ФЕДЕРАЛЬНОЕ АГЕНТСТВО

ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

ГОСТ Р 72515— 2026 (ИСО/МЭК 12792: 2025)

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

Таксономия прозрачности систем искусственного интеллекта

(ISO/IEC 12792:2025, Information technology — Artificial intelligence (AI) — Transparency taxonomy of AI systems, MOD)

Издание официальное

Москва
Российский институт стандартизации
2026

gost.e-ecolog.ru

Предисловие

1 ПОДГОТОВЛЕН Федеральным государственным автономным образовательным учреждением высшего образования «Волгоградский государственный университет» (ФГАОУ ВО ВолГУ) и 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 30 января 2026 г. № 65-ст

4 Настоящий стандарт является модифицированным по отношению к международному стандарту ИСО/МЭК 12792:2025 «Информационные технологии. Искусственный интеллект (ИИ). Таксономия прозрачности систем ИИ» [ISO/IEC 12792:2025 «Information technology — Artificial intelligence (AI) — Transparency taxonomy of AI systems», MOD] путем внесения технических отклонений, объяснение которых приведено во введении к настоящему стандарту.

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

Сведения о соответствии ссылочных национальных стандартов международным стандартам, использованным в качестве ссылочных в примененном международном стандарте, приведены в дополнительном приложении ДА.

Сопоставление структуры настоящего стандарта со структурой указанного международного стандарта приведено в дополнительном приложении ДБ

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, 2025
© IEC, 2025
© Оформление. ФГБУ «Институт стандартизации», 2026

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 2
4 Сокращения 4
5 Обзор 4
5.1 Общие положения 4
5.2 Организация и использование таксономии 4
5.3 Ограничения на раскрытие информации о прозрачности 5
5.4 Понятие прозрачности 5
6 Потребности заинтересованных сторон и цели прозрачности 7
6.1 Общие положения 7
6.2 Цели и задачи обеспечения прозрачности 8
6.3 Отдельные роли заинтересованных сторон в обеспечении прозрачности 9
7 Таксономия контекста 10
7.1 Общие положения 10
7.2 Социальный контекст 11
7.3 Экологический контекст 14
8 Таксономия системы 17
8.1 Общие положения 17
8.2 Основные сведения 17
8.3 Организационные процессы 18
8.4 Применимость 19
8.5 Обзор технических характеристик 20
8.6 Доступ к внутренним компонентам 22
8.7 Качество и показатели производительности 22
9 Таксономия модели 24
9.1 Общие положения 24
9.2 Основные сведения 24
9.3 Использование и взаимодействие с другими моделями 24
9.4 Технические характеристики 25
9.5 Используемые данные 27
9.6 Функциональная корректность 28
10 Таксономия наборов данных 28
10.1 Общие положения 28
10.2 Основные сведения 29
10.3 Происхождение данных 29
10.4 Свойства данных 31
10.5 Предметная область и назначение набора данных 31
10.6 Предвзятость и ограничения данных 32
10.7 Социальные аспекты 33
10.8 Подготовка данных 33
10.9 Обслуживание наборов данных 34

Приложение А (справочное) Примеры ролей заинтересованных сторон в раскрытии прозрачности 35
Приложение ДА (справочное) Сведения о соответствии ссылочных национальных стандартов международным стандартам, использованным в качестве ссылочных в примененном международном стандарте. 37
Приложение ДБ (справочное) Сопоставление структуры настоящего стандарта со структурой примененного в нем международного стандарта 38
Библиография 39

Введение

Целями настоящего стандарта являются:

- повышение доверия, подотчетности и коммуникаций между различными заинтересованными в искусственном интеллекте (ИИ) сторонами, включая партнеров по цепочке поставок, клиентов, пользователей, общество и регулирующие органы, путем создания согласованной терминологии, относящейся к прозрачности систем ИИ (*СИИ*);
- предоставление заинтересованным в ИИ сторонам информации о различных элементах прозрачности с указанием их значимости и возможных ограничений для различных вариантов использования и целевой аудитории;
- обеспечение основы для разработки стандартов прозрачности СИИ для конкретных технологий, отраслей и территорий.

Прозрачность *СИИ* — свойство, при котором заинтересованные стороны получают надлежащую информацию о системе. К ней может относиться информация о функциональных возможностях системы, ограничениях, данных, проектировании системы и проектных решениях (см. *ГОСТ Р 71476—2024*, пункт 5.15.8).

Повышенная прозрачность предоставляет соответствующим заинтересованным сторонам информацию, позволяющую лучше понять, как разрабатывается и используется *СИИ*. Например, это позволяет клиенту ИИ (например, пользователю ИИ) определить, подходит ли *СИИ* для его ситуации, а аудитору ИИ помогает оценить, отвечает ли она требованиям соответствия.

Стандартизированная таксономия прозрачности *СИИ* позволяет представителям различных профессий лучше понимать друг друга за счет использования единой терминологии. Это, в свою очередь, способствует лучшему пониманию *СИИ* и создает основу для разработки интероперабельных и согласованных стандартов, связанных с прозрачностью.

СИИ часто влияют на общество и окружающую среду (см. раздел 7). Однако верно и то, что общество и окружение могут влиять на показатели производительности *СИИ*. К ним относятся следующие аспекты:

- условия окружения, влияющие на функционирование и долговечность *СИИ*;
- внедрение или усиление нежелательной предвзятости;
- организационные практики, которые могут привести к неудовлетворительным результатам;
- такие эффекты, как дрейф концепции (дрейф концепции происходит при изменении функциональной связи между входными и выходными данными модели) и дрейф данных (дрейф данных — изменение статистических свойств и характеристик входных данных);
- возникновение негативной замкнутой обратной связи (например, усиление дискриминационных шаблонов), что может быть особенно проблематичным в случае непрерывного обучения.

Несмотря на то, что данные вопросы являются важными для рассмотрения, их подробный анализ в настоящем стандарте не приводится. Некоторые аспекты более подробно описаны в *[1]—[3].*

В настоящий стандарт внесены следующие изменения по отношению к международному стандарту ИСО/МЭК 12792:2025:

- *изменены ссылки;*
- *исключены отдельные положения, которые дублируются по тексту стандарта;*
- *из библиографии исключена информация о документах, ссылки на которые не использованы при изложении настоящего стандарта;*
- *включены сокращения;*
- *термины и определения приведены в соответствие с терминологией, установленной в документах по стандартизации Российской Федерации;*
- *структура стандарта изменена для приведения в соответствие с правилами, установленными в ГОСТ 1.5—2001 (подразделы 4.2 и 4.3);*
- *в текст стандарта внесена дополнительная сноска для пояснения текста оригинала.*

Все изменения и дополнения в тексте стандарта выделены курсивом.

Внесение указанных технических отклонений направлено на учет целесообразности использования ссылочных национальных стандартов вместо ссылочных международных стандартов, а также особенностей национальной стандартизации, технологий работы с большими данными и искусственным интеллектом.

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

Таксономия прозрачности систем искусственного интеллекта

Artificial intelligence. Transparency taxonomy of artificial intelligence systems

Дата введения — 2026—05—01

1 Область применения

Настоящий стандарт устанавливает таксономию информационных элементов, которая позволит заинтересованным в искусственном интеллекте сторонам определить и удовлетворить потребности в обеспечении прозрачности системы искусственного интеллекта. Настоящий стандарт устанавливает семантику информационных элементов и их соответствие целям разных заинтересованных сторон.

Настоящий стандарт предназначен для применения в организациях любого типа, использующих любые типы систем искусственного интеллекта.

2 Нормативные ссылки

В настоящем стандарте использованы нормативные ссылки на следующие стандарты:

ГОСТ Р 71476—2024 (ИСО/МЭК 22989:2022) Искусственный интеллект. Концепции и терминология искусственного интеллекта

ГОСТ Р 71484.4—2024 (ИСО/МЭК 5259-4:2024) Искусственный интеллект. Качество данных для аналитики и машинного обучения. Часть 4. Структура процесса управления качеством данных

ГОСТ Р 71539 (ИСО/МЭК 5338:2023) Искусственный интеллект. Процессы жизненного цикла системы искусственного интеллекта

ГОСТ Р ИСО 50001 Системы энергетического менеджмента. Требования и руководство по применению

ГОСТ Р ИСО/МЭК 25000 Системная и программная инженерия. Требования и оценка качества систем и программных средств (SQuaRE). Руководство

ГОСТ Р ИСО/МЭК 27037 Информационная технология. Методы и средства обеспечения безопасности. Руководства по идентификации, сбору, получению и хранению свидетельств, представленных в цифровой форме

ГОСТ Р ИСО/МЭК 42001—2024 Искусственный интеллект. Система менеджмента

Примечание — При пользовании настоящим стандартом целесообразно проверить действие ссылочных стандартов в информационной системе общего пользования — на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет или по ежегодному информационному указателю «Национальные стандарты», который опубликован по состоянию на 1 января текущего года, и по выпускам ежемесячного информационного указателя «Национальные стандарты» за текущий год. Если заменен ссылочный стандарт, на который дана недатированная ссылка, то рекомендуется использовать действующую версию этого стандарта с учетом всех внесенных в данную версию изменений. Если заменен ссылочный стандарт, на который дана датированная ссылка, то рекомендуется использовать версию этого стандарта с указанным выше годом утверждения (принятия). Если после утверждения настоящего стандарта в ссылочный стандарт, на который дана датированная ссылка, внесено изменение, затрагивающее положение, на которое дана ссылка, то это положение рекомендуется применять без учета данного изменения. Если ссылочный стандарт отменен без замены, то положение, в котором дана ссылка на него, рекомендуется применять в части, не затрагивающей эту ссылку.

Издание официальное

3 Термины и определения

В настоящем стандарте применены термины по *ГОСТ Р 71476*, а также следующие термины с соответствующими определениями:

3.1 **аттестация** (attestation): Выдача заявления, основанного на решении о том, что выполнение заданных требований продемонстрировано.

Примечания

1 Окончательное заявление, которое подтверждает, что заданные требования выполнены. Такое подтверждение само по себе не дает договорных или каких-либо других правовых гарантий.

2 Аттестация первой и третьей сторонами различается в соответствии с терминами «декларация», «сертификация» и «аккредитация», но для деятельности по аттестации второй стороной специального термина не существует.

3 См. [4] (пункт 7.3); за исключением слов в примечании 1: «о котором в настоящем стандарте идет речь как о «заявлении о соответствии».

3.2

углеродный след продукции (carbon footprint of a product): Сумма выбросов ПГ и поглощений ПГ в продукционной системе, выраженная в единицах-эквивалента и основанная на оценке жизненного цикла с использованием одной категории воздействия — изменение климата.

[Адаптировано из *ГОСТ Р ИСО 14050—2023*, пункт 3.11.1]

3.3 **конфиденциальность информации** (confidential data): *Обязательное для выполнения лицом, получившим доступ к определенной информации, требование не передавать такую информацию третьим лицам без согласия ее обладателя.*

Примечание — См. [5] (статья 2).

3.4 **орган по оценке соответствия** (conformity assessment body): Орган, осуществляющий деятельность по оценке соответствия за исключением аккредитации.

Примечание — См. [4] (пункт 4.6).

3.5

компрометация данных (data breach): Нарушение безопасности, которое приводит к случайному или незаконному разрушению, потере, изменению, несанкционированному раскрытию или доступу к защищаемым данным, передаваемым, хранимым или иным образом обрабатываемым.

[ГОСТ Р ИСО/МЭК 27018—2020, пункт 3.1]

3.6

защита данных (data protection): Техническая и социальная система мероприятий по согласованию, управлению и обеспечению неприкосновенности, конфиденциальности и защиты информации.

[ГОСТ 33707—2016, статья 4.390]

3.7 **цифровой отпечаток документа** (digital fingerprint): Последовательность битов, сформированная на основе электронного документа с помощью алгоритма, который однозначно идентифицирует исходный документ.

Примечания

1 Любая модификация электронного документа приведет к изменению отпечатка.

2 См. [6] (пункт 3.15).

3.8 **развивающаяся система** (evolutive system): Система искусственного интеллекта, поведение которой может меняться без явного вмешательства со стороны ее разработчика.

Примечания

1 Примерами развивающихся систем являются система искусственного интеллекта с запоминающим компонентом, система искусственного интеллекта, использующие действия и обратную связь со стороны пользователя, и система искусственного интеллекта с непрерывным обучением.

2 Изменение поведения не подразумевает изменения цели, но может привести к лучшему или худшему достижению поставленной цели.

3.9 **признак** *(машинное обучение)* (feature)*:* Измеримое свойство объекта или события, связанное с заданным набором характеристик.

Примечания

1 Признаки играют роль в обучении и генерации результатов.

2 Признаки предоставляют машиночитаемый способ описания соответствующих объектов. Поскольку алгоритм не будет возвращаться к самим объектам или событиям, представления признаков разработаны таким образом, чтобы содержать всю полезную информацию.

3 См. [7] (пункт 3.3.3).

3.10

оценка воздействия жизненного цикла (life cycle impact assessment): Стадия оценки жизненного цикла, направленная на определение и оценку величины и значимости возможных воздействий на окружающую среду продукционной системы на протяжении жизненного цикла продукции.

[Адаптировано из *ГОСТ Р ИСО 14050—2023*, пункт 3.6.5]

3.11 **объект оценки соответствия** (объект) (object of conformity assessment; object): Сущность, к которой применимы заданные требования.

Пример — Продукция, процесс, услуга, система, установка, проект, данные, проектная документация, материал, утверждение, лицо, орган или организация либо любая их комбинация.

Примечания

1 Термин «орган» применяется в данном контексте только в тех случаях, когда речь идет об органах по оценке соответствия и органах по аккредитации. Термин «организация» используется в его общем значении и, в соответствии с контекстом, может включать органы.

2 См. [4] (пункт 4.2).

3.12 **коэффициент энергоэффективности;** PUE (power usage effectiveness, PUE): Отношение общего энергопотребления центра обработки данных к энергопотреблению ИТ-оборудования, вычисленному, измеренному или оцененному за тот же период времени.

Примечания

1 Иногда используется обратное значение PUE, называемое коэффициентом энергоэффективности инфраструктуры центра обработки данных (DCiE).

2 См. *[8] (пункт 3.1.3).*

3.13 **область аттестации** (scope of attestation): Диапазон или характеристики объектов оценки соответствия, охватываемых аттестацией.

Примечание — См. [4] (пункт 7.4).

3.14 **заданное требование** (specified requirement): Заявленная потребность или ожидание.

Примечания

1 Заданные требования могут быть изложены нормативными документами, такими как регламенты, стандарты и технические условия.

2 Заданные требования могут быть четко определенными или общими.

3 См. [4] (пункт 5.1).

3.15

таксономия (taxonomy): Системная классификация объектов в группы на основе параметров, являющихся общими для нескольких объектов.

[ГОСТ Р 70841—2023, пункт 3.84]

3.16

водный след (water footprint): Количественные показатели, которые определяют потенциальные воздействия на окружающую среду, связанные с водой.

[*ГОСТ Р ИСО 14050—2023*, пункт 3.10.1]

4 Сокращения

ИИ — искусственный интеллект;

МО — машинное обучение;

ПДн — персональные данные;

ПО — программное обеспечение;

СИИ — система искусственного интеллекта;

API — прикладной программный интерфейс;

JPEG — стандарт сжатия цифрового изображения, определенный в [9];

JSON — текстовый формат обмена данными, основанный на языке JavaScript;

REST — передача состояния представления;

URL — унифицированный указатель ресурса.

5 Обзор

5.1 Общие положения

В настоящем разделе представлен обзор настоящего стандарта и описана концепция прозрачности *СИИ*, в частности, в отношении других характеристик доверия к таким системам.

5.2 Организация и использование таксономии

Настоящий стандарт описывает фасетную таксономию прозрачности *СИИ*. В разделе 6 рассматривается, как потребности в прозрачности могут меняться в зависимости от контекста (обстоятельств разработки и использования) *СИИ* и роли различных заинтересованных сторон. Далее в настоящем стандарте рассматриваются различные уровни и точки зрения на систему. Данное рассмотрение включает четыре части:

- раздел 7 описывает элементы таксономии прозрачности (далее — элементы таксономии), содержащие информацию о контексте *СИИ* (например, требования к прозрачности участников трудовых отношений);
- в разделе 8 описывается раскрытие информации в рамках обеспечения прозрачности на уровне *СИИ* (например, рекомендуемое и недопустимое использование, а также документация по API);
- раздел 9 посвящен внутреннему функционированию системы, на уровне компонентов модели и алгоритмических процессов;
- в разделе 10 приводятся рекомендации по документированию наборов данных как самостоятельных элементов с учетом того, что жизненный цикл *СИИ* включает эти наборы данных (что также требует обеспечения прозрачности), а также то, что наборы данных могут повторно использоваться в неизменном виде в нескольких *СИИ* для различных вариантов использования.

