



ПРЕДВАРИТЕЛЬНЫЙ
НАЦИОНАЛЬНЫЙ
СТАНДАРТ
РОССИЙСКОЙ
ФЕДЕРАЦИИ

**ПНСТ
842—
2023
(ИСО/МЭК 25059:
2023)**

Программная инженерия

**ТРЕБОВАНИЯ И ОЦЕНКА КАЧЕСТВА СИСТЕМ
И ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ (SQuaRE)**

**Модель качества для систем
искусственного интеллекта**

(ISO/IEC 25059:2023, MOD)

Издание официальное

Москва
Российский институт стандартизации
2023

Предисловие

1 ПОДГОТОВЛЕН Федеральным государственным автономным образовательным учреждением высшего образования «Национальный исследовательский университет «Высшая школа экономики» (НИУ ВШЭ) на основе собственного перевода на русский язык англоязычной версии документа, 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 7 ноября 2023 г. № 50-пнст

4 Настоящий стандарт является модифицированным по отношению к международному стандарту ИСО/МЭК 25059:2023 «Программная инженерия. Требования и оценка качества систем и программного обеспечения (SQuaRE). Модель качества для систем на основе искусственного интеллекта» [ISO/IEC 25059:2023 «Software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Quality model for AI systems», MOD] путем изменения отдельных фраз (слов, значений, показателей, ссылок), которые выделены в тексте курсивом.

Внесение указанных технических отклонений направлено на учет особенностей российской национальной стандартизации.

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

Правила применения настоящего стандарта и проведения его мониторинга установлены в ГОСТ Р 1.16—2011 (разделы 5 и 6).

Федеральное агентство по техническому регулированию и метрологии собирает сведения о практическом применении настоящего стандарта. Данные сведения, а также замечания и предложения по содержанию стандарта можно направить не позднее чем за 4 мес до истечения срока его действия разработчику настоящего стандарта по адресу: info@tc164.ru и/или в Федеральное агентство по техническому регулированию и метрологии по адресу: 123112 Москва, Пресненская набережная, д. 10. стр. 2.

В случае отмены настоящего стандарта соответствующая информация будет опубликована в ежемесячном информационном указателе «Национальные стандарты» и также будет размещена на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет (www.rst.gov.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

II

Содержание

1 Область применения	1
2 Нормативные ссылки	1
3 Термины и определения	2
4 Сокращения	3
5 Модель качества продукта	3
6 Модель качества при использовании	6
Приложение А (справочное) Требования к качеству систем и программного обеспечения SQuaRE . . .	9
Приложение В (справочное) Соотношение риск-ориентированного подхода с качеством-ориентированным подходом и с моделями качества	10
Приложение С (справочное) Производительность	12
Приложение ДА (справочное) Сведения о соответствии ссылочных национальных стандартов международным стандартам, использованным в качестве ссылочных в примененном международном стандарте	13
Библиография	14

Введение

Высококачественные программные продукты и компьютерные системы представляют значимость для заинтересованных сторон. Модели качества, требования к качеству, измерение качества и оценка качества были стандартизированы и вошли в серию стандартов SQuaRE (см. приложение А).

Системы искусственного интеллекта (ИИ) требуют учета дополнительных свойств и характеристик систем, а заинтересованные стороны имеют разные потребности. Системы ИИ имеют разные свойства и характеристики. Например, системы ИИ могут:

- самостоятельно принимать решения;
- основываться на искаженных или неполных данных;
- являться вероятностными;
- адаптироваться в процессе эксплуатации.

Согласно [1], под надежностью понимается и рассматривается непрерывный организационный процесс, а также нефункциональное требование, которое уточняет возникающие свойства системы, т. е. набор ее характеристик с атрибутами в контексте качества использования, описанных в *ГОСТ Р ИСО/МЭК 25010*.

В [1] рассматривается применимость по отношению к системам ИИ, которые были разработаны для обычного программного обеспечения. Согласно [1], непредсказуемый характер систем ИИ на основе данных учитывается в недостаточной степени. Принимая во внимание существующий объем работ на данную тему, в [1] определяется необходимость разработки новых стандартов для систем ИИ, в которых будут рассмотрены не только характеристики и требования для разработки традиционного программного обеспечения.

В [1] рассматриваются различные подходы к тестированию и оценке систем ИИ. В отчете отмечается, что для тестирования системы ИИ, модифицированной версии существующего программного и аппаратного обеспечения необходимо применять методы верификации и валидации. Также там определяются несколько концептуальных различий между многими системами ИИ и обычными системами и делается вывод о том, что «способность системы [ИИ] достичь запланированного и желаемого результата не всегда поддается измерению с помощью традиционных подходов тестирования программного обеспечения». Тестирование систем ИИ описывается в [2].

В настоящем стандарте рассматривается расширение модели качества SQuaRE, описанной в *ГОСТ Р ИСО/МЭК 25010*, для прикладной системы ИИ.

Системы ИИ выполняют задачи. Для одной системы ИИ можно поставить одну или несколько задач. Для оценки выполнения задачи могут быть установлены требования к качеству.

