ФЕДЕРАЛЬНОЕ АГЕНТСТВО

ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

ГОСТ Р 70889— 2023 (ИСО/МЭК 8183:2023)

Информационные технологии

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

Структура жизненного цикла данных

(ISO/IEC 8183:2023, MOD)

Издание официальное

Москва
Российский институт стандартизации
2023

gost.e-ecolog.ru

Предисловие

1 ПОДГОТОВЛЕН Научно-образовательным центром компетенций в области цифровой экономики Федерального государственного бюджетного образовательного учреждения высшего образования «Московский государственный университет имени М.В. Ломоносова» (МГУ имени М.В.Ломоносова) 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 2 ноября 2023 г. № 1327-ст

4 Настоящий стандарт является модифицированным по отношению к международному стандарту ИСО/МЭК 8183:2023 «Информационные технологии. Искусственный интеллект. Структура жизненного цикла данных» (ISO/IEC 8183:2023 «Information technology — Artificial intelligence — Data life cycle framework», MOD) путем изменения отдельных фраз (слов, значений показателей, ссылок), которые выделены в тексте курсивом.

Внесение указанных технических отклонений направлено на учет особенностей национальной стандартизации технологий работы с большими данными и искусственного интеллекта

5 ВВЕДЕН ВПЕРВЫЕ

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

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

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

Содержание

1 Область применения 1
2 Нормативные ссылки 1
3 Термины и определения 1
4 Сокращения 1
5 Обзор жизненного цикла данных 2
6 Структура жизненного цикла данных 2
7 Стадии и процессы жизненного цикла данных 7
Библиография 10

Введение

Системы искусственного интеллекта внедряются организациями всех типов, размеров и целей. Данные необходимы для разработки и эксплуатации систем искусственного интеллекта.

В области систем искусственного интеллекта имеется большое число жизненных циклов данных, используемых и рассматриваемых для различных целей (например, для оценки качества данных, систематических ошибок в данных, управления данными, разработки и использования систем искусственного интеллекта). Без всеобъемлющей структуры эти различные жизненные циклы данных могут быть сложны для правильной интерпретации теми, кто не имеет предварительных знаний, контекста и опыта. Существует риск того, что многочисленные жизненные циклы данных не будут применяться должным образом.

В настоящем стандарте приведен обзор стадий жизненного цикла данных в составе жизненного цикла системы искусственного интеллекта (см. раздел 5), описаны данные стадии (см. раздел 6) и предоставлены дополнительные сведения о процессах, связанных с ними (см. раздел 7).

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

Информационные технологии

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

Структура жизненного цикла данных

Information technology.
Artificial intelligence.
Data life cycle framework

Дата введения — 2024—01—01

1 Область применения

Настоящий стандарт определяет возможные стадии и связанные с ними действия по обработке данных на всем жизненном цикле системы искусственного интеллекта, включая комплектование, создание, разработку, развертывание, техническую поддержку и вывод из эксплуатации. Настоящий стандарт не определяет конкретные сервисы, платформы или инструменты. Данный стандарт применим ко всем организациям, независимо от их типа, размера и характера, которые используют данные при разработке и использовании систем искусственного интеллекта.

2 Нормативные ссылки

В настоящем стандарте нормативные ссылки *отсутствуют*.

3 Термины и определения

В настоящем стандарте применены термины по *[1]*.

4 Сокращения

В настоящем стандарте применены следующие сокращения:

ИИ — искусственный интеллект;
МО — машинное обучение;
ПДн — персональные данные;
DPIA — оценка воздействия обработки ПДн на права и свободы субъекта данных (data protection impact assessment);
JSON — текстовый формат описания объектов JavaScript (JavaScript object notation);
OWL — язык веб-онтологий (web ontology language);
XML — расширяемый язык разметки (extensible markup language).

Издание официальное

5 Обзор жизненного цикла данных

Жизненный цикл данных для систем искусственного интеллекта включает в себя процессы работы с данными от первоначального замысла создания новой системы ИИ до окончательного вывода этой системы из эксплуатации и разделен на несколько отдельных стадий. Каждая из этих стадий — часто, но не всегда — будет частью жизненного цикла данных в системе ИИ.

