DOI 10.37882/2223-2966.2024.7.13

ПОНЯТИЙНЫЙ БАЗИС И СПЕЦИФИКА ИСПОЛЬЗОВАНИЯ СИСТЕМЫ ИИ В КОНТУРЕ УПРАВЛЕНИЯ ИС

CONCEPTUAL BASIS AND SPECIFICS OF USING AN AI SYSTEM IN AN IS CONTROL LOOP

B. Goryachkin
K. Grishin

Summary. *Problem Statement*. The development in the field of information technology has covered a large number of tools, fulfilling another round in the formation of artificial intelligence (hereinafter AI). At the moment, its use is not limited to any one area, including such as the intelligent control loop of information systems. Having influenced most areas, the very concept of AI and its technical embodiment in the context of the intelligent control loop is pushing towards deeper and more fundamental concepts, the provisions for its use. What we «put» into an AI system (starting from the process of creation, realization and further operation) will determine the output itself. In addition to the technical component, accurately formulated and described issues of trust and ethics play a key role in this. However, with all the advantages of such intelligence (rapid performance of routine, difficult-to-predict tasks, creation of «hybrid» interaction with humans, etc.), it contains some pitfalls (possible threats coming from the artificial intelligence, error, inefficiency in solving some tasks) that require detailed consideration.

Purpose. To consider normative and technical standards related to AI regulation to consider the issues of trust, friendliness, objectivity of AI, determination of essential characteristics of such a system, possible threats.

Results. Regulatory and technical standards related to the regulation of AI are analyzed to consider the issues of trust, friendliness, objectivity, determine the essential characteristics, threats. It is concluded that AI should have the characteristics of friendliness, trust in the structure of the information system control loop for widespread, convenient and efficient use.

Practical significance. The presented information in the future can be used as a source for creating a full-fledged intelligent control loop (hereinafter referred to as ICL), standards, unifying regulations governing its operation.

Keywords: intelligent control loop, artificial intelligence, essential characteristics, classification of artificial intelligence systems, features, methods and process of trust assurance, system quality.

Горячкин Борис Сергеевич
кандидат технических наук, доцент,
Московский государственный технический университет им. Н.Э. Баумана
bsgor@mail.ru
Гришин Кирилл Павлович
аспирант, Московский государственный технический университет им. Н.Э. Баумана
kirillgrish2014@yandex.ru

Аннотация. *Постановка проблемы*. Развитие в сфере информационных технологий охватило большое количество инструментов, выполнив очередной виток в становлении искусственного интеллекта (далее ИИ). На данный момент, его использование не ограничивается какой-либо одной областью, включая такую как интеллектуальный контур управления информационных систем. Подвергнув влиянию большинство сфер, сама концепция ИИ и его техническое воплощение в контексте интеллектуального контура управления (далее ИКУ) подталкивают к более глубинным и фундаментальным понятиям, положениям его использования. От того, что мы «вкладываем» в систему ИИ (начиная от процесса создания, реализации и дальнейшей эксплуатации), будут зависеть и сами выходные данные. Помимо технической составляющей, ключевую роль в этом играют точно сформулированные и описанные вопросы доверия, этики. В свою очередь преимущества при использовании подобного интеллекта (быстрота выполнения рутинных, сложнопрогнозируемых задач, создание «гибридного» взаимодействия с человеком и т.д.) могут содержать «подводные камни» (возможные угрозы, исходящие со стороны искусственного интеллекта, погрешность, неэффективность в решении некоторых задач), требующие детального рассмотрения.

Цель. Рассмотреть нормативно-технические стандарты, связанные с регулированием ИИ, для рассмотрения вопросов доверии, дружелюбности, объективности ИИ, определение существенных характеристик подобной системы, возможных угроз.

Результаты. Проанализированы нормативно-технические стандарты, связанные с регулированием ИИ, для рассмотрены вопросы доверия, дружелюбности, объективности, определены существенные характеристики, угрозы. Сделан вывод о необходимости наличия у ИИ характеристик дружественности, доверия в структуре контура управления информационной системы для повсеместного, удобного и результативного использования.

Практическая значимость. Представленную информацию в перспективе возможно использовать в качестве источника для создания полноценного ИКУ, стандартов, унифицирующих положений, регулирующих его работу.

Ключевые слова: интеллектуальный контур управления, искусственный интеллект, существенные характеристики, классификация систем искусственного интеллекта, особенности, способы и процесс обеспечения доверия, качество системы.

Введение

В наше время системы ИИ становятся неотъемлемой частью нашей повседневной жизни, играя важнейшую роль в научной, производственной, сфере и др. Они проникают в самые различные области, включая управление информационными системами, внося тем самым значительный вклад в их развитие.

В том числе, при использовании ИИ возможно создание контура управления закрытого типа [6], однако, изучение особенностей применения и интеграции ИИ в контур управления представляет собой важную и в тоже время нетривиальную задачу. Особенно остро стоит вопрос о доверии, этики ИИ, интегрированного в контур управления. Ответ на них в перспективе поможет повысить эффективность функционирования предприятий, организаций, улучшить взаимодействие с человеком, а также качество принимаемых им решений.

