ЗАДАЧИ НОРМАТИВНО-ТЕХНИЧЕСКОГО РЕГУЛИРОВАНИЯ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

Гарбук С.В.[1]

Цель исследования. *Повышение эффективности решения задач информационной безопасности за счёт устранения нормативно-технических барьеров, препятствующих применению в перспективных системах информационной безопасности технологий искусственного интеллекта.*

Метод исследования. *В статье применён метод функциональной декомпозиции интеллектуальных задач информационной безопасности, основанный на аналогии задач искусственного и естественного интеллекта. Применительно к полученной функциональной структуре выполнена декомпозиция интеллектуальной системы информационной безопасности по процессам её жизненного цикла с выявлением частных задач нормативно-технического регулирования, специфичных для каждого из процессов, и последующим агрегированием задач в группы, соответствующие основным направлениям стандартизации таких систем.*

Полученные результаты. *Приведён структурированный перечень задач информационной безопасности, качество решения которых может быть повышено с использованием технологий искусственного интеллекта. Показано, что основные нормативно-технические барьеры, препятствующие эффективному созданию и применению интеллектуальных систем информационной безопасности, связаны с недостатками метрологического обеспечения интеллектуальных систем и с особенностями обеспечения конфиденциальности обрабатываемой в таких системах информации. Проведён анализ современного состояния работ по подготовке национальных и международных стандартов, регулирующих вопросы создания и применения интеллектуальных систем информационной безопасности, и показано, что работы в данном направлении носят начальный, постановочный характер. Обоснован перечень частных задач стандартизации, направленных на преодоление выявленных нормативно-технических барьеров при реализации отдельных процессов жизненного цикла интеллектуальных систем. Частные задачи сгруппированы по основным направлениям стандартизации, для каждого из которых подготовлены предложения по корректировке существующих и разработке новых нормативно-технических документов в области искусственного интеллекта и информационной безопасности.*

Ключевые слова: *искусственный интеллект, прикладные задачи искусственного интеллекта, интеллектуальные задачи информационной безопасности, жизненный цикл системы, оценка функциональных характеристик интеллектуальных систем, интеллометрия, качество интеллектуальных систем, информационная безопасность интеллектуальных систем.*

DOI:10.21681/2311-3456-2021-3-68-83

Интеллектуальные системы информационной безопасности

Под интеллектуальными системами информационной безопасности (ИСИБ) понимается подкласс информационных систем[2] или автоматизированных систем[3], предназначенных для решения различных задач обеспечения информационной безопасности, создание которых в обязательном порядке связано с использованием специальным образом подготовленных наборов данных (НД), описывающих представительную совокупность примеров решения этих задач. Примеры решения прикладных задач ИБ могут быть подготовлены внешней системой («учителем», чаще всего – квалифицированным человеком-оператором, обладающим необходимыми компетенциями в решении данной задачи), либо сформированы системой ИИ самостоятельно методом проб и ошибок. Во втором случае предполагается, что в системе ИИ изначально заложено критериальное правило, позволяющее отличать «хорошие» решения от «плохих».

Алгоритм работы интеллектуальных систем ИБ, сформированный в процессе обобщения обучающих примеров, не обладает, в общем случае, свойством объяснимости (понятности, интерпретируемости) для человека, что является основным отличием ИСИБ от прочих информационных систем, основанных на понятных человеку правилах и аналитических моделях [5].

1 Гарбук Сергей Владимирович, кандидат технических наук, старший научный сотрудник, директор по научным проектам НИУ «Высшая школа экономики», председатель ТК 164 «Искусственный интеллект», Москва, Россия. Email: garbuk@list.ru.

2 Информационная система – ГОСТ 33707-2016 (ISO/IEC 2382:2015) Информационные технологии (ИТ). Словарь.

3 Автоматизированная система – ГОСТ 34.003-90 Информационная технология. Комплекс средств на автоматизированные системы. Автоматизированные системы. Термины и определения.

Таблица 1

Задачи искусственного интеллекта в сфере информационной безопасности

№	Класс интеллектуальных задач ИБ	Частные задачи ИБ
1	Распознавание образов	a. применение нейросетевых алгоритмов в криптографии и стеганографии [2]; b. обнаружение вариативных компьютерных атак на информационные системы; c. выявление фрагментов программного кода, реализующего недекларированные возможности, в том числе, в условиях полиморфизма кода; d. обнаружение нетипичной активности легитимных пользователей информационных систем; e. обнаружение в исходящем трафике признаков несанкционированной передачи конфиденциальной информации
2	Построение моделей типовых объектов и процессов	a. моделирование поведения злоумышленников, соответствующих различным категориям угроз (государственные организации, террористические организации, организованные преступные сообщества, одиночные нарушители); b. моделирование информационного трафика при проведении нагрузочных испытаний средств ИБ (например, в условиях воздействия DOS-атак); c. моделирование поведения типовых пользователей информационных систем
3	Поиск решений, в том числе, в непредвиденных ситуациях	a. поиск типовых решений по обеспечению ИБ информационных систем в условиях реализации комплексных угроз информационной безопасности; b. поддержка принятия решений при проектировании и создании инфраструктуры информационной безопасности; c. оценка легитимности управляющего информационного трафика на основе прогнозирования и оценки допустимости последствий от его выполнения на физическом уровне объекта управления (реализация т.н. метода «предиктивной защиты» [3]) d. оптимизированный перебор (фаззинг) компьютерных атак при оценке эффективности используемых средств защиты информации; e. динамическое управления правами доступа пользователей к информационным ресурсам с учётом местоположения пользователей и их сетевых адресов
4	Социальные коммуникации	a. биометрическая идентификация и аутентификация; b. формирование на основе биометрических данных ключевой криптографической информации; c. визуализация объектов информационной инфраструктуры и параметров информационного трафика при поддержке работы оператора ИБ с использованием технологий виртуальной реальности