Модель качества рассматривается в двух плоскостях: качество продукта, описанное в пункте 5, и качество при использовании, описанное в пункте 6. Объясняется релевантность данных терминов и приводятся ссылки на результаты в других стандартах (например, *ГОСТ Р 70462.1*, [3]).

В [4] описывается модель качества данных, которая дополняет модель, описанную в настоящем стандарте. [4] расширяется для систем ИИ в [5].

ПРЕДВАРИТЕЛЬНЫЙ НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

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

Срок действия — с 2024—01—01
до 2027—01—01

1 Область применения

Настоящий стандарт устанавливает модель качества для систем ИИ и продолжает серию стандартов «Требования и оценка качества систем и программного обеспечения» (SQuaRE). В характеристиках и подхарактеристиках (субхарактеристиках), описанных в данной модели качества, приводится терминология для определения, измерения и оценки качества системы ИИ. Там же приводятся и сопоставляются характеристики качества с требованиями к качеству для обстоятельного и многостороннего охвата изучаемого предмета.

2 Нормативные ссылки

В настоящем стандарте использованы нормативные ссылки на следующие стандарты:

ГОСТ Р ИСО/МЭК 25010—2015 Информационные технологии. Системная и программная инженерия. Требования и оценка качества систем и программного обеспечения (SQuaRE). Модели качества систем и программных продуктов

ГОСТ Р 70462.1/ISO/IEC TR 24029-1:2021 Информационные технологии. Интеллект искусственный. Оценка робастности нейронных сетей. Часть 1. Обзор

ПНСТ 835—2023 (ISO/IEC TS 4213:2022) Искусственный интеллект. Оценка эффективности моделей и алгоритмов машинного обучения в задаче классификации

ПНСТ 836—2023 (ISO/IEC TR 5469)¹⁾ Искусственный интеллект. Функциональная безопасность и системы искусственного интеллекта

ПНСТ 838—2023 (ИСО/МЭК 23053:2022) Искусственный интеллект. Структура описания систем искусственного интеллекта, использующих машинное обучение

ПНСТ 839—2023 (ISO/IEC TR 24027:2021) Искусственный интеллект. Смещенность в системах искусственного интеллекта и при принятии решений с помощью искусственного интеллекта

ПНСТ 840—2023 (ISO/IEC TR 24368:2022) Искусственный интеллект. Обзор этических и общественных аспектов

Примечание — При использовании 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 «На-

¹⁾ Международный документ ISO/IEC TR 5469 «Искусственный интеллект. Функциональная безопасность и системы искусственного интеллекта» находится в стадии разработки.

циональные 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.

3 Термины и определения

В настоящем стандарте применены термины по [6] и ПНСТ 838—2023, а также следующие термины с соответствующими определениями.

3.1 Общие положения

3.1.1

показатель (measure): Переменная, которой присваивается какое-то значение как конкретный результат измерения.

Примечание — Форма множественного числа «показатели» используется для ссылки на основные показатели, производные показатели и индикаторы.

[ГОСТ Р 58606—2019, пункт 3.15]

3.1.2

измерять (measure): Проводить измерения (ИСО/МЭК 14598-1).

[ГОСТ Р ИСО/МЭК 25010—2015, пункт 4.4.6]

3.1.3

показатель качества программного обеспечения (software quality measure): Показатель внутреннего качества, показатель внешнего качества программного обеспечения или показатель качества программного обеспечения при использовании.

Примечание — Показатель внутреннего качества, показатель внешнего качества программного обеспечения или показатель качества программного обеспечения при использовании определены в модели качества в ИСО/МЭК 25010.

[ГОСТ Р ИСО/МЭК 25040—2014, пункт 4.61]

3.1.4

мера по уменьшению риска (risk treatment measure): защитная мера, действие или средство, используемые для исключения опасностей или уменьшения рисков.

[Адаптировано из [7], пункт 3.13, — «понижение» заменено на «уменьшение»]

3.1.5 прозрачность (transparency): Степень, с которой заинтересованные стороны получают требуемую информацию о системе ИИ.

Примечание — Требуемая для прозрачности системы ИИ информация может включать такие аспекты как: возможности, компоненты, процедуры, меры, цели проектирования, выбор дизайна и допущения.

3.2 Качество продукта

3.2.1 управляемость со стороны пользователя (user controllability): Степень, с которой пользователь может своевременно вмешаться в работу системы.

3.2.2 функциональная адаптируемость (functional adaptability): Степень, с которой система ИИ может точно получать информацию из данных или из результатов предыдущих действий и использовать эту информацию для будущих прогнозов.

3.2.3

функциональная корректность (functional correctness): Степень обеспечения продуктом или системой необходимой степени точности корректных результатов.

Примечание — Системы ИИ, в особенности те, которые используют модели машинного обучения, обычно не обеспечивают функциональную корректность во всех условиях эксплуатации.

[Адаптировано из *ГОСТ Р ИСО/МЭК 25010—2015*, пункт 4.2.1.2, — Добавлено примечание]

3.2.4 **возможность вмешательства** (intervenability): Степень, с которой оператор может своевременно вмешаться в работу системы для того, чтобы избежать опасности или ущерба.

3.2.5 **робастность** (robustness): Степень, с которой система ИИ может поддерживать свой уровень функциональной корректности при любых обстоятельствах.

