимые политики хранения и удаления данных;
- известные или потенциальные проблемы, связанные с появлением предвзятости в данных;
- подготовку данных.

B.4.4 Инструментальные ресурсы

Меры управления

В рамках идентификации ресурсов организация должна задокументировать информацию об инструментальных средствах, используемых для системы ИИ.

Руководство по внедрению

Инструментальные средства для системы ИИ и, в частности, для машинного обучения, могут включать в себя следующее, но не ограничиваются этим:

- типы алгоритмов и модели машинного обучения;
- инструменты или процессы обработки данных;
- методы оптимизации;
- методы оценивания;
- инструменты предоставления ресурсов;
- инструменты, способствующие разработке моделей;
- программное и аппаратное обеспечение для проектирования, разработки и развертывания систем ИИ;
- рекомендации по поводу различной семантики чисел с плавающей запятой в аппаратном обеспечении разработки и развертывания.

Дополнительная информация

Стандарт [13] содержит подробное руководство по типам, методам и подходам к различным инструментальным средствам для машинного обучения.

B.4.5 Система и вычислительные ресурсы

Меры управления

В рамках идентификации ресурсов организация должна документировать информацию о системе и вычислительных ресурсах, используемых для системы ИИ.

Руководство по внедрению

Информация о системе и вычислительных ресурсах для системы ИИ может содержать следующее, но не ограничивается этим:

- требования к ресурсам системы ИИ (для обеспечения возможности работы системы на устройствах с ограниченными ресурсами);
- данные о расположении системы и вычислительных ресурсов (например, локальные, облачные вычисления или периферийные вычисления);
- ресурсы обработки (включая сеть и хранилище);
- данные о влиянии аппаратного обеспечения, используемого для выполнения рабочих нагрузок системы ИИ (например, воздействие на окружающую среду в результате использования или производства аппаратного обеспечения или стоимость использования аппаратного обеспечения).

Необходимо учитывать, что для обеспечения постоянного улучшения систем ИИ могут потребоваться различные ресурсы. Разработка, развертывание и эксплуатация системы могут иметь различные системные потребности и требования.

Примечание — В ИСО/МЭК 22989 описываются различные аспекты использования системных ресурсов.

[1)] Данная серия включает в себя: ИСО/МЭК 5259-1:2024, ИСО/МЭК 5259-3:2024, ИСО/МЭК 5259-4:2024. В стадии разработки находятся: ISO/IEC FDIS 5259-2, ISO/IEC FDIS 5259-5.

B.4.6 Человеческие ресурсы

Меры управления

В рамках идентификации ресурсов организация должна документировать информацию о человеческих ресурсах и их компетенциях, используемых для разработки, развертывания, эксплуатации, управления изменениями, технического обслуживания, ввода в эксплуатацию и вывода из эксплуатации, а также проверки и интеграции системы ИИ.

Руководство по внедрению

Организация должна учитывать потребность в различных экспертных знаниях и определять типы ролей, необходимых для функционирования системы. Например, организация может включать определенные демографические группы, связанные с наборами данных, используемыми для подготовки моделей машинного обучения, если их включение является необходимым компонентом проектирования системы. К необходимым человеческим ресурсам относятся следующие, но не ограничиваются ими:

- специалисты по обработке данных;
- специалисты, связанные с человеческим надзором за системами ИИ;
- эксперты в области надежности, включающей безопасность, защиту и неприкосновенность частной жизни;
- исследователи и специалисты в области ИИ, а также эксперты в предметной области, имеющие отношение к системам ИИ.

На разных стадиях жизненного цикла системы ИИ могут потребоваться различные ресурсы.

B.5 Оценка воздействия систем искусственного интеллекта

B.5.1 Цель

Оценить воздействие системы ИИ на отдельных лиц или группы лиц, или и тех, и других, а также социальные группы, затронутые системой ИИ на протяжении всего ее жизненного цикла.

Б.5.2 Процесс оценки воздействия системы искусственного интеллекта

Меры управления

Организация должна разработать процесс оценки потенциальных последствий для отдельных лиц или групп лиц, или и тех, и других, а также социальных групп, которые могут возникнуть в результате внедрения системы ИИ на протяжении всего ее жизненного цикла.

Руководство по внедрению

Поскольку системы ИИ потенциально оказывают значительное влияние на отдельные лица, группы лиц, или и тех, и других, и социальные группы, организация, предоставляющая и использующая такие системы, должна, исходя из предполагаемого назначения и использования этих систем, оценить воздействие этих систем на эти группы.

Организация должна рассмотреть вопрос о том, влияет ли система ИИ:

- на правовое положение или жизненные возможности отдельных лиц;
- физическое или психологическое благополучие отдельных лиц;
- всеобщие права человека;
- социальные группы.

Организации следует включить следующие процедуры, но не ограничиваться ими:

a) обстоятельства, при которых следует проводить оценку воздействия системы ИИ, и которые могут включать в себя, но не ограничиваться следующим:
   1) критичность предполагаемой цели и среды, в которой используется система ИИ, или любые существенные изменения в них;
   2) сложность технологии ИИ и уровень автоматизации систем ИИ или какие-либо существенные изменения в этом;
   3) чувствительность типов данных и источников, обрабатываемых системой ИИ, или любые существенные изменения в них;

b) элементы, являющиеся частью процесса оценки воздействия системы ИИ, который может включать:
   1) идентификацию (например, источников, событий и результатов);
   2) анализ (например, последствий и вероятности);
   3) оценку (например, принятие решений и расстановку приоритетов);
   4) обработку (например, меры по смягчению последствий);
   5) документацию, отчетность и информирование (см. 7.4, 7.5 и B.3.3);

c) кем будет проводиться оценка воздействия системы ИИ;

d) каким образом можно использовать оценку воздействия системы ИИ (например, какое воздействие она может оказывать на проектирование или использование системы (см. B.6 и B.9), может ли она инициировать проверки и предоставлять разрешения);

е) отдельные лица и социальные группы, на которые потенциально воздействует система в зависимости от ее предполагаемого назначения, использования и характеристик (например, оценка отдельных лиц, групп лиц или социальных групп).

При оценке воздействия следует учитывать различные аспекты системы ИИ, включая данные, используемые для разработки системы ИИ, используемые технологии ИИ и функциональность всей системы.

Процессы могут варьироваться в зависимости от роли организации и области применения ИИ, а также в зависимости от конкретных дисциплин, для которых оценивается воздействие (например, защита, конфиденциальность и безопасность).

Дополнительная информация

Для некоторых дисциплин или организаций детальное рассмотрение воздействия на отдельных лиц или группы лиц, или и тех, и других, а также социальные группы является частью управления рисками, особенно в таких дисциплинах, как информационная безопасность, охрана труда и экологический менеджмент. Организация должна определить, в достаточной ли степени оценки воздействия конкретной дисциплины, выполняемые в рамках такого процесса управления рисками, учитывают соображения использования ИИ для этих конкретных аспектов (например, конфиденциальности).

Примечание — В [14] описывается, как организация может проводить анализ воздействия для самой организации, а также для отдельных лиц или групп лиц или и тех и других, а также социальных групп в рамках общего процесса управления рисками.

B.5.3 Документация по оценке воздействия системы искусственного интеллекта

Меры управления

Организация должна документировать результаты проведения оценок воздействия системы ИИ и сохранять результаты в течение определенного периода.

Руководство по внедрению

Сохранение документации имеет большое значение при определении информации, которая должна быть доведена до сведения пользователей и других заинтересованных сторон.

Результаты проведения оценок воздействия системы ИИ должны сохраняться и актуализироваться по мере необходимости в соответствии с элементами оценки воздействия системы ИИ, задокументированными в B.5.2. Сроки хранения документации могут определяться графиками хранения организации или регламентироваться юридическими или другими требованиями.

Тем не менее, существует ряд требований, соответствие которым в рамках системы менеджмента ИИ организация может демонстрировать посредством разработки ряда документов. Следует выделить некоторые из них, но не ограничиваться ими:

- предполагаемое использование системы ИИ и обоснованно прогнозируемое неправильное использование системы ИИ;
- положительное и отрицательное воздействие системы ИИ на соответствующих отдельных лиц или группы лиц, или и тех, и других, а также социальные группы;
- предсказуемые сбои, их потенциальные последствия и меры, принимаемые для их смягчения;
- соответствующие демографические группы, к которым применима система;
- сложность системы;
- роль людей во взаимоотношениях с системой, включая возможности человеческого надзора, процессы и инструменты, доступные для предотвращения негативных воздействий;
- трудоустройство и повышение компетентности персонала.