Особенности интеллектуальной составляющей в контексте контура управления

Под ИКУ следует понимать трехкомпонентную (дополненную) эргатическую систему («человек — машина — ИИ»), в которой ИИ способен выступать в качестве партнера, «цифрового помощника» [7] при выполнении тех или иных операций, задач, брать на себя отдельно взятые функции под свой контроль и т.д. (рис 1):

Рис. 1. Интеллектуальный контур управления типа «Человек — машина — ИИ»

Так, например, главными теоретическими вопросами при создании ИКУ могут быть выделены следующие:

1. Кто занимает верховенство в принятии решений?
2. Кто из элементов первичен/вторичен при разработке и создании и дальнейшей эксплуатации ИКУ? (Человек-оператор или ИИ)?
3. Каким образом будет происходить взаимодействие между элементами?
4. Как «доверить» интеллектуальной части ИКУ выполнение действий? Сохранить этический баланс и др.;
5. Рационально ли использование ИИ в контуре управления? Если да, то каковы преимущества/недостатки?

Частично, определение ИКУ даёт ответ на поставленные выше вопросы, но так или иначе, большинство из них относятся к ИИ, что неудивительно, ведь используемая технология, особенно в контексте взаимодействия с человеком требует открытости, решенных вопросов доверия. Сложность в отслеживании логики (как например с «черным ящиком» ИИ) работы, вынуждает ориентироваться на «прозрачные» способы обучения нейронных сетей, позволяя наблюдать за процессом обучения ИИ и направлять ход обучения в зависимости от полученных результатов [1]. Это в свою очередь относится к вопросам «объяснимости», «понятности» и «предсказуемости» ИИ в ходе его обучения и выполнения поставленных задач.

Под доверием понимается состояние уверенности субъектов, использующих данную систему (пользовате-

Таблица. 1.
Стадии и этапы жизненного цикла системы ИИ

№	Этапы ЖЦ системы ИИ	Пример адаптации действий на каждом этапе
1	Формулирование конечного видения системы ИИ	Обозначение критериев, причин, указывающие на необходимость создания ИИ; Проведение исследований в данной области; Создание некоторого количества вариантов концепции ИИ, соответствующий нуждам пользователя.
2	Разработка	Формирование и утверждение требований к ИИ в контуре управления; Формулирование ТЗ проекта и документации для ИИ; Разработка, проектирование рабочей документации, ПО или их адаптация под ИИ.
3	Верификация и валидация	Выполнение подготовки материальной (подключение оборудования) и программной составляющей (проведение тестирования работы ИИ в контексте контура управления) и др.
4	Эксплуатация и сопровождение	Выполнение типовых задач при помощи ИИ, проведение дополнительно обучения ИИ, оптимизация ПО и др.
5	Вывод из эксплуатации	Прекращение работы системы ИКУ

ли, организации, регулирующие положения о создании и применении подобных систем и др.) в выполнении с определенным уровнем качества тех или иных задач, функций, которые на нее возложены [1]. Также здесь определены существенные характеристики, адаптированные в табл. 4 с иными характеристиками «ГОСТ Р 59898–2021 Оценка качества систем искусственного интеллекта» [2]; стадии и этапы жизненного цикла (далее ЖЦ) системы ИИ (табл. 1), факторы снижения ее качества на этапах создания, эксплуатации и способы обеспечение доверия к системам ИИ на соответствующих стадиях жизненного цикла (табл. 2). Табл.1 возможно интерпретировать с точки зрения её адаптации к контуру управления и ИИ:

Помимо этого, данный стандарт формулирует процесс обеспечения доверия и оценки качества систем ИИ (рис. 2).

Для того, чтобы осуществить проверку доверия, необходимо сопоставить выбранный набор характеристик и требования, предъявляемые к данной системе ИИ. Подобные действия могут быть выполнены определённым кругом лиц:

— Разработчиком, как в контексте открытых (проверка на соответствие любой из заинтересованных сторон), так и закрытых, внутренних требований (выполняются исключительно самим разработчиком);
— Потребителем, в ходе тестирования системы;
— Организацией, выполняющая функции регулирования создания и применения систем ИИ в соответствии с принятыми регулятивными нормами (например, национальными стандартами). Подобная проверка обязательна в случаях, если неправильная работа системы способна повлечь за собой возникновение угрозы техногенного характера, чрезвычайной ситуации и др.

Таблица. 2.

Факторы снижения качества и способы обеспечения доверия на стадиях создания, эксплуатации систем ИИ

Стадия ЖЦ		Факторы, негативно сказывающееся на качестве системы ИИ	Способы по обеспечению доверия
Создание системы ИИ	Концепция	Неполный/неточный выбор необходимых характеристик системы ИИ (безопасность, надежность, функциональные характеристики и др.)	Определение и использование корректных правил для набора характеристик системы и их дальнейшего выбора
	Разработка	Отсутствие представительности в данных, предназначенные для обучения системы ИИ	Выполнение действий, по созданию представительной обучающей выборки
		Наличие смещенности в обучающей выборке и как следствие «необъективность» выходных данных	Проведение статистического анализа над набором исходных данных с дальнейшей оценкой их представительности Выполнение метода кросс-валидации выборки, полученной при ручной разметке данных. Сужение области применения системы ИИ путем внесения ограничений и др.
	Производство	Посредственная надежность системы Высокая стоимость владения	Рациональное использование ресурсов ИТ-инфраструктуры
		Несоответствие системы необходимому уровню понятности, объяснимости, предсказуемости, защищенности информации о модели данных и др.	Использование улучшенных/оптимальных алгоритмов обработки данных Осуществление процедур по защите сведений о модели данных
Эксплуатация системы ИИ	Использование	Нецелевое использование системы ИИ Недостаточная представительность выборки, периодичность самого тестирования системы ИИ. Невозможность проведения автоматического тестирования после каждого этапа обучения системы ИИ Отсутствие должной защищенности с точки зрения информационной безопасности (персональные данные, особенности функционирования системы, используемая модель данных и др.).	Учитывание факторов, среды в которой выполняется взаимодействие системы ИИ и пользователя Установление и подтверждение функциональных требований и дальнейший отбор тестовой выборки Проведение мероприятий, способствующие повышению защищенности системы в ходе эксплуатации
	Поддержка	Модель данных теряет свою актуальность	Выполнение ряда действий по актуализации модели данных
	Прекращение использования	Вывод системы из эксплуатации влечет нарушение приватности информации, данных	Принятие и соблюдение должных мер по защите от «утечек» данных.