Для конкретной *СИИ* могут рассматриваться все четыре раздела таксономии (разделы 7—10), поскольку они содержат взаимодополняющую информацию. Структура настоящего стандарта графически представлена на рисунке 1.

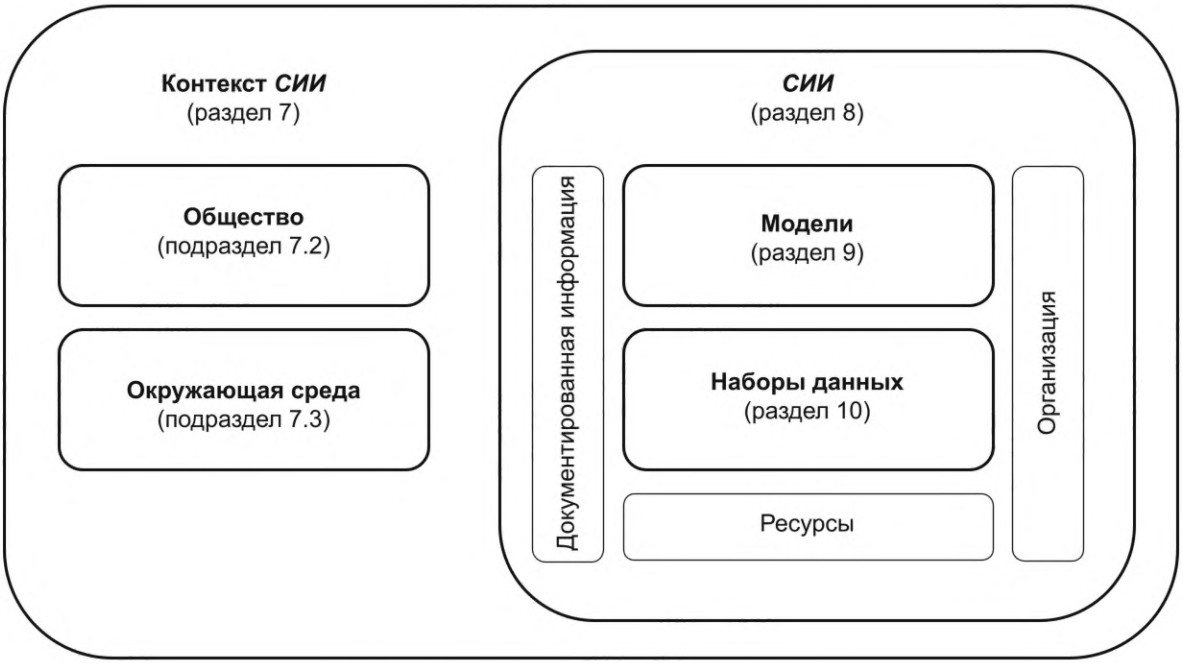

Рисунок 1 — Обзор структуры документа

Поскольку в разработку *СИИ* могут быть вовлечены различные заинтересованные стороны (поставщик данных, разработчик ИИ, поставщик платформы ИИ и т. д.), а также поскольку *СИИ* может включаться как подсистема в более крупную, данная таксономия разработана по модульному принципу. В частности, документация *СИИ* может содержать ссылки на документацию (полностью или только на отдельные части) другой *СИИ*.

Например, для моделей общего назначения с открытым исходным кодом целесообразно документировать их рекомендуемое и недопустимое использование в соответствии с разделами 8 и 9. Другая организация, интегрируя модель общего назначения в качестве внутреннего компонента более сложной *СИИ*, может создать новую документацию для этой *СИИ*, включающую:

- собственную документацию в соответствии с разделом 8 (например, рекомендуемое и недопустимое использование всей *СИИ*);
- ссылку на документацию модели общего назначения в соответствии с разделом 9;
- новый раздел, разработанный в соответствии с разделом 9, для любой специализированной модели, разработанной для этой *СИИ*;
- однако допускается исключение информации по разделу 8 о модели общего назначения, если данные сведения нерелевантны в контексте конкретной *СИИ*. Примером нерелевантных сведений в данном контексте является информация о рекомендуемом и недопустимом использовании модели общего назначения, включающая альтернативные варианты использования, которые выходят за рамки единственного варианта использования модели общего назначения в конкретной *СИИ*.

5.3 Ограничения на раскрытие информации о прозрачности

Элементы таксономии, описанные во всех четырех разделах настоящего стандарта (см. разделы 7—10), представляют собой элементы, которые полезно учитывать по крайней мере в некоторых случаях, но которые не обязательно применимы ко всем вариантам использования и системам. В частности:

- существуют элементы таксономии, раскрытие которых нецелесообразно в определенных вариантах использования. Например, публикация исходного кода недопустима, если он содержит закрытый (проприетарный) код (см. 8.6). Указание источника финансирования для создания набора данных может быть невыполнимым, если набор данных создан добровольцами или раскрытие информации запрещается договорными обязательствами (см. 10.3). Аналогичным образом, раскрытие размера вознаграждения разметчиков данных может противоречить принципам конфиденциальности (см. 10.8);

- некоторые элементы таксономии применимы только к определенным типам систем. Например, сведения о диалекте языка не всегда релевантны для систем компьютерного зрения (см. 10.5);

- отдельные элементы таксономии описывают процессы, которые невозможны или нежелательны в конкретных вариантах использования. Например, сравнительная оценка с аналогичной системой может быть невыполнимой при создании нового типа *СИИ*, у которой нет существующей альтернативы (см. 8.7.4);

- аналогичным образом, требования по обеспечению разнообразия и инклюзивности не всегда целесообразны при создании набора данных, репрезентативного для конкретной группы населения, например пожилых женщин с нарушениями речи (см. 10.6);

- для определения полезности элементов обеспечения прозрачности организация также может оценить выгоду и затраты, связанные с каждым из соответствующих элементов обеспечения прозрачности, чтобы найти возможный компромисс и определить надлежащий уровень прозрачности для конкретного варианта использования. Например, организация может оценить, могут ли элементы обеспечения прозрачности в некоторых случаях принести отдельным лицам и обществу больше вреда, чем пользы, и может ли раскрытие информации нарушать действующее законодательство, нормативные документы или договорные требования.

5.4 Понятие прозрачности

Прозрачность является широким понятием, однако в настоящем стандарте основное внимание уделяется прозрачности как свойству системы и, в меньшей степени, прозрачности как свойству организации. Прозрачность *СИИ* означает обеспечение заинтересованных сторон надлежащей информацией о системе (см. *ГОСТ Р 71476—2024, пункт 5.15.8*). В отличие от этого прозрачность организации (в настоящем стандарте именуемая организационной прозрачностью) связана с тем, как надлежащие мероприятия и решения доводятся до сведения соответствующих заинтересованных сторон (см. *[10]*).

Организационная прозрачность актуальна в той степени, в которой организационные принципы и процессы влияют на *СИИ* на различных стадиях ее жизненного цикла, как описано в *ГОСТ Р 71476—2024* (раздел 6).

Примерами надлежащей информации для обеспечения прозрачности являются цели, известные ограничения, проектные решения и допущения, характеристики, модели, алгоритмы, методы обучения, сведения об используемых данных и процессах обеспечения качества (для *СИИ*, для данных, используемых в обучении или оценке). Связанные с обеспечением прозрачности потребности могут различаться для разных заинтересованных сторон, а также в зависимости от варианта использования. Следовательно, состав надлежащей информации, подлежащей раскрытию в рамках обеспечения прозрачности *СИИ*, зависит не только от того, кто именно является целевой заинтересованной стороной, но и от варианта использования. Например, потребности аудитора ИИ или регулирующего органа в сфере ИИ отличаются по аспектам или уровню детализации от потребностей пользователя *СИИ*. В некоторых случаях аудит может не понадобиться; потребности отдельных лиц и обществ в прозрачности различаются при использовании *СИИ* для выбора фильма, прогнозирования ремонта турбин и опознания преступников в рамках полицейского расследования. Кроме того, для определенного элемента таксономии (например, используемой модели) прозрачность может обеспечиваться разными способами в зависимости от заинтересованных сторон: для одних достаточно предоставления документации по выбору используемой архитектуры модели, а для других будет целесообразно предоставление доступа к внутренним состояниям модели, таким как ее параметры.

Значимость, достоверность, полнота доступности и понятность информации влияет на степень ее соответствия ожиданиям. Степень соответствия зависит от варианта использования и типа заинтересованной стороны. Например, обычному пользователю не требуется та же степень детализации информации, что техническому эксперту-криминалисту. При этом степень доступности информации для определенных заинтересованных сторон, особенно при запросе доступа к данным, коду или моделям, может вступать в противоречие с требованиями к безопасности, защите *ПДн* и конфиденциальности. В таких случаях, чтобы определить возможность обеспечения прозрачности информации для сторон, заинтересованных в последующих стадиях жизненного цикла *СИИ*, защищая при этом конфиденциальную информацию о компонентах предыдущих стадий, может потребоваться обеспечение безопасности, защиты или использования информации в соответствии с внутренними политиками организации в отношении прав интеллектуальной собственности, договорных обязательств, конфиденциальной информации или коммерческой тайны, чтобы найти компромиссное решение.

Помимо определений, используемых в настоящем стандарте, понятие прозрачности также можно рассматривать как особенность представления информации или как открытость и готовность сообщать о действиях и решениях. На рисунке 2 представлен обзор различных основных концепций прозрачности и соответствующих стандартов.

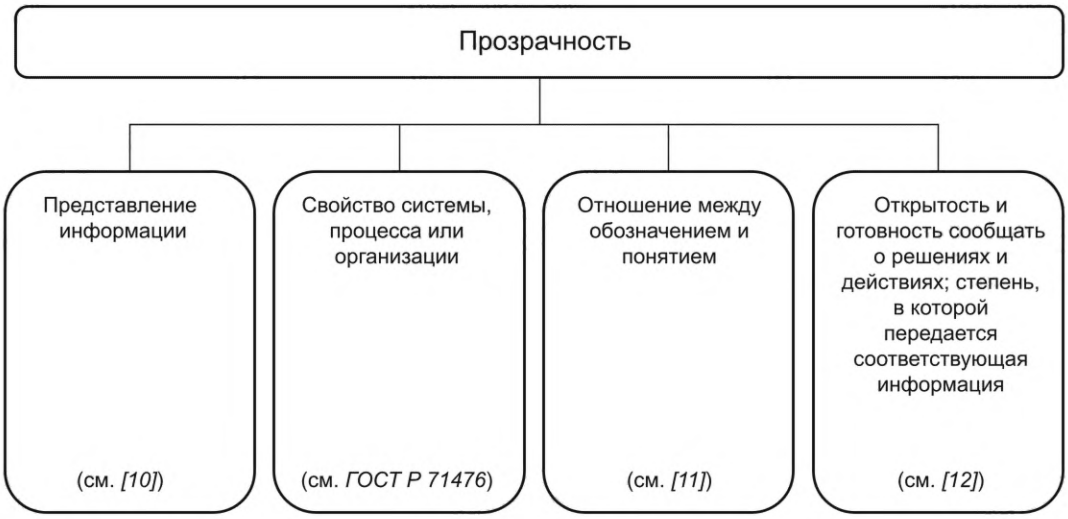

Рисунок 2 — Соответствующие высокоуровневые концепции прозрачности

Свойство прозрачности *СИИ* сочетается с рядом других свойств, таких как воспроизводимость, наблюдаемость, проверяемость, удобство использования, объяснимость и интерпретируемость, подотчетность, качество и доверие (см., например, *ГОСТ Р ИСО/МЭК 27037*, *[10]—[12]*), а также во многих случаях предваряет эти свойства.

Например, если необходимо учитывать этические и социальные аспекты, то прозрачность и объяснимость являются ключевыми факторами для принятия таких систем. Прозрачность *СИИ* позволяет заинтересованным сторонам понимать принципы ее работы, оценивать функциональность и осознавать ее сильные стороны и ограничения. Это помогает разрабатывать *СИИ*, которые приносят пользу обществу, принимая при этом во внимание этические и социальные аспекты их использования (см. *[13]*).

С точки зрения модели качества систем и ПО, определенной в *[12]*, прозрачность является подхарактеристикой удобства использования. Она помогает потенциальным пользователям выбрать *СИИ*, соответствующую их требованиям, повышает уровень знаний заинтересованных сторон о применимости и ограничениях *СИИ*, а также способствует ее объяснимости. Прозрачность также является подхарактеристикой удовлетворенности в модели качества использования (см. *[12]*).

С точки зрения объяснимости, прозрачность является дополнительным и полезным свойством для достижения объяснимости. Например, некоторые методы объяснимости требуют определенной степени прозрачности для доступа к внутренним состояниям или промежуточным результатам из *СИИ* либо для получения информации о моделях, используемых в *СИИ* (см. *[14]*).

6 Потребности заинтересованных сторон и цели прозрачности

6.1 Общие положения

Прозрачность является характеристикой, способствующей обеспечению доверия к *СИИ*. В рамках усилий по обеспечению и поддержанию доверия к *СИИ* в качестве заинтересованных сторон следует рассматривать как организации, так и отдельные лица и группы лиц. В число таких заинтересованных сторон входят те из них, которые изначально могли не иметь потребностей или ожиданий в отношении *СИИ*. Потребности и ожидания у этих заинтересованных сторон (например, у субъектов *ПДн*, не осведомленных об использовании их данных до момента уведомления) с большой вероятностью могут возникнуть в результате раскрытия информации в рамках обеспечения прозрачности. Как и в случае с другими характеристиками доверия к *СИИ*, преимущества прозрачности могут быть усилены за счет

явной контекстуализации и пояснения, какие заинтересованные стороны стремятся установить и поддерживать доверие других заинтересованных сторон в отношении конкретных аспектов разработки и использования *СИИ* (см. *[10]*).

Прозрачность подразумевает доведение соответствующей информации до заинтересованных сторон в полном, доступном и понятном виде. Прозрачность может относиться как к предоставлению соответствующим заинтересованным сторонам надлежащей информации о *СИИ*, так и к передаче им сведений о принимаемых организационных мерах и решениях.

Прозрачность может включать передачу информации о целях проектирования, целевом использовании, известных ограничениях, определениях, проектных решениях и допущениях, характеристиках, компонентах, процедурах, метриках, моделях, алгоритмах, методах обучения, процессах обеспечения качества (для *СИИ*, для данных, используемых в обучении или оценке), а также о характеристиках и происхождении данных, используемых при обучении и оценке. Также может включаться информация о характере обработки и автоматизации, связанных с решениями, имеющими отношение к заинтересованным сторонам (см. *ГОСТ Р 71476—2024*, пункт 5.15.8).

Роли и суброли заинтересованных сторон в сфере ИИ определены в *ГОСТ Р 71476—2024* (рисунок 2). От различающихся целей и задач заинтересованных сторон зависит тип релевантной информации. Например, цели и задачи устойчивого развития, которые могут быть поддержаны данными анализа углеродного следа (см. 7.3), отличаются от целей и задач закупок или снабжения, для которых актуальна информация о затратах на эксплуатацию и развитие (см. 9.4.7).

Учет потребностей и целей заинтересованных сторон также предполагает адаптацию языка и уровня детализации информации о прозрачности. Если, например, таким заинтересованным сторонам, как выполняющие определенные роли пользователи ИИ или субъекты ИИ, требуется четко сформулированная простым языком информация, описывающая общие принципы работы и воздействие *СИИ* (см. 8.2—8.4), то для выступающих в качестве производителя и/или интегратора *СИИ*, обычно требуется бóльшая детализация (как это предусмотрено, например, в 8.5—8.7). В [*15*] содержатся соответствующие рекомендации о том, как добиться простого языка, который может быть важен для достижения прозрачности.