B.5.4 Оценка воздействия системы искусственного интеллекта на отдельных лиц и группы лиц

Меры управления

Организация должна оценивать и документировать потенциальное воздействие систем ИИ на отдельных лиц или группы лиц на протяжении всего жизненного цикла системы.

Руководство по внедрению

При проведении оценки воздействия на отдельных лиц или группы лиц, или и тех, и других, а также социальные группы, организация должна учитывать присущие ей принципы управления, политики и цели в области ИИ. Лица, использующие систему ИИ, или лица, чьи персональные данные обрабатываются системой ИИ, могут предъявлять требования к надежности системы ИИ. Следует принимать во внимание особую потребность в защите таких групп, как дети, инвалиды, пожилые люди и работники. Организация должна провести оценку ожиданий и рассмотреть средства их реализации в рамках оценки воздействия на систему.

В зависимости от назначения и области применения системы ИИ в рамках оценки следует учитывать следующие области воздействия, но не ограничиваться ими:

- справедливость;

- подотчетность;
- прозрачность и объяснимость;
- защита и конфиденциальность;
- безопасность и гигиена труда;
- финансовые последствия;
- доступность;
- права человека.

Дополнительная информация

При необходимости организация должна консультироваться с экспертами (например, с исследователями, экспертами в предметной области и пользователями), чтобы, насколько это возможно, получить полное представление о потенциальном воздействии системы ИИ на отдельных лиц или группы лиц, или и тех, и других, а также социальные группы.

B.5.5 Оценка воздействия систем искусственного интеллекта на общество

Меры управления

Организация должна оценивать и документировать потенциальное социальное воздействие систем ИИ на общество на протяжении всего их жизненного цикла.

Руководство по внедрению

Воздействие на общество может сильно варьироваться в зависимости от среды организации и типов систем ИИ. Воздействие систем ИИ на общество может быть как позитивным, так и негативным. Примерами таких потенциальных социальных воздействий могут служить:

- экологическая устойчивость (включая воздействие на природные ресурсы и выбросы парниковых газов);
- экономика (включая доступ к финансовым услугам, возможности трудоустройства, налоги, торговлю и коммерцию);
- правительство (включая законодательные процессы, дезинформацию в политических целях, системы национальной безопасности и уголовного правосудия);
- здоровье и безопасность (включая доступ к медицинскому обслуживанию, медицинской диагностике и лечению, а также потенциальный физический и психологический вред);
- нормы, традиции, культура и ценности (включая дезинформацию, которая приводит к предубеждениям или причиняет вред отдельным лицам или группам лиц, или и тем и другим, а также социальным группам).

Дополнительная информация

Разработка и использование систем ИИ может потребовать значительных вычислительных ресурсов, что может оказать соответствующее воздействие на экологическую устойчивость (например, выбросы парниковых газов из-за увеличения энергопотребления, воздействие на воду, землю, флору и фауну). Аналогичным образом, системы ИИ могут использоваться для повышения экологической устойчивости других систем (например, сокращения выбросов парниковых газов, связанных со строительством и транспортировкой). Организация должна учитывать воздействие систем ИИ в контексте своих общих целей и стратегий в области экологической устойчивости.

Организации следует рассмотреть вопрос о возможности использования ее систем ИИ не по назначению для причинения вреда обществу и о возможности их использования для устранения ранее нанесенного ущерба. Например, могут ли системы ИИ препятствовать доступу к финансовым услугам, таким как кредиты, гранты, страхование и инвестиции и могут ли системы ИИ улучшить доступ к этим инструментам?

Системы ИИ использовались для оказания влияния на результаты выборов и создания дезинформации (например, дипфейки цифровых медиа), которые могут привести к политическим и социальным волнениям. При применении правительством систем ИИ в целях уголовного правосудия был выявлен риск алгоритмической предвзятости по отношению к социальным группам, отдельным лицам или группе лиц, обусловленный использованием ИИ. Следует также рассмотреть вопрос о том, каким образом злоумышленники могут злоупотреблять системами ИИ и каким образом системы ИИ могут усиливать нежелательные исторически сложившиеся социальные предубеждения.

Системы ИИ могут использоваться для диагностики и лечения заболеваний, а также для оценки критериев определения тех граждан, которые имеют право на медицинские льготы. Системы ИИ также развертываются в сценариях, где сбои в работе (машин) могут привести к летальному исходу или травмам людей (например, в случае с беспилотными автомобилями, взаимодействия человека и машины). Организация должна учитывать как положительные, так и отрицательные результаты при использовании систем ИИ в сценариях, связанных со здоровьем и безопасностью людей.

Примечание — В [17] содержится высокоуровневый обзор этических и социальных проблем, связанных с системами и приложениями ИИ.

B.6 Жизненный цикл системы искусственного интеллекта

B.6.1 Руководство по управлению разработкой системы искусственного интеллекта

B.6.1.1 Цель

Обеспечить определение и документирование целей, а также внедрение процессов ответственного проектирования и разработки надежных систем ИИ.

B.6.1.2 Цели ответственной разработки системы искусственного интеллекта

Меры управления

Организация должна определить и задокументировать цели, которыми следует руководствоваться при ответственной разработке систем ИИ, а также учитывать эти цели и интегрировать меры по их достижению в жизненный цикл разработки.

Руководство по внедрению

Организация должна определить цели (см. 6.2), оказывающие влияние на процессы проектирования и разработки системы ИИ, и учитывать их. Например, если организация определяет «справедливость» как одну из целей, это должно быть включено в спецификацию требований при сборе данных, их обработке, обучении модели, верификации, валидации и т. д. Организация должна предоставить требования и руководящие принципы, необходимые для обеспечения интеграции мер на различных этапах (например, требование об использовании конкретного инструмента или метода тестирования для устранения несправедливости или нежелательной предвзятости) для достижения таких целей.

Дополнительная информация

Методы ИИ используются для усиления мер безопасности, таких как обнаружение и прогнозирование угроз или предотвращение атак на систему безопасности. Имеется в виду применение методов ИИ, которые можно использовать для усиления мер безопасности для защиты как систем ИИ, так и обычных программных систем, не основанных на ИИ. В приложении С приведены примеры организационных целей по управлению рисками, которые могут быть полезны при определении целей разработки системы ИИ.

B.6.1.3 Процессы ответственного проектирования и разработки систем искусственного интеллекта

Меры управления

Организация должна определить и задокументировать конкретные процессы ответственного проектирования и разработки системы ИИ.

Руководство по внедрению

Ответственная разработка системных процессов ИИ включает в себя рассмотрение, помимо прочего, следующих вопросов:

- стадии жизненного цикла (общая модель жизненного цикла системы ИИ представлена в ИСО/МЭК 22989, однако организация может указать свои собственные стадии жизненного цикла);
- требования к тестированию и предполагаемые ресурсы тестирования;
- требования человеческого надзора, включая процессы и инструменты, особенно в случае, если система ИИ может воздействовать на физические лица;
- на каких этапах следует проводить оценку воздействия системы ИИ;
- требования и правила к данным для обучения (например, какие данные можно использовать, утвержденные поставщики данных и маркировка);
- требуемый опыт (в предметной или другой области) или обучение для разработчиков систем ИИ или и то, и другое;
- критерии выпуска;
- согласования и подписи, необходимые на различных этапах;
- управление изменениями;
- удобство использования и управляемость;
- вовлечение заинтересованных сторон.

Конкретные процессы проектирования и разработки зависят от функциональности и технологий ИИ, которые предполагается использовать в системе ИИ.

B.6.2 Жизненный цикл системы искусственного интеллекта

B.6.2.1 Цель

Определить критерии и требования для каждой стадии жизненного цикла системы ИИ.

B.6.2.2 Требования и спецификация к системе искусственного интеллекта

Меры управления

Организация должна определить и задокументировать требования к новым системам ИИ или существенным усовершенствованиям существующих систем.

Руководство по внедрению

Организация должна задокументировать обоснование разработки системы ИИ и ее цели. К вопросам, которые следует учитывать, документировать и понимать, можно отнести следующие:

a) чем обусловлена разработка системы ИИ (например, экономическое обоснование, запрос клиента или политика правительства);

b) каким образом можно обучить модель и добиться удовлетворения требованиям к данным.