3.3 Качество при использовании

3.3.1 **смягчение отрицательных последствий социального и этического риска** (societal and ethical risk mitigation): Степень, с которой система ИИ смягчает потенциальные риски для общества.

Примечание — Смягчение отрицательных последствий социального и этического риска включает в себя: отслеживаемость, непредвзятость, прозрачность и объяснимость, профессиональную ответственность, продвижение человеческих ценностей, конфиденциальность, контроль человека над технологией, участие в жизни общества и их развитии, соблюдение верховенства закона, соблюдение норм поведения и трудовых практик.

4 Сокращения

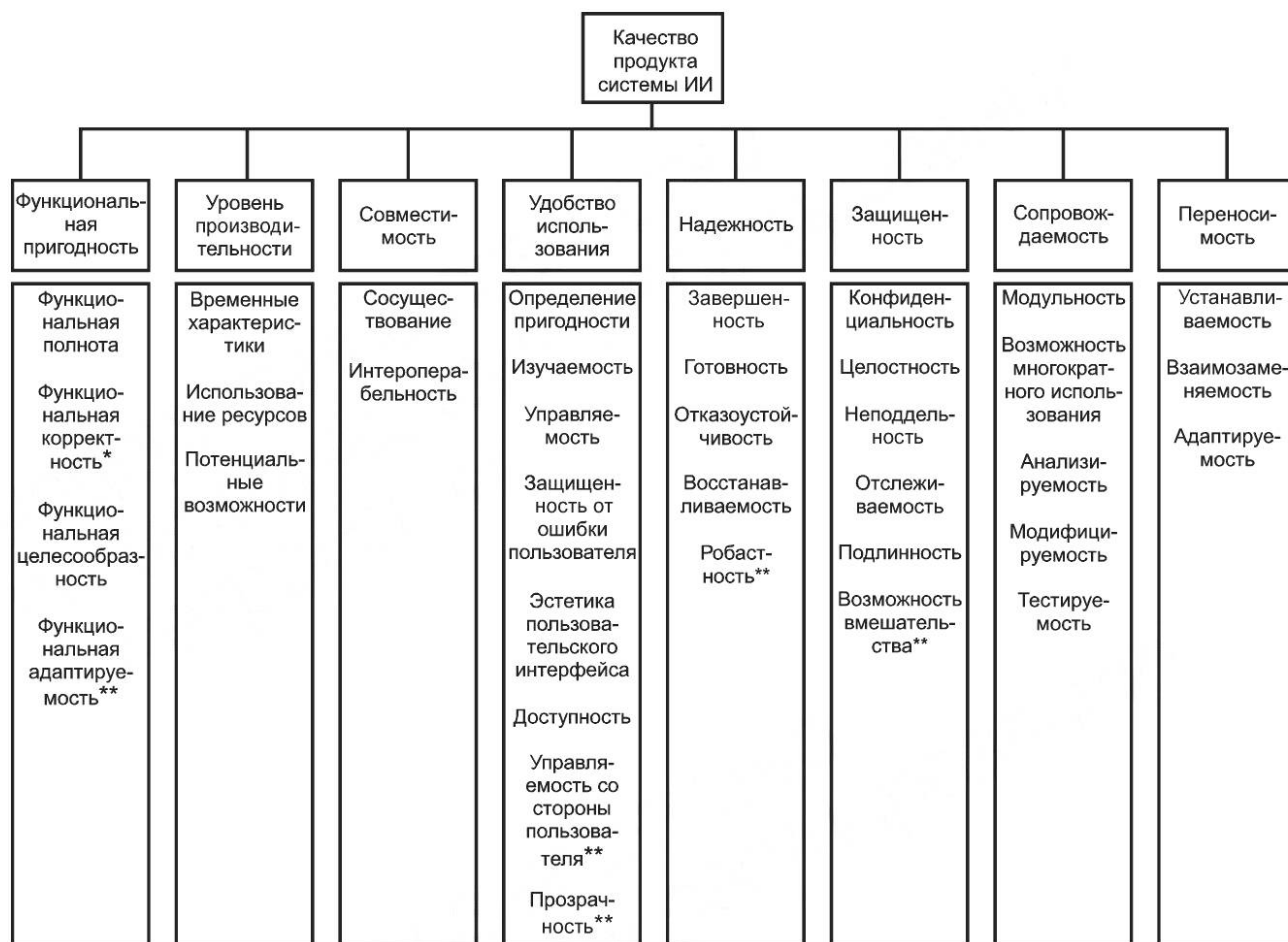
ИИ — искусственный интеллект;

МО — машинное обучение.

5 Модель качества продукта

5.1 Общие положения

На рисунке 1 представлена модель качества продукта системы ИИ. Данная модель основана на измененной версии основной модели качества системы, описанной в *ГОСТ Р ИСО/МЭК 25010*. Новые и измененные подхарактеристики обозначены сноской. Некоторые подхарактеристики имеют другое значение или область употребления по сравнению с моделью по *ГОСТ Р ИСО/МЭК 25010*. Все изменения, дополнения и отличия описаны в настоящем разделе. Исходные неизмененные характеристики являются частью модели качества продукта системы ИИ и должны толковаться в соответствии с *ГОСТ Р ИСО/МЭК 25010*.