В *ГОСТ Р 71476—2024* (подраздел 5.19) определены заинтересованные стороны, которые классифицируются в соответствии с их ролью и субролью в жизненном цикле ИИ.

В обеспечении прозрачности у пары заинтересованных сторон может быть две роли: одна сторона предоставляет содержимое, а другая получает его. Для демонстрации обязанностей и требований к обеспечению прозрачности, которые могут быть у разных заинтересованных сторон, в 6.3 представлен перечень, содержащий для каждой из них:

- цель обеспечения прозрачности;
- выбранные элементы прозрачности;
- ссылку на раздел, в котором рассматривается этот элемент.

Информация, представленная в 6.3, приводится исключительно в качестве примеров для иллюстрации:

- вариативности потребностей заинтересованных сторон в сфере ИИ;
- роли различных заинтересованных сторон в предоставлении требуемых элементов таксономии.

Приведенные примеры не являются исчерпывающими и не охватывают все возможные роли заинтересованных сторон и все их потенциальные потребности, а также аспекты, которые также могут варьироваться в зависимости от целевого использования и контекста конкретной *СИИ*.

6.2 Цели и задачи обеспечения прозрачности

Примеры целей и задач обеспечения прозрачности включают:

- право на информацию — владение информацией о том, что люди взаимодействуют с *СИИ*, а также владение достаточной информацией о содержании синтеза ИИ, возможностях и ограничениях *СИИ*, а также о правах затронутых лиц. Это право может включать в себя автономные решения, взаимодействие с *СИИ*, содержимое, генерируемое ИИ, и соблюдение законодательства;
- надлежащее использование — помогает пользователям понимать обязательства, варианты правильного и недопустимого использования *СИИ*, тем самым позволяя пользователям надлежащим образом использовать *СИИ*. Надлежащее использование может включать в себя законное использование *СИИ*, правильное использование *СИИ*, предотвращение неправомерного использования и снижение рисков;

- прослеживаемость — данная цель обеспечения прозрачности реализуется на протяжении всего жизненного цикла *СИИ* (проектирование и разработка, верификация и валидация, развертывание, текущий контроль и т. д.), чтобы помочь заинтересованным сторонам лучше использовать *СИИ* и управлять ей.

Следует отметить, что описанные здесь цели не являются взаимоисключающими (т. е. возможно, что элемент таксономии одновременно удовлетворяет нескольким целям обеспечения прозрачности).

6.3 Отдельные роли заинтересованных сторон в обеспечении прозрачности

Примеры возможных обязанностей и требований к обеспечению прозрачности различных пар заинтересованных сторон:

a) от поставщика ИИ к стороне, развертывающей ИИ:

1) с целью обеспечения права на информацию содержимое прозрачности может включать основную информацию о системе (см. 8.2);

2) с целью обеспечения надлежащего использования содержимое прозрачности может включать:

- применимость (см. 8.4);
- обзор технических характеристик (см. 8.5);
- доступ к внутренним компонентам (см. 8.6);

3) с целью обеспечения прослеживаемости содержимое прозрачности может включать:

- основную информацию о системе (см. 8.2);
- качество и показатели производительности (см. 8.7);

b) от поставщика ИИ к аудитору ИИ с целью обеспечения прослеживаемости содержимое прозрачности может включать:

- доступ к внутренним компонентам (см. 8.6);
- местоположение обработки и хранения данных (см. 8.2, 8.5.5);

c) от поставщика ИИ к релевантным уполномоченным органам с целью обеспечения прослеживаемости содержимое прозрачности может включать:

- соответствие применимым требованиям законодательства или действующих стандартов (см. 8.2);
- местоположение обработки и хранения данных (см. 8.2, 8.5.5);

d) от поставщика ИИ к конечному пользователю ИИ, аудиторам ИИ с целью обеспечения прослеживаемости содержимое прозрачности может включать аспекты используемого набора данных (см. 9.5), см. также *ГОСТ Р ИСО/МЭК 42001—2024* (таблица А.1, пункт А.7);

e) от стороны, развертывающей ИИ, к конечному пользователю ИИ:

1) с целью обеспечения права на информацию содержимое прозрачности может включать основную информацию о системе (см. 8.2);

2) с целью обеспечения надлежащего использования содержимое прозрачности может включать:

- применимость (см. 8.4);
- обзор технических характеристик (см. 8.5);
- доступ к внутренним компонентам (см. 8.6);

3) с целью обеспечения прослеживаемости содержимое прозрачности может включать:

- основную информацию о системе (см. 8.2);
- качество и показатели производительности (см. 8.7);

f) от стороны, развертывающей ИИ, к субъекту ИИ с целью обеспечения права на информацию содержимое прозрачности может включать:

- основную информацию о системе (см. 8.2);
- сведения о том, какая транзакционная информация сохраняется и как она используется (см. 8.5.4);
- порядок обращения для устранения недостатков или исключения данных о субъекте ИИ (см. 8.2, 8.3.2);

g) от производителя ИИ к поставщику ИИ:

1) с целью обеспечения права на информацию содержимое прозрачности может включать основную информацию о модели (см. 9.2);

2) с целью обеспечения надлежащего использования содержимое прозрачности может включать:

- проектную документацию *СИИ* (см. 8.5.5);
- полезную информацию для оптимизации показателей производительности (см. 8.7.3);
- вопросы устойчивого развития (см. 8.4.3);
- технические характеристики модели (см. 9.4);

3) с целью обеспечения прослеживаемости содержимое прозрачности может включать:

- основную информацию о модели (см. 9.2);
- качество и показатели производительности (см. 8.7);
- использование модели (см. 9.3);
- используемые данные (см. 9.5);
- функциональную корректность (см. 9.6);

h) от производителя ИИ к аудитору ИИ с целью обеспечения прослеживаемости содержимое прозрачности может включать:

- основную информацию о модели (см. 9.2);
- использование модели (см. 9.3);
- информацию о конкретных компонентах системы, например релевантных технических характеристиках модели или моделей, входящих в систему ИИ (см. 9.4);
- используемые данные (см. 9.5);
- функциональную корректность (см. 9.6);

i) от производителя ИИ к релевантным уполномоченным органам с целью обеспечения прослеживаемости содержимое прозрачности может включать:

- основную информацию о модели (см. 9.2);
- использование модели (см. 9.3);
- используемые данные (см. 9.5);
- функциональную корректность (см. 9.6);

j) от поставщика данных к производителю ИИ:

1) с целью обеспечения права на информацию содержимое прозрачности может включать основную информацию о наборе данных (см. 10.2);

2) с целью обеспечения надлежащего использования содержимое прозрачности может включать:

- свойства данных (см. 10.4);
- предметную область и назначение набора данных (см. 10.5);
- обслуживание наборов данных (см. 10.9);

3) с целью обеспечения прослеживаемости содержимое прозрачности может включать:

- основную информацию о наборе данных (см. 10.2);
- происхождение данных (см. 10.3).

См. также табличное представление данных примеров в приложении *А*.

Примечание — Поставщик ИИ включает поставщика услуг или продуктов ИИ и поставщика платформы ИИ.

Согласно *ГОСТ Р 71476—2024* (пункт 5.19.1), организация или лицо могут выполнять несколько ролей или субролей. Например, организация, занимающаяся настройкой *СИИ*, может быть клиентом ИИ при ее приобретении, а также разработчиком *СИИ* при ее настройке.

7 Таксономия контекста

7.1 Общие положения

В разделе описываются элементы таксономии, которые предоставляют информацию о контексте разработки и использования *СИИ*, подразделяемом в настоящем стандарте на социальный и экологический контекст.

7.2 Социальный контекст

7.2.1 Общие положения

7.2.1.1 Основные социальные аспекты раскрытия информации

В настоящем подразделе описываются элементы таксономии, которые предоставляют информацию о каждом из социальных аспектов и их применимости к ролям заинтересованных сторон в сфере ИИ, представленным в *ГОСТ Р 71476—2024* (подраздел 5.19). Социальные аспекты *СИИ* рассматриваются в [*13*].

СИИ встроены в свой социальный контекст и взаимодействуют с ним: отдельные лица, группы и общество в целом воздействуют на них и, в свою очередь, подвергаются их влиянию. По мере расширения использования *СИИ* становится все более важным оценивать последствия взаимного воздействия, как положительные, так и отрицательные, чтобы направлять его в позитивное русло и минимизировать вред. В настоящем подразделе, посвященном социальному контексту, описываются различные элементы таксономии, которые могут помочь заинтересованным сторонам в решении данной задачи на всех стадиях жизненного цикла *СИИ*.

7.2.1.2 Раскрытие информации на ранних стадиях жизненного цикла

Следующая базовая информация актуальна на начальной стадии и может помочь в определении требований, а также обеспечить основу для оценки воздействия и рисков на последующих стадиях жизненного цикла:

a) целевое назначение системы;

b) место или места предполагаемого развертывания (географическое местоположение);

c) контекст развертывания (например, в рамках коммуникационной платформы на рабочем месте или в социальных сетях с преимущественно подростковой аудиторией);

d) заинтересованные стороны, их роли и степень вовлеченности на различных стадиях жизненного цикла;

e) перечень требований законодательства с учетом местоположения поставщика СИИ, а также предполагаемого места ее развертывания и контекста.

7.2.1.3 Документирование социальных рисков и требований

На основе базовой информации можно идентифицировать и документировать ряд требований и рисков для СИИ. К ним могут относиться:

a) культурные и физические характеристики, способные влиять на показатели производительности *СИИ* и имеющие значение для ее обучения и оценки. Возможными примерами являются формы вежливости, которые могут являться важным аспектом для системы машинного перевода, или культурные привычки, которые могут влиять на программное обеспечение по распознаванию эмоций;

b) используемые языки или языковые подмножества (например, социолекты, диалекты, варианты языка, регистры). Например, это может влиять на системы обработки естественного языка: система может поддерживать одни языки, но не другие. Это также определяет, на каком языке предоставляется документация или обучение;

c) группы риска (см., например, [*16*] и *[17]*). Понимание исторически сложившихся предубеждений и характеристик, дискриминация по которым законодательно запрещена, помогает определить, какие группы подвержены риску нежелательной предвзятости (например, приводящей к дискриминации) со стороны *СИИ*. Такие группы могут быть приняты во внимание при проведении оценки воздействия и в рамках обсуждений с заинтересованными сторонами по вопросам требований к системе, включая требования справедливости и доступности;

d) конкретные обстоятельства, при которых использование *СИИ* может привести к рискам для отдельных лиц;

e) действия пользователей, способные повлиять на поведение и показатели производительности системы.

7.2.1.4 Сведения о воздействии на общество

Прозрачность в отношении социального контекста может также включать сведения о соответствующих оценках воздействия. Эти оценки могут охватывать как положительные, так и отрицательные воздействия (более подробный анализ см. в [*18*]). К соответствующей информации, которая может быть предоставлена, относятся:

a) дата оценки воздействия;

b) стадии жизненного цикла *СИИ*, подлежащие оценке, а также версия и конфигурация оцениваемой *СИИ*;

c) ФИО лиц и/или название организаций, отвечающих за оценку, и их контактные данные;
d) краткое изложение результатов и ссылка на полный текст документа;
e) последствия оценки и целевая аудитория;
f) даты предыдущих оценок со ссылками на предыдущие оценки;
g) периодичность проведения будущих (запланированных) оценок;
h) аспекты, охватываемые оценкой, включая такие элементы, как:
1) здоровье;
2) безопасность;
3) контроль со стороны человека и автономность;
4) подотчетность;
5) нежелательная предвзятость и дискриминация; например:
- определение групп риска и обоснование;
- принятые меры по устранению нежелательной предвзятости и их обоснование;
- используемые критерии принятия и их обоснование;
- тесты на предвзятость, выполненные на обучающих данных: описание методов и результатов. Например, в случае систем на основе МО, выполняющих задачи классификации и регрессии, некоторые методы выявления нежелательной предвзятости в обучающих данных описаны в [*17*] (раздел 5);
- тесты на предвзятость, выполненные на выходных данных системы: описание методов и результатов;
- любые решения о продолжении внедрения системы с учетом результатов обработки нежелательной предвзятости в отношении критериев принятия и их обоснование;
6) защита данных, в том числе персональных;
7) кибербезопасность.

7.2.1.5 Сведения о мерах снижения вреда

К социальному контексту также относятся сведения о мерах, принимаемых для улучшения показателей производительности *СИИ* и снижения потенциального вреда, например:

a) перечень реализованных мер (на уровне аппаратного и/или программного обеспечения) для устранения источников вреда, максимального сокращения остаточного вреда и минимизации потенциального остаточного вреда. Раскрытие может включать информацию о мерах:
1) по защите данных или моделей от краж или взлома;
2) по защите *ПДн* (например, федеративное обучение в системе МО);
3) по снижению риска нежелательной предвзятости и дискриминации. В этом случае может предоставляться более подробная информация, например:
- какие методы устранения нежелательной предвзятости используются на различных стадиях жизненного цикла (подробную информацию о классификации и регрессии систем МО см. в [*17*]);
- определения справедливости и критерии принятия или используемые пороговые значения, а также их обоснование;
4) применяемые процедуры мониторинга, включая процедуры ведения журналов о событиях и мониторинга производительности;
5) выбранные показатели робастности и производительности, а также результаты измерений;
6) определение подходящих показателей тестирования и процессов тестирования;
b) принятые остаточные риски (связанные с обеспечением справедливости или другие риски и нежелательные результаты) и их обоснование;
c) описание мер, принятых организацией, таких как:
1) виды информации, раскрываемой внутри организации и за ее пределами (включая данные об остаточных источниках вреда или потенциального вреда);
2) четкое распределение обязанностей на различных стадиях жизненного цикла;
3) стороны, ответственные и подотчетные за функционирование системы ИИ;
4) механизмы обратной связи и обмена информацией об инцидентах или сбоях системы;
5) меры избыточности для предотвращения «единых точек отказа» в деятельности по надзору;
6) управление изменениями, включая обновления системы и другие меры, связанные с техническим обслуживанием;
7) иные принимаемые организацией меры, направленные на устранение источников вреда или снижение потенциального вреда;

d) описание принимаемых мер контроля со стороны оператора и человека. К ним могут относиться:

1) требования к навыкам и компетенциям операторов-людей;

2) выходные данные системы, содержащие пояснения, достаточные для принятия обоснованного решения пользователем или оператором;

3) инструкции по использованию;

4) проектирование интерфейсов между человеком и *СИИ* с учетом принципов когнитивной эргономики, адаптированных под различные потребности, навыки и компетенции операторов, а также уровень когнитивной вовлеченности в задачу (для предотвращения автоматизированной предвзятости);

5) целенаправленное обучение, включая желаемый уровень грамотности в сфере ИИ, которого необходимо достичь (например, обучение, дающее как минимум базовые знания о *СИИ*, их функционировании, различных типах продуктов и применений, об их рисках и преимуществах).

7.2.2 Трудовые отношения

Компании все чаще инвестируют в *СИИ* для мониторинга, анализа и принятия решений с целью информировать, консультировать, дополнять, а в некоторых случаях и полностью автоматизировать принятие решений.

Работники могут выполнять различные роли заинтересованных сторон на протяжении жизненного цикла *СИИ*:

- поставщик ИИ, пользователь ИИ, партнер ИИ или аудитор ИИ, учитывающие требования прозрачности для выполнения своей работы;

- клиент ИИ, который может в рамках процесса закупки запросить надлежащие элементы обеспечения прозрачности для используемых продуктов;

- субъект ИИ, когда работники непосредственно подвержены влиянию *СИИ* в своей профессиональной деятельности, например при подборе персонала, оценке, организации работы и использовании инструментов управления персоналом.