Требования к системе ИИ должны быть определены и охватывать весь ее жизненный цикл. Такие требования следует пересматривать в случаях, когда разработанная система ИИ не функционирует должным образом или появляется новая информация, которая может быть использована для изменения и улучшения требований. Например, разработка системы ИИ может стать невыполнимой с финансовой точки зрения.

Дополнительная информация

Процессы для описания жизненного цикла системы ИИ предусмотрены в [10]. Для получения дополнительной информации об ориентированном на человека проектировании интерактивных систем см. [3].

B.6.2.3 Документация по проектированию и разработке системы искусственного интеллекта

Меры управления

Организация должна документировать проектирование и разработку системы ИИ на основе целей организации, документированных требований и критериев спецификации.

Руководство по внедрению

Варианты проектирования, необходимые для функционирования системы ИИ, включают, но не ограничиваются следующими:

- подход к машинному обучению (например, контролируемый или неконтролируемый);
- алгоритм обучения и тип используемой модели машинного обучения;
- методы обучения модели и качество данных (см. B.7);
- проведение оценки и улучшение моделей;
- аппаратные и программные компоненты;
- угрозы безопасности, зафиксированные на протяжении всего жизненного цикла системы ИИ; угрозы безопасности, характерные для систем ИИ, включающие отравление данных, кражу моделей или атаки с инверсией моделей;
- интерфейс и представление выходных данных;
- способы взаимодействия людей и системы;
- вопросы функциональной совместимости и переносимости.

Между проектированием и разработкой может быть несколько итераций, при этом на данном этапе документация должна поддерживаться в актуальном состоянии, и должна быть доступна конечная документация по архитектуре системы.

Дополнительная информация

Для получения дополнительной информации об ориентированном на пользователя проектировании интерактивных систем см. [3].

B.6.2.4 Верификация и валидация системы искусственного интеллекта

Меры управления

Организация должна определить и задокументировать методы верификации и валидации для системы ИИ и указать критерии для их использования.

Руководство по внедрению

Методы верификации и валидации могут включать следующее, но не ограничиваться этим:

- методологии и инструменты тестирования;
- выбор тестовых данных и их репрезентативность в отношении предполагаемой области использования;
- требования к критериям выпуска.

Организация должна определить и задокументировать следующие критерии оценки, но не ограничиваться ими:

- этапы проведения оценки компонентов системы ИИ и всей системы ИИ в целом на предмет рисков, связанных с воздействием на отдельных лиц или группы лиц, или и тех, и других, а также социальные группы;
- этапы оценки могут включать следующие критерии:
  1) требования к надежности и безопасности системы ИИ, включая допустимую частоту ошибок для показателей деятельности/производительности системы ИИ;

2) ответственные цели разработки и использования систем ИИ, подобные указанным в В.6.1.2 и В.9.3;
3) эксплуатационные факторы, такие как качество данных, предполагаемое использование, включая допустимые диапазоны каждого эксплуатационного фактора;
4) любые предполагаемые виды применения, которые могут потребовать определения более строгих эксплуатационных факторов, включая различные допустимые диапазоны эксплуатационных факторов или более низкую частоту ошибок;

- методы, рекомендации или показатели, используемые для оценки того, могут ли заинтересованные стороны, которые принимают решения или подпадают под действие решений, основанных на результатах работы системы ИИ, адекватно интерпретировать результаты работы системы ИИ. Периодичность проведения оценки должна быть определена и может зависеть от результатов оценки воздействия системы ИИ;

- любые приемлемые факторы, которые могут объяснить неспособность достичь целевого минимального уровня производительности, особенно когда система ИИ оценивается на предмет воздействия на отдельных лиц или группы лиц, или и тех, и других, или социальные группы (например, низкое разрешение изображения для систем компьютерного зрения или фоновый шум, влияющий на системы распознавания речи). Мероприятия по борьбе с низкой производительностью системы ИИ, обусловленной вышеперечисленными факторами, также должны быть задокументированы.

Систему ИИ следует оценивать в соответствии с задокументированными критериями оценки.

В случаях, когда система ИИ не может соответствовать задокументированным критериям оценки, особенно в отношении целей ответственной разработки и использования системы ИИ (см. В.6.1.2 и В.9.3), организация должна пересмотреть или устранить недостатки предполагаемого использования системы ИИ, свои требования к производительности и то, как организация может эффективно реагировать на воздействие на отдельных лиц или группы лиц, или и тех, и других, а также социальные группы.

Примечание — Дополнительная информация о том, как обеспечить надежность нейронных сетей, представлена в [16].

В.6.2.5 Развертывание системы искусственного интеллекта

Меры управления

Организация должна задокументировать план развертывания и обеспечить выполнение соответствующих требований до инициирования процесса развертывания.

Руководство по внедрению

Системы ИИ могут разрабатываться в одних средах и развертываться в других (например, разрабатываться локально и развертываться с использованием облачных вычислений), и организация должна учитывать эти различия при разработке плана развертывания. Также следует рассмотреть вопрос о том, развертываются ли компоненты отдельно (например, программное обеспечение и модель могут быть разработаны независимо друг от друга). Кроме того, организация должна установить набор требований, которые должны быть выполнены до выпуска и развертывания (иногда называемых «критериями выпуска»). Критерии выпуска могут включать в себя: принятые методы верификации и валидации, выполненные показатели производительности, пройденное пользовательское тестирование, а также полученные согласования руководства и подписи. План развертывания должен учитывать перспективы соответствующих заинтересованных сторон и воздействие на них.

В.6.2.6 Эксплуатация и мониторинг системы искусственного интеллекта

Меры управления

Организация должна определить и задокументировать необходимые элементы для непрерывной работы системы ИИ. Как минимум, это должно включать мониторинг системы и производительности, ремонт, обновления и поддержку.

Руководство по внедрению

Любое действие для эксплуатации и мониторинга может учитывать, например, следующие аспекты:

- мониторинг системы и производительности может включать мониторинг общих ошибок и сбоев, а также проверку того, работает ли система с производственными данными должным образом. Технические критерии эффективности могут включать показатели успешности в решении проблем и в выполнении задач, а также уровни доверия. Другие критерии могут быть связаны с выполнением обязательств или ожиданий и потребностей заинтересованных сторон, включая, например, постоянный мониторинг для обеспечения соответствия требованиям заказчика или применимым законодательным требованиям;

- некоторые развернутые системы ИИ повышают свою эффективность в результате машинного обучения, при применении которого производственные и выходные данные используются для дальнейшего обучения модели МО. При применении непрерывного обучения организации следует осуществлять контроль производительности системы ИИ для обеспечения гарантии, что она продолжает соответствовать целям проектирования и оперирует производственными данными по назначению;

- производительность некоторых систем ИИ может измениться, даже если такие системы не используют непрерывное обучение. Как правило это происходит из-за концепции или смещения данных в производственных данных. В таких случаях мониторинг может выявить необходимость в переобучении для обеспечения гарантии, что система ИИ продолжает соответствовать целям проектирования и оперирует производственными данными по назначению (см. в [13]);

- ремонт может включать в себя устранение ошибок и сбоев в системе. Организации следует внедрить процессы реагирования на эти факторы и их устранения. Кроме того, обновления могут быть необходимы по мере развития системы, выявления меньшего количества критических факторов или в результате внешних выявленных факторов (например, несоответствие ожиданиям клиентов или юридическим требованиям). Необходимо внедрить процессы обновления системы, включая затронутые обновлением компоненты, график обновления, информацию для пользователей о том, что подлежит обновлению;

- системные обновления также могут включать изменения в работе системы, новые или модифицированные виды использования по назначению или другие изменения в функциональности системы. В организации должны быть внедрены процедуры реагирования на операционные изменения, включая информирование пользователей;

- поддержка системы может быть внутренней, внешней или той и другой в зависимости от потребностей организации и способа приобретения системы. Процессы поддержки должны учитывать то, каким образом осуществляется обращение пользователей за соответствующей помощью, сообщается о проблемах и инцидентах, а также поддерживаются соглашения об уровне обслуживания и показателях;

- если системы ИИ используются для целей, отличных от тех, для которых они были разработаны, или способами, которые не предполагались, следует рассмотреть целесообразность такого использования;

- организация, применяющая или разрабатывающая системы ИИ, должна выявить характерные для них угрозы информационной безопасности. Угрозы информационной безопасности, характерные для ИИ, включают, но не ограничиваются следующими: отравление данных, кража моделей и атаки с инверсией моделей.