* Измененные подхарактеристики.

** Новые подхарактеристики.

Рисунок 1 — Модель качества продукта системы ИИ

В настоящем разделе приведено описание измененных и новых подхарактеристик.

5.2 Управляемость со стороны пользователя

Управляемость со стороны пользователя является новой подхарактеристикой удобства использования. Управляемость со стороны пользователя — это свойство системы ИИ, при котором человек или другой внешний агент может своевременно вмешаться в работу системы. Если непредвиденное поведение системы не может быть полностью исключено и может привести к негативным последствиям, повышенная подконтрольность будет полезна.

Управляемость со стороны пользователя связана с управляемостью, которая описана в [6].

5.3 Функциональная адаптируемость

Функциональная адаптируемость является новой подхарактеристикой функциональной пригодности. Функциональная адаптируемость системы ИИ — это способность данной системы адаптировать себя к меняющейся динамической среде, в которой она функционирует. Системы ИИ могут обучаться на основе новых данных для обучения, производственных данных или результатов, полученных на основе предыдущего опыта работы системы. Концепция функциональной адаптируемости основана на концепции непрерывного обучения, которая представлена в [6].

Непрерывное обучение не является обязательным требованием для функциональной адаптируемости. Например, система, которая дискретно переключается между различными моделями классификации, с учетом событий в окружающей среде, может рассматриваться как функционально адаптируемая.

Функциональная адаптируемость в системах ИИ отличается от других характеристик качества тем, что результаты этих систем не могут быть интерпретированы линейно (например, от плохого к хорошему). Как правило, наличие высокой степени функциональной адаптируемости может привести к улучшению результатов, выданных системами ИИ.

В некоторых системах высокая степень функциональной адаптируемости может привести к получению дополнительных нежелательных результатов, которые будут в большей степени основываться на предыдущем выборе системы. Взвешивание пути принятия решений с относительно высокой неопределенностью, усиленное предыдущим решением системы ИИ, может привести к повышенной вероятности получения непреднамеренных отрицательных результатов. В этом случае функциональная адаптируемость может повысить степень негативной предвзятости.

В то время как традиционные алгоритмы как правило выдают одинаковый результат при одном и том же наборе входных данных, системы ИИ, благодаря непрерывному обучению, могут демонстрировать разное поведение и, следовательно, выдавать разные результаты.

5.4 Функциональная корректность

Функциональная корректность описана в *ГОСТ Р ИСО/МЭК 25010*. В модель качества продукта системы ИИ внесены поправки в описание, так как системы ИИ, в частности вероятностные методы МО, обычно не обеспечивают полной функциональной корректности, поскольку в их выходных данных ожидается определенная частота появления ошибок. В связи с этим, необходимо тщательно измерять корректность и неточность. Для этого существуют многочисленные измерения в области методов МО с примерами, которые можно применить к модели классификации, описанной в *ПНСТ 835—2023*.

Вдобавок может существовать трейд-офф (одновременная недостижимость) между характеристиками: уровень производительности, робастность и функциональная корректность.

В приложении С представлена дополнительная информация о том, почему термину функциональная корректность отдается предпочтение по сравнению с другими более распространенными, такими как общая производительность, для описания корректности модели.

5.5 Робастность

Робастность является новой подхарактеристикой характеристики надежности. Она используется для описания способности системы поддерживать определенный уровень функциональной корректности (см. приложение С, в котором описан термин производительность) при различных обстоятельствах, таких как:

- наличие ввода непредвиденных, смещенных, конфликтующих или некорректных данных;
- внешнее вмешательство;
- способность работы в различных условиях среды, включая способность к обобщению, отказоустойчивость, надежность;
- атрибуты, связанные с надлежащей работой системы, заложенные разработчиками.

Надлежащая работа системы необходима для защищенности системы и безопасности ее заинтересованных сторон в данной среде или области. Информация о функциональной безопасности в области систем ИИ содержится в *ПНСТ 836—2023*.

Робастность рассматривается в [1], методы оценки описываются в *ГОСТ Р 70462.1* и определяются в [3].

5.6 Прозрачность

В модели качества продукта прозрачность является подхарактеристикой удобства использования, а в модели качества при использовании — подхарактеристикой удовлетворенности.

Она отражает степень, с которой до заинтересованных сторон доводится релевантная информация о системе ИИ.

Прозрачность систем ИИ помогает потенциальным пользователям выбрать ту систему, которая соответствует их требованиям, а также повышает знания заинтересованных сторон о применимости и ограничениях системы ИИ, а также объяснимости систем ИИ.

Прозрачность может включать в себя информацию о функциональности системы ИИ, ее составных частях, об интерфейсах, об используемых моделях МО, о данных для обучения, данных для верификации и валидации, о бенчмаркинге (результатах сравнения производительности на эталонных набо-

рах данных с другими системами), о логах (журналах событий) и практиках управления в организации, ответственной за данную систему.