Жизненный цикл данных представляет собой все стадии, через которые данные могут проходить в системе ИИ, оперирующей данными произвольного типа. Он предназначен для поддержки достижения целей, связанных с управлением системой, полезностью системы, качеством данных и безопасностью данных, путем обеспечения того, чтобы обработке данных уделялось должное внимание при планировании, разработке, использовании и выводе системы из эксплуатации.

Детальное предназначение и последовательность прохождения этих стадий на протяжении жизненного цикла зависят от множества факторов, включая социальные, коммерческие, организационные и технические соображения, каждый из которых может меняться или иногда сочетаться с другими стадиями в течение жизненного цикла системы ИИ. Настоящий стандарт описывает следующие 10 стадий:

- стадия 1 — замысел;
- стадия 2 — формирование деловых требований;
- стадия 3 — планирование работы с данными;
- стадия 4 — комплектование наборов данных;
- стадия 5 — подготовка наборов данных;
- стадия 6 — построение модели ИИ;
- стадия 7 — развертывание системы ИИ;
- стадия 8 — эксплуатация системы ИИ;
- стадия 9 — вывод данных из эксплуатации;
- стадия 10 — вывод системы ИИ из эксплуатации.

Сведения о жизненном цикле данных при их использовании приведены в *[2]*.

6 Структура жизненного цикла данных

6.1 Общие положения

Структура жизненного цикла данных на рисунке 1 идентифицирует множество концептуально различных стадий оперирования данными в системе ИИ — от планирования работы с данными до вывода данных из эксплуатации. Рисунок 1 также включает замысел, формирование деловых требований и вывод системы ИИ из эксплуатации, которые относятся к стадиям жизненного цикла системы.

Сведения о наборах данных приведены в *[3]* (пункт 6.5). С каждой стадией жизненного цикла соотносятся процессы, соответствующие определенной задаче. Процессы жизненного цикла описывают действия с данными на конкретной стадии.

Стадия 9 (вывод данных из эксплуатации) и стадия 10 (вывод системы ИИ из эксплуатации) относятся к выводу из эксплуатации, но стадия 9 конкретно описывает, что происходит с данными (например, безопасное удаление, архивирование, перепрофилирование), а стадия 10 описывает, что происходит с системой независимо от того, что происходит с обрабатываемыми данными.