Дополнительная информация

Организации следует учитывать эксплуатационные характеристики, которые могут повлиять на заинтересованные стороны, и учитывать это при разработке и определении критериев эффективности.

Критерии эффективности для действующих систем ИИ должны определяться рассматриваемой задачей, такой как классификация, регрессия, ранжирование, кластеризация или уменьшение размерности.

Критерии эффективности могут включать статистические аспекты, такие как частота ошибок и продолжительность обработки. Для каждого критерия организация должна определить все соответствующие показатели, а также взаимозависимость между показателями. Для каждого показателя организация должна рассмотреть приемлемые значения, основанные, например, на рекомендациях эксперта в предметной области и анализе ожиданий заинтересованных сторон относительно существующих практик, не связанных с ИИ.

Например, организация может определить, что оценка является подходящим показателем эффективности, основываясь на оценке влияния ложноположительных и ложноотрицательных результатов, как описано в [8]. Затем организация может установить значение F_1, которому, как ожидается, будет соответствовать система ИИ. Следует оценить возможность решения этих проблем с помощью существующих мер. В противном случае следует рассмотреть возможность внесения изменений в существующие меры или определить дополнительные меры для выявления этих проблем и их устранения.

Организация должна учитывать эффективность действующих систем или процессов, не связанных с ИИ, и использовать их в качестве потенциально релевантного контекста при установлении критериев эффективности.

Организация должна дополнительно обеспечить, чтобы средства и процесс, используемые для оценки системы ИИ, включая, где применимо, отбор оценочных данных и управление ими, повышали полноту и надежность оценки ее эффективности в соответствии с определенными критериями.

Разработка методологий оценки эффективности может основываться на критериях, показателях и ценностях/значениях. Описанные критерии должны отражать объем данных и типы процессов, используемых при оценке, а также роли и опыт персонала, проводящего оценку.

Методологии оценки эффективности должны максимально точно отражать атрибуты и характеристики функционирования и использования для обеспечения полезности и актуальности результатов оценки. Некоторые аспекты оценки эффективности могут потребовать контролируемого введения ошибочных или ложных данных или процессов для оценки влияния на эффективность.

Для определения критериев эффективности может быть использована модель качества по [19].

B.6.2.7 Техническая документация по системе искусственного интеллекта

Меры управления

Организация должна определить, какая техническая документация, касающаяся системы ИИ, необходима для каждой соответствующей категории заинтересованных сторон, таких как пользователи, партнеры, контролирующие органы, и предоставить им техническую документацию в соответствующей форме.

Руководство по внедрению

Техническая документация системы ИИ может включать следующие элементы, но не ограничиваться ими:

- общее описание системы ИИ, включая ее предполагаемое назначение;
- инструкции по эксплуатации;
- технические допущения о развертывании и эксплуатации системы ИИ (среда выполнения, соответствующие программные и аппаратные возможности, предположения, сделанные на основе данных, и т. д.);
- технические ограничения (например, допустимая частота ошибок, точность, надежность, робастность);
- возможности мониторинга и функции, позволяющие пользователям или операторам влиять на работу системы.

Документация, относящаяся ко всем стадиям жизненного цикла системы ИИ (как определено в ИСО/МЭК 22989), может включать, но не ограничиваться следующими элементами:

- спецификация проектирования и архитектуры системы;
- проектирование и меры по обеспечению качества, принятые в процессе разработки системы;
- информация о данных, используемых при разработке системы;
- сделанные допущения и принятые меры по обеспечению качества данных (например, предполагаемые статистические распределения);
- управленческие действия (например, управление рисками), осуществляемые в ходе разработки или эксплуатации системы ИИ;
- записи о верификации и валидации;
- записи об изменениях, вносимых в систему ИИ во время ее эксплуатации;
- документация по оценке воздействия в соответствии с B.5.

Организация должна документировать техническую информацию, связанную с ответственной эксплуатацией системы ИИ. Это может включать следующее, но не ограничиваться этим:

- документирование плана по управлению сбоями, в т. ч. необходимость описания плана отката для системы ИИ, отключения функций системы ИИ, процесса обновления или плана уведомления клиентов, пользователей и т. д. об изменениях в системе ИИ, актуализированной информации о системных сбоях и способах их устранения;
- документирование процессов мониторинга работоспособности системы ИИ (т. е. использование системы ИИ по назначению и в пределах ее нормальных эксплуатационных возможностей, также называемое наблюдаемостью) и процессов устранения сбоев в системе ИИ;
- документирование стандартных операционных процедур для системы ИИ, включая то, какие события необходимо отслеживать и каким образом журналы событий расставляются по приоритетам и просматриваются; это также может включать в себя способы анализа и предотвращения сбоев;
- документирование ролей персонала, ответственного за работу системы ИИ, а также ролей лиц, ответственных за подотчетность использования системы, особенно в отношении устранения последствий сбоев системы ИИ или управления обновлениями системы ИИ;
- документирование обновлений системы, которое также может включать изменения в работе системы, новые или измененные виды использования по назначению или другие изменения в функциональности системы.

В организации должны быть внедрены процедуры для реагирования на операционные изменения, включая информирование пользователей и проведение внутренней оценки типа изменений.

Документация должна быть актуальной и точной. Документация должна быть одобрена соответствующим руководством организации.

При предоставлении пользовательской документации следует принимать во внимание меры управления, приведенные в таблице A.1.

B.6.2.8 Ведение журналов событий системой искусственного интеллекта

Меры управления

Организации следует определить, на каких стадиях жизненного цикла системы ИИ следует включить ведение журналов событий. Как минимум, ведение журналов необходимо при непосредственном использовании системы ИИ.

Руководство по внедрению

Организация должна обеспечить ведение журналов для систем ИИ, которые она развертывает, для автоматического сбора и записи журналов событий, связанных с определенными событиями, происходящими во время работы. Ведение журнала для систем ИИ может включать следующее, но не ограничиваться этим:

- отслеживание функциональности системы ИИ для обеспечения надлежащей/правильной работы системы ИИ;
- обнаружение функционирования системы ИИ за пределами предполагаемых условий ее эксплуатации, что может привести к нежелательному функционированию на производственных данных или к воздействиям на соответствующие заинтересованные стороны посредством мониторинга работы системы ИИ.

Журналы событий системы ИИ могут включать в себя такую информацию, как время и дата каждого использования системы ИИ, производственные данные, с которыми она работает; выходные данные, которые выходят за рамки предполагаемой работы системы ИИ, и т. д.

Журналы событий должны храниться до тех пор, пока это требуется для предполагаемого использования системы ИИ и в соответствии с политиками хранения данных организации. Также могут применяться юридические требования, связанные с хранением данных.

Дополнительная информация

В зависимости от законодательства некоторые системы ИИ, такие как системы биометрической идентификации, могут предъявлять дополнительные требования к ведению журнала, и организации должны быть осведомлены об этих требованиях.

B.7 Данные для систем искусственного интеллекта

B.7.1 Цель

Обеспечить понимание организацией роли и влияния данных в системах ИИ при применении и разработке, предоставлении или использовании систем ИИ на протяжении их жизненного цикла.

B.7.2 Данные для разработки и усовершенствования системы искусственного интеллекта

Меры управления

Организация должна определять, документировать и внедрять процессы управления данными, связанные с разработкой систем ИИ.

Руководство по внедрению

Управление данными может включать нижеследующие темы, но не ограничиваться ими:

- последствия для конфиденциальности и защиты в связи с использованием данных, некоторые из которых могут носить конфиденциальный характер;
- угрозы безопасности, которые могут возникнуть в результате разработки системы ИИ, зависящей от данных;
- аспекты прозрачности и объяснимости, включая происхождение данных и возможность предоставить объяснение того, как данные используются для определения выходных данных системы ИИ, если система требует прозрачности и объяснимости;
- репрезентативность обучающих данных по сравнению с рабочей областью использования;
- точность и целостность данных.

Примечание — Подробная информация о жизненном цикле системы ИИ и концепциях управления данными приведена в ИСО/МЭК 22989.

B.7.3 Сбор данных

Меры управления

Организация должна определить и задокументировать подробную информацию о сборе и отборе данных, используемых в системах ИИ.