Прозрачные системы ИИ документируют, регистрируют или отображают свои внутренние процессы при помощи инструментов внутреннего контроля и файлов данных. Поток данных можно отслеживать на каждом этапе, при этом принимаемые решения, обработка исключений и правила принятия решений документируются. Содержание логов (журнала событий) отслеживает прохождение отдельных этапов в пайплайнах (цепочке вычислений), в которых преобразуются данные, а также отслеживает вызовы на системном уровне. Ошибки регистрируются подробно, особенно на этапах преобразования данных. Системы ИИ с высокой степенью прозрачности состоят из хорошо задокументированных подкомпонентов, интерфейсы которых подробно описаны. Прозрачность всех систем ИИ облегчает исследование системных сбоев.

В системе с низкой степенью прозрачности трудно изучить внутренние процессы извне. Отсутствие подробной регистрации процессов может негативно влиять на тестируемость, оценку социального и этического фактора и обработку риска.

В итоге прозрачность систем ИИ способствует установлению доверия, подотчетности и взаимодействия между заинтересованными сторонами. Некоторые аспекты прозрачности описаны в [1].

5.7 Возможность вмешательства

Степень возможности вмешательства может быть определена в зависимости от области, в которой может использоваться система ИИ. Решающим фактором использования возможности вмешательства является наблюдение за состоянием системы ИИ, переход от опасного к безопасному состоянию. Под управляемостью понимается состояние, в котором система ИИ имеет атрибуты, позволяющие осуществлять управление и контроль, что подчеркивает важность пользовательского интерфейса системы ИИ. По сравнению с управляемостью, возможность вмешательства имеет большую значимость и предназначена предотвращать возникновение опасности или ущерба со стороны системы ИИ.

Возможность вмешательства относится к управляемости, которая описана в [6].

6 Модель качества при использовании

6.1 Общие положения

На рисунке 2 представлена модель качества при использовании системы ИИ. Данная модель представляет собой измененную версию модели качества при использовании, описанной в *ГОСТ Р ИСО/МЭК 25010*. Новые подхарактеристики обозначены сноской. Некоторые подхарактеристики имеют другое значение или область употребления по сравнению с моделью по *ГОСТ Р ИСО/МЭК 25010*. Все изменения, дополнения и отличия описаны в настоящем разделе. Неизменные характеристики являются частью модели качества при использовании и должны толковаться в соответствии с *ГОСТ Р ИСО/МЭК 25010*.



* Новые подхарактеристики

Рисунок 2 — Модель качества при использовании системы ИИ

6.2 Смягчение отрицательных последствий этического и социального риска

Смягчение отрицательных последствий этического и социального риска является новой подхарактеристикой характеристики свобода от риска. В ПНСТ 840—2023 наряду с данным понятием описываются следующие:

- подотчетность;
- непредвзятость и отсутствие дискриминации;
- прозрачность и объяснимость;
- профессиональная ответственность;
- продвижение человеческих ценностей;
- конфиденциальность;
- безопасность и защищенность;
- контроль человека над технологией;
- повышение и развитие вовлеченности общества;
- ориентированный на человека дизайн;
- соблюдение верховенства закона;
- соблюдение норм поведения;
- устойчивое развитие окружающей среды;
- соблюдение трудовых практик.

Данные термины следует рассматривать в соответствии с ПНСТ 840—2023, наряду с желаемыми характеристиками системы для смягчения отрицательных последствий этического и социального риска.

Безопасность и защищенность, а также устойчивое развитие окружающей среды включены в существующие характеристики SQuaRE, такие как смягчение отрицательных последствий риска здоровья и безопасности и экологического риска, и поэтому не рассматриваются в настоящем стандарте.

В приложении В приводится сравнение риск-ориентированного подхода с качеством-ориентированным подходом, а также с моделями качества. На примерах демонстрируется, как оба подхода могут дополнять друг друга.

В ПНСТ 839—2023 в контексте непредвзятости и недискриминации рассматривается проблема смещенности оценок моделей МО.

6.3 Прозрачность

В модели качества продукта прозрачность является подхарактеристикой удовлетворенности (см. 5.6).

Приложение А (справочное)

Требования к качеству систем и программного обеспечения SQuaRE

А.1 Организация серии стандартов SQuaRE

На рисунке А.1 (адаптирован из *ГОСТ Р ИСО/МЭК 25010*) представлена серия стандартов SQuaRE.