Рис. 2. Процессы обеспечения доверия и оценки качества (подтверждения соответствия требованиям)

В свою очередь процесс подтверждения доверия к системе ИИ включает:

1. Выбор исчерпывающего набора характеристик, заинтересованными сторонами (Разработчик, потребитель, организация);
2. Формулирование и регламентирование требований, предъявляемых к характеристикам, выбранным в п.1;
3. Выполнение организационных процессов по подтверждению характеристик из п.1 требованиям, обозначенных п.2;
4. Фактическое доведение до требуемого уровня путем реализации мероприятий по обеспечению соответствия набора критически необходимых характеристик системы ИИ установленным требованиям (устранение причин, приводящих к снижению качества).

В свою очередь реализация способов в табл. 2 может протекать на любой стадии ЖЦ, при инициативе любой из ранее упомянутых сторон. Охват большей части способов доверия с последующей реализацией способно повысить качество работы всей системы в целом, поскольку каждый из них содержит те или иные преимущества и недостатки на разных стадиях (таблица 3):

Таблица 3.

Особенности способов обеспечения доверия на разных стадиях жизненного цикла систем ИИ

Тип характеристик	Пример преимуществ	Примеры недостатков
Способы обеспечения доверия на стадии создания системы ИИ	Предварительное «обогащение» системы способствует снабжению необходимыми свойствами (на этапе проектирования и разработки) Высокая степень воспроизводимости свойств разрабатываемой системы	Сложность в получении полной «прозрачности» процесса разработки системы (относится прежде всего к открытым требованиям, проверяемые заинтересованной стороной). Отсутствие информативности для потребителя системы (в отличии от разработчика)
Способы обеспечения доверия на стадии эксплуатации системы ИИ	Возможность более понятной трактовки в контексте потребителя	Тестовые изменения характеристик системы не всегда могут быть точно прогнозируемы для использования в реальных условиях

Система ИИ предполагает наличие составляющих элементов, образующих многообразную структуру или архитектуру. Она поделена на конкретные уровни:

1. Физический (или сенсорный уровень, где выполняется взаимодействие со средой, в которую интегрирована система);
2. Инфраструктурный (Достаточно необходимая совокупность ресурсов, средств, способная хранить, обрабатывать и передавать данные);
3. Прикладной (т.е заключительный «абстрактный» уровень, где выполняется построение интеллектуального алгоритма для обрабатывания получаемых данных).

Каждый из представленных архитектурных уровней может содержать собственный перечень существенных характеристики, а также «основания» для доверия ко всей совокупности уровней.

Так, например, на «первом» уровне совокупность требований, связанных с надежностью, безопасностью и функциональностью является приоритетной в связи с выполнением задач по физическим измерениям, проведением тестов и т.д. Показательным примером в данном случае может быть успешное выполнение технического контроля, результаты которого возможно интерпретировать как доверительные.

Реализация доверия к следующему уровню заключается в выполнении требований безопасности (выполнение мероприятий по сохранению конфиденциальности данных, подлежащих обработке, принятие и дальнейшая поддержка целостности и доступности системы в целом).

И наконец, заключительный уровень требует «подтверждения надежности и безопасности ПО» [1] как и у «обычных» информационных систем. Однако особенность отчетливо прослеживается, в ситуациях, где задействовано машинное обучение в системе ИИ. Здесь, отсутствие предвзятости или «необоснованного смещения формулируемых оценок» является одним из способов обеспечения доверия [1, с. 11].

Важно отметить, что доверие к системе ИИ будет достигнуто лишь в том случае, если все характеристики на каждом из уровней соответствуют поставленным требованиям. Иными словами, в зависимости от целей, возможна гибкая «настройка» требований как для каждого отдельного уровня, так и для всей системы ИИ целиком с учетом подбора существенных параметров (характеристик), в конечном счете способствуя повышению качества системы.

Логичным продолжением предыдущего стандарта является «ГОСТ Р 59898—2021 Оценка качества систем искусственного интеллекта. Общие положения» [2], раскрывающую часть структуры рис.1 в области оценки качества.

Как и любой другой технический или программный продукт представляет собой сложноорганизованную структуру, система ИИ в этом плане не является исключением. По этой причине, на каждом этапе жизненного цикла (см. табл. 1) необходимо удостовериться в том, что приемлемый уровень качества был достигнут. Подобная процедура позволяет выполнить несколько поставленных задач:

— Выявить состояние параметров системы (надежность, производительность, востребованности и др.) с дальнейшим их улучшением;
— Выполнение функции проверки выходных данных системы на предмет их приемлемости и обоснованности;

Таблица 4.