Руководство по внедрению

Организации могут потребоваться различные категории данных из разных источников в зависимости от области применения их систем ИИ.

Детали сбора данных могут включать:

- категории данных, необходимых для функционирования системы ИИ;
- количество необходимых данных;
- источники данных (например, внутренние, приобретенные, совместно используемые, открытые, синтетические данные);
- характеристики источника данных (например, статические, потоковые, собранные, сгенерированные автоматически данные);
- демографические данные и характеристики субъекта данных (например, известные или потенциальные случаи возникновения предвзятости или другие систематические ошибки);
- предварительная обработка данных (например, информация о предыдущем использовании, соответствие требованиям конфиденциальности и безопасности);
- права на базы данных (например, личная идентифицирующая информация, авторские права);
- связанные метаданные (например, подробности маркировки и расширения данных);
- происхождение данных.

Дополнительная информация

Категории данных и структура использования данных, представленные в [12], могут быть использованы для документирования подробных сведений о сборе и использовании данных.

B.7.4 Качество данных для систем искусственного интеллекта

Меры управления

Организация должна определить и задокументировать требования к качеству данных и обеспечить соответствие данных, используемых для разработки и эксплуатации системы ИИ, этим требованиям.

Руководство по внедрению

Качество данных, используемых для разработки и эксплуатации систем ИИ, потенциально оказывает значительное влияние на достоверность результатов работы системы. В [18] качество данных определяется как степень, в которой характеристики данных удовлетворяют заявленным и подразумеваемым потребностям при использовании в определенных условиях. Для систем ИИ, в которых используется контролируемое или полуконтролируемое машинное обучение, важно, чтобы качество обучающих, валидационных, тестовых и производственных данных было определено, измерено и улучшено, насколько это возможно. Организация должна гарантировать, что данные соответствуют своему прямому назначению. Организации также следует учитывать влияние предвзятости на производительность и справедливость системы и вносить необходимые коррективы в модель и данные, используемые для повышения производительности и справедливости для того, чтобы сделать их приемлемыми для каждого конкретного варианта использования.

Дополнительная информация

Дополнительная информация о качестве данных представлена в серии стандартов [9], посвященных качеству данных для аналитики и машинному обучению. Информация о различных формах искажения данных, используемых в системах ИИ, представлена в [15].

B.7.5 Происхождение данных

Меры управления

Организация должна определить и задокументировать процесс записи происхождения данных, используемых в ее системах ИИ, на протяжении жизненного цикла данных и системы ИИ.

Руководство по внедрению

Согласно [1], запись о происхождении данных может включать информацию о создании, обновлении, транскрипции, абстрагировании, валидации и передаче управления данными. Кроме того, обмен данными (без передачи управления) и преобразование данных могут рассматриваться как происхождение данных. В зависимости от таких факторов, как источник данных, их содержание и контекст их использования, организациям следует рассмотреть вопрос о необходимости принятия мер по проверке происхождения данных.

B.7.6 Подготовка данных

Меры управления

Организация должна определить и задокументировать свои критерии для отбора методов подготовки данных и сами методы подготовки данных, которые будут использоваться.

Руководство по внедрению

Данные, используемые в системе ИИ, обычно требуют подготовки для применения в конкретной задаче ИИ. Например, алгоритмы машинного обучения иногда проявляют нетерпимость к отсутствующим или неправильным записям, ненормальному распределению и широко варьирующимся масштабам. Для повышения качества данных можно использовать методы подготовки и преобразования. Неспособность должным образом подготовить данные потенциально может привести к ошибкам системы ИИ. Распространенные методы подготовки и преобразования данных, используемые в системах ИИ, включают:

- статистическое исследование данных (например, распределение, среднее значение, медиана, стандартное отклонение, диапазон, стратификация, выборка) и спецификация статистических метаданных (например, инициатива по документированию данных [28]);
- очистка данных (т. е. исправление записей, устранение отсутствующих записей);
- вменение/условное исчисление (т. е. методы заполнения недостающих записей);
- нормализация;
- масштабирование;
- обозначение целевых переменных;
- кодирование (например, преобразование категориальных переменных в числа).

Для выполнения конкретной задачи ИИ организация должна задокументировать свои критерии для отбора конкретных методов подготовки данных и преобразований, а также сами конкретные методы подготовки данных и преобразования, которые будут использоваться.

Примечание — Дополнительная информация о подготовке данных, специфичных для машинного обучения, приведена в [9] и [13].

B.8 Информация для заинтересованных сторон

B.8.1 Цель

Обеспечить предоставление соответствующим заинтересованным сторонам необходимой информации для понимания и оценки рисков и их последствий (как положительных, так и отрицательных).

B.8.2 Системная документация и информация для пользователей

Меры управления

Организация должна определить и предоставить необходимую информацию пользователям системы.

Руководство по внедрению

В зависимости от контекста информация о системе ИИ может включать в себя как технические детали и инструкции, так и общие уведомления пользователей о том, что они взаимодействуют с системой ИИ. К этому можно отнести саму систему, а также потенциальные результаты работы системы (например, уведомление пользователей о том, что изображение создано с помощью ИИ).

Хотя системы ИИ могут быть сложными, крайне важно, чтобы пользователи при взаимодействии с системой ИИ могли понимать принципы ее работы. Пользователям также необходимо понимать ее предполагаемое назначение и предполагаемое использование, ее потенциальную возможность нанести ущерб или принести пользу пользователю. Определенная системная документация может быть предназначена для более технического использования (например, системными администраторами), и организация должна понимать потребности различных заинтересованных сторон и то, что каждая из заинтересованных сторон имеет свой определенный уровень понимания информации. Информация также должна быть доступной как с точки зрения простоты ее поиска, так и для пользователей, которым могут потребоваться дополнительные меры по обеспечению доступности.

Информация, которая может быть предоставлена пользователям, включает в себя помимо прочего:

- назначение системы;
- информацию о том, что пользователь взаимодействует с системой ИИ;
- способы взаимодействия с системой;
- сроки и способы переопределения системы;
- технические требования к работе системы, включая необходимые вычислительные ресурсы и ограничения системы, а также ожидаемый срок ее службы;
- потребность в человеческом надзоре;
- информацию о точности и производительности;
- соответствующую информацию, полученную в результате проведения оценки воздействия, включая потенциальные преимущества и вред, особенно если они применимы в конкретных средах или определенных демографических группах (см. B.5.2 и B.5.4);
- пересмотр утверждений о преимуществах системы;
- обновления и изменения в методах работы системы, а также любые необходимые меры по техническому обслуживанию, включая частоту их применения;
- контактную информацию;
- учебные материалы для эксплуатации системы.

Критерии, используемые организацией для определения того, должна ли она предоставлять информацию и какую именно, должны быть задокументированы. Соответствующие критерии включают, но не ограничиваются нижеследующими: предполагаемое использование и обоснованно прогнозируемое неправильное использование системы ИИ, опыт пользователя и конкретное воздействие системы ИИ.

Информация может предоставляться пользователям различными способами, включая документированные инструкции по эксплуатации, оповещения и другие уведомления, встроенные в саму систему, информацию на вебстранице и т. д. В зависимости от того, какие методы использует организация для предоставления информации, она должна подтвердить, что пользователи имеют доступ к этой информации и что предоставленная информация является полной, актуальной и точной.

B.8.3 Внешняя отчетность

Меры управления

Организация должна предоставить возможность заинтересованным сторонам сообщать о неблагоприятных воздействиях системы.

Руководство по внедрению

Организации следует отслеживать работу системы на предмет сообщений о выявленных проблемах и сбоях, а также предоставлять пользователям или другим внешним сторонам возможность сообщать о неблагоприятных воздействиях (например, о несправедливости).

B.8.4 Информирование об инцидентах

Меры управления

Организация должна определить и задокументировать план информирования пользователей системы об инцидентах.

Руководство по внедрению

Инциденты, связанные с системой ИИ, могут быть специфичными для самой системы ИИ или связаны с информационной безопасностью или конфиденциальностью (например, утечка данных). Организация должна понимать свои обязательства по уведомлению пользователей и других заинтересованных сторон об инцидентах в зависимости от среды, в которой работает система. Например, к инциденту с компонентом ИИ, который является частью продукта и влияет на безопасность, могут предъявляться иные требования к уведомлению, чем к системам других типов. Могут применяться юридические требования (такие как контракты) и регулирующая деятельность, которые могут устанавливать требования в отношении:

- типов инцидентов, о которых необходимо сообщать;
- сроков уведомления;
- необходимость уведомления соответствующих органов и каких именно;
- деталей, которые необходимо сообщить.