Заинтересованные в ИИ стороны, участвующие в жизненном цикле *СИИ* (включая работников в их различных ролях заинтересованных в ИИ сторон), могут обеспечивать доступ к информации и обучению, способствующим удовлетворению потребностей заинтересованных сторон в прозрачности. В этом контексте может раскрываться следующая информация:

a) перечень заинтересованных сторон, которые были проинформированы и с которыми были проведены консультации, или они каким-либо другим образом были вовлечены в разработку или развертывание *СИИ*; а также описание относительной роли заинтересованных сторон и степень их влияния на различные стадии жизненного цикла *СИИ*;

b) описание основных каналов коммуникации с персоналом и его представителями для их информирования о внедрении *СИИ*;

c) описание механизмов и процедур, применяемых для выявления и снижения случаев нежелательной предвзятости и возможной дискриминации работников. Эти процедуры могут также включать специальный аудит обучающих и эксплуатационных данных;

d) разъяснение процедур, применяемых для рассмотрения возможных случаев дискриминации или нежелательной предвзятости, возникающих в результате использования *СИИ*. Например, это могут быть имя и контакты уполномоченного по правам человека или отдела, занимающегося рассмотрением подобных случаев;

e) разъяснение процедур, обеспечивающих защиту данных, создаваемых работниками, и данных о работниках, от доступа или распространения среди неуполномоченных сторон (как внутри, так и вне организации);

f) описание порядка доступа работников к своим *ПДн*, а также их исправления или удаления по запросу;

g) объяснение работникам и их представителям процедур, используемых для мониторинга деятельности работников, и мер, принимаемых для предотвращения неправомерного наблюдения и нарушений неприкосновенности частной жизни;

h) пояснение степени автоматизации (см. ГОСТ Р 71476—2024, таблица 1) любой *СИИ*, используемой для формирования рекомендаций, которые могут оказать неблагоприятное воздействие на работников;

i) описание процедур, гарантирующих эффективный человеческий надзор за любыми решениями, которые могут оказать существенное негативное воздействие на работников;

j) четкий и понятный перечень правил об ответственности и обязательствах организации или разработчиков, предоставляемый всем работникам и их представителям;

k) разработка плана по внедрению программы цифровой грамотности в сфере ИИ для работников, затронутых внедрением и использованием *СИИ* в организации.

7.2.3 Потребности потребителя

7.2.3.1 Общие положения

Потребители, выступая в роли клиентов ИИ или пользователей ИИ, являются одной из ключевых заинтересованных сторон, участвующей в принятии *СИИ*. Удовлетворение их потребностей в прозрачности имеет важное значение для устойчивого развития и распространения технологий ИИ.

Для обеспечения защиты потребителей, а также удовлетворения потребностей клиентов ИИ, связанных с использованием *СИИ*, на различных стадиях жизненного цикла *СИИ* могут применяться соответствующие меры по обеспечению прозрачности.

7.2.3.2 Документация до развертывания системы искусственного интеллекта

Например, на ранних стадиях жизненного цикла полезная документация может включать:

a) информацию о целевом назначении *СИИ* в соответствии с 8.4.2;

b) анализ возможных рисков и воздействий, которые могут возникнуть в результате:

1) использования клиентом ИИ *СИИ* в соответствии с ее рекомендуемым использованием;

2) использования клиентом *СИИ* вне рекомендуемого использования;

3) злонамеренного использования системы третьими сторонами.

Подобная документация может способствовать реализации конструктивных особенностей и мер по устранению или снижению частоты, серьезности и вероятности причинения вреда.

7.2.3.3 Документация после развертывания системы искусственного интеллекта

После развертывания *СИИ* и начала ее эксплуатации клиентам ИИ может быть полезна следующая информация:

a) руководство по установке и настройке для конечного пользователя;

b) инструкции по эксплуатации или руководство по использованию;

c) информация по техническому обслуживанию, включая:

1) ожидаемый срок службы *СИИ*,

2) необходимые меры по техническому обслуживанию и уходу для обеспечения надлежащего функционирования,

3) сведения об ожидаемой продолжительности поддержки *СИИ* (включая аппаратное и программное обеспечение);

4) периодичность обновлений;

d) рекомендуемое и недопустимое использование;

e) потенциальные опасности, возникающие при использовании *СИИ* вне рекомендуемого использования;

f) для систем на основе МО, информация о неопределенности, присущей выходным данным систем МО;

g) информация о сборе данных (где это применимо), включая их использование, хранение, защиту, совместное использование и контроль потребителей за своими данными;

h) сервисные центры;

i) горячие линии или контактная информация для предоставления отзывов и жалоб;

j) гарантии и возможные обязательства по поддержке и обновлениям;

k) информация о сертификации продукта и организации;

l) информация о выводе продукта из эксплуатации.

7.3 Экологический контекст

7.3.1 Общие положения

СИИ используют вычислительные ресурсы в разной степени и на разных стадиях.

Экологическая устойчивость *СИИ* напрямую связана с воздействием технологического развития на окружающую среду (например, углеродный и водный след, более подробный анализ см. также в [*18*]), а также с тем, как можно снизить негативное воздействие на окружающую среду, связанное с развитием *СИИ* (см. [*19*]).

Воздействие *СИИ* на окружающую среду может быть описано с помощью соответствующих раскрытий информации о прозрачности, таких как:

a) оценка воздействия на окружающую среду, или оценка воздействия на жизненный цикл (см. [*20*]); этот элемент может включать в себя такую информацию, как:

1) краткое изложение результатов;
2) ссылка на полный текст отчета;
3) сведения об ответственном за проведение оценки: ФИО и контактные данные;
4) дата проведения оценки;
5) версия и конфигурация системы, которая была оценена;
6) периодичность оценки, со ссылкой на предыдущие оценки;
7) планируемая дата следующей оценки.

7.3.2 Конкретная раскрываемая информация о влиянии на окружающую среду

В частности, следующие элементы таксономии могут рассматриваться как часть общей оценки воздействия на протяжении жизненного цикла или как самостоятельные информационные элементы:

a) энергопотребление: элемент таксономии предоставляет оценку энергии, потребляемой *СИИ*. В свою очередь его можно разделить:

1) на энергию, потребляемую в процессе разработки, например:
   - показатели, используемые для измерения энергопотребления;
   - методика измерения энергопотребления;
   - вклад различных компонентов системы в общее потребление;
   - при необходимости, основание для повторного использования субкомпонентов, включая модели, данные, программные инструменты и методы (но не ограничиваясь ими) (см. *ГОСТ Р 71476—2024*, рисунок 6), с разделением на субкомпоненты, разработанные исключительно для конкретной СИИ, и субкомпоненты с более широкой областью применения. Это особенно актуально, например, для моделей, разработка которых требует значительных вычислительных ресурсов, но которые могут быть повторно использованы во множестве приложений: в зависимости от объема повторного использования и потребностей в энергии в ходе эксплуатации, такая ситуация может приводить к общему ограниченному энергопотреблению;
2) оценку энергопотребления при развертывании (см. *ГОСТ Р ИСО 50001*). К ней могут относиться:
   - показатели, используемые для измерения энергопотребления;
   - методика измерения энергопотребления;
   - вариативность энергопотребления для разных методов развертывания;
   - указание на то, для какого метода развертывания предполагается оценка энергопотребления;
3) рекомендации по энергоэффективным условиям эксплуатации;
4) условия, приводящие к повышенному энергопотреблению;
5) использование центра обработки данных, в том числе:
   - показатели, используемые для измерения энергопотребления;
   - методика измерения энергопотребления;
6) предпринимаемые шаги по оптимизации эффективности моделей, такие как использование разреженных моделей;

b) элементы устойчивого развития, такие как:

1) углеродный след: общий углеродный след *СИИ* во время ее разработки и как он определяется;
2) водный след: общий водный след *СИИ* и как он измеряется;
3) модульность и цикличность:
   - аппаратные субкомпоненты системы, которые были повторно использованы, перепрофилированы или в значительной степени состоят из переработанных материалов (в сравнении с общим объемом материала, используемого системой);
   - аппаратные субкомпоненты системы, которые могут быть повторно использованы или перепрофилированы в будущем (в сравнении с общим объемом материала, используемого системой);
   - программные субкомпоненты, которые могут быть повторно использованы или перепрофилированы в других *СИИ*, а также их ожидаемая или наблюдаемая частота повторного использования;
   - доля первичных, повторно используемых или перепрофилированных материалов, а также переработанных материалов в процентном отношении к общему объему материалов для аппаратных и программных компонентов;

c) воздействие аппаратного обеспечения *СИИ* на окружающую среду. Например, полезная информация может относиться к производству продуктов и вредным выбросам в процессе производства (подробнее см. *[21]*, подраздел 11.4). Другие элементы, относящиеся к аппаратному обеспечению *СИИ*, которые могут быть раскрыты, включают в себя:

1) использование процессоров и систем, оптимизированных для машинного обучения;

2) коэффициент энергоэффективности (PUE, см. [*8*]) местоположения или местоположений, в которых развернута *СИИ*. Также может указываться период, за который рассчитан коэффициент энергоэффективности и частота обновления данного коэффициента;

3) возможность оптимизации схем размещения высоконагруженных *СИИ* в местоположениях с низким углеродным следом;

d) техническое обслуживание *СИИ*:

1) как долго будет осуществляться поддержка *СИИ* (включая аппаратное и программное обеспечение), в том числе ее обновление (например, для противодействия навязываемому обновлению ПО и оборудования),

2) наличие сервисных центров для ремонта оборудования;

3) наличие запасных частей, доступ к исходному коду и наборам данных, доступ к проектным решениям аппаратного обеспечения (т. е. использование «открытого аппаратного обеспечения»), ремонтопригодность, модернизируемость, модульность проектного решения и расширяемость ПО;

e) план вывода из эксплуатации. Этот пункт содержит информацию об обращении с *СИИ* по окончании ее жизненного цикла и может включать:

1) указания на то, что происходит с данными;

2) указания на то, что происходит с ПО;

3) инструкции по работе с различными аппаратными компонентами, включая, по возможности, применимые программы:

- для модернизации и ремонта компонентов;
- повторного использования;
- ремонта;
- переработки;
- вывода из эксплуатации;

4) информацию о том, что происходит с выведенным из эксплуатации оборудованием;

f) информация о цепочке поставок, например об усилиях по управлению цепочкой поставок (см. *[21]—[24]*).

7.3.3 Организационное раскрытие информации

Описанные в 7.3.2 элементы таксономии могут быть дополнены организационными и процедурными раскрытиями информации в отношении усилий организаций по ограничению воздействия *СИИ* на устойчивость окружающей среды:

a) подтверждение внедрения экологически ответственных закупок, таких как внутренние процедуры оценки поставщиков продуктов и услуг ИИ с учетом аспектов их воздействия на окружающую среду;

b) подтверждение внедрения методов и принципов экономики замкнутого цикла на протяжении всего жизненного цикла *СИИ* (см. *[25]*):

1) сокращение отходов;

2) повторное использование продукции;

3) переработка материалов.

Раскрытие информации о прозрачности также может включать обоснованно прогнозируемое воздействие *СИИ*, выходящее за рамки энергии и материалов, используемых при ее непосредственной эксплуатации (более подробную информацию об оценке воздействия *СИИ* см. в *[18]*). Косвенные эффекты возникают в результате применения ИИ и могут быть как положительными (например, возможная экономия энергии благодаря технологии «умных» энергосетей), так и отрицательными (например, неустойчивые изменения в моделях потребления). По возможности производитель ИИ может предоставлять информацию о потенциальном возникновении косвенных эффектов в результате развертывания *СИИ*.

8 Таксономия системы

8.1 Общие положения

С технической точки зрения *СИИ* может состоять из нескольких компонентов различной природы. Например, голосовой помощник обычно строится как комбинация компонентов распознавания речи, семантического анализа, поиска информации, генерирования (для формулирования ответа) и синтеза голоса (для его воспроизведения). Более подробную информацию см. в *[26]*.

СИИ также может включать компоненты, построенные вокруг модели ИИ, но которые сами по себе не относятся к дисциплине ИИ. Например, создание миниатюр для постов в социальных сетях обычно основано на модели обнаружения объектов, позволяющей идентифицировать основный элемент изображения, но также включает в себя код для фактической генерации изображения на основе выбранного фокуса и для его воспроизведения. Ожидания по прозрачности также могут распространяться на эту более широкую систему, а не только на модель обнаружения объектов. Если *СИИ* состоит из компонентов, которые можно четко разделить, таксономия может быть применена к каждому из них. Это особенно полезно, если компоненты получены от разных поставщиков или предполагается их повторное использование в другом контексте. Как указано в 5.2, документация системы может ссылаться на документацию компонентов.

Модели ИИ, встроенные в *СИИ*, могут в равной степени основываться как на машинном обучении, так и на других технологиях ИИ. Например, система машинного перевода внутри может быть основана на машинном обучении или на системе, основанной на правилах (которая действительно применяется в некоторых высокопроизводительных системах), не представляя никакой разницы для пользователя. Таким образом, если некоторые элементы таксономии могут зависеть от используемой технологии (например, от обучающих данных, используемых для моделей машинного обучения), то ряд других компонентов таксономии применим независимо от внутренней работы системы. Например, это относится к элементам таксономии, связанным с оценкой, которые скорее относятся к документированию качества выходных данных системы, чем к способу их получения.

8.2 Основные сведения

Прозрачность *СИИ* включает в себя основные элементы информации, аналогичные тем, что требуются для любой другой информационной системы. К ним могут относиться:

a) название: название *СИИ* в ее текущей конфигурации;

b) производитель ИИ или интегратор *СИИ*: название и контактные данные организации, которая произвела или собрала *СИИ* в ее текущей конфигурации. Везде, где это возможно, необходимо указать ФИО и контактные данные местных или региональных уполномоченных представителей организации;

c) поставщик услуги или продукта ИИ, если отличается от производителя ИИ, название и контактные данные организации, которая делает *СИИ* доступной для использования в качестве (части) продукта или услуги. Везде, где это возможно, необходимо указать ФИО и контактные данные местных или региональных уполномоченных представителей организации;

d) версия: номер версии или сборки *СИИ* в ее текущей конфигурации. Данная информация может быть дополнена в разделе об управлении конфигурацией сведениями о том, чем текущая версия отличается от предыдущих версий;

e) информация о лицензировании: по какой лицензии, если таковая имеется, может использоваться *СИИ*;

f) местоположение: раскрытие данной информации, например о геополитической юрисдикции, может быть полезно, если *СИИ* развернута в одном месте;

g) заявления о соответствии. *СИИ* может сопровождаться одним или несколькими заявлениями о соответствии. Информация, подлежащая раскрытию, включает:

1) объект оценки соответствия;

2) область аттестации: например, если объектом соответствия является *СИИ*, то оценивается версия и конфигурация *СИИ*;

3) заданные требования: например, международный стандарт или регламент, который был применен;

4) орган по оценке соответствия: включает в себя указание на то, является ли аттестация внутренней (выполненной первой стороной), внешней (выполненной второй стороной) или независимой

(выполненной третьей стороной), а также может включать название и контактные данные ответственной организации;

h) обратная связь: название и контактные данные организации, ответственной за обработку вопросов и отзывов по *СИИ*;

i) документация: тип и объем предоставляемой документации по системе может варьироваться в зависимости от целевого получателя. Элемент таксономии может включать сведения:

1) о том, где можно найти полную документацию по *СИИ*;

2) типе доступной документации (например, справочные документы, интерактивные онлайн-дисплеи или машиночитаемые файлы);

3) структуре документации;

4) наличии связанных с *СИИ* научных результатов, включая детали научных публикаций.

Название и версия *СИИ* релевантны уже на начальной стадии. Другие элементы таксономии приобретают значимость после развертывания системы.

8.3 Организационные процессы

8.3.1 Общие положения

Прозрачность распространяется не только на продукт ИИ или услугу ИИ, которые поставщик ИИ предлагает клиентам. Она также имеет отношение к организационным процессам, которые связаны и характерны для определенной *СИИ*. Если рассматривать жизненный цикл *СИИ*, то процессы, обеспечивающие ее создание, заслуживают не меньшего, а в некоторых случаях и большего внимания, чем сам производственный процесс. Это особенно актуально для вариантов использования с повышенным уровнем риска.

8.3.2 Стратегическое управление

Согласно *[27]*, прозрачность достигается путем предоставления заинтересованным сторонам организации соответствующего уровня явной и понятной информации о деятельности, процессах и принятии решений. Прозрачность организационных процессов может быть достигнута за счет:

a) документального подтверждения принятия общих или специфических для ИИ процессов обеспечения прозрачности и руководств *[27]*, по управлению защитой ПДн *[2]*, по управлению информационной безопасностью *[28]* и по управлению рисками *[1]*;

b) документации по *СИИ*, содержащей сведения:

1) о контексте *СИИ* (в соответствии с разделом 7);

2) *СИИ* в целом (в соответствии с разделом 8);

3) дополнительной документации по компонентам *СИИ*: моделям (см. раздел 9) и данным (см. раздел 10);

c) документации по жизненному циклу *СИИ*, например с использованием структуры, представленной в *ГОСТ Р 71539*;

d) документации о действующих механизмах обратной связи и подачи жалоб заинтересованными сторонами. К ним могут относиться:

1) ссылка на адрес веб-сайта, на котором можно оставить отзыв или жалобу;

2) описание пользовательского интерфейса, позволяющего пользователям ИИ оставлять отзывы;

3) ФИО и контактные данные уполномоченного по правам человека или отдела по работе с жалобами.