Существенные характеристики и субхарактеристики системы ИИ для модели качества продукта

Категория	Характеристики из ГОСТ Р 59276	Существенная характеристика	Специализированная характеристика
Функциональная группа	Функциональные возможности	Функциональные возможности	Полезность функционала; Функциональная точность; Соответствие требованиям; Функциональная полнота; Потенциал в самообучении.
		Способность к взаимодействию	Соответствие; Функциональная совместимость; Управляемость.
	Эффективность	Производительный уровень	Поведение в течение времени; Ресурсопотребление; Производительные возможности
	Мобильность	Мобильность	Адаптируемость; Простота внедрения; взаимозаменяемость
	Практичность	Практичность	Объяснимость; Обучаемость; Простота в использовании; Защита от ошибок пользователя; Эстетика пользовательского интерфейса; Доступность; Возможность взаимодействия (сотрудничества)
	Сопровождаемость	Сопровождаемость	Анализируемость; Изменяемость; Стабильность; Тестируемость; Модульность; Настраиваемость
Надежность	Надежность	Надежность	Стабильность; Отказоустойчивость; Восстанавливаемость; Невосприимчивость к изменениям (робастность)
Безопасность	–	Защищенность	Конфиденциальность; Целостность; Безотказность; Подотчетность; Подлинность; Приватность

— Проверка при заданном уровне параметров (точности, надежности, достоверности) выходных данных, позволяющая убедиться в том, что система ИИ достигает поставленных целей;
— Отождествление характеристик системы и требований, указанные в нормативных актах, технической документации.

Кроме этого, данный стандарт своим содержанием раскрывает «фундаментальные» понятия (ИИ, качество, его оценка и т.д.), приводит методологическую составляющую оценки качества на стадиях ЖЦ, её модель, показатели, а также классификацию тех характеристик, которые являются существенными. В таблице 4 представлен их подробный перечень, соотнося с уже существующими в ГОСТ Р 59276-2020:

Данный стандарт, как и предыдущий невозможно использовать для «сильного» и «общего» ИИ, однако его реализация возможна на всех системах, которые используют те или иные методы ИИ (в т.ч. алгоритмы на основе машинного, дедуктивного обучения) в целях решения практических поставленных задач.

Вступивший в силу с начала года национальный стандарт 838-2023/ИСО/МЭК 23053:2022 «Искусственный интеллект. Структура описания систем искусственного интеллекта, использующих машинное обучение» систематизирует и описывает структуры систем ИИ, которые используют машинное обучение, т.е. «вычислительные методы, дающие системе возможность обучаться на основе входных данных и опыта» [3]. Данный стандарт возможно применить в любом из типов организаций (например, некоммерческие, государственные организации), ставящей своей целью внедрить и в дальнейшем использовать системы ИИ.

В данном стандарте представлена система машинного обучения, определяющая конкретную роль и функции, специфичные для машинного обучения. Не исчерпывающий перечень элементов подобной системы представлен на рис. 3.

Также обозначены алгоритмы машинного обучения. Каждый из них прежде всего определяет модели машинного обучения и дальнейший подход к ее обучению. Структура модели и ее подход к обучению также будет зависеть от т.н. гиперпараметров — характеристики алгоритма машинного обучения, которые влияют на процесс обучения. Так или иначе, данные алгоритмы возможно использовать как в контексте предоставления информации (стадия подготовки данных) для выделения признаков, так и для полноценного создания модели машинного обучения (рис. 4,5).

Представлены различные подходы к машинному обучению, каждый из которых имеет структуру, способы взаимодействия с информацией и специфику применения при решении конкретных задач. Некоторый перечень обозначенных подходов представлен на рис. 6:

Кроме этого, процессы машинного обучения представляется возможным сопоставить с ЖЦ ИИ (рис. 7).

Рис. 3. Элементы системы машинного обучения

Рис. 4. Различные категорий алгоритмов машинного обучения

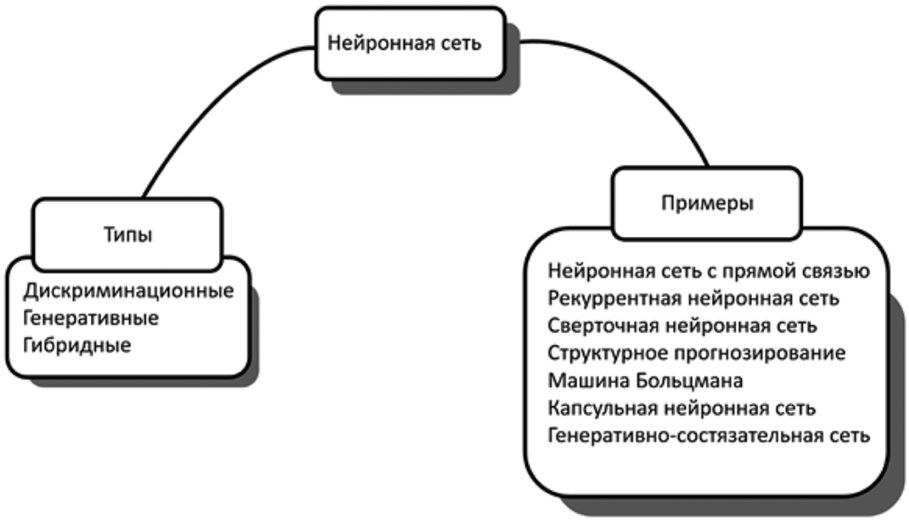

Рис. 5. Примеры различных категорий и типов алгоритмов машинного обучения (нейронных сетей)