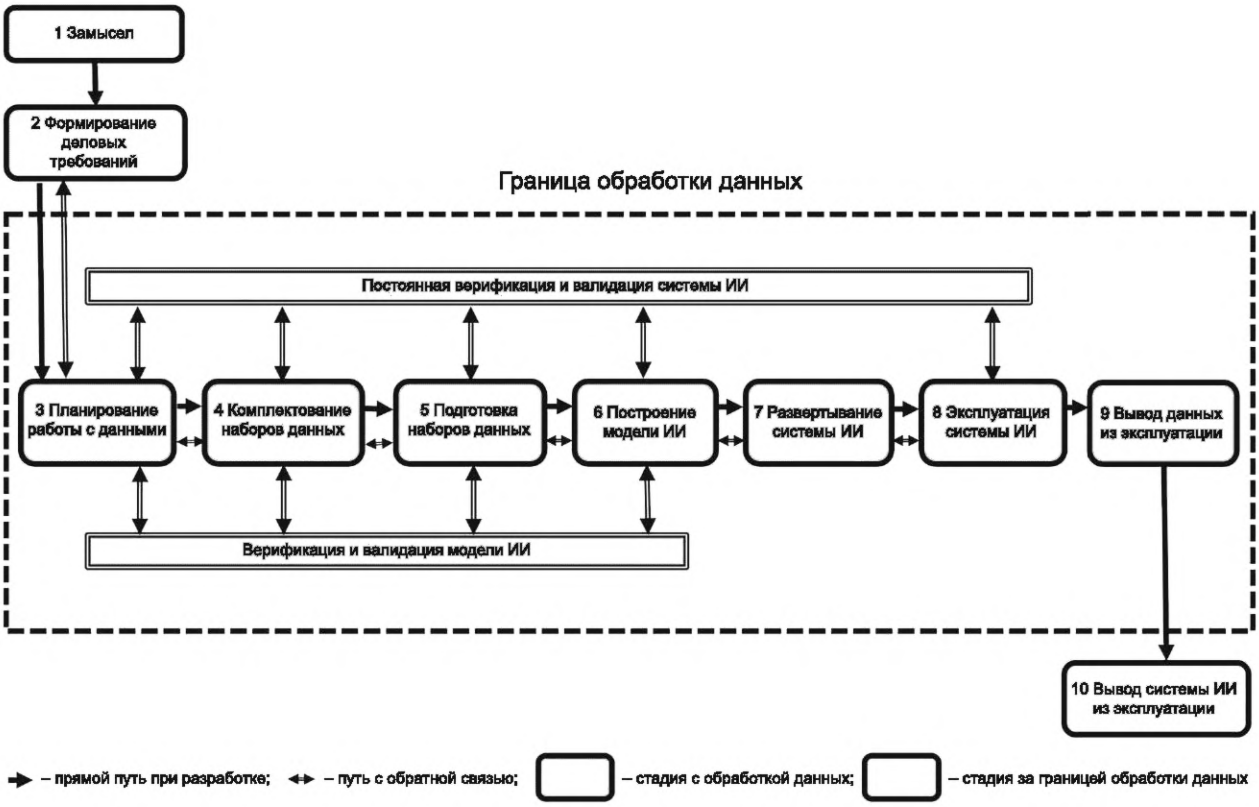

Рисунок 1 — Структура жизненного цикла данных

Примечания

1 Однонаправленные стрелки на рисунке 1 изображают прямой путь по стадиям жизненного цикла, а двунаправленные стрелки показывают путь с обратной связью между стадиями жизненного цикла.

2 Верификация и валидация модели ИИ относятся к внутреннему процессу разработки, результатом которого является модель. Верификация и валидация системы ИИ относятся к системе в целом в ходе ее эксплуатации.

6.2 Стадия 1. Замысел

Замысел включает в себя процесс осознания потребности или требований к новой или модернизируемой системе ИИ. Система ИИ может использоваться как частичное или полное решение существующей или предполагаемой задачи или проблемы, с которой сталкивается организация.

Замысел также может быть обусловлен более широкими потребностями организации (например, экономическими, техническими, стратегическими, рыночными или нормативными правовыми требованиями). В конечном счете замысел может быть зафиксирован в виде одного или нескольких вопросов, на которые может дать ответ система ИИ. Решаемые системой ИИ вопросы целесообразно сопоставлять и согласовывать с деловыми целями и показателями.

6.3 Стадия 2. Формирование деловых требований

На стадии формирования деловых требований могут участвовать представители одной или нескольких заинтересованных сторон, обладающих соответствующими полномочиями или влиянием, и решающих выяснить, можно ли превратить замысел в функционирующую систему ИИ и следует ли вкладывать в него дополнительные средства. Эта стадия включает:

- определение намерений проекта (например, формулирование видения, целей и стратегии);
- определение имеющихся и подлежащих приобретению активов;
- спецификацию требований к данным, ключевому элементу системы ИИ, исходя из целей и потребностей конечных пользователей;

- идентификацию факторов, влияющих на реализацию проекта, включая внутренние компетенции и знания, организационную структуру, технологии и внешние ресурсы;
- обеспечение возможности реализации проекта в соответствии с политикой и процедурами (или процессами) организации, включая такие, как:
  - соответствие (например, требования к защите ПДн),
  - этика (например, справедливость результатов),
  - культура,
  - лидерство,
  - процессы управления.

Стадия формирования деловых требований может завершаться подготовкой заключения об осуществимости проекта.

Примечание — На стадии формирования деловых требований никакие данные не обрабатываются.