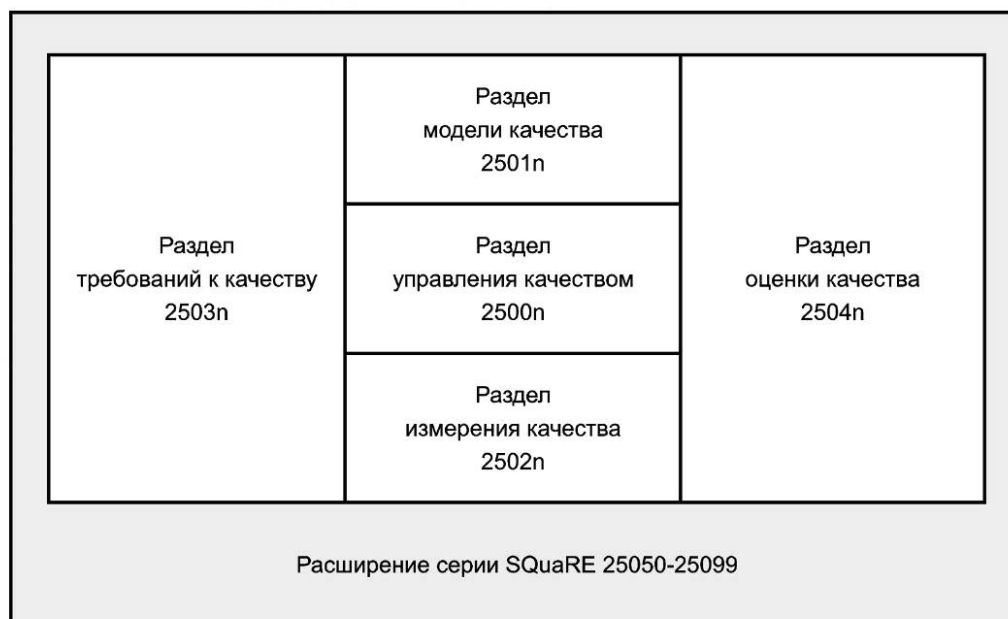


Рисунок А.1 — Организация серии стандартов SQuaRE

В серию SQuaRE входят следующие разделы:

- ИСО/МЭК 2500n — раздел «Управление качеством». Стандарты, входящие в этот раздел, определяют все общие модели, термины и определения, используемые всеми другими стандартами серии SQuaRE. В разделе также представлены требования и методические материалы по поддерживающей функции, отвечающей за управление требованиями к программной продукции, спецификацию и оценку;

- ИСО/МЭК 2501n — раздел «Модель качества». Стандарты, которые входят в этот раздел, представляют детализированные модели качества программной продукции, качества при использовании и качества данных. Кроме того, представлено практическое руководство по использованию модели качества;

- ИСО/МЭК 2502n — раздел «Измерение качества». Стандарты, входящие в этот раздел, включают в себя эталонную модель измерения качества программной продукции, математические определения показателей качества и практическое руководство по их использованию. В этом разделе представлены показатели внутреннего качества программной продукции, показатели внешнего качества программной продукции и показатели качества при использовании. Кроме того, определены и представлены элементы показателей качества — ЭПК, формирующие основу для вышеперечисленных показателей;

- ИСО/МЭК 2503n — раздел «Требования к качеству». Стандарты, которые входят в этот раздел, помогают задать требования к качеству на основе моделей качества и показателей качества. Такие требования к качеству могут использоваться в процессе формирования требований к качеству при сборе информации перед разработкой программной продукции или как исходные данные для процесса оценки;

- ИСО/МЭК 2504n — раздел «Оценка качества». Стандарты, которые входят в этот раздел, формулируют требования, рекомендации и методические материалы для оценки программного продукта, выполняемой как независимыми оценщиками, так и приобретателями или разработчиками. Кроме того, там представлена поддержка документирования измерения как модуля оценки;

- номера ИСО/МЭК 25050 — ИСО/МЭК 25099 зарезервированы для расширения серии стандартов SQuaRE и/или технических отчетов, которые касаются конкретных областей применения и/или могут быть использованы в дополнение к одному или нескольким стандартам SQuaRE.

Приложение В
(справочное)**Сравнение риск-ориентированного подхода с качеством-ориентированным подходом
и с моделями качества****В.1 Общие положения**

В риск-ориентированном и качеством-ориентированном подходах используются две взаимодополняющие методологии, в которых определяются требования к системам, к процессу разработки систем в соответствии с этими требованиями и оценка функциональности систем. В большинстве случаев для охвата всех аспектов определенного поведения системы необходим комбинированный подход. Применение и использование по назначению каждой методологии зависит от многих факторов, включая помимо прочего, зрелость используемых технологий, роль системы в цепочке создания стоимости и область ее применения. Эти факторы учитываются при определении требований к товару или нормативных требований или структуры для разработки и оценки определенной системы с целью достижения практических результатов и надлежащего управления ожиданиями пользователей или потребителей.

Использование модели качества наиболее эффективно в ситуациях, когда для всех характеристик, так же, как и для многих подхарактеристик, устанавливаются количественные или четко определяемые показатели системы. Для зрелых технологий в определенной области применения значения показателей могут быть определены как часть системных требований.

Частичное использование модели качества может оставаться эффективным в случае оценки новых технологий или приложений, в ситуациях с высокой степенью риска негативных последствий. В таких случаях определяются и оцениваются только значения показателей подхарактеристик, отражающие высокую степень риска негативных последствий.

Для новых или развивающихся технологий показатели качества программного обеспечения часто являются предметом активных исследований и сотрудничества научных кругов и индустрии.

Риск-ориентированный подход больше подходит в ситуациях, когда количественные показатели для многих характеристик или важных подхарактеристик не установлены. К тому же риск-ориентированный подход, по определению, преимущественно подходит для реагирования на непредвиденные события.

В.2 Взаимосвязь с другими стандартами

В [8] определена структура и процесс менеджмента рисками, которые применимы к любому типу организации и деятельности, включая принятие решений на всех уровнях. [9] разработан на основе [8]. Он адаптирует структуру и процесс менеджмента рисками для организаций и проектов, связанных с ИИ.