На определенных её этапах возможно выполнение некоторой цепочки процессов (т.н. конвейер), которые важны для определения задачи и дальнейшего выбора необходимого набора данных, предназначенного для обучения модели машинного обучения. В некотором смысле, сравнение модели и этапов позволяет выявить проблемные моменты в процессе разработки ИИ для предложения способов их решения. Также, описание стандарта содержит следующую формулировку: «Данные процессы не зависят друг от друга и могут быть осуществлены в любом порядке или одновременно … в зависимости от конкретного варианта использования могут существовать дополнительные технические ограничения, которые устанавливают определенный порядок процессов или создают необходимость их повтора», что говорит о вариативности и гибкости настройки модели машинного обучения с учетом возникновения тех или иных ограничений.

Таким образом, стандарт аккумулирует информацию о машинном обучении, представляя её как неотъемлемую часть ИИ, где оптимальный выбор алгоритма, инструментов, техник, а также дальнейшее построение модели машинного обучения позволяют создать сбалансированную структуру, способную решать круг поставленных перед ней задач.

Также следует упомянуть немаловажный стандарт ГОСТ Р 59277-2020 «Системы искусственного интеллекта. Классификация систем искусственного интеллекта» [4]. При помощи него установлены принципы классификации систем ИИ для сравнения подобных систем по некоторым видам параметров (деятельность, структура знаний, функции контура управления, безопасность, степень автоматизации и др.), что позволит повысить эффективность использования данных систем ИИ для решения прикладных задач (в т.ч. и во взаимодействии с человеком-оператором).

В данном стандарте и схеме классификации отображены лишь некоторые классы. К «базовым» классам возможно группирование по принципам:

Рис. 6. Подходы машинного обучения

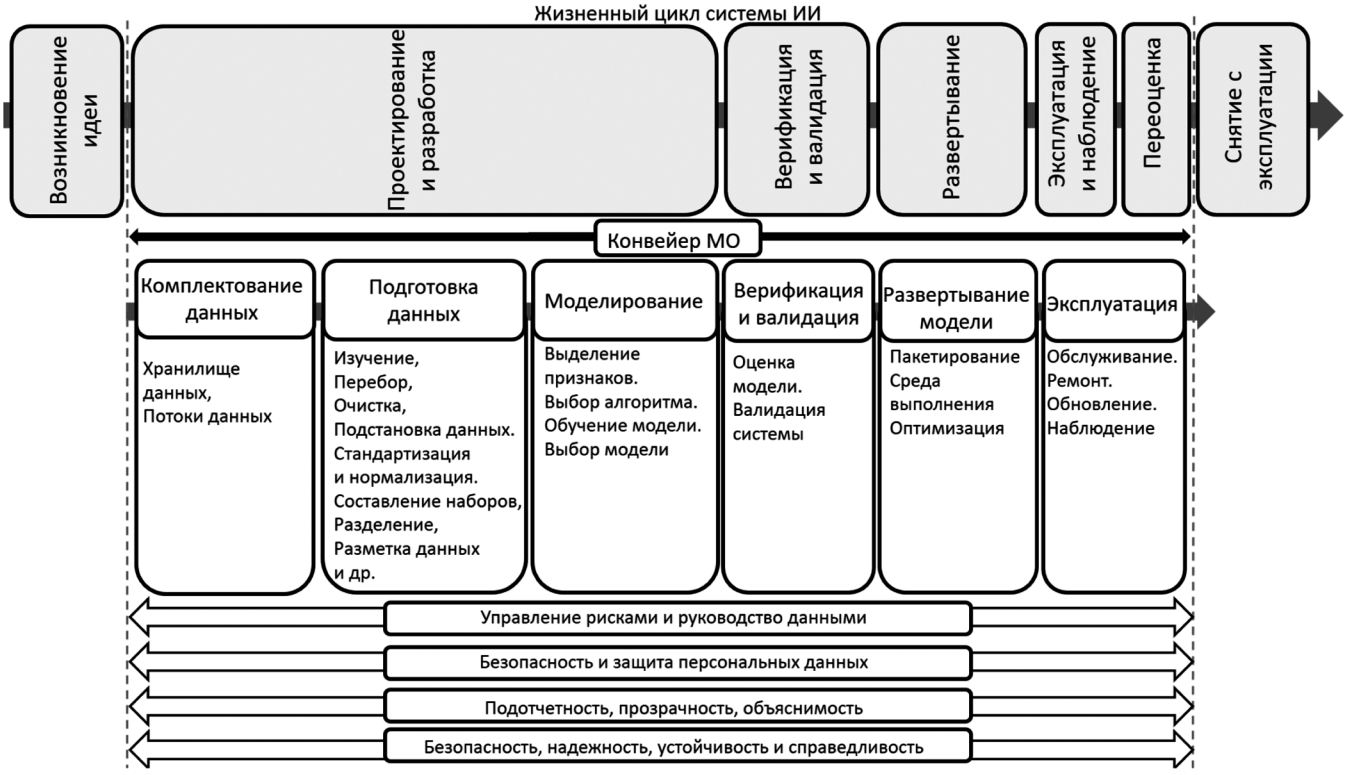

Рис. 7. Конвейер машинного обучения и его сопоставление с жизненным циклом системы ИИ

1. Самих классов и категорий объектов в управлении;
2. Технологий построения, получения и реализации знаний;
3. Функционального назначения в контексте выполнения в контуре управления;
4. Использования методов и технологий в самой системе ИИ;
5. Задействования методов и средств в системе ИИ с иными системами и с человеком-оператором.