8.3.3 Система менеджмента

ГОСТ Р ИСО/МЭК 42001 — это стандарт системы менеджмента ИИ, содержащий требования и рекомендации по созданию, внедрению, поддержанию в рабочем состоянии и постоянному улучшению системы менеджмента ИИ в организации. Стандарт определяет набор средств контроля и передовых практик, направленных на повышение качества *СИИ* и снижение рисков. Этот элемент прозрачности может описывать, каким образом организация внедрила набор средств контроля и передовых практик, изложенных в *ГОСТ Р ИСО/МЭК 42001*, включая (при наличии) заявление о соответствии требованиям этого стандарта.

8.3.4 Управление рисками

В *[1]* содержится руководство для любых организаций независимо от размера, типа и рода деятельности, которые управляют рисками, связанными с разработкой и использованием ИИ. В *[1]* также содержатся рекомендации о том, как организации могут интегрировать управление рисками в свою деятельность, основанную на ИИ.

Организации могут предоставить подробную информацию о том, как было реализовано это руководство, включая (но не ограничиваясь):

a) дату принятия руководства по управлению рисками;

b) дату обновления системы управления рисками в соответствии с данным руководством;

c) дату последней проверки эффективности внедрения руководства по управлению рисками.

Если организация применяет другие руководства по управлению рисками, она также может указать и сослаться на формализованные, стандартизированные и публично доступные руководства по управлению рисками *СИИ*.

8.3.5 Управление качеством

Высококачественные программные продукты и компьютерные системы имеют решающее значение для заинтересованных сторон. Модели качества, требования к качеству, измерение качества и оценка качества стандартизированы в рамках серии стандартов о требованиях и оценке качества систем и программных средств (SQuaRE) (см. *ГОСТ Р ИСО/МЭК 25000*). В *[12]* и *[29]* представлена модель, характеристики качества и руководство по оценке качества, которые позволяют организациям определять и документировать качество продукта *СИИ* и качество его использования.

Настоящий элемент таксономии может описывать, как организация внедрила руководство и процессы управления качеством, изложенные в *[12]* и *[29]*.

8.4 Применимость

8.4.1 Общие положения

В данной части таксономии содержится информация о целевом назначении, возможностях, функциональных ограничениях, а также рекомендуемом и недопустимом использовании *СИИ*. Информация предназначена для широкого круга заинтересованных сторон ИИ, как определено в *ГОСТ Р 71476*. Такая информация может быть полезна, например, клиентам ИИ или пользователям ИИ при выборе подходящего решения ИИ для ответственного удовлетворения их потребностей.

8.4.2 Целевое назначение

Настоящий элемент таксономии предоставляет информацию о целях, для которых была разработана и протестирована *СИИ*. Элемент сфокусирован на целях и задачах системы (например, на потребностях пользователя, которые она может удовлетворить) и описывает, как технология ИИ способствует достижению этих целей и задач.

8.4.3 Возможности

Настоящий элемент таксономии описывает возможности системы и то, как они способствуют удовлетворению потребностей заинтересованных сторон. Возможности представляют собой функциональные и нефункциональные характеристики, для достижения которых *СИИ* разработана, протестирована и валидирована, и которые способствуют обработке данных в соответствии с целевым назначением системы. Назначение системы может быть различных типов: выполнение определенной задачи (подтвержденные возможности системы, например, машинный перевод), аспекты выполнения данных задач (характеристики того, как она их выполняет, например способность обрабатывать сарказм в машинном переводе) и другие аспекты обработки, выходящие за рамки задач (например, способность адаптировать машинный перевод к новым возникающим темам и лексике, в соответствии с характеристикой качества «функциональная адаптивность», определенной в *[12]*).

Полезная информация может включать:

a) описание возможностей;

b) подтверждение или критерии, используемые для их оценки (например, с помощью ссылок на другие разделы стандарта, такие как информация, указанная в подразделе 8.7);

c) взаимосвязь возможностей между собой и с удовлетворением потребностей.

8.4.4 Функциональные ограничения

Настоящий элемент таксономии содержит информацию об ограничениях возможностей *СИИ*. К ней могут относиться ограничения, связанные с точностью, робастностью и кибербезопасностью. Ограничения могут включать подробную информацию о контекстах использования и об обстоятельствах, которые могут привести к рискам, например для физических лиц.

8.4.5 Рекомендуемое использование

Настоящий элемент таксономии предоставляет информацию о рекомендуемом использовании, для которого была разработана и протестирована *СИИ*. Элемент таксономии описывает, кто может использовать систему, для какой задачи или назначения, а также при каких обстоятельствах (географическое местоположение, физическая среда, на каких типах устройств).

8.4.6 Недопустимое использование

Настоящий элемент таксономии содержит информацию о недопустимых вариантах использования *СИИ*. К такому недопустимому использованию относяся:

a) преждевременное использование *СИИ* с учетом текущего уровня развития технологии;

b) варианты использования, которые специально не поддерживаются;

c) варианты использования, которые запрещены соответствующими требованиями законодательства (действующими в настоящее время или находящимися в процессе подготовки к вступлению в силу);

d) иные варианты использования, которые производитель ИИ или поставщик продукта/услуги ИИ считают необходимым исключить, например на основании оценки рисков.

Список недопустимых вариантов использования должен регулярно пересматриваться в соответствии с изменяющимися технологиями и правовыми рамками.

8.5 Обзор технических характеристик

8.5.1 Общие положения

Настоящий элемент таксономии описывает, как функционирует *СИИ*. Это описание может включать различные компоненты системы, целевую рабочую среду системы и то, как пользователи и другие заинтересованные стороны могут взаимодействовать с системой, включая входные и выходные данные.

В *ГОСТ Р 71476—2024* (раздел 8) представлены функциональные уровни *СИИ* и описывается совокупность технологий, используемых для построения таких систем. В настоящем разделе рассматривается соответствующее подмножество этих уровней.

8.5.2 Ожидаемые входные и выходные данные

Настоящий элемент таксономии описывает характер входных данных, обрабатываемых *СИИ* (например, изображения, текст, числа, сигналы и т. д.), а также характер выходных данных, производимых *СИИ* (например, теги, классы конечных множеств, структурированная информация, изображения, текст, числа). Необходимый уровень информации о входных и выходных данных может включать любую пояснительную информацию (например, текстовое описание, словарь значений, контекстная информация), которая необходима для правильного понимания пользователями ИИ и другими заинтересованными сторонами входных и выходных данных *СИИ*.

8.5.3 Эксплуатационные данные

Настоящий элемент таксономии содержит информацию о критериях и процедуре, используемых для выборки или отбора данных, обрабатываемых развернутой системой в процессе эксплуатации (эксплуатационные данные), из потенциально более широкого набора доступных данных во время работы развернутой системы.

Например, для *СИИ*, предлагающей варианты текста при написании электронных писем, может быть полезно указать, формируются ли предложения на основе ранее отправленных пользователем писем или учитываются все письма в почтовом ящике пользователя, включая как отправленные, так и полученные.

8.5.4 Ведение журнала о событиях и хранении

Настоящий элемент таксономии содержит сведения о любых данных или информации, хранящихся в системе после обработки. Эти сведения включают в себя:

a) информацию о технических журналах о событиях, позволяющих осуществлять мониторинг и анализ функционирования и сбоев системы, в том числе с указанием их местоположения и содержания;

b) информацию о том, хранятся ли эксплуатационные данные для повторного использования в будущем (например, для повторного обучения, статистики или как источника данных) и при каких условиях, включая указание на любую релевантную процедуру обеспечения защиты *ПДн* и конфиденциальности (такую как удаление данных), например в случае персональных или конфиденциальных данных;

c) инструкции пользователей и соответствующие инструменты для сбора, хранения и интерпретации журналов событий.

8.5.5 Декомпозиция системы

СИИ может состоять из различных аппаратных и программных компонентов. Настоящий элемент таксономии описывает различные компоненты, из которых состоит система. Это описание может включать зависимости между компонентами и, при необходимости, версии, используемые при построении системы.

Высокоуровневое описание *СИИ* может включать:

a) используемые модели: различные модели, используемые в *СИИ*, включая ссылки на более подробные описания моделей, приведенные в разделе 9;

b) используемые наборы данных: различные наборы данных, используемые на различных стадиях жизненного цикла *СИИ*, включая ссылки на более подробные описания наборов данных, приведенные в разделе 10;

c) предварительную и последующую обработку: различные преобразования, применяемые *СИИ* до, после или в промежутке между обработкой моделями;

d) поток данных и информации между моделями: указание моделей, выходные данные которых служат входными данными для другой модели, возможно, связанные между собой с помощью программных компонентов, не использующих ИИ, которые выполняют, например, операции преобразования форматов или фильтрации;

e) поток данных между местоположениями: раскрытие информации о том, как данные перемещаются между местоположениями, может быть полезным (например, если местоположения находятся в разных геополитических юрисдикциях);

f) диаграмму системы.

8.5.6 Прикладной программный интерфейс

СИИ может предоставлять точки входа (API) для автоматизированного доступа к своим функциональным возможностям. Доступ к API может быть полезен пользователям, которые хотят интегрировать *СИИ* с другими системами и продуктами. Этот доступ также может быть особенно важен для оценки рисков и контроля качества, а также для сторонних оценщиков и аудиторов, которым требуется проверить показатели производительности системы. Возможные элементы прозрачности включают:

a) набор предоставленных API;

b) набор допустимых форматов входных данных и формат выходных данных;

c) метаданные, информацию, обогащающую выходные данные, и информацию о внутренних процессах системы, которая может быть извлечена через API;

d) условия получения доступа и ограничения на использование;

e) документацию по API;

f) технический тип API (например, REST, JSON API);

g) синтаксис API (например, допустимые ключевые слова для JSON API и их значение);

h) порядок использования API (например, обязательные поля, такие как имя пользователя или токен API).

8.5.7 Человеческий фактор

Настоящий элемент таксономии описывает, как пользователи системы могут взаимодействовать с системой. Он важен для пользователей системы, но также может быть полезен для оценщиков и аудиторов системы. Данный компонент предоставляет описание доступного пользовательского интерфейса (подробную информацию см., например, в *[30]*) и может включать такие элементы, как:

a) тип предлагаемого интерфейса (ручной ввод текста, графический интерфейс пользователя, чат-бот и т. д.);

b) ограничения доступности;

c) контекст использования;

d) доступные руководства пользователя.

8.5.8 Методы развертывания

Информация о методах развертывания может включать:

a) облачные или периферийные вычисления, развертывание на месте;

b) необходимое оборудование;

c) необходимое ПО.

8.5.9 Управление конфигурацией

Настоящий элемент таксономии содержит сведения, помогающие понять состояние текущей версии *СИИ* и то, как в целом осуществляется управление конфигурациями. Многие из описанных здесь элементов не обязательно предназначены для всех заинтересованных сторон (или даже не подходят, например, с точки зрения безопасности системы), но могут быть полезны для внутреннего тестирования и контроля качества, а также могут быть необходимы сторонним аудиторам или оценщикам системы. К данным элементам относятся:

a) версия системы:

1) текущая версия (частично дублирует информацию 8.2);

2) предыдущая версия: в зависимости от того, для кого предназначена информация, это может быть предыдущая внутренняя рабочая версия или, если версия системы уже находится в эксплуатации, предыдущая промышленная версия системы;

3) различия между текущей и предыдущей версией;

b) текущая стадия жизненного цикла: например, вся *СИИ* может попадать под управление конфигурацией на стадии развертывания, при верификации и валидации, в процессе эксплуатации или при выводе из эксплуатации;

c) конфигурация системы: элемент таксономии содержит информацию о требованиях к *СИИ* (например, использование памяти, операционная система) и зависимостях (например, библиотеки ПО, включая конкретную версию или сборку). Возможно предоставление этой информации в определенных форматах, например в удобочитаемом формате для сериализации данных и конфигураций.

8.6 Доступ к внутренним компонентам

Настоящий элемент таксономии соответствует способности заинтересованных сторон получать доступ к внутренним компонентам *СИИ*. Такой доступ может варьироваться в зависимости от стадии жизненного цикла *СИИ*. Он также может различаться для разных типов заинтересованных сторон — например, для аудиторов ИИ и пользователей ИИ могут предусматриваться различные уровни доступа.

Доступ к внутренним элементам *СИИ* может предоставляться для различных целей, например для проверки, лучшего понимания результатов ее работы, для объединения с другой системой (например, в качестве инструмента анализа) или для повторного использования третьей стороной. Такой доступ предоставляется только в рамках ограничений соответствующих лицензий, если таковые имеются (см. 8.2).

Примерами внутренних элементов, доступ к которым может быть рассмотрен, являются:

a) доступ к отдельным компонентам *СИИ* в качестве самостоятельных элементов, которыми можно манипулировать и с которыми можно взаимодействовать напрямую (возможно, через специальные API);

b) доступ к частям исходного кода (например, определенной модели или других частей *СИИ*);

c) доступ к содержимому моделей, в частности:

1) для символьной *СИИ*: список правил, присутствующих в модели, основанной на правилах, или элементы знаний, встроенные в граф знаний;

2) для машинного обучения: параметры модели МО;

3) для эволюционных моделей (см. 9.4.9): можно рассмотреть, является ли такой доступ более актуальным для выпущенной версии модели или для ее текущего содержимого. Это рассмотрение включает в себя потенциальные внутренние базы данных, которые поддерживаются внутри *СИИ* (независимо от их расположения внутри или за пределами модели);

d) доступ к данным и информации, хранящимся в *СИИ* согласно 8.5.4;

e) доступ к внутренним значениям и промежуточным результатам при обработке заданных входных данных:

f) например, доступ к баллу, присвоенному каждому возможному классу, помимо класса, рассчитанного классификатором;

g) если *СИИ* включает компоненты предварительной или последующей обработки, это может обеспечивать доступ к предварительно обработанной версии входных данных или необработанным выходным данным перед применением преобразований последующей обработки;

h) если в *СИИ* используется несколько моделей с потоком данных между ними, целесообразно рассмотреть доступ к элементам данных, которые передаются от одной модели к другой, чтобы лучше понять эффект каждой модели.

8.7 Качество и показатели производительности

8.7.1 Общие положения

Модель качества систем и программного обеспечения определена в *[31]*. Для *СИИ* данная модель расширена в *[12]*. Модели качества определяют потенциально желательные характеристики и подхарактеристики систем и могут использоваться для определения требований и выявления рисков.

Одной из ключевых подхарактеристик качества является функциональная корректность, т. е. степень, в которой система обеспечивает правильные результаты с необходимой степенью точности. Другие характеристики качества, которые не являются специфичными для *СИИ*, такие как время обработки

и эффективность, могут потребовать специальных для ИИ способов отчетности, например если они опираются на специфичные для ИИ показатели.

8.7.2 Процессы верификации и валидации

Настоящий элемент таксономии описывает процессы и методы, которые используются для верификации и валидации *СИИ*.

Примеры полезной информации включают:

a) модульные тесты, которые были выполнены для проверки того, что *СИИ* реализует требуемую функцию. К данной проверке относится перечисление входных данных модульных тестов и ожидаемое соответствующее поведение;

b) тесты интеграции системы, выполненные, если *СИИ* является частью более крупной системы;

c) полевые испытания, проведенные для проверки обеспечения *СИИ* ожидаемых результатов в реальных условиях. К ним относятся описание условий проведения испытания, участие в нем людей-операторов или испытуемых, критерии и меры, используемые для оценки успешности испытания, а также особенности условий проведения испытания, влияющие на репрезентативность будущих результатов (например, воздействие на характеристики данных);

d) пороговые значения функциональной корректности модели, используемые при принятии решения об ее интеграции в систему. Это предполагает наличие документации по функциональной корректности соответствующих моделей (см. 9.6);

e) в случае наличия нескольких моделей в *СИИ* — информация о сквозной оценке функциональной корректности объединенных моделей, документированная таким же образом, как и для отдельных моделей (см. 9.6). При необходимости это может также включать информацию о наборах данных, используемых для оценки (документируемых согласно 9.5).