6.4 Стадия 3. Планирование работы с данными

Стадия планирования работы с данными включает решение о составе наборов данных, которые нужны для ответов на вопросы, сформулированные на стадии формирования деловых требований. Основные аспекты, которые рассматриваются на этой стадии, включают:

- сведения о том, существуют ли необходимые данные, доступны ли они для повторного использования, нужно ли их приобретать, собирать, преобразовывать, создавать, курировать или сочетать некоторые из перечисленных выше действий;
- требуемый объем данных;
- источник данных;
- сведения о том, можно ли создать синтетические, т. е. искусственные, данные для дополнения имеющихся;
- сведения о том, какие выходные данные будут созданы и как система ИИ будет оперировать ними;

формат данных;

- сведения о том, что данные представляют;
- свойства данных, которые могут повлиять на выбор алгоритма для построения модели ИИ;
- требования к лицензированию данных;
- требования к безопасности, защите персональных данных и жизнестойкости данных;
- требования к комплектованию данных, в том числе к сбору данных;
- требования к защите данных;
- достоверность данных и наличие в них систематических ошибок;
- дискретность записи данных (в основном для временных рядов);
- тип данных (структурированные/частично структурированные/неструктурированные);
- требования к срокам и способам хранения данных (например, тип, стоимость, емкость, производительность, встроенные возможности, сроки удаления).

Целью этой стадии является обеспечение того, чтобы цели, требования и потребности, установленные на стадиях замысла и формирования деловых требований, могли быть выполнены. Потенциально необходимые наборы данных могут быть идентифицированы, получены и проверены. Наборы данных могут быть внутренними или получены из общедоступных источников, государственных органов, поставщиков или сторонних организаций.

6.5 Стадия 4. Комплектование наборов данных

Стадия комплектования наборов данных включает создание или обеспечение доступа к данным, определенным на стадии планирования работы с данными. Приобретенные данные могут поступать из внутренних источников, от третьих лиц или сообществ (например, открытые данные, общедоступные данные).

Получение данных от третьих лиц может основываться на согласиях, договорах или лицензиях, как описано в 6.4.

Данные могут быть в разных формах (например, статические данные, потоковые данные, данные интернета вещей в реальном времени) и в разных форматах (таких, как XML, JSON, текст с разделите-

лями, бинарный формат). Данные могут быть структурированными, полуструктурированными или неструктурированными.

В процессах комплектования наборов данных должны использоваться передовые практики управления данными, обеспечивающие безопасность, защиту ПДн и качество.

6.6 Стадия 5. Подготовка наборов данных

Стадия подготовки наборов данных включает в себя обработку данных, собранных на стадии комплектования. Стадия подготовки наборов данных может включать выполнение следующих операций.

- Расшифровка: преобразование зашифрованных данных в состояние, при котором их можно использовать в системе ИИ, когда это необходимо и возможно.
- Очистка: включает в себя преобразования и операции, такие как валидация релевантности, дедупликация, удаление выбросов, устранение систематических ошибок, дополнение недостающих значений, корректировка записей и исправление форматов данных.
- Инженерия признаков: чтобы повысить параметры производительности МО, для использования и возможных преобразований можно отобрать подходящие признаки. Имеющиеся признаки можно комбинировать и обрабатывать с целью выработки новых признаков, которые способны улучшить процессы обучения и вывода.
- Нормализация и масштабирование: может возникнуть необходимость преобразования данных из сильно различающихся диапазонов в определенный диапазон безразмерных величин (например, между 0 и 1). Также может быть необходимо масштабировать выборки данных, чтобы они соответствовали стандартному распределению или другому заданному распределению.
- Организация данных: данные могут быть реорганизованы без изменения их значения или смысла. Для получения полного набора данных может понадобиться еще добавить наборы данных или объединить таблицы. Может потребоваться объединить или разделить столбцы или поля для достижения необходимой структуры набора данных.
- Разметка (маркировка): значения целевых переменных должны быть установлены с помощью подходящего ручного или автоматического процесса. Например, при обучении с учителем значения можно определять вручную, при полуконтролируемом обучении значения могут быть определены с помощью автоматизированных методов. Примером ручной разметки данных для программы распознавания изображений является привлечение к анализу людей при определении видов животных в множестве цифровых изображений.
- Обогащение: запуск инструментов для связывания различных источников данных и добавления дополнительного контекста к данным. Например, неструктурированные данные могут быть обогащены средствами обработки естественного языка для извлечения именованных сущностей. Названия местностей и адреса могут быть идентифицированы и геокодированы с помощью географического справочника, чтобы впоследствии можно было проводить анализ на основе местоположения;
- Обезличивание: может возникнуть необходимость для удаления персональных данных из наборов данных для защиты неприкосновенности частной жизни субъектов данных.
- Повторная выборка и балансировка: например, может быть полезно использовать частичную выборку из больших наборов данных, чтобы улучшить согласованность статистической значимости различных классов данных или сократить время, необходимое для построения и тестирования модели ИИ, при этом получая полезные результаты. Точно так же можно пополнить выборку (т. е. расширить с заменой данных), чтобы повысить согласованность статистической значимости различных классов данных.
- Кодирование: может потребоваться кодирование данных, используемых для построения модели ИИ. Например, может потребоваться кодирование текстовых значений для категориальных переменных (такое как преобразование текстовых признаков в числовые или оцифровка аналоговых сигналов).
- Верификация целостности: применение процесса, специфичного для конкретного типа данных с целью проверки общей целостности набора данных. Скорее всего, это применимо к структурированным и частично структурированным данным, для которых уже может существовать структурная модель (например, схема базы данных, формальная онтология).
- Происхождение данных: обновление записи происхождения каждого набора данных для записи изменений и предпринятых операций.
- Анонимизация данных или псевдонимизация.