Организация может интегрировать мероприятия по реагированию на инциденты и отчетности для ИИ в свою более широкую деятельность по управлению инцидентами. При этом организации следует учитывать уникальные требования, связанные с системами ИИ или отдельными компонентами систем ИИ (например, могут быть различные требования к отчетности, связанные с утечкой персональных данных и конфиденциальностью данных в обучающих данных для системы).

Дополнительная информация

В [22] и [21] представлены дополнительные сведения об управлении инцидентами в целях обеспечения безопасности и конфиденциальности соответственно.

B.8.5 Информация для заинтересованных сторон

Меры управления

Организация должна определить и задокументировать свои обязательства по предоставлению информации о системе ИИ заинтересованным сторонам.

Руководство по внедрению

В некоторых случаях юрисдикция может потребовать предоставления информации о системе регулирующим органам. Информация может быть доведена до сведения заинтересованных сторон, таких как клиенты или регулирующие органы, в соответствующие сроки. Сюда могут относиться, например:

- техническая документация по системе, включая, но не ограничиваясь этим, наборы данных для обучения, валидации и тестирования, а также обоснования выбора алгоритмов и записи о верификации и валидации;
- риски, связанные с использованием системы ИИ;
- результаты оценок воздействия системы ИИ;
- журналы событий и другие системные записи.

Организация должна понимать свои обязательства в этом отношении и обеспечивать передачу надлежащей информации уполномоченным органам власти. Кроме того, предполагается, что организация осведомлена о требованиях юрисдикции в отношении информации, передаваемой правоохранительным органам.

B.9 Использование систем искусственного интеллекта

B.9.1 Цель

Обеспечение организацией ответственного использования систем ИИ и в соответствии с политиками организации.

B.9.2 Процессы ответственного использования искусственного интеллекта

Меры управления

Организация должна определить и задокументировать процессы ответственного использования систем ИИ.

Руководство по внедрению

В зависимости от среды организации может быть множество рекомендаций для определения того, какую систему ИИ лучше использовать. Независимо от того, кто является разработчиком системы ИИ — организация или третья сторона, в задачи организации входит проанализировать эти рекомендации и разработать политики для их учета, например:

- требуемые утверждения;
- затраты (включая расходы на текущий мониторинг и техническое обслуживание);
- утвержденные требования к поставщикам;

- юридические требования, применимые к организации.

Если организация приняла политики использования других систем, активов и т. д., при желании могут быть использованы и эти политики.

B.9.3 Цели ответственного использования системы искусственного интеллекта

Меры управления

Организация должна определить и задокументировать цели, которыми следует руководствоваться при ответственном использовании систем ИИ.

Руководство по внедрению

Организация, действующая в разных средах, может иметь разные ожидания и цели в отношении того, что представляет собой ответственное развитие систем ИИ. В соответствии со средой, организации следует определить свои цели в отношении надежного использования. Некоторые цели включают в себя:

- справедливость;
- подотчетность;
- прозрачность;
- объяснимость;
- надежность;
- безопасность;
- робастность и избыточность;
- конфиденциальность и защиту;
- доступность.

После определения своих целей организация должна внедрить механизмы для их достижения внутри организации. Это может включать определение того, соответствует ли стороннее решение/решение сторонних разработчиков/производителей целям организации или применимо ли решение, разработанное внутри организации, для предполагаемого использования. Организация должна определить, на каких стадиях жизненного цикла системы ИИ следует внедрять значимые цели человеческого надзора. К этому можно отнести:

- привлечение рецензентов для проверки выходных данных системы ИИ, в том числе наделение полномочиями отменять решения, принимаемые системами ИИ;
- обеспечение внедрения человеческого надзора, если это требуется для осуществления допустимого использования системы ИИ в соответствии с инструкциями или другой документацией, связанной с предполагаемым развертыванием системы ИИ;
- мониторинг производительности системы ИИ, включая точность выходных данных системы ИИ;
- оповещение соответствующих заинтересованных сторон о проблемах, связанных с выходными данными системы ИИ, и их воздействии;
- оповещение о проблемах, связанных с изменениями в производительности или способности системы ИИ выдавать правильные выходные данные на основе производственных данных;
- рассмотрение вопроса о том, подходит ли автоматизированное принятие решений для ответственного подхода к использованию систем ИИ и предполагаемого использования системы ИИ.

Необходимость внедрения человеческого надзора может быть обоснована оценками воздействия системы ИИ (см. B.5). Персонал, участвующий в деятельности по человеческому надзору за системой ИИ, должен быть проинформирован и понимать инструкции и другую документацию, связанную с системой ИИ, а также обязанности, которые он выполняет для достижения целей человеческого надзора. При сообщении о проблемах с производительностью человеческий надзор может дополнить автоматизированный мониторинг.

Дополнительная информация

В приложении C приведены примеры целей по управлению рисками в организации, которые могут быть полезны при определении целей использования системы ИИ.

B.9.4 Предполагаемое использование системы искусственного интеллекта

Меры управления

Организация должна гарантировать, что система ИИ используется в соответствии с предполагаемым использованием системы ИИ и сопровождающей ее документацией.

Руководство по внедрению

Система ИИ должна быть развернута в соответствии с инструкциями и другой документацией, связанной с системой ИИ (см. B.8.2). Для развертывания могут потребоваться конкретные ресурсы для поддержки развертывания, включая необходимость обеспечения надлежащего контроля со стороны персонала (см. B.9.3). Для приемлемого использования и обеспечения корректной работы системы ИИ необходимо обеспечить согласованность данных, используемых в работе системы ИИ с документацией, связанной с системой ИИ.

Следует контролировать работу системы ИИ (см. B.6.2.6). В тех случаях, когда правильное развертывание системы ИИ в соответствии с надлежащими инструкциями вызывает обеспокоенность в отношении воздействия

на соответствующие заинтересованные стороны или юридические требования, организация должна сообщать о своих опасениях соответствующему персоналу внутри организации, а также соответствующим сторонним поставщикам системы ИИ.

Организация должна вести журналы событий или другую документацию, связанную с развертыванием и эксплуатацией системы ИИ, которую можно использовать для демонстрации того, что система ИИ используется по назначению, или для информирования о проблемах, связанных с предполагаемым использованием системы ИИ. Журналы событий и другую документацию следует хранить в течение определенного срока в соответствии с предполагаемым использованием системы ИИ, политиками хранения данных, принятыми в организации, и соответствующими юридическими обязательствами по хранению данных.

B.10 Взаимоотношения с третьими сторонами и клиентами

B.10.1 Цель

Обеспечение гарантии того, что организация понимает свои обязанности и сохраняет отчетность, а риски распределяются соответствующим образом при вовлечении третьих сторон на любой стадии жизненного цикла системы ИИ.

B.10.2 Распределение обязанностей

Меры управления

Организации следует обеспечить распределение обязанностей в рамках жизненного цикла системы ИИ между организацией, ее партнерами, поставщиками, клиентами и третьими сторонами.

Руководство по внедрению

В жизненном цикле системы ИИ обязанности могут быть разделены между сторонами, предоставляющими данные, сторонами, предоставляющими алгоритмы и модели, сторонами, разрабатывающими или использующими систему ИИ и несущими ответственность перед некоторыми или всеми заинтересованными сторонами. Организации следует документально зафиксировать все стороны, участвующие в жизненном цикле системы ИИ, и их роли, а также определить их требования.

При поставке системы ИИ третьей стороне организация должна обеспечить соблюдение ею ответственного подхода к разработке систем ИИ. Меры управления и рекомендации представлены в B.6. Организация должна предоставить необходимую документацию (см. B.6.2.7 и B.8.2) по системе ИИ соответствующим заинтересованным сторонам и третьей стороне, которой организация поставляет системы ИИ.

В случаях, если обрабатываемые данные включают ПДн, обязанности обычно распределяются между обработчиками ПДн и операторами ПДн. Дополнительная информация об обработчиках ПДн и операторах ПДн содержится в [23]. При необходимости сохранить конфиденциальность ПДн следует рассмотреть средства контроля, подобные представленным в [21]. Исходя из деятельности организации и деятельности системы ИИ по обработке данных ПДн и роли организации в применении и разработке систем ИИ на протяжении всего их жизненного цикла, организация может выполнять роль обработчика ПДн (или совместного обработчика ПДн), оператора ПДн или и того, и другого.