Процессы верификации и валидации можно сделать прозрачными с помощью двух ключевых документов: плана испытаний и отчета о завершении испытаний.

В плане испытаний объясняется, как конкретные методы тестирования применяются в процессе тестирования для снижения риска, что может быть задокументировано в соответствии с установленными требованиями (см. *[32]*, подраздел 7.2). План испытаний может содержать информацию об интервалах и процедурах повторного проведения конкретных тестов для развернутой *СИИ* с целью постоянного мониторинга качества системы. План испытаний также может содержать набор и обоснование соответствующих показателей.

В отчете о завершении испытаний описываются результаты испытаний, что может быть задокументировано в соответствии с *[32]* (подраздел 7.4).

8.7.3 Измерение времени выполнения

Настоящий элемент таксономии содержит информацию об измерениях времени обработки, которые были проведены в *СИИ*. Компонент содержит:

a) сведения об используемых показателях (например, задержка, пропускная способность или более специализированные показатели в зависимости от функции, выполняемой *СИИ*). Например, в случае *СИИ* для классификации некоторые примеры показателей, связанных с задержкой, пропускной способностью и эффективностью, описаны в *[33]* (подраздел 6.6);

b) условия проведения испытания;

c) входные данные, использованные для измерения, вместе с конечными результатами и доступными данными о погрешности измерений.

8.7.4 Сравнение с альтернативными системами

Настоящий элемент таксономии содержит информацию о количественных и качественных сравнениях, которые были проведены по отношению к другим существующим *СИИ* с аналогичными функциями.

Компонент может включать измерения, выполненные в аналогичных условиях (на одинаковых данных с использованием одинаковых показателей) на этой и других *СИИ*. Также можно сообщать об измерениях, выполненных в частично отличающихся условиях (например, одинаковые показатели на немного отличающихся данных или на подмножестве данных), при условии, что данные различия в условиях указываются вместе с полученными результатами.

9 Таксономия модели

9.1 Общие положения

В разделе описываются элементы таксономии для раскрытия информации о моделях как внутренних компонентах, которые могут входить в состав *СИИ*. *СИИ* может состоять из более чем одной модели; в этом случае каждая модель будет иметь свои собственные данные для раскрытия, как описано в настоящем разделе. Кроме того, возможно, что некоторые модели, интегрированные в систему, ранее были выпущены как самостоятельные *СИИ*.

Пример — Предварительно обученная языковая модель, которая затем может быть настроена и интегрирована в систему анализа настроения.

Модели ИИ в *СИИ* могут в равной степени основываться на различных технологиях ИИ, включая символьный ИИ или машинное обучение. Указанные здесь элементы таксономии могут применяться к моделям ИИ в целом, независимо от технологии, на которой они основаны; или же они могут применяться только к некоторым видам моделей ИИ. Например, как указано далее в 9.5, информация, относящаяся к обучающим данным, релевантна только для моделей, основанных на МО, а информация, относящаяся к тестовым данным, актуальна в целом.

9.2 Основные сведения

Основная информация предназначена для однозначной идентификации конкретной модели и указания ответственных за ее разработку. Полезная информация для раскрытия включает:

a) наименование модели;

b) сведения о разработчиках: ФИО и контактные данные лица или команды, ответственных за разработку модели на текущем этапе;

c) внешнюю документацию: если модель основана на ранее выпущенной *СИИ* — ссылка на документацию этой *СИИ*;

d) номер модели (номер версии, номер сборки или оба номера):

1) текущая версия: версия или номер сборки модели в ее текущем состоянии;

2) предыдущая версия: в зависимости от того, для кого предназначена информация, это может быть предыдущая внутренняя рабочая версия или, если версия модели уже интегрирована в действующую *СИИ*, то предыдущая рабочая версия модели;

3) различия между текущей и предыдущей версией;

e) конфигурацию системы: элемент таксономии содержит информацию о требованиях к модели и зависимостях. Возможно предоставление данной информации в определенных форматах, например в удобочитаемом формате для сериализации данных и конфигураций.

9.3 Использование и взаимодействие с другими моделями

9.3.1 Обработка, выполняемая моделью

Настоящий элемент таксономии описывает обработку, выполняемую моделью в текущей *СИИ*, подробно описывая характер входных и выходных данных модели (например, являются ли они предложениями, документами, записями о банковских счетах, ограничивающими рамками вокруг лиц, тематическими тегами), а также связи между ними (т. е. задача модели).

9.3.2 Зависимость от других моделей

Настоящий элемент таксономии содержит информацию об интерфейсах с другими моделями, которые необходимы для выполнения данной моделью своих функций.

Например, модель может быть адаптирована к особенностям выходных данных другой модели (например, выполняя предварительную обработку) и может некорректно функционировать при замене компонента.

9.3.3 Согласованность с целевым назначением системы искусственного интеллекта

Настоящий элемент таксономии сообщает об известных сходствах и расхождениях между техническими характеристиками модели и целевым назначением *СИИ*, в состав которой входит данная модель.

Примерами являются различия в распределении данных, полный или частичный охват предметной области или проектное решение модели, оптимизированное для аспектов, нерелевантных для использования в *СИИ*.

Другим примером является использование на уровне модели другой целевой переменной в качестве прокси для фактически значимой информации системы (например, если она не поддается прямому наблюдению или количественной оценке). В таком случае подробное описание обоснования выбора данного прокси, а также возможных условий согласования или рассогласования между ними может дать полезную информацию для лучшего понимания последствий использования данной модели.

9.4 Технические характеристики

9.4.1 Тип используемой технологии

Настоящий элемент таксономии предоставляет информацию о типе технологии, на которой основана модель, на высоком уровне (например, на уровне детализации машинного обучения, нейронных сетей, гибридного ИИ, эвристической оптимизации или системы на основе правил).

9.4.2 Признаки, извлеченные из входных данных

Настоящий элемент таксономии описывает, каким образом входные данные подаются на обработку алгоритму внутри модели (см. 9.4.3), например, обрабатываются ли данные непосредственно в их исходном виде, или из них извлекаются признаки, какие именно, комбинируются ли они для формирования сложных признаков.

9.4.3 Алгоритм обработки

Настоящий элемент таксономии содержит информацию о конкретном алгоритме, используемом для обработки входных данных модели.

Пример — Метод опорных векторов, Байесовские сети (включая описание структуры сети), сети с долгой краткосрочной памятью, сверточные нейронные сети, конкретные варианты архитектуры трансформеров, алгоритм A*, логическая сеть Маркова.

9.4.4 Порядок построения модели

Настоящий элемент таксономии описывает процесс создания модели после выбора алгоритмов.

Для моделей, основанных на правилах, описываются критерии, используемые для создания и выбора правил.

Для моделей машинного обучения описываются:

a) функция потерь (или функция вознаграждения в случае обучения с подкреплением), как описано в *[7]* (подпункт 6.5.4.1):

1) примерами являются перекрестная энтропия и дивергенция Кульбака-Лейблера;

2) возможны комбинации нескольких функций потерь (например, для регуляризации или в подходах, основанных на вспомогательных затратах);

b) процедура оптимизации, описанная в *[7]* (пункт 6.5.4):

1) примерами являются такие процедуры, как Adagrad и другие методы градиентного спуска (см. *[7]*, подпункт 6.5.4.2), метод Ньютона (*[7]*, подпункт 6.5.4.3), сопряженный градиент (*[7]*, подпункт 6.5.4.4), оценка по методу максимального правдоподобия (*[7]*, подпункт 6.5.4.7) или максимизации ожидания (*[7]*, подпункт 6.5.4.8), а также другие стратегии, такие как метод Пауэлла, метод имитации отжига и т. д.;

c) стратегия итерации по обучающим данным в процессе обучения модели (см. *[7]*, подраздел 8.4):

1) например, происходит ли выборка данных с помощью мини-пакетов (одновременная обработка нескольких образцов), пакетного обучения (все данные одновременно) или чистого онлайн-обучения (по одному образцу за один раз);

2) кроме того, если обучение включает мини-пакеты или отдельные образцы, описание того, как образцы извлекаются из набора данных посредством приоритизации выборок на основе заданного свойства, наложения ограничений на состав мини-пакетов (например, одинаковые длина или язык) или выполнения случайных выборок из набора данных с заменой или без нее;

d) другие, возможно, более сложные процедуры:

1) такие как тонкая настройка предварительно обученных моделей (*[7]*, подраздел 7.7), полуконтролируемое обучение (*[7]*, подраздел 7.4), обучение по учебной программе, дистилляция знаний и т. д.;

2) подробные условия, в которых они выполняются.

9.4.5 Гиперпараметры

Настоящий элемент таксономии содержит информацию о гиперпараметрах, задействованных в модели и ее создании, а также о том, как они были выбраны. Такая информация особенно (но не ис-

ключительно) полезна для моделей машинного обучения, на поведение которых может существенно влиять выбор гиперпараметров. Примеры полезной информации включают:

a) гиперпараметры алгоритма обработки (см. 9.4.3), и их значения, например:

1) количество и размер каждого слоя в нейронной сети или других компонентов нейронной сети (например, количество и размер фильтров, количество механизмов внимания в слое нейронной сети);

2) шаг (пространственное смещение между повторными применениями фильтра) в сверточной нейронной сети;

3) использование заполнения, усечения, разбиения входных данных на блоки и их размер;

4) максимальная глубина дерева решений;

5) предопределенное количество кластеров в модели кластеризации (при наличии);

6) числовая точность параметров;

7) пороговые значения, применяемые к уровням доверия;

8) использование конкретных функций внутри модели, таких как функции активации для нейронных сетей; формула механизма внимания для нейросетей, основанных на внимании; ядро для метода опорных векторов;

9) количество моделей в случае ансамблевых моделей (несколько моделей, используемых вместе с голосованием или другими более сложными схемами объединения);

b) гиперпараметры процедуры построения модели (см. 9.4.4) и их значения, например:

1) количество эпох или итераций процедуры, предопределенная продолжительность обучения;

2) скорость обучения и снижение веса для методов градиентного спуска;

3) размер мини-пакета;

4) процент отсева при обучении нейронной сети;

5) распределение, используемое для случайной инициализации параметров модели;

6) весовые значения классов в случае применения методов балансировки во время обучения;

7) веса, применяемые к каждой части потерь, в случае составных потерь;

8) любые алгоритмо-специфичные гиперпараметры, используемые на определенных этапах обучения (например, границы, пороги, критерии);

c) процедура выбора значений гиперпараметров:

1) ручное или автоматизированное, систематическое или случайное исследование потенциальных значений-кандидатов или по заданному критерию, использование метода оптимизации значений гиперпараметров;

2) набор или диапазон значений, рассматриваемых в качестве кандидатов;

3) показатель, используемый для принятия решения;

4) подмножество данных, используемых для вычисления показателя: отложенные данные или перекрестная проверка обучающих данных (размер блоков для перекрестной проверки, количество повторений и т. д.).

9.4.6 Форматы входных и выходных данных

Настоящий элемент таксономии описывает ожидаемый формат входных данных, обрабатываемых моделью (включая описание формата исходных данных и способа их пополнения дополнительной информацией), а также формат выходных данных.

Примеры полезной информации включают:

a) для текста: «Вводятся ли данные в виде необработанного текста? Представляются ли они в виде списка уже разделенных слов или подслов (лексем)? Предоставляется ли, например, информация о частях речи (положение глаголов, существительных и т. д.)?»;

b) для аудио: «Закодирован ли он в виде сигнала, в виде мел-частотных кепстральных коэффициентов и т. д.?», «Обеспечивается ли идентификация говорящего для каждого высказывания и каким образом?»;

c) для изображений: «Используется ли цветовая модель (красный, зеленый и синий); применяется ли модель (оттенок, насыщенность, светлота) и т. д.?», «Используется ли сжатие, например, в формате JPEG?».

9.4.7 Вычислительное оборудование

Настоящий элемент таксономии содержит сведения о типе вычислительного оборудования, на котором работает модель. Существуют различные уровни описания оборудования, начиная от общих описаний систем до конкретных моделей продуктов или деталей на уровне компонентов оборудования.

Полезная информация может включать:

a) оборудование, которое было использовано для построения модели, например для целей машинного обучения;

b) оборудование, которое может использоваться для работы модели, включая возможные ограничения, связанные с применением данного оборудования. Это полезно, например, для заинтересованных сторон, чтобы лучше понимать компромиссы при выборе варианта оборудования или заранее планировать его поставку в случае моделей с ограниченными аппаратными возможностями;

c) оборудование, на котором модель эффективно работает в контексте конкретной системы ИИ, если она поставляется вместе с системой.

9.4.8 Вычислительные затраты

Настоящий элемент таксономии содержит сведения о затратах на вычисления, связанных с использованием модели. Для определения затрат используются различные типы показателей в зависимости от целей и задач заинтересованных сторон.

Полезная информация может включать:

a) вычислительные затраты, потраченные на построение модели, которые могут включать стоимость данной конкретной модели, а также дополнительные затраты на предварительные исследования и эксперименты, приведшие к созданию данной модели. Данные затраты могут быть выражены, например, в общем процессорном времени, графическом процессорном времени или в комбинации обоих показателей (при необходимости);

b) ожидаемые вычислительные затраты при запуске модели в эксплуатацию. Данные затраты могут быть выражены, например, в процессорном или графическом процессорном времени на один входной образец, на объем входных данных, на пользовательское время и в других единицах измерения. Данная информация помогает заинтересованным сторонам более точно прогнозировать объем вычислительных ресурсов, которые необходимо предоставить.

9.4.9 Модели в развивающихся системах

Настоящий элемент таксономии описывает механизмы, посредством которых поведение модели может изменяться в рамках заданной версии. Примеры включают:

a) наличие внутренней базы данных, которая аккумулирует новую информацию для использования (неизменяемой) моделью. Например, модель распознавания лиц может быть связана с растущей базой цифровых отпечатков пальцев для распознавания известных лиц и добавления нового отпечатка пальца для нового лица, с которым сталкивается модель. Аналогично модель, адаптированная под пользователя, может включать создание и поддержку профилей пользователей, которые влияют на поведение модели без ее непосредственного изменения;

b) использование эксплуатационных данных для внесения изменений в модель в режиме реального времени. Для машинного обучения является непрерывным. Для символьного ИИ примером является добавление новых узлов в граф знаний на основе информации, извлеченной из эксплуатационных данных;

c) сохранение и использование действий пользователя (например, корректировка решения модели) или других форм обратной связи для влияния или изменения поведения модели. Например, случаи обучения на протяжении всего жизненного цикла (возможно, инкрементального, но не обязательно непрерывного), а также специализированные модули постредактирования.

9.5 Используемые данные

Прозрачность модели *СИИ* также может включать прозрачность данных, которые были использованы для ее создания или оценки. Соответствующая информация для документирования зависит от характера и источника этих данных. Например, если данные взяты из существующего общедоступного набора, ссылки на название, версию и репозиторий набора могут покрыть большинство потребностей. Если это недавно созданный или аннотированный набор данных, может потребоваться значительно больше информации, такой как сведения об источниках сбора данных, процессе аннотирования, статистические данные о содержимом или условия использования данных (см. также раздел 10). В любом случае указание того, какая часть данных была использована, как и с какой целью, крайне полезно для обеспечения прозрачности.

В частности, для большинства *СИИ* процесс оценки включает в себя тестовые данные (также известные как данные оценки, см. *ГОСТ Р 71476—2024*, подраздел 5.10). Информация о предметной области, объеме, содержании, разнообразии тестовых данных помогает заинтересованным сторонам

определить надежность проведенной оценки, а также лучше интерпретировать результаты работы и сделать значимые выводы о возможных ограничениях *СИИ*. Например, если потенциальный пользователь считает, что система решает проблему в определенной области, а используемые тестовые данные не охватывают данную область, нет никаких гарантий ее результирующей производительности. Другим примером полезной информации является список проверок, которые были проведены, чтобы подтвердить, что тестовые данные не использовались на этапах проектирования и разработки (даже косвенно). Например, для модели обработки изображений проверяется, что при разработке не использовалось тестовое изображение, а также то, что тестовые изображения не принадлежат тем же видеопоследовательностям, что и изображения, использованные в других этапах. Прозрачность в таких аспектах придает больше смысла заявленным показателям производительности системы.