ГОСТ Р ИСО/МЭК 25010 описывает две модели качества. В модели качества продукта представлены характеристики и подхарактеристики программного обеспечения или компьютерной системы. В модель качества при использовании вошли характеристики и подхарактеристики, определяющие взаимодействие системы с окружающей средой. Это можно рассматривать как двухуровневый подход, при котором критерии качества продукта обеспечивают основу для многих критериев, существующих при использовании продукта.

В модели качества при использовании одной из качественных характеристик является «свобода от риска». Ее подхарактеристики называются «смягчение отрицательных последствий риска», они отражают основные серьезные озабоченности общества в области экономики, здравоохранения, безопасности, окружающей среды и этики. В риск-ориентированном подходе, для сравнения, основные показатели, связанные с обществом, экономикой, здравоохранением, безопасностью, окружающей средой и этикой называются «целями». К тому же, в данном подходе цели не определяются общественным мнением. Они также определяются разработчиками, например, цели, связанные с защищенностью системы, надежностью и прозрачностью. Впоследствии для каждой установленной цели определяются и смягчаются потенциальные риски путем применения выбранных мер по обработке рисков или «элементов управления». Исходя из этого, каждое соответствующее свойство (социальное или системное и т. д.) может быть выражено в виде характеристики или цели.

В.3 Сравнение подходов

Основная трудность в определении или оценке систем с использованием только качеством-ориентированного подхода, заключается в том, что подхарактеристики качества недостаточно определены и не соответствуют принятым качественным и количественным показателям.

Рассмотрим «непредвзятость» как подхарактеристику модели качества ИИ. Быть «непредвзятым» может совсем по-разному толковаться в различных прикладных областях или системах для разных целей (см. ПНСТ 839—2023). Не существует ни одного показателя непредвзятости для измерения «непредвзятости» системы. Более того, доступность и зрелость показателей непредвзятости зависит от типа системы и используемой

технологии. Например, показатели непредвзятости систем, основанных на моделях МО для классификации, более понятны по сравнению с показателями непредвзятости систем на основе обучения с подкреплением.

Риск-ориентированный подход позволяет определять и оценивать цели, включая те, для которых нет прямых показателей измерения. В риск-ориентированном подходе, потенциальные негативные последствия, называемые иногда реперкуссиями или результатами, идентифицируются, количественно оцениваются (где это возможно) или квалифицируются. В рамках обработки риска исследуются приводящие к негативным последствиям источники риска и выбираются меры обработки риска, направленные на уменьшение (или устранение) источника риска. Итеративный процесс выбора и реализации мер по обработке риска и оценки остаточного риска продолжается до тех пор, пока риск не будет снижен до приемлемого уровня (как определено в требованиях к системе).

По сравнению с двумя упомянутыми моделями качества, управление риском представляет собой многоуровневый нисходящий подход. Это означает, что меры по смягчению отрицательных последствий риска (или «элементы управления») на верхнем уровне, выбранные в процессе управления риском, становятся целями управления риском на нижнем уровне. В свою очередь эти новые цели соотносятся с набором разных подхарактеристик.

Предположим, что «непредвзятость» была заявлена как одно из требований к системе.

В результате оценки рисков были выявлены следующие потенциальные источники рисков:

- отсутствие опыта в прикладной области;
- смещенные данные для обучения.

Выбранные меры обработки риска (или «элементы управления») для смягчения отрицательных последствий включали:

- анализ функциональных характеристик экспертами в прикладной области;
- уменьшение степени смещенности в выборке данных для обучения, используя соответствующие методы контроля качества данных.

Следующим уровнем целей является:

- выполнение рекомендаций по итогам анализа экспертов;
- поддержание уровня смещенности данных на приемлемом уровне, как предписано, с использованием выбранных статистических показателей качества данных.

Риск-ориентированный подход может помочь при разработке новых показателей качества программного обеспечения. Его можно использовать для оценки поведения или ценности нового показателя при различных условиях и уровнях риска как функцию существующих «косвенных» показателей.

Предположим, что в процессе уточнения требований к системе классификации ИИ, используемой для найма людей, демографический паритет и равенство возможностей (см. ПНСТ 839—2023), были определены как подходящие показатели для оценки непредвзятости системы. Однако, в настоящее время не накоплен достаточный опыт для определения значения требуемого показателя. В такой ситуации риск-ориентированный подход позволит внедрить систему ИИ с определением соответствующих показателей обработки риска путем смягчения отрицательных последствий риска и оценки их результатов.

Впоследствии опыт эксплуатации системы и новые полученные данные помогут найти взаимосвязь между уровнями риска и показателями производительности системы. Эти знания можно будет использовать для спецификации аналогичных систем в будущем.

В заключение следует отметить, что использование риск-ориентированного подхода повышает гибкость при определении требований. Характеристики и подхарактеристики системы в качество-ориентированном подходе, соответствуют высокоуровневым целям в риск-ориентированном подходе. Каждое свойство системы может быть выражено как в виде характеристики, так и в виде цели. Риск-ориентированный подход позволяет определять, внедрять и проводить верификацию поведения системы для достижения целей, в том числе для тех, для которых показатели качества программного обеспечения отсутствуют или их значения не определены. Кроме того, риск-ориентированный подход может помочь в определении подходящих значений для новых показателей.