Кроме этого, возможно создание дополнительных классификаций в силу той или иной специфики, способной содержать определенные требования к объектам, процессам (Интероперабельность, безопасность и др.). Гибкость и вариативность в классификации классов позволяет применять к некоторым системам ИИ несколько классов, дополнить классификацию по новым основаниям, подробнее раскрыть сами классы путем введения новых свойств или подклассов (внешнее наблюдение со

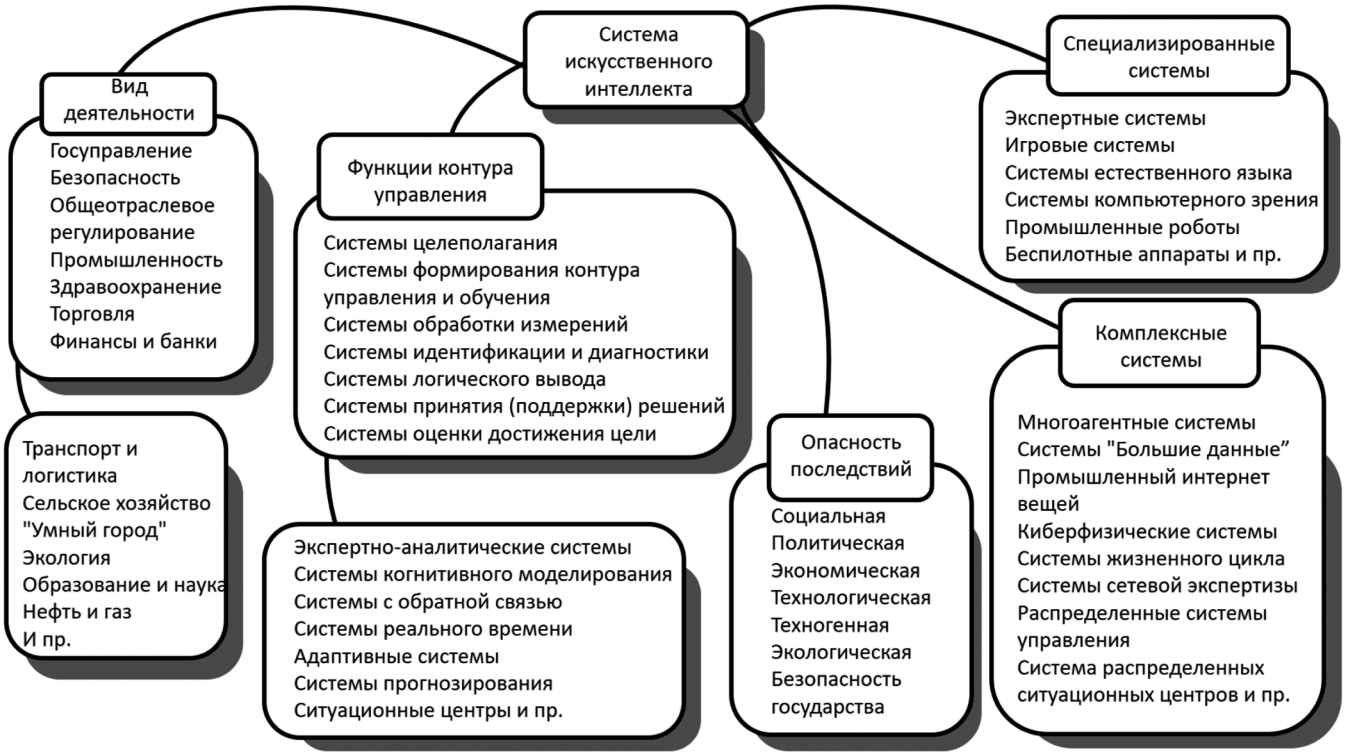

Рис. 8. Схема классификации систем ИИ

стороны автоматизированной системы/человека; степень понимания системы ИИ; степень надежности, безопасности; способность принятия управленческих решений/ выполнения планирования и др.).

На данный момент некоторые из представленных в классификации классов представляется возможным дополнить или заменить функциями ИИ. Для этого было выполнено цветовое обозначение некоторого перечня «прикладных» классов (сфер), где уже возможна интеграция ИИ. Данная информация представлена на рис. 8 и в табл. 5:

Тем не менее, ИИ, являясь результатом деятельности человека, способен вбирать в себя те или иные качества, присущие его создателю, в т.ч. такую как предвзятость. В контексте взаимодействия с системами ИИ, подобная характеристика может интерпретироваться как предвзятость подтверждения — «тип когнитивной предвзятости человека, который предпочитает прогнозы систем ИИ, подтверждающие ранее существовавшие убеждения или гипотезы», наряду с предвзятостью автоматизации — «Склонностью человека отдавать предпочтение предложениям автоматизированных систем принятия решений и игнорировать противоречивую информацию, полученную без применения автоматизации, даже если она верна» и др.