Некоторые наиболее значимые прикладные задачи ИБ, эффективность решения которых может быть повышена с использованием методов и технологий ИИ, представлены в табл.1. В приведённой таблице задачи структурированы в соответствии с классификацией, предложенной в [1] и основанной на аналогии искусственного и естественного интеллекта. Два класса интеллектуальных задач – реализация физических воздействий, а также автономное движение и позиционирование в пространстве – не являются характерными для приложений ИБ и, соответственно, не рассматриваются.

Эффективное использование в сфере ИБ новых информационных технологий, включая технологии искусственного интеллекта, невозможно без соответствующей нормативной базы[4]. Необходимо отметить, что к

4 ГОСТ Р 52633.5-2011 Защита информации. Техника защиты информации. Автоматическое обучение нейросетевых преобразователей биометрия-код доступа.

системам ИБ, использующим интеллектуальные алгоритмы обработки данных, применимы универсальные нормативно-технические документы в области информационных технологий, инвариантные к способам реализации информационных систем. Тем не менее, для ИСИБ таких документов оказывается недостаточно, что приводит, в частности, к следующим проблемам (нормативно-техническим «барьерам»):

- отсутствие унифицированных механизмов объективной оценки функциональных характеристик интеллектуальных систем ИБ в условиях недостаточной понятности (прозрачности, интерпретируемости) используемых алгоритмов обработки данных ограничивает возможность применения таких систем при обеспечении информационной безопасности на ответственных объектах;
- отсутствие требований к процедурам гарантированной «деклассификации» и автоматизированным средствам подтверждения невозможности бесконтрольного повышения уровня конфиденциальности информации, обрабатываемой в системах ИИ, препятствует формированию общедоступных обучающих наборов данных и, соответственно, созданию передовых интеллектуальных систем ИБ с привлечением широкого круга разработчиков.

Мировой опыт нормативно-технического регулирования вопросов применения технологий ИИ в системах информационной безопасности

Перечисленные выше проблемы, вызванные недостатками нормативно-технического регулирования вопросов применения ИИ в системах информационной безопасности, в той или иной мере рассматриваются в разных странах и на международном уровне.

Наиболее активно разработка соответствующих стандартов ведётся в Китае. Китайскими специалистами считается, что применение технологий ИИ позволит повысить эффективность решения таких задач информационной безопасности, как обнаружение уязвимостей, оценка защищённости информационных систем, предупреждение об угрозах, обнаружение атак и реагирование на чрезвычайные ситуации. При этом особое внимание должно уделяться ключевым продуктам и приложениям, таким как умные автомобили и умные дома. К 2020 году планировалось повысить общий уровень кибербезопасности в стране за счёт создания интеллектуальной платформы обеспечения информационной безопасности, включающей ситуационную осведомленность, тестирование, обмен информацией об угрозах и реагирование на чрезвычайные ситуации [1].

Особая роль в Китае отводится также вопросам обеспечения кибербезопасности для защиты все более взаимосвязанных производственных процессов от компьютерных сетевых атак. В частности, к 2025 году планируется построить базовую систему интеллектуального производственного оборудования и систему технологических инноваций с независимой управляемостью, безопасностью, надежностью и производительностью. К этому же году (второй этап реализации национального плана по развитию ИИ) будут разработаны законы и нормативные акты в области ИИ, этические нормы, нормативно-методические принципы оценки возможностей и контроля безопасности ИИ.

В обзоре, подготовленном специалистами Института инженеров электротехники и электроники (IEEE)[5], приведены следующие особенности нормативного регулирования ИБ в условиях развития технологий искусственного интеллекта:

1) поддержка создания усовершенствованных регуляторных структур, наиболее полно учитывающих изменения аспектов ИБ, связанные с развитием технологий ИИ. Например, в США правоприменение во многом подпадает под юрисдикцию Федеральной торговой комиссии;

2) создание дополнительных контуров технической обратной связи для регуляторов. В США основная часть вопросов ИБ, связанных с развитием интеллектуальных технологий, будет рассматриваться в Конгрессе и соответствующих регулирующих органах. Отмечается, однако, что в настоящее время у Конгресса нет финансируемого офиса по оценке технологий ИИ и необходимо возродить этот технический орган в виде соответствующего управления;