Приложение С
(справочное)

Производительность

Термин производительность используется по-разному в области программного обеспечения и системной инженерии, по сравнению с областью ИИ.

В области программного обеспечения и системной инженерии под производительностью понимается то, насколько быстро определенная часть программного обеспечения выполняет свою функцию и насколько она эффективна. Данные свойства входят в существующие характеристики SQuaRE уровня производительности.

В области ИИ производительность означает то, как хорошо определенная система ИИ выполняет поставленную задачу. Производительность поставленной задачи системы ИИ можно измерить, определив подходящую оценку показателей, релевантных к данному типу системы. Например, в ПНСТ 835—2023 определяются методы для оценки классификации производительности МО.

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 ИСО/МЭК 25010—2015	IDT	ISO/IEC 25010:2011 «Системная и программная инженерия. Требования и оценка качества систем и программного обеспечения (SQuaRE). Модели качества систем и программного обеспечения»
ГОСТ Р 70462.1—2022/ ISO/IEC TR 24029-1:2021	IDT	ISO/IEC TR 24029-1:2021 «Искусственный интеллект (ИИ). Оценка робастности нейронных сетей. Часть 1. Обзор»
ПНСТ 835—2023 (ISO/IEC TS 4213:2022)	MOD	ISO/IEC TS 4213:2022 «Информационные технологии. Искусственный интеллект. Оценка эффективности классификации машинного обучения»
ПНСТ 836—2023 (ISO/IEC TR 5469)	MOD	ISO/IEC TR 5469 «Искусственный интеллект. Функциональная безопасность и системы искусственного интеллекта»
ПНСТ 838—2023 (ИСО/МЭК 23053:2022)	MOD	ISO/IEC 23053:2022 «Платформа разработки систем искусственного интеллекта (AI) с использованием машинного обучения (ML)»
ПНСТ 839—2023 (ISO/IEC TR 24027:2021)	MOD	ISO/IEC TR 24027:2021 «Информационные технологии. Искусственный интеллект (ИИ). Смещенность в системах ИИ и при принятии решений с помощью ИИ»
ПНСТ 840—2023 (ISO/IEC TR 24368:2022)	MOD	ISO/IEC 24368:2022 «Информационная технология. Искусственный интеллект. Обзор этических и социальных проблем»
Примечание — В настоящей таблице использованы следующие условные обозначения степени соответствия стандартов: - IDT — идентичные стандарты; - MOD — модифицированные стандарты.		

Библиография

- [1] ISO/IEC TR 24028:2020 Информационная технология. Искусственный интеллект. Обзор достоверности систем искусственного интеллекта (Information technology — Artificial intelligence — Overview of trustworthiness in artificial intelligence)
- [2] ISO/IEC TR 29119-11:2020 Системная и программная инженерия. Тестирование программного обеспечения. Часть 11. Руководящие указания по тестированию систем искусственного интеллекта (AI) (Software and systems engineering — Software testing — Part 11: Guidelines on the testing of AI-based systems)
- [3] ИСО/МЭК 24029-2 Искусственный интеллект (ИИ). Оценка робастности нейронных сетей. Часть 2. Методология использования формальных методов (Artificial intelligence (AI) — Assessment of the robustness of neural networks — Part 2: Methodology for the use of formal methods)
- [4] ИСО/МЭК 25012:2008 Программная инженерия. Требования и оценка качества программного продукта (SQuaRE). Модель качества данных (Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Data quality model)
- [5] ИСО/МЭК 5259 (все части) Искусственный интеллект. Качество данных для аналитики и машинного обучения [Artificial intelligence — Data quality for analytics and machine learning (ML)]
- [6] ИСО/МЭК 22989:2022 Информационная технология. Искусственный интеллект. Концепции и терминология искусственного интеллекта (Information technology — Artificial intelligence — Artificial intelligence concepts and terminology)
- [7] Руководство ИСО/МЭК 51:2014 Аспекты безопасности. Рекомендации по их включению в стандарты (Safety aspects — Guidelines for their inclusion in standards)
- [8] ИСО 31000 Менеджмент риска. Принципы и руководство (Risk management — Guidelines)
- [9] ИСО/МЭК 23894 Информационная технология. Искусственный интеллект. Руководство по менеджменту риска (Information technology — Artificial intelligence — Guidance on risk management)

УДК 004.01:006.354

ОКС 35.080

Ключевые слова: программное обеспечение, качество систем и программного обеспечения, модель качества, искусственный интеллект

Редактор *Е.Ю. Митрофанова*
Технический редактор *И.Е. Черепкова*
Корректор *Р.А. Ментова*
Компьютерная верстка *А.Н. Золотаревой*

Сдано в набор 08.11.2023. Подписано в печать 28.11.2023. Формат 60×84%. Гарнитура Ариал.
Усл. печ. л. 2,32. Уч.-изд. л. 1,90.

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.
www.gostinfo.ru info@gostinfo.ru