Примечание — Дополнительную информацию о подготовке наборов данных для МО см. в *[3]*.

6.7 Стадия 6. Построение модели ИИ

Построение модели ИИ включает принятие решения об организации, хранении и доступе к данным таким образом, чтобы их можно было обработать для построения модели, выполняющей отдельные функции для достижения целей. Процесс построения модели ИИ может либо завершиться созданием фиксированной модели, либо непрерывно продолжаться с постоянным пересмотром модели (непрерывное обучение). В любом случае модель ИИ может возникнуть как:

a) результат обучения алгоритма МО с использованием обучающих данных. Построение модели ИИ может происходить централизованно или через сеть ресурсов (например, федеративное обучение, раздельное обучение). Примеры обученных моделей ИИ включают деревья решений, индуктивное логическое программирование и различные типы нейронных сетей;

b) результат объединения инженерно-технических знаний (например, декларативных или процедурных) с процессом логического вывода. Примеры форм инженерных знаний человека включают предложения Хорна (например, используемые в языке программирования Prolog), разновидности логики описания (например, используемые в языке веб-онтологий OWL и OWL2) и программирование набора ответов.

Примечания

1 Индуктивное логическое программирование — это форма МО символьных структур, в которой логическая программа автоматически модифицируется для удовлетворения заданным целевым условиям.

2 Программирование набора ответов — это форма логического программирования, в которой используется алгоритм решения набора ответов для построения символической модели ИИ, в которой все переменные заменены литералами.

Данные используются для обучения и калибровки модели ИИ наряду с человеческим опытом, а также для верификации того, что результаты функционирования и производительность системы ИИ соответствуют ожиданиям заинтересованных сторон. Оценка воздействия обработки ПДн на права и свободы субъекта данных (Data Processing Impact Assessment, DPIA) может быть выполнена для решения таких вопросов, как соблюдение требований к защите персональных данных в полученных результатах. Модель ИИ также может быть оценена на наличие других потенциальных проблем (таких как предвзятость, справедливость, другие этические проблемы) и предпринятых корректирующих действий.

На стадии построения следует оценить модель ИИ и убедиться, что она соответствует требованиям, установленным на предыдущих стадиях (таких, как формирование деловых требований, планирование работы с данными, комплектование наборов данных). Аналогичным образом результаты и эффективность модели ИИ следует оценивать в сравнении с ожиданиями соответствующих заинтересованных сторон, включая их способность использовать модель на практике. В некоторых случаях для систем ИИ, связанных с общественной безопасностью, может потребоваться независимая оценка безопасности (Independent Safety Assessment, ISA).