Валидационные данные, также известные как проверочные данные (см. *ГОСТ Р 71476—2024*, подраздел 5.10), включают любые данные, используемые на этапе проектирования и разработки для настройки некоторых гиперпараметров и других алгоритмических свойств модели. Эти данные часто используются для устранения основных причин низкой производительности и ненадежности *СИИ*, чтобы информация о существовании и использовании валидационных данных позволяла заинтересованным сторонам проверить, что модель разработана с максимальным использованием ее потенциала. К полезной информации относится объем данных, а также выявление сходств и различий этих данных с остальными данными, участвующими в жизненном цикле, что дает ценную информацию о способности системы к обобщению.

Наконец, если модель использует машинное обучение, часть данных применяется в качестве обучающих данных. В этом случае прозрачность в отношении характера и характеристик этих обучающих данных (например, предметной области, объема, распределения) позволяет заинтересованным сторонам выявить потенциальную предвзятость и другие ограничения в поведении *СИИ*, вызванные выбором обучающих данных.

Настоящий элемент таксономии включает:

a) ссылки на документацию используемых наборов данных, как описано в разделе 10;

b) подробную информацию об использовании наборов данных в контексте данной модели, включая операции фильтрации или разделения данных, и назначение каждого подмножества.

9.6 Функциональная корректность

Настоящий элемент таксономии предоставляет доступную информацию о том, насколько эффективно модель выполняет предполагаемую функцию (см. *[12]*, подраздел 5.4), полученную либо путем оценки показателей производительности, либо посредством других, не количественных процессов.

Примеры полезной для документирования информации включают:

a) измерения показателей производительности, включая указание оцениваемых аспектов, используемых показателей, условий измерения, полученных результатов и контрольных точек для сравнения — бенчмаркинг может быть полезен как часть процесса валидации для сравнения результатов с другими *СИИ*. Различные организации предоставляют наборы бенчмаркинга для конкретных сценариев;

b) качественную оценку, включая субъективные суждения одного или нескольких испытуемых человек, протокол, используемый для их сбора и агрегации, а также оцениваемые критерии;

c) создание отчетов о поведении модели на заданном наборе входных данных пограничных случаев;

d) анализ ошибок, проведенный на выходных данных модели, включая информацию о критериях, результатах и интерпретации.

10 Таксономия наборов данных

10.1 Общие положения

Наборы данных являются неотъемлемой частью надлежащего функционирования *СИИ* — от тестовых наборов данных, используемых для верификации и валидации функционирования *СИИ* и ее различных компонентов, до наборов для обучения и валидации, применяемых в сочетании с моделями МО. Таким образом, предоставление информации об используемых наборах данных является важной составляющей прозрачности *СИИ*.

10.2 Основные сведения

Основные сведения о наборе данных могут включать:

a) наименование, под которым выпущен набор данных;

b) информацию о поставщике данных: наименование и контактные данные организации, которая предоставляет доступ к набору данных;

c) версию: если набор данных изменяется или поддерживается, номер версии текущего выпуска;

d) информацию о лицензировании: по какой лицензии, если таковая имеется, выпущен набор данных;

e) сертификаты:

1) какие сертификаты, если таковые имеются, получены для набора данных. К ним также относятся сертификаты по отдельным аспектам набора данных, например, защите *ПДн*;

2) соответствие требованиям законодательства;

3) по каждой проведенной сертификации или оценке (более подробную информацию см. в *[34]*):

- аудиторы: наименование и контактные данные организации, проводившей оценку набора данных, с указанием, являлась ли она третьей стороной;

- краткое изложение результатов сертификации или оценки, а также указание, где можно найти дополнительную информацию;

f) обратную связь: наименование и контактные данные организации, ответственной за обработку вопросов и обратную связь по набору данных;

g) документацию: тип и объем предоставляемой документации по набору данных может варьироваться в зависимости от предполагаемого получателя. Настоящий элемент таксономии может включать информацию:

1) о том, где найти более полную документацию по набору данных;

2) типе доступной документации (например, справочные документы, интерактивные онлайн-дисплеи, машиночитаемые файлы);

3) структуре документации;

h) правовой контекст:

1) законодательная база, в соответствии с которой был создан набор данных, например указание страны, в которой были собраны данные, или страны, в которой они были промаркированы;

2) физическое расположение серверов, на которых размещены данные;

3) экспортный контроль или иные правовые требования, применимые к набору данных или его отдельным элементам.

10.3 Происхождение данных

Настоящий элемент таксономии предоставляет информацию об источнике набора данных, способе и времени их сбора, создания или преобразования, а также соблюдении требований законодательства в отношении авторских прав и иных применимых прав на данные. Информационные элементы могут включать:

a) сведения о том, кто создал набор данных, кто финансировал его создание;

b) цель создания набора данных;

c) сведения о том, включало ли создание набора данных сбор уже существующих данных. Если да:

1) как были получены данные, напрямую или от третьей стороны?

2) если данные собирались напрямую, то кто их собирал? Если получено от третьей стороны, то от какой именно?

3) были ли получены соответствующие разрешения (интеллектуальная собственность, согласие субъекта данных)?

4) какое использование данных разрешено?

5) где были собраны данные (регион, среда и т. д.)?

6) когда были собраны данные?

d) сведения о том, были ли созданы новые данные при создании набора данных? Если да:

1) являются ли эти данные реальными (например, измеренными, записанными) или сгенерированными (например, разработанные вручную, полученные с помощью алгоритма)?

2) насколько они близки к данным естественного происхождения и в чем их отличия? Например, созданы ли они с привлечением актеров вместо реальных людей? Используется ли в них упрощенная физическая модель?

3) какие методы и процессы применялись для создания данных (например, ручной, автоматизированный, используемые алгоритмы или протоколы)?

Если вручную, то какие инструкции давались лицам, участвующим в создании (создателям или субъектам)?

Если автоматизировано, использовались ли шаблоны, формальные модели (например, автоматизированное проектирование, 3D-модели, физические модели), статистические модели, генерация на основе ИИ (возможно, на основе базового набора данных)?

4) если применимо, наименование, ссылки и документация инструментов, оборудования, систем, использованных при создании, а также других наборов данных, задействованных в процессе;

5) можно ли проследить происхождение созданных данных до средств их создания (инструменты, наборы данных и т. д.)? Например, если в процессе создания используются конфиденциальные или *ПДн*, можно ли получить эту информацию из новых созданных данных?

Важно отметить, что в случае объединения генерируемых данных, защищающих *ПДн*, с собранными данными, защита *ПДн* гарантирована быть не может;

e) сведение о том, было ли создание набора данных связано с преобразованием новых или существующих данных? Если да:

1) какие преобразования были применены (например, добавление шума, объединение нескольких экземпляров в один, удаление некоторой информации)?

2) в чем заключалась цель данных преобразований?

3) какие методы и процессы использовались для осуществления преобразований (например, ручные или автоматизированные, какой алгоритм)?

4) если применимо, название, ссылки и документация инструментов, оборудования, систем, использованных при преобразовании, а также других наборов данных, задействованных в процессе?

5) какие свойства были сохранены (например, распределения признаков, корреляции)?

6) каковы пороговые значения для размера группы, при наличии?

7) были ли изменены какие-либо свойства и почему (например, для достижения справедливости)?

8) какие свойства или признаки были изменены и каким образом?

9) задействованы ли при преобразованиях какие-либо механизмы или меры по защите *ПДн*, такие как защита редких категорий, защита экстремальных значений (например, люди старше 105 лет), минимизация переобучения или другие меры защиты конфиденциальной информации? Например, может быть предпочтительно не сохранять пограничные случаи, если они считаются уязвимыми с точки зрения защиты *ПДн*; в этом случае возможно, что хвостовое распределение данных также может быть недостаточно точно сохранено;

f) качество исходных собранных данных, исходных созданных данных или исходных преобразованных данных:

1) методы оценки. Например, статистика признаков, корреляции, сравнение моделей МО на собранных и созданных данных, сравнение моделей МО на преобразованных и непреобразованных данных;

2) результаты оценки;

g) метки для собранных, созданных и преобразованных данных (т. е. указание, какие конкретно экземпляры к ним относятся);

h) процентное соотношение собранных, созданных и преобразованных данных в наборе;

i) субъекты данных: если набор данных содержит информацию о людях (более подробную информацию см. *[2]*):

1) содержат ли данные персональную информацию, позволяющую идентифицировать человека?

2) содержат ли данные конфиденциальные характеристики?

3) можно ли использовать эти данные для получения конфиденциальных характеристик?

4) могут ли субъекты данных исправить свою информацию или потребовать ее удаления из набора данных?

5) могут ли субъекты данных контролировать цели, для которых используются их данные?

10.4 Свойства данных

Настоящий элемент таксономии описывает свойства данных в наборах данных и может включать следующую информацию:

a) сведения о том, из чего состоят экземпляры набора данных (например, изображения, текст, структурированные данные);

b) размер набора данных:

1) количество экземпляров;

2) размер каждого экземпляра (например, длительность аудио или видео, длина текста, разрешение изображения, частота дискретизации звука, размерность встраиваемых данных) или средний размер по всем экземплярам в наборе данных, в зависимости от ситуации;

c) сведения о том, содержит ли набор данных полную статистическую совокупность или выборку. Если да:

1) какую совокупность?

2) если это выборка, как она была создана?

d) сведения о том, существуют ли связи между экземплярами набора данных. Например, несколько экземпляров поступают от одного и того же пользователя; или экземпляры представляют узлы в сети с различными связями с другими узлами;

e) сведения о том, как предоставляется доступ к набору данных, как получить доступ к набору данных (посредством URL-адреса или API, а также формата файла);

f) сведения о том, доступен ли набор данных в качестве самостоятельного ресурса или он опирается на контент третьей стороны (например, набор данных, в котором содержатся только идентификаторы сообщений, а не сами онлайн-сообщения);

g) формат набора данных:

1) для табличных данных — словарь данных с описанием каждого признака, его названия в наборе данных и его свойств:

- номинальный, порядковый, дискретный, непрерывный;
- возможные значения или диапазон значений;
- распределение переменной;
- корреляции между выбранными переменными, особенно с чувствительными признаками и с целевой переменной;

h) статистические свойства набора данных:

1) есть ли пропущенные значения?

2) есть ли выбросы или точки с высоким уровнем влияния?

3) есть ли в наборе данных известные ложные корреляции?

i) параметры защиты *ПДн*[1)]: меры, предпринимаемые для оценки уязвимости к атакам на *ПДн* или другим нежелательным раскрытиям *ПДн*, например сведения о компрометации данных (подробнее см. в *[36]* и *[37]*).

10.5 Предметная область и назначение набора данных

10.5.1 Общие положения

Настоящий элемент таксономии содержит информацию о предметной области данных в наборе данных и целях его создания. Информация может включать:

a) первоначальное намерение при создании этого набора данных;

b) варианты использования, для которых изначально предназначен набор данных;

c) ситуации, в которых набор данных успешно применялся. Примеры со ссылками и цитатами могут предоставить полезные рекомендации, в частности, для ситуаций, которые изначально не были предусмотрены создателями набора данных;

d) сведения о том, к какой предметной области относятся данные (например, данные о левшах старше 65 лет);

1) в 10.5.2 содержится дополнительная информация о предметной области для языковых данных;

2) в 10.5.3 содержится информация о предметной области для визуальных данных;

[1)] *Обращение с ПДн должно осуществляться в соответствии с [35].*

e) сведения о том, какие задачи учитывались при подготовке набора данных (например, семантическая сегментация изображений);

f) сведения о том, в каких ситуациях набор данных успешно применялся? Примеры со ссылками и цитатами могут предоставить полезные рекомендации, в частности, для ситуаций, которые изначально не были предусмотрены создателями набора данных.

10.5.2 Информация о предметной области для языковых данных

Для наборов данных, содержащих языковые данные, раскрываемая информация о предметной области может включать следующее:

a) указание на диалект языка (например, язык, страна, региональный или местный вариант, социолект, повседневный язык или другие регистры или стили языка);

b) любая демографическая информация о говорящих или пишущих, которая может влиять на их использование или понимание языка (например, возраст, пол, происхождение, родной язык, социально-экономический статус), количество представленных говорящих или пишущих, наличие речевых нарушений (заикание, шепелявость и т. д.), использование других форм существования языка (например, язык жестов, шрифт Брайля), наличие нетрадиционного языка (например, клингонский);

c) ситуация, в которой был создан языковой контент: время и место создания, способ (например, устный, жестовый, письменный, напечатанный), спонтанная или подготовленная речь, синхронное или асинхронное взаимодействие, целевая аудитория;

d) для речевых данных — тип средства записи (аналоговый, цифровой, тип устройства) и качество записи (например, тип помещения, окружение, расстояние до записывающих устройств);

e) жанр (например, юридический, новостной, литературный, социальные сети, научный) и тематика данных.

10.5.3 Информация о предметной области для визуальных данных

Для наборов данных, включающих данные визуального восприятия или изображения, раскрываемая информация о предметной области может включать:

a) указание темы и жанра (например, медицинские изображения, предметы повседневного обихода, платежные документы);

b) сведения о том, присутствуют ли на изображениях люди или связаны ли изображения с конкретными лицами или группами (например, медицинские изображения), а также соответствующая демографическая информация (может быть особенно релевантно для медицинских изображений);

c) информация о месте получения изображений;

d) средства записи (например, тип устройств, используемых для записи) и качество записи (например, разрешение, профессиональная или любительская запись).

10.6 Предвзятость и ограничения данных

Настоящий элемент таксономии предоставляет информацию о возможных ограничениях набора данных. К ним относятся:

a) сведения о том, есть ли в наборе данных некоторые случаи, совокупности, типы объектов, которые недостаточно представлены, чрезмерно представлены или представлены таким образом, что могут создавать нежелательную предвзятость. Примерами таких предвзятых представлений являются наборы данных, которые содержат преимущественно женские лица, или наборы данных, в которых большинство людей определенной национальности указаны как имеющие одну и ту же профессию;

b) сведения о том, отсутствуют ли в данных определенные конфигурации. Примером отсутствующей конфигурации является ситуация, когда набор данных для ответов на вопросы не содержит ни одного вопроса, на который невозможно ответить (т. е. вопросы, на которые невозможно ответить на основе предоставленных входных данных);

c) сведения о том, какие ограничения в информации, содержащейся в наборе данных, выявлены. Например, если набор данных содержит данные только о взрослых в возрасте от 20 до 65 лет, в нем отсутствует информация, специфичная для других возрастных групп;

d) сведения о том, запрещено ли какое-либо использование набора данных законом. Например, если речь идет о *ПДн*, это могут быть варианты использования, для которых не было получено согласия.

10.7 Социальные аспекты

Социальные аспекты могут включать:

a) сведения о том, проводилась ли оценка воздействия на защиту *ПДн* (более подробную информацию см. в *[3]*):

1) кто проводил оценку;
2) что было исследовано;
3) какие риски были выявлены и какие меры по их снижению были предприняты;
4) ссылка на соответствующую публикацию, при наличии;
5) кто утвердил или одобрил оценку;

b) сведения о том, содержатся ли в данных известные социальные предубеждения:

1) какие предубеждения исследовались и каким образом;
2) кто проводил оценку;
3) ссылка на соответствующую публикацию, при наличии;

c) сведения о том, проводилась ли оценка воздействия на окружающую среду, с учетом того, что хранение и обработка больших объемов данных могут иметь последствия для окружающей среды:

1) кто проводил оценку;
2) что было исследовано;
3) какие риски были выявлены и какие меры по их снижению были предприняты;
4) ссылка на соответствующую публикацию, при наличии;

d) сведения о том, содержит ли набор данные, которые можно считать оскорбительными, токсичными или небезопасными. Если да, то кто принял такое решение и какие критерии использовались (это можно раскрыть таким же образом, как указано в 10.8 для маркировки данных).