B.10.3 Поставщики

Меры управления

Организация должна разработать процесс, гарантирующий, что использование ею услуг, продуктов или материалов, предоставляемых поставщиками, соответствует подходу организации к ответственной разработке и использованию систем ИИ.

Руководство по внедрению

Организации, разрабатывающие или использующие систему ИИ, могут использовать услуги поставщиков несколькими способами: от поиска наборов данных, алгоритмов или моделей машинного обучения или других компонентов системы, таких как библиотеки программного обеспечения, до целой системы ИИ для использования самостоятельно или как часть другого продукта (например, транспортного средства).

При выборе поставщиков и требований, предъявляемых к этим поставщикам, а также уровней постоянного мониторинга и оценки, необходимых для поставщиков, организациям следует учитывать различные типы поставщиков, поставляемые ими продукты или услуги, а также различный уровень риска, который может возникнуть для системы и организации в целом.

Организации должны документировать, как система ИИ и ее компоненты интегрируются в системы ИИ, разработанные или используемые организацией.

Если организация считает, что система ИИ или компоненты системы ИИ от поставщика не работают должным образом или могут привести к негативным последствиям для отдельных лиц или групп лиц, или и тех, и других, а также социальных групп, что не соответствует ответственному подходу к системам ИИ, принятому организацией, организация вправе потребовать от поставщика принятия корректирующих мер. Организация может принять решение о сотрудничестве с поставщиком для достижения этой цели.

Организация должна обеспечить предоставление поставщиком системы ИИ надлежащей и адекватной документации, относящейся к системе ИИ (см. В.6.2.7 и В.8.2).

В.10.4 Заказчики

Меры управления

При ответственном подходе к разработке и использованию систем ИИ должны быть учтены ожидания и потребности клиентов организации.

Руководство по внедрению

При поставке продуктов или услуг, связанных с системой ИИ, организация должна понимать ожидания и потребности клиентов (т. е. в том случае, когда она сама является поставщиком). Они могут проявляться в форме требований к самому продукту или услуге на этапе разработки или проектирования, либо в форме контрактных требований или общих соглашений об использовании. Одна организация может устанавливать много разных типов отношений с клиентами, и все клиенты могут иметь разные потребности и ожидания.

Организация должна хорошо понимать сложную природу отношений с поставщиками и клиентами и понимать, в каких случаях ответственность лежит на поставщике системы ИИ, а в каких — на клиенте, при этом удовлетворяя их потребности и ожидания.

Например, организация может определить риски, связанные с использованием ее продуктов и услуг ИИ клиентом, а также принять решение об обработке выявленных рисков путем предоставления соответствующей информации заказчику для последующей обработки им соответствующих рисков.

Ограничение возможной области применения системы ИИ необходимо довести до сведения заказчика (см. В.6.2.7 и В.8.2).

Приложение С
(справочное)

Потенциальные организационные цели и источники рисков, связанные с применением искусственного интеллекта

С.1 Общие положения

В настоящем приложении излагаются потенциальные организационные цели, источники рисков и описания, которые организация может учитывать при управлении рисками. Настоящее приложение не является исчерпывающим или применимым к каждой организации. Организация должна определить релевантные цели и источники риска. В [14] представлена более подробная информация относительно целей и источников рисков, а также их взаимосвязи с управлением рисками. Оценка систем ИИ, первоначальная, регулярная и при необходимости, предоставляет доказательства того, что система ИИ оценивается на соответствие целям организации.

С.2 Цели

С.2.1 Подотчетность

Использование ИИ может изменить существующие системы подотчетности. Если ранее ответственность за свои действия несли люди, то теперь эти действия могут выполняться системой ИИ.

С.2.2 Опыт в области искусственного интеллекта

Необходимо выделить группу специалистов с междисциплинарными компетенциями и опытом в оценке, разработке и развертывании систем ИИ.

С.2.3 Доступность и качество обучающих данных и данных для тестирования

Системам ИИ, основанным на машинном обучении, необходимы данные для обучения, валидации и тестирования, чтобы обучать и верифицировать системы на предмет предполагаемого поведения.

С.2.4 Воздействие на окружающую среду

Использование ИИ может оказывать как положительное, так и отрицательное воздействие на окружающую среду.

С.2.5 Справедливость

Ненадлежащее применение систем ИИ для автоматизированного принятия решений может привести к необъективности по отношению к конкретным лицам или группам лиц.

С.2.6 Удобство сопровождения

Удобство сопровождения — способность системы ИИ изменяться для исправления дефектов и адаптироваться к новым требованиям.

С.2.7 Конфиденциальность

Неправильное использование или разглашение личных и конфиденциальных данных (например, медицинских записей) может привести к негативным последствиям для субъектов данных.

С.2.8 Робастность

В ИИ свойства робастности демонстрируют способность (или неспособность) системы обеспечивать сопоставимую производительность на новых данных с данными, на которых она была обучена, или с данными типичных операций.

С.2.9 Безопасность

Безопасность понимается в данном контексте как гарантия того, что система при определенных условиях не приведет к состоянию, при котором чья-то жизнь, здоровье или имущество, а также окружающая среда окажутся под угрозой.

С.2.10 Защита

В контексте ИИ и, в частности, в отношении систем ИИ, основанных на подходах машинного обучения, следует рассматривать новые факторы защиты, выходящие за рамки классических проблем информационной и системной безопасности.

С.2.11 Прозрачность и объяснимость

Прозрачность относится как к характеристикам организации, эксплуатирующей системы ИИ, так и к самим этим системам. Объяснимость относится к объяснениям важных факторов, влияющих на результаты работы системы ИИ, которые предоставляются в форме, понятной и доступной для заинтересованных сторон.

С.3 Источники риска

С.3.1 Сложность рабочей среды

Когда системы ИИ работают в сложных условиях с широким диапазоном ситуаций, может возникнуть неопределенность в отношении производительности и, следовательно, источник риска (например, в области автоматизации управления движением).

С.3.2 Отсутствие прозрачности и объяснимости

Неспособность предоставить соответствующую информацию заинтересованным сторонам может быть источником риска (например, с точки зрения надежности и подотчетности организации).

С.3.3 Уровень автоматизации

Уровень автоматизации может оказывать влияние на различные проблемные области, такие как безопасность жизнедеятельности, справедливость или кибербезопасность.

С.3.4 Источники риска, связанные с машинным обучением

Качество данных, используемых для машинного обучения, и процесс, используемый для сбора данных, также могут быть источником риска, поскольку могут повлиять на такие цели, как безопасность и робастность (например, из-за проблем с качеством данных или их искажения).

С.3.5 Проблемы с аппаратным обеспечением системы

Источники риска, связанные с аппаратным обеспечением, включают аппаратные ошибки, основанные на дефектных компонентах или переносом обученных моделей машинного обучения между разными системами.

С.3.6 Проблемы жизненного цикла системы

Источники риска могут появляться на протяжении всего жизненного цикла системы ИИ (например, недостатки в проектировании, неадекватное развертывание, отсутствие технического обслуживания, проблемы с выводом из эксплуатации).

С.3.7 Технологическая готовность

Источники риска могут быть связаны как с менее зрелой технологией из-за неизвестных факторов, так и с более зрелой технологией из-за избыточной уверенности в применяемой технологии.

Приложение D
(справочное)

Использование системы менеджмента искусственного интеллекта в разных областях и сферах деятельности

D.1 Общие положения

Представленная система менеджмента применима к любой организации, разрабатывающей, предоставляющей или использующей продукты или услуги, применяющие системы ИИ. Таким образом, система потенциально применима к большому разнообразию продуктов и услуг в различных областях, на которые распространяются обязательства, передовая практика, ожидания или договорные обязательства по отношению к заинтересованным сторонам. Примерами сфер деятельности являются:

- здоровье;
- оборона;
- транспорт;
- финансы;
- занятость;
- энергетика.

Для ответственной разработки и использования систем ИИ следует рассмотреть различные организационные задачи (см. возможные цели в приложении С). Настоящий стандарт содержит требования и рекомендации для конкретной технологии ИИ. Для нескольких потенциальных целей существуют общие или отраслевые стандарты системы менеджмента. Эти стандарты системы менеджмента обычно рассматривают цель с технологически нейтральной точки зрения, в то время как система менеджмента ИИ учитывает специфику технологии ИИ.