6.8 Стадия 7. Развертывание системы ИИ

Развертывание системы ИИ включает в себя «запуск» системы в целевой среде. Данная стадия не обязательно представляет собой простое включение системы ИИ; вместо этого она может включать ряд процессов, гарантирующих, что система работает должным образом. На этой стадии необходимо проверить потоки данных, чтобы убедиться, что они работают так, как предполагалось, особенно если целевая среда включает новые системы ИИ или соединения.

6.9 Стадия 8. Эксплуатация системы ИИ

Эксплуатация системы ИИ включает использование модели ИИ, генерирующей выходные данные из входных (или произведенных) данных, *с возможностью дополнительного обучения на обогащенных данных (при необходимости).*

Входные данные для дополнительного обучения моделей ИИ, находящихся в составе эксплуатируемой системы, должны:

- соответствовать всем требованиям к данным, согласованным на стадиях 4 и 5;

- иметь возможность быть использованными при дополнительном обогащении набора данных, который использовался для обучения модели на стадии 6, без возможности изменения структурной модели данных этого набора.

Дополнительные обогащенные данные для обучения моделей ИИ должны иметь возможность использования при дополнительном обучении моделей на стадии эксплуатации системы ИИ.

Выходные данные также могут быть обработаны несколькими способами, такими как:

- прием или извлечение данных;
- псевдонимизация или анонимизация данных;
- манипулирование данными или комбинирование данных;
- анализ данных;
- визуализация данных;
- передача данных;
- хранение данных.

Эта стадия может включать авторизацию доступа к данным, аутентификацию и предполагаемое использование.

Произведенные в ходе эксплуатации данные должны постоянно контролироваться, чтобы гарантировать, что качество данных поддерживается, а система ИИ не используется для целей, которые изначально не предполагались.

Проведение непрерывной верификации и валидации системы ИИ может снизить риски, но не всегда является возможным или целесообразным. В случае, когда это возможно и целесообразно, рекомендуется постоянно верифицировать и валидировать систему ИИ, чтобы удовлетворить требования бизнеса и ожидания заинтересованных сторон. Система ИИ может улучшаться по мере необходимости и использовать новые обучающие данные для смягчения снижения производительности. Управление новыми обучающими данными должно осуществляться на стадии 4 (комплектование наборов данных) и стадии 5 (подготовка наборов данных).

6.10 Стадия 9. Вывод данных из эксплуатации

Вывод данных из эксплуатации предусматривает решение судьбы данных, которые больше не используются системой ИИ (например, их безопасное удаление, архивирование, перепрофилирование). Категории данных должны быть определены, а некоторые категории данных должны быть сохранены для целей аудита (например, данные журналирования для подтверждения соответствия).

Эксплуатация модели ИИ может быть также прекращена, если в ней остались элементы обучающих данных, или из-за других требований (например, безопасности, приватности и конфиденциальности). Требования лицензирования данных могут предполагать удаление данных определенного уровня. Кроме того, может потребоваться удаление ПДн, например, в соответствии с требованиями законодательства или условиями договора.

6.11 Стадия 10. Вывод системы ИИ из эксплуатации

Вывод системы ИИ из эксплуатации включает в себя прекращение обработки данных и утилизацию компонентов системы, на которые не распространяется вывод данных из эксплуатации, таких как компоненты целевой среды. Данные, не относящиеся к модели ИИ, такие как системные журналы, можно сохранить для дальнейшего изучения или безопасно удалить.

7 Стадии и процессы жизненного цикла данных

На рисунке 2 показаны процессы и сквозные аспекты, важные для жизненного цикла данных.