10.8 Подготовка данных

Информация о подготовке данных может включать:

a) сведения о том, доступен ли исходный необработанный набор данных;

b) указание выполненных этапов предварительной обработки:

1) объединялись ли данные из нескольких источников;
2) какие изменения, при наличии, были применены к исходным данным. Возможна обработка, связанная с пропущенными значениями, выбросами, дисбалансом данных, справедливостью, анонимизацией или псевдонимизацией;

c) сведения о том, была ли автоматически добавлена дополнительная информация поверх исходных данных:

1) какая информация была добавлена (например, автоматические метки, метаданные, связи между отдельными экземплярами, структурированная версия исходных данных);
2) какой инструмент, метод или *СИИ* использовались;

d) сведения о том, доступен ли скрипт предварительной обработки;

e) сведения о том, был ли набор данных размечен вручную. Если да, могут быть также актуальны следующие сведения:

1) кто выполнял разметку набора данных? Полезная для раскрытия информация:
- была ли она выполнена посредством внутренней, аутсорсинговой или краудсорсинговой разметки (см. *ГОСТ Р 71484.4—2024*, пункт 8.4.6); в случае аутсорсинга или краудсорсинга информация об используемой платформе также может быть полезна;
- были ли маркировщики специалистами в предметной области?
- если речь идет о маркировке диалектов языка, любая демографическая информация о маркировщиках, которая может повлиять на их использование или понимание языка (например, возраст, пол, происхождение, родной язык, социально-экономический статус, языковая подготовка и т. д.);
- при необходимости, аспекты разнообразия и инклюзивности среди маркировщиков;
2) уровень оплаты труда маркировщиков и принципы ее расчета;
3) стратегия разметки данных:
- какие инструкции получили маркировщики данных?
- была ли разметка полностью ручной или полуавтоматической (если да, с использованием каких инструментов или *СИИ*)?
- какие механизмы контроля качества применяются (например, проверка согласованности результатов одного разметчика данных, проверка согласованности результатов нескольких разметчи-

ков данных, профессиональные тесты разметчиков данных, арбитраж третьей стороной результатов нескольких разметчиков данных);

f) были ли заранее определены разделения данных? Какие именно и на каком основании?

10.9 Обслуживание наборов данных

Информация о поддержке набора данных может включать:

a) сопровождающий набора данных: ФИО и контактные данные;

b) способ хранения набора данных, включая:

1) географическое расположение центров хранения данных или серверов;

2) аспекты обеспечения безопасности данных, включая контроль доступа;

3) аспекты обеспечения защиты *ПДн*;

4) энергоресурсы;

c) обработку ошибок:

1) описание механизма сообщения об ошибках;

2) указание на то, публикуются ли исправления; если да, то где и как часто они обновляются;

3) указание на то, поддерживаются ли старые версии данных; если да, то где и как;

d) сведения о том, является ли набор данных статичным или постоянно обновляется;

e) если набор данных обновляется:

1) какая система управления конфигурацией используется?

2) поддерживаются ли более старые версии данных?

3) в чем разница между текущей и предыдущей версиями набора данных?

f) если набор данных содержит информацию о физических лицах, сведения о том, могут ли субъекты данных требовать удаления своих данных или ограничивать цели их использования;

g) сведения о том, есть ли у набора данных срок действия, по истечении которого весь набор или его элементы будут автоматически удалены.

Приложение А
(справочное)

Примеры ролей заинтересованных сторон в раскрытии прозрачности

Таблица А.1 — Примеры прозрачности системы искусственного интеллекта

Пары заинтересованных сторон		Цели/задачи	Содержимое
Поставщик содержимого	Получатель содержимого		
Поставщик ИИ	Сторона, развертывающая ИИ	Право на информацию	Основная информация о системе (см. 8.2)
		Надлежащее использование	Применимость (см. 8.4). Обзор технических характеристик (см. 8.5). Доступ к внутренним компонентам (см. 8.6)
		Прослеживаемость	Основная информация (см. 8.2). Качество и показатели производительности (см. 8.7)
Поставщик ИИ	Аудитор ИИ	Прослеживаемость	Доступ к внутренним компонентам (см. 8.6)
Поставщик ИИ	Конечный пользователь ИИ, аудиторы ИИ	Прослеживаемость	Аспекты используемого набора данных (см. 9.5), см. также *ГОСТ Р ИСО/МЭК 42001—2024* (таблица А.1, пункт А.7)
Сторона, развертывающая ИИ	Конечный пользователь ИИ	Право на информацию	Основная информация о системе (см. 8.2)
		Надлежащее использование	Применимость (см. 8.4). Обзор технических характеристик (см. 8.5). Доступ к внутренним компонентам (см. 8.6)
		Прослеживаемость	Основная информация (см. 8.2). Качество и показатели производительности (см. 8.7)
Сторона, развертывающая ИИ	Субъект ИИ	Право на информацию	Основная информация о системе (см. 8.2). Какая транзакционная информация сохраняется и как она используется (см. 8.5.4). Порядок обращения для устранения недостатков или исключения данных о субъекте ИИ (см. 8.2, 8.3.2)
Производитель ИИ	Поставщик ИИ	Право на информацию	Основная информация о модели (см. 9.2)
		Надлежащее использование	Проектная документация *СИИ* (см. 8.5.5). Полезная информация для оптимизации показателей производительности (см. 8.7.3). Вопросы устойчивого развития (см. 8.4.3). Технические характеристики модели (см. 9.4)
		Прослеживаемость	Основная информация о модели (см. 9.2). Качество и показатели производительности (см. 8.7). Использование модели (см. 9.3). Используемые данные (см. 9.5). Функциональная корректность (см. 9.6)

Окончание таблицы А.1

Пары заинтересованных сторон		Цели/задачи	Содержимое
Поставщик содержимого	Получатель содержимого		
Производитель ИИ	Аудитор ИИ	Прослеживаемость	Информация о конкретных компонентах системы, например, релевантные технические характеристики модели или моделей, входящих в *СИИ* (см. 9.4)
Производитель ИИ	Релевантные уполномоченные органы	Прослеживаемость	Основные сведения о модели (см. 9.2). Использование модели (см. 9.3). Используемые данные (см. 9.5). Функциональная корректность (см. 9.6)
Поставщик данных	Производитель ИИ	Право на информацию	Основная информация о наборе данных (см. 10.2)
		Надлежащее использование	Свойство данных (см. 10.4). Предметная область и назначение набора данных (см. 10.5). Обслуживание наборов данных (см. 10.9)
		Прослеживаемость	Основная информация о наборе данных (см. 10.2). Происхождение данных (см. 10.3)

Приложение ДА
(справочное)

Сведения о соответствии ссылочных национальных стандартов международным стандартам, использованным в качестве ссылочных в примененном международном стандарте

Таблица ДА.1

Обозначение ссылочного национального стандарта	Степень соответствия	Обозначение и наименование ссылочного международного стандарта
ГОСТ Р 71476—2024 (ИСО/МЭК 22989:2022)	MOD	ISO/IEC 22989:2022 «Информационные технологии. Искусственный интеллект. Концепции и терминология искусственного интеллекта»

Примечание — В настоящей таблице использовано следующее условное обозначение степени соответствия стандарта:

- MOD — модифицированный стандарт.

Приложение ДБ
(справочное)

Сопоставление структуры настоящего стандарта со структурой примененного в нем международного стандарта

Таблица ДБ.1

Структура настоящего стандарта			Структура международного стандарта ISO/IEC 12792:2025		
Разделы	Пункты	Подпункты	Раздел	Пункты	Подпункты
Приложения	—		A		
	A		B		
	ДА		—		
	ДБ		—		
Библиография			Библиография		

Примечания

1 Сопоставление структуры стандартов приведено начиная с элемента «Приложение», т. к. предыдущие разделы стандартов и их иные структурные элементы (за исключением предисловия) идентичны.

2 Приложение А исключено, т. к. его положения не соответствуют требованиям основополагающих документов по стандартизации в Российской Федерации.

Библиография

[1] ИСО/МЭК 23894:2023 Информационная технология. Искусственный интеллект. Руководство по менеджменту риска (Information technology — Artificial intelligence — Guidance on risk management)

[2] ИСО/МЭК 27701:2025 Информационная безопасность, кибербезопасность и защита конфиденциальности — Системы управления информацией о конфиденциальности — Требования и руководство (Information security, cybersecurity and privacy protection — Privacy information management systems — Requirements and guidance)

[3] ИСО/МЭК 29134:2023 Информационная технология. Методы и средства обеспечения безопасности. Руководящие указания по оценке воздействия на конфиденциальность (Information technology — Security techniques — Guidelines for privacy impact assessment)

[4] ИСО/МЭК 17000:2020 Оценка соответствия. Словарь и общие принципы (Conformity assessment — Vocabulary and general principles)

[5] *Федеральный закон от 27 июля 2006 г. № 149-ФЗ «Об информации, информационных технологиях и о защите информации»*

[6] *ИСО 14641:2018* *Управление электронными документами. Проектирование и эксплуатация информационной системы для обеспечения долговременной сохранности электронных документов. Технические требования (Electronic document management — Design and operation of an information system for the preservation of electronic documents — Specifications)*

[7] ИСО/МЭК 23053:2022 Экосистема разработки систем искусственного интеллекта (ИИ) с использованием машинного обучения (МО) [Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML)]

[8] *ИСО/МЭК 30134-2:2016* *Информационные технологии. Центры обработки данных. Ключевые показатели эффективности. Часть 2. Коэффициент энергоэффективности (PUE) [Information technology — Data centres — Key performance indicators — Part 2: Power usage effectiveness (PUE)]*

[9] *ИСО/МЭК 10918-1:1994* *Информационная технология. Цифровое сжатие и кодирование неподвижных полутоновых изображений. Часть 1. Требования и руководящие принципы (Information technology — Digital compression and coding of continuous-tone still images — Part 1: Requirements and guidelines)*

[10] ISO/IEC TR 24028:2020 Информационная технология. Искусственный интеллект. Обзор достоверности систем искусственного интеллекта (Information technology — Artificial intelligence — Overview of trustworthiness in artificial intelligence)

[11] ИСО/МЭК 27036-1:2021 Кибербезопасность. Отношения с поставщиками. Часть 1. Обзор и основные понятия (Cybersecurity — Supplier relationships — Part 1: Overview and concepts)

[12] ИСО/МЭК 25059:2023 Программная инженерия. Требования и оценка качества систем и программного обеспечения (SQuaRE). Модель качества для систем искусственного интеллекта [Software engineering — Systems and software Quality Requirements and Evaluation (SquaRE) — Quality model for AI systems]

[13] ISO/IEC TR 24368:2022 Информационные технологии. Искусственный интеллект. Обзор этических и социальных проблем (Information technology — Artificial intelligence — Overview of ethical and societal concerns)

[14] ISO/IEC TS 6254:2025 Информационные технологии. Искусственный интеллект. Цели и подходы для обеспечения объяснимости и интерпретируемости моделей машинного обучения (МО) и систем искусственного интеллекта (ИИ) [Information technology — Artificial intelligence — Objectives and approaches for explainability and interpretability of machine learning (ML) models and artificial intelligence (AI) systems]

[15]	ИСО 24495-1:2023	Простой язык. Часть 1. Руководящие принципы и директивы (Plain language — Part 1: Governing principles and guidelines)
[16]	ИСО 22458:2022	Уязвимость потребителя. Требования и руководящие указания по разработке и предоставлению инклюзивных услуг (Consumer vulnerability — Requirements and guidelines for the design and delivery of inclusive service)
[17]	ISO/IEC TS 12791:2024	Информационная технология. Искусственный интеллект. Устранение нежелательных отклонений в задачах машинного обучения классификации и регрессии (Information technology — Artificial intelligence — Treatment of unwanted bias in classification and regression machine learning tasks)
[18]	ИСО/МЭК 42005:2025	Информационные технологии. Искусственный интеллект (ИИ). Оценка воздействия системы искусственного интеллекта [Information technology — Artificial intelligence (AI) — AI system impact assessment]
[19]	ISO/IEC TR 20226:2025	Информационные технологии. Искусственный интеллект. Аспекты экологической устойчивости систем искусственного интеллекта (Information technology — Artificial intelligence — Environmental sustainability aspects of AI systems)
[20]	ИСО 14044:2006	Экологический менеджмент. Оценка жизненного цикла. Требования и рекомендации (Environmental management — Life cycle assessment — Requirements and guidelines)
[21]	ИСО/МЭК 17360:2023	Технологии автоматической идентификации и сбора данных. Применение радиочастотной идентификации в цепи поставок. Продукция, снабженная радиочастотными метками, упакованная продукция, транспортируемые единицы, возвратные транспортные упаковочные средства и возвратные упаковочные средства (Automatic identification and data capture techniques — Supply chain applications of RFID — Product tagging, product packaging, transport units, returnable transport units and returnable packaging items)
[22]	ИСО 28000:2022	Безопасность и устойчивость. Системы менеджмента безопасности. Требования (Security and resilience — Security management systems — Requirements)
[23]	ИСО 22301:2019	Безопасность и устойчивость. Системы менеджмента непрерывности деятельности. Требования (Security and resilience — Business continuity management systems — Requirements)
[24]	ИСО 20400:2017	Устойчивые закупки. Руководство (Sustainable procurement — Guidance)
[25]	ИСО 59004:2024	Экономика замкнутого цикла. Словарь, принципы и руководство по внедрению (Circular economy — Vocabulary, principles and guidance for implementation)
[26]	ИСО/МЭК 24661:2023	Информационная технология. Пользовательские интерфейсы. Полнодуплексное речевое взаимодействие (Information technology — User interfaces — Full duplex speech interaction)
[27]	ИСО/МЭК 38507:2022	Информационные технологии. Стратегическое управление ИТ. Последствия влияния стратегического управления при использовании искусственного интеллекта организациями (Information technology — Governance of IT — Governance implications of the use of artificial intelligence by organizations)
[28]	ИСО/МЭК 27001:2022	Информационная безопасность, кибербезопасность и защита конфиденциальности. Системы менеджмента информационной безопасности. Требования (Information security, cybersecurity and privacy protection — Information security management systems — Requirements)
[29]	ISO/IEC TS 25058:2024	Системная и программная инженерия. Требования и оценка качества систем и программной продукции (SQuaRE). Руководство по оценке качества систем искусственного интеллекта [Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Guidance for quality evaluation of artificial intelligence (AI) systems]

[30] ИСО 9241-210:2019 Эргономика взаимодействия человек-система. Часть 210. Человеко-ориентированное проектирование интерактивных систем (Ergonomics of human-system interaction Part 210: Human-centred design for interactive systems)

[31] ИСО/МЭК 25010:2023 Системная и программная инженерия. Требования и оценка качества систем и программного обеспечения (SQuaRE). Модели качества систем и программного обеспечения [Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Product quality model]

[32] ISO/IEC/IEEE 29119-3:2021 Системная и программная инженерия. Тестирование программного обеспечения. Часть 3. Документация тестирования [Software and systems engineering — Software testing — Part 3: Test documentation]

[33] ISO/IEC TS 4213:2022 Информационные технологии. Искусственный интеллект. Оценка классификации машинного обучения (Information technology — Artificial intelligence — Assessment of machine learning classification performance)

[34] ИСО/МЭК 42006:2025 Информационные технологии. Искусственный интеллект. Требования к органам, осуществляющим аудит и сертификацию систем управления искусственным интеллектом (Information technology — Artificial intelligence — Requirements for bodies providing audit and certification of artificial intelligence management systems)

[35] *Федеральный закон от 27 июля 2006 г. № 152-ФЗ «О персональных данных»*

[36] ISO/IEC TR 27563:2023 Безопасность и конфиденциальность в примерах использования искусственного интеллекта. Лучшие практики (Security and privacy in artificial intelligence use cases — Best practices)

[37] ISO/IEC TR 27550:2019 Информационная технология. Методы и средства обеспечения безопасности. Техника обеспечения конфиденциальности процессов жизненного цикла системы (Information technology — Security techniques — Privacy engineering for system life cycle processes)

УДК 004.8:006.354

ОКС 01.040.35
35.020

Ключевые слова: искусственный интеллект, система искусственного интеллекта, таксономия прозрачности системы искусственного интеллекта

Редактор *М.В. Митрофанова*
Технический редактор *И.Е. Черепкова*
Корректор *О.В. Лазарева*
Компьютерная верстка *Е.А. Кондрашовой*

Сдано в набор 11.02.2026. Подписано в печать 24.02.2026. 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