Подобная особенность может быть заложена на этапе разработки, проектировании или дальнейшей эксплуатации в самой системе, влияющая на «объективную», непредвзятую работу системы. С этой точки зрения, подобное свойство именуется как смещенность в системах ИИ — «систематическое различие в обработке определенных объектов», направленное на характеристику «входных данных и составных элементов систем ИИ» [5, с. 3].

Возникновение смещенности связано прежде всего с процессом обучения систем ИИ на основе реальных данных и выбора самой модели машинного обучения. В зависимости от этих и иных параметров возможно влияние смещенности на результат и как следствие, порождение как положительного (если в представленной профессии доминирующее большинство представители мужского пола, то система ИИ для компенсации смещенности (неравенства) будет отбирать больше представителей женского пола на подобные позиции), так и отрицательного эффекта (ИИ не распознает голосовые команды людей, с дефектами речи). При определенном регулировании, также возможно достижение баланса (нейтрального эффекта) в смещенности, однако в некоторых случаях (например, в типе обучения без учителя), без такового функционирование и дальнейшее обучение не представляется возможным [5, с. 4]. Воздействие на снижение или наоборот, устранение одного вида смещенности может повлечь за собой появление или увеличение другого.

Иными словами, для достижения оптимального результата работы системы ИИ необходима «тонкая» на-

Таблица 5.

Перечень «прикладных» классов (сфер), где возможна интеграция ИИ

Основания для классификации	Классы	Заменяемые функции (Пример возможности замены)
По видам деятельности	Государственное управление, Безопасность, Общеотраслевое регулирование, Промышленность, Здравоохранение, Торговля, Финансы и банки, Транспорт и логистика, Сельское хозяйство, “Умный город”, Экология, Образование и наука, Нефть и газ	Автоматизация процессов подачи и обработки заявлений; Системы видеонаблюдения с распознаванием лиц; Автоматизация процесса лицензирования; Роботизированные конвейерные линии; Анализ медицинских изображений для диагностики; Персонализированные рекомендации товаров; Оптимизация маршрутов доставки и т.п.
По функциям контура управления	Системы с обратной связью, Системы реального времени, Адаптивные системы, Системы формирования цели, Системы формирования контура управления и обучения, Системы обработки измерений, Системы идентификации и диагностики, Системы когнитивного моделирования, Системы логического вывода Системы принятия решений, Экспертно-аналитические системы, Системы оценки достижения цели, Ситуационные центры, Системы прогнозирования	Адаптивное управление климат-контролем в зданиях; Мониторинг и управление энергосетями; Адаптация интерфейсов в зависимости от нужд пользователя; Обучение роботов для выполнения новых задач; Оптимизация бизнес-процессов; Мониторинг и управление чрезвычайными ситуациями; Прогнозирование погоды и климатических изменений и т.п.
По специализации систем	Экспертные системы (управление знаниями), Игровые системы, Системы естественного языка, Системы компьютерного зрения, Промышленные роботы, Беспилотные аппараты,	Поддержка принятия решений в сложных областях; Автоматический перевод текстов и документов; Анализ и обработка изображений и видео; Автономное проведение разведывательных операций и т.п.
По комплексности и сложности систем	Многоагентные системы, Системы “Большие данные”, Промышленный интернет вещей, Киберфизические системы, Системы жизненного цикла, Системы сетевой экспертизы, Распределенные системы управления, Система распределенных ситуационных центров,	Координация действий роботов на склад; Анализ и обработка больших объемов данных для выявления трендов; Мониторинг и управление производственными процессами; Интеграция физических объектов с компьютерными системами для повышения эффективности и т.п.
По опасности последствий	Социальная, Политическая, Экономическая, Технологическая, Техногенная, Экологическая, Безопасность государства	Анализ социальных сетей для выявления общественных трендов; Прогнозирование экономических показателей и рыночных тенденций; Мониторинг и анализ экологической обстановки Кибербезопасность и защита информационной инфраструктуры и др.

Рис. 9. Виды нежелательной смещенности

стройка входных данных, подходящий выбор модели машинного обучения, разработка и применение методов для устранения «нежелательной» смещенности (выбор и «интеграция» моральных и культурных ценностей, установление критериев приемлемости, тестирование внутренней валидности и т.д.). На рис. 9 представлены виды нежелательной смещенности:

Данную особенность возможно представить в качестве угрозы, под которой следует понимать перечень обстоятельств и действий, способных привести нарушению безопасности системы. Угрозы, исходящие от ИИ, способны проявляться в различных сферах деятельности, где ему находится соответствующее применение (информационные технологии, юриспруденция, промышленность, здравоохранение и т.д.).

Примерами могут быть этические (предвзятость в данных, влекущая за собой дискриминацию отдельной группы людей, общества и т.д.), социальные (Замена ИИ рабочих мест; непредсказуемость работы ИИ, предвзятость его алгоритмов при решении определенных задач в производственной сфере), а также угрозы, связанные со сферой безопасности (Использование недостатков системы ИИ при выполнении атак на информационную систему; генерация контента (текст, изображения) с целью нанесения ущерба (дезинформация, DeepFake и т.д.); ненадежность или недостаточная защищенность системы ИИ, приводящая к компрометации конфиденциальных данных и т.д.).