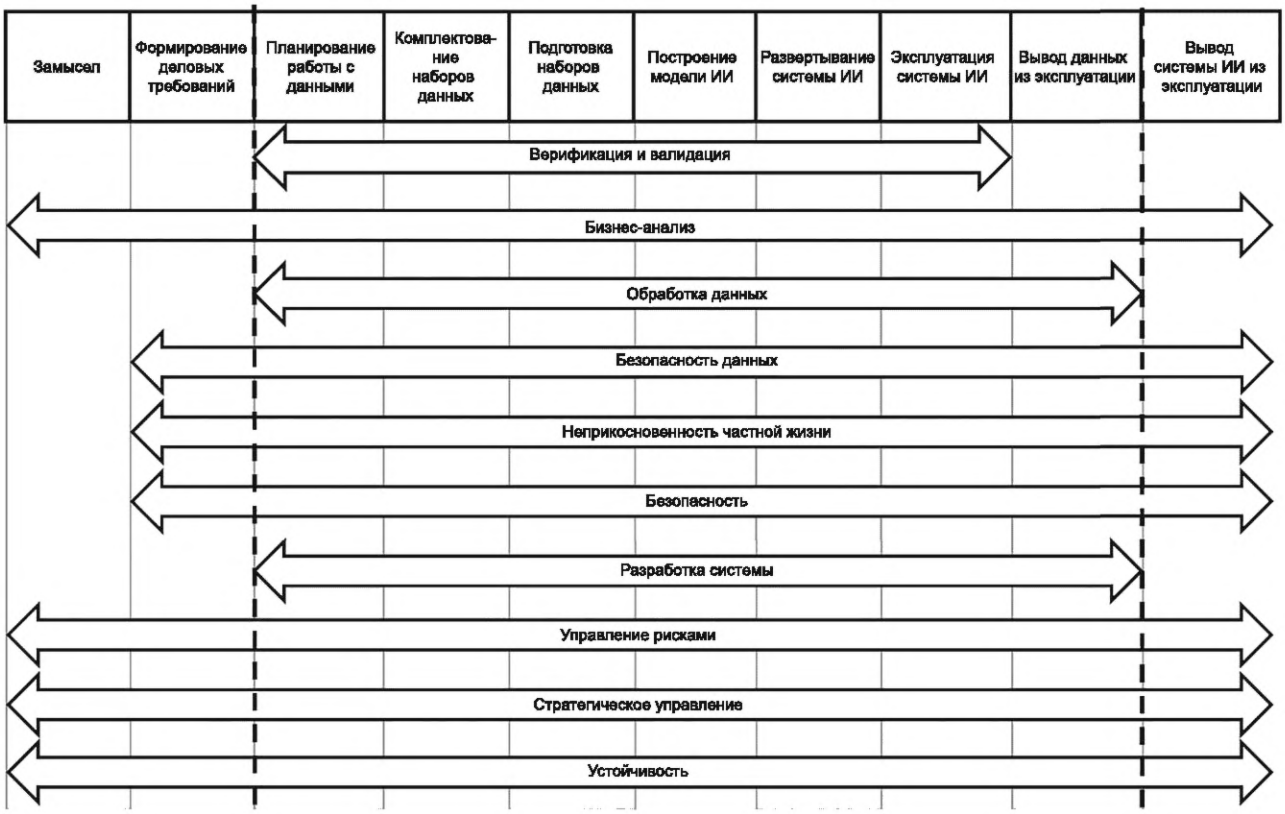

Рисунок 2 — Соотношение процессов и стадий жизненного цикла данных

К ним относятся следующие процессы:

- валидация и верификация: два вопроса «Построили ли мы правильную систему?» и «Построили ли мы систему правильно?» постоянно проверяются путем тестирования выходных результатов системы ИИ на соответствие деловым требованиям;
- бизнес-анализ: взаимосвязь между требованиями организации и способностью системы ИИ и особенно используемых в ней данных, соответствовать этим требованиям анализируется на различных стадиях жизненного цикла данных;
- обработка данных: под обработкой понимаются любые операции с данными, включая сбор, сохранение, передачу, удаление или иные действия. В связи с этим возросла важность защиты персональных данных и особенно специальных категорий персональных данных. Сведения о происхождении данных, используемых в системах ИИ, должны актуализироваться в соответствии с требованиями организации;
- безопасность данных: данные в системе ИИ должны быть конфиденциальными (т. е. защищенными от несанкционированного доступа), доступными (т. е. доступными для санкционированного доступа) и целостными (т. е. защищенными от несанкционированного изменения).