Системы ИИ состоят не только из компонентов, использующих технологию ИИ, но могут использовать самые разные технологии и компоненты. Таким образом, при ответственной разработке и использовании систем ИИ требуется принимать во внимание не только специфику ИИ, но и систему в целом со всеми используемыми технологиями и компонентами. Даже в части, касающейся технологии ИИ, следует принимать во внимание другие, не связанные с ИИ аспекты. Например, поскольку ИИ представляет собой технологию обработки информации, к нему в целом применяются соображения информационной безопасности. Такие цели, как безопасность, защита, конфиденциальность и воздействие на окружающую среду, должны управляться комплексно, а не отдельно для ИИ и других компонентов системы. Таким образом, для ответственной разработки и использования систем менеджмента ИИ важное значение имеет интеграция системы менеджмента ИИ с общими или отраслевыми стандартами систем управления по соответствующим темам.

D.2 Интеграция системы менеджмента искусственного интеллекта с другими стандартами систем менеджмента

При предоставлении или использовании систем ИИ у организации могут быть цели или обязательства, связанные с аспектами, которые являются темами других стандартов системы менеджмента. К ним могут относиться, например, безопасность, конфиденциальность, качество и соответственно темы, описанные в [22], [21] и [2]. Стандарты системы менеджмента ИСО разработаны таким образом, чтобы облегчить их интегрированное использование.

При предоставлении, использовании или разработке систем ИИ потенциально соответствующими общими стандартами системы менеджмента, являются следующие (но не ограничиваются ими):

- [22] — в большинстве случаев безопасность является ключом к достижению целей организации с помощью системы ИИ. Способы достижения цели обеспечения безопасности зависят от среды и собственных политик организации. Если организация определяет необходимость внедрения системы менеджмента ИИ и решения задач безопасности аналогичным тщательным и систематическим образом, она также может внедрить систему менеджмента информационной безопасности в соответствии с [22]. Учитывая, что в [22] и системах менеджмента ИИ используют структуру высокого уровня, их комплексное использование упрощается и приносит большую пользу организации. В таком случае способ внедрения мер управления, которые (частично) относятся к информационной безопасности в настоящем стандарте (см. В.6.1.2), может быть интегрирован с внедрением организацией стандарта [22].

- [21] — во многих средах и прикладных областях ПДн обрабатываются системами ИИ. После этого организация может соблюдать применимые обязательства в отношении конфиденциальности, а также свои собственные политики и цели. Аналогичным образом, что касается [22], организация может извлечь выгоду из интеграции стандарта [21] с системой управления ИИ. Цели и меры управления системы управления ИИ, связанные с конфиденциальностью (см. В.2.3 и В.5.4), могут быть интегрированы с внедрением организацией стандарта [21].

- [2] — для многих организаций соответствие стандарту [2] является ключевым признаком того, что они ориентированы на клиента и придают большое значение обеспечению о внутренней результативности. Независимая оценка соответствия стандарту [2] облегчает ведение бизнеса во всех организациях и вселяет доверие клиентов к продуктам и услугам. Уровень доверия клиентов к организации или системе ИИ может быть значительно повышен,

если система менеджмента ИИ внедряется совместно с [2] при использовании технологий ИИ. Система менеджмента ИИ может дополнять требования стандарта [2] (управление рисками, разработка программного обеспечения, согласованность цепочки поставок и т. д.), помогая организации достигать своих целей.

Помимо общих стандартов системы менеджмента, упомянутых выше, система менеджмента ИИ также может использоваться совместно с системой менеджмента, предназначенной для конкретной сферы деятельности. Например, и стандарт [5], и система менеджмента ИИ актуальны для систем ИИ, которые используются в пищевой отрасли. Другим примером является [4]. Внедрение системы менеджмента ИИ может поддерживать требования, относящиеся к программному обеспечению медицинских изделий [4] или требования других международных стандартов сферы здравоохранения, таких как [6].

Приложение ДА
(справочное)

Сведения о соответствии ссылочных международных стандартов национальным стандартам

Таблица ДА.1

Обозначение ссылочного международного стандарта	Степень соответствия	Обозначение и наименование соответствующего национального стандарта
ISO/IEC 22989:2022	MOD	ГОСТ Р 71476—2024 (ИСО/МЭК 22989:2022) «Искусственный интеллект. Концепции и терминология искусственного интеллекта»
Примечание — В настоящей таблице использовано следующее условное обозначение степени соответствия стандарта: - MOD — модифицированный стандарт.		

Библиография

[1]	ISO 8000-2	Data quality — Part 2: Vocabulary
[2]	ISO 9001	Quality management systems — Requirements
[3]	ISO 9241-210	Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems
[4]	ISO 13485	Medical devices — Quality management systems — requirements for regulatory purposes
[5]	ISO 22000	Food safety management systems — Requirements for any organization in the food chain
[6]	IEC 62304	Medical device software — Software life cycle processes
[7]	ISO/IEC Guide 51	Safety aspects — Guidelines for their inclusion in standards
[8]	ISO/IEC TS 4213	Information technology — Artificial Intelligence — Assessment of machine learning classification performance
[9]	ISO/IEC 5259 (all parts)	Data quality for analytics and Machine Learning (ML)
[10]	ISO/IEC 5338	Information technology — Artificial intelligence — AI system life cycle process
[11]	ISO/IEC 17065:2012	Conformity assessment — Requirements for bodies certifying products, processes and services
[12]	ISO/IEC 19944-1	Cloud computing and distributed platforms — Data flow, data categories and data use — Part 1: Fundamentals
[13]	ISO/IEC 23053	Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML)
[14]	ISO/IEC 23894	Information technology — Artificial intelligence — Guidance on risk management
[15]	ISO/IEC TR 24027	Information technology — Artificial intelligence (AI) — Bias in AI systems and AI aided decision making
[16]	ISO/IEC TR 24029-1	Artificial intelligence (AI) — Assessment of the robustness of neural networks — Part 1: Overview
[17]	ISO/IEC TR 24368	Information technology — Artificial intelligence — Overview of ethical and societal concerns
[18]	ISO/IEC 25024	Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaR E) — Measurement of data quality
[19]	ISO/IEC 25059	Software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Quality Model for AI systems
[20]	ISO/IEC 27000:2018	Information technology — Security techniques — Information security management systems — Overview and vocabulary
[21]	ISO/IEC 27701	Security techniques — Extension to ISO/IEC 27001 and ISO/IEC 27002 for privacy information management — Requirements and guidelines
[22]	ISO/IEC 27001	Information technology — Security techniques — Information security management systems — Requirements
[23]	ISO/IEC 29100	Information technology — Security techniques — Privacy framework
[24]	ISO 31000:2018	Risk management — Guidelines
[25]	ISO 37002	Whistleblowing management systems — Guidelines
[26]	ISO/IEC 38500:2015	Information technology — Governance of IT for the organization

[27] ISO/IEC 38507 Information technology — Governance of IT — Governance implications of the use of artificial intelligence by organizations

[28] Lifecycle D.D.I. 3.3, 2020-04-15. Data Documentation Initiative (DDI) Alliance. [viewed on 2022-02-19]. Available at: gost.e-ecolog.ru

[29] Risk Framework N.I.S.T.-A.I. 1.0, 2023-01-26. National Institute of Technology (NIST) [viewed on 2023-04-17] gost.e-ecolog.ru

УДК 004.8:006.354 ОКС 35.020

Ключевые слова: информационные технологии, искусственный интеллект, система менеджмента

Технический редактор *И.Е. Черепкова*
Корректор *С.И. Фирсова*
Компьютерная верстка *М.В. Малеевой*

Сдано в набор 11.11.2024. Подписано в печать 25.11.2024. Формат 60×84⅛. Гарнитура Ариал.
Усл. печ. л. 5,58. Уч.-изд. л. 4,74.

Подготовлено на основе электронной версии, предоставленной разработчиком стандарта

Создано в единичном исполнении в ФГБУ «Институт стандартизации»
для комплектования Федерального информационного фонда стандартов,
117418 Москва, Нахимовский пр-т, д. 31, к. 2.

gost.e-ecolog.ru gost.e-ecolog.ru

gost.e-ecolog.ru