Принимая во внимание тот факт, что ИКУ состоит из отдельных элементов, очевидно, что каждый из элементов подвержен нормативному регулированию. Поскольку ИКУ состоит из отдельных и в тоже время взаимосвязанных элементов, возможно выполнить объединение «регулятивной» и «практической» составляющей, именуемые «внешним» и «внутренним» контуром. Подобная иллюстрация представлена на рис. 10:

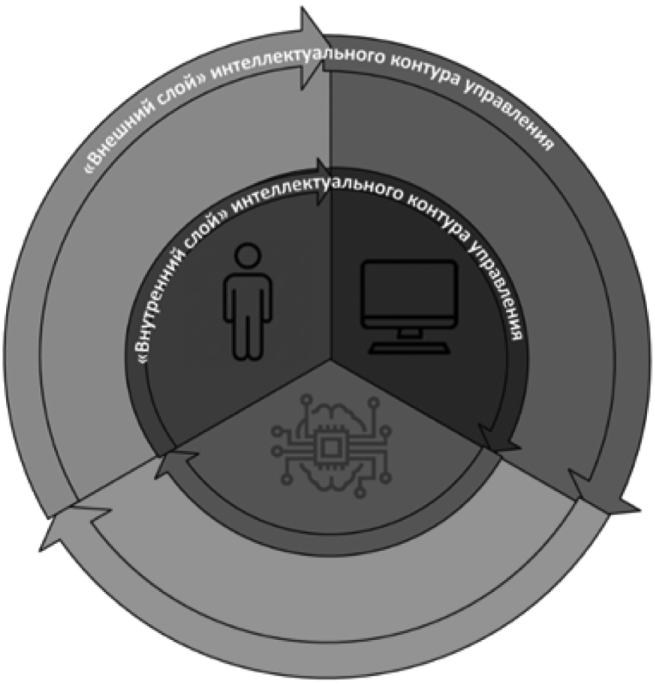

Рис. 10. Структура внешнего и внутреннего слоя интеллектуального контура управления

Внешний слой представлен в виде стандартов для интеллектуального контура управления, выполняющие

регулятивные, «эталонные» функции каждой отдельной области. В свою очередь внутренний, представляет собой цельную структуру, а именно сам ИКУ, используемый для решения тех или иных поставленных задач.

Заключение

Таким образом, исследование возможностей использования различных систем ИИ, замещающих компоненты системы «Человек-Машина-ИИ» показало, что интегрированный ИИ в структуру контура управления информационной системы для повсеместного, удобного и результативного использования должен быть дружественным, доверительным. Выше представленные особенности исходят из запроса прежде всего человека, в последующем приобретающие форму в тех или иных требованиях, стандартах. Следовательно, соответствие ИИ подобным «запросам» в рамках ИКУ делает его таковым.

ЛИТЕРАТУРА

1. ГОСТ Р 59276-2020. «Системы искусственного интеллекта. СПОСОБЫ ОБЕСПЕЧЕНИЯ ДОВЕРИЯ. Общие положения: Национальный стандарт Российской Федерации: дата введения 2021-01-03 / Федеральное агентство по техническому регулированию и метрологии. — Изд. Официальное. — Москва: Стандартинформ, 2021— 16 с.
2. ГОСТ Р 59898–2021. Оценка качества систем искусственного интеллекта. Общие положения: Национальный стандарт Российской Федерации: дата введения 2022-01-03 / Федеральное агентство по техническому регулированию и метрологии. — Изд. Официальное. — Москва: Российский институт стандартизации, 2021 — 24 с.
3. ПНСТ 838-2023/ИСО/МЭК 23053:2022. Искусственный интеллект. Структура описания систем искусственного интеллекта, использующих машинное обучение. Национальный стандарт Российской Федерации: дата введения 2024-01-01 / Федеральное агентство по техническому регулированию и метрологии. — Изд. Официальное. — Москва: Российский институт стандартизации, 2023 — 40 с.
4. ГОСТ Р 59277-2020. Системы искусственного интеллекта. Классификация систем искусственного интеллекта. Национальный стандарт Российской Федерации: дата введения 2021-01-03 / Федеральное агентство по техническому регулированию и метрологии. — Изд. Официальное. — Москва: Стандартинформ, 2021 — 16 с.
5. ПНСТ 839-2023. Искусственный интеллект. Смещенность в системах искусственного интеллекта и при принятии решений с помощью искусственного интеллекта. Национальный стандарт Российской Федерации: дата введения 2024-01-01 / Федеральное агентство по техническому регулированию и метрологии. — Изд. Официальное. — Москва: Российский институт стандартизации, 2023 — 42 с.
6. Остроух, А.В. Интеллектуальные системы: монография / А.В. Остроух. — Красноярск: Научно-инновационный центр, 2020. — 316 с.
7. Никитенко С.В. Международно-правовое регулирование искусственного интеллекта: анализ текущего состояния и перспективы развития // Вестник ВУиТ. 2021. №2 (98). URL: https://cyberleninka.ru/article/n/mezhdunarodno-pravovoe-regulirovanie-iskusstvennogo-intellekta-analiz-tekuschego-sostoyaniya-i-perspektivy-razvitiya (дата обращения: 25.04.2024).