Примечание — Для обеспечения безопасности данных необходимо учитывать категории данных (особенно специальные категории персональных данных) и особые требования при работе с ними в отдельных сферах деятельности;

- неприкосновенность частной жизни: ПДн должны быть защищены, а их целостность и доступность должны поддерживаться на основе принципов конфиденциальности, целостности, доступности. Процессы должны включать защиту ПДн, а также предотвращать случайное или преднамеренное создание новых ПДн;
- безопасность: необходимо обеспечить эффективное и результативное оперирование данными системой ИИ для снижения потенциальных рисков или вреда, которые могут быть ею вызваны.

Пример 1 — Разметка данных: система ИИ, которая обучается для корректной разметки визуального контента для ограничения просмотра по возрасту. Если система ИИ не была обучена избегать

неправильной идентификации элементов визуального контента по возрасту, то в процессе эксплуатации она подвергает будущую аудиторию риску.

Пример 2 — Мониторинг данных: непрерывный мониторинг данных, которые вводятся для анализа, необходим для выявления изменений в данных, которые не могут быть обработаны в процессе анализа. Если начальные условия неверны, то возможными исходами работы системы ИИ станут незавершенность или бессмысленные результаты. Анализ может быть ошибочным, потому что система ИИ оперирует данными, которые не удовлетворяют свойствам, которые она должна была учитывать.

Пример 3 — Нормативные правовые требования и обеспечение их исполнения: если данные могут использоваться в приложениях ИИ, на которые распространяются конкретные требования законодательства или требования общественной безопасности (например, в железнодорожной, авиационной, атомной, медицинской или нефтегазовой отраслях), то к данным могут возникнуть дополнительные требования.

Примечание — Исполнение требований к наборам данных согласовывается с регулирующим органом или оценивается им. Это может быть сделано как в рамках оценки функциональной безопасности, так и в более широком контексте. Требования могут применяться как для статических данных, так и для текущих или динамических данных;

- разработка системы ИИ: создание системы, которую можно использовать для выполнения деловых требований организации, включая их непрерывную валидацию и верификацию;
- управление рисками: идентификация и управление любыми рисками, связанными с системой ИИ, на основе склонности организации к риску. Сюда включены риски, связанные с защитой персональных данных и безопасностью системы и данных в ней;
- стратегическое управление: роль руководящего органа и менеджеров в эксплуатации и выводе из эксплуатации системы ИИ и данных, используемых в ней. Это включает рассмотрение этических проблем, соблюдение требований законодательства, стандартов и передовой практики, подотчетности, управления рисками и фидуциарных обязательств (см. *[4]*);
- социальная и экологическая устойчивость: учитываются социальные и экологические последствия разработки и использования системы ИИ, чтобы данные, обрабатываемые в системе, можно было обрабатывать максимально устойчивым образом. Это относится как к физическому функционированию системы ИИ (например, путем обеспечения возможности работы в периоды низкого энергопотребления или предотвращения ненужного использования системы), так и к результатам функционирования системы (таким, как соблюдение этических норм).

Библиография

[1]	ИСО/МЭК 22989:2022	Information technology — Artificial intelligence — Artificial intelligence concepts and terminology
[2]	ИСО/МЭК 5212[1)]	Information technology — Data usage — Guidance for data usage
[3]	ИСО/МЭК 23053	Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML)
[4]	ИСО/МЭК 38507	Information technology — Governance of IT — Governance implications of the use of artificial intelligence by organizations

[1)] На момент публикации в разработке находится международный стандарт ISO/IEC DIS 5212:2023.

УДК 004.01:006.354 ОКС 35.020

Ключевые слова: информационные технологии, искусственный интеллект, рамочная структура, жизненный цикл данных, большие данные, набор данных, модель данных, комплектование наборов данных, подготовка наборов данных, вывод данных из эксплуатации, жизненный цикл системы искусственного интеллекта

Редактор *М.В. Митрофанова*
Технический редактор *И.Е. Черепкова*
Корректор *Р.А. Ментова*
Компьютерная верстка *М.В. Малеевой*

Сдано в набор 07.11.2023. Подписано в печать 21.11.2023. Формат 60×84⅛. Гарнитура Ариал.
Усл. печ. л. 1,86. Уч.-изд. л. 1,58.

Подготовлено на основе электронной версии, предоставленной разработчиком стандарта

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

gost.e-ecolog.ru