3) совершенствование нормативной базы и финансовая поддержка исследований в области ИБ систем ИИ, устранение соответствующих терминологических неопределённостей;

4) устранение проблем, связанных с необходимостью международной гармонизации законодательства в области информационной безопасности систем ИИ в условиях существенных различий регуляторных норм в разных странах;

5) необходимость совершенствования инструментов, предназначенных для выявления и судебного преследования преступлений, совершаемых с помощью ИИ;

6) необходимость понимания ограниченных функциональных способностей систем ИИ и повышения уровня доверия к системам ИИ, некорректная работа которых может привести к тяжким последствиям;

7) необходимость использования неизбежного временного интервала между развитием новых технологий ИИ и разработкой нормативных правовых документов, направленных на их регулирование, для создания технических стандартов. Такие стандарты могут создаваться, например, на базе IEEE путём формирования международного консенсуса;

8) необходимость совершенствования процедур сбора сведений о выявленных уязвимостях систем ИИ и обновления форматов представления этих сведений. В условиях технологий ИИ сбор сведений должен осуществляться более оперативно, полнота собираемой информации также должна возрасти;

9) предъявление требований в области ИБ к системам ИИ ещё на этапе их проектирования и разработки;

10) обязательный учёт этических принципов при разработке и применении систем ИИ.

5 Artificial Intelligence and Machine Learning Applied to Cybersecurity. The result of an intensive three-day IEEE Confluence 6-8 October 2017.

Таким образом, в вышеупомянутом обзоре IEEE, ориентированном в основном на специфику Соединённых Штатов, аспекты нормативно-технического регулирования вопросов применения ИИ в сфере информационной безопасности хотя и упоминаются, но не детализируются.

На международном уровне нормативно-технические документы, регулирующие вопросы создания и применения систем ИИ разрабатываются в рамках профильного подкомитета SC42 “Artificial Intelligence” объединённого технического комитета ISO/IEC JTC 1 “Information Technologies”. При этом документы, непосредственно регулирующие особенности применения систем ИИ в средствах ИБ, в настоящее время не разрабатываются.

Общие вопросы обеспечения несмещённости (объективности, справедливости) обучающих наборов данных для систем ИИ, актуальные также и для ИСИБ, рассматриваются в международном стандарте[6]. В частности, вводятся следующие определения:

- предвзятость (bias) – систематические различия в отношении к определённым объектам, людям и их группам по сравнению друг с другом;
- когнитивные заблуждения человека (human cognitive bias) - человеческая предвзятость, которая может повлиять на разработку и применение системы ИИ;
- заблуждение автоматизации (automation bias) – когнитивные заблуждения человека, вызванные чрезмерным доверием к рекомендациям системы ИИ;
- стереотипные заблуждения (confirmation bias) – тип когнитивных заблуждений человека, заключающийся в предпочтении тем решениям системы ИИ, которые подтверждают существовавшие ранее гипотезы или ожидания;
- удобная выборка (convenience sample) – выборка данных, сформированная из соображений удобства процедуры сбора данных, а не из соображений их репрезентативности;
- смещение данных (data bias) – особенности данных, игнорирование которых приводит к тому, что системы ИИ, созданные с их помощью, работают лучше или хуже для разных типов объектов.

Также в рамках подкомитета SC42 в 2019–2020 годах рассмотрены различные варианты использования технологий ИИ и в результате был составлен специальный технический отчёт[7], в котором были выделены следующие приложения в области безопасности:

1) системы анализа настроения и поведения людей для предотвращения противоправных действий, предупреждения суицидальных происшествий, словесного описания состояния и поведения человека и решения других задач в области защиты прав человека;

2) системы на основе роевого ИИ, обеспечивающие выявление атак на объекты интернета вещей. Предназначены для предотвращения кражи электроэнергии на основе выявления аномальных событий в «умных» счётчиках с использованием роевого ИИ;

3) робот-гуманоид для управления дорожным движением. Данное решение позволит исключить нахождение человека-регулировщика в сильно загрязнённой среде;

4) роботизированная система для замены человека при производстве работ в опасных условиях, включая шахты, доменные печи, котловое оборудование и т.п.;

5) средства выявления вредоносного ПО на мобильных устройствах, основанные на анализе таких действий устройства, как использование аккумуляторной батареи, данных, служб оценки местоположения, микрофона и т.п. Ожидается, что система будет выявлять различные типы кибератак без доступа к персональным данным пользователя, однако может оказаться неэффективна в отношении выявления наиболее сложных атак.

Таким образом, в ISO/IEC DTR 24030 вопросам ИБ из 132 приложений целиком посвящено одно (5) и частично – ещё одно (2).

В контексте вопросов информационной безопасности обращает на себя внимание документ[8], разработка которого осуществляется в рамках смежного для SC42 подкомитета SC27 “Information security, cybersecurity and privacy protection”. В документе ISO/IEC CD 20547-4 основное внимание уделено вопросам обеспечения безопасности (в соответствии со стандартом[9]) и приватности (в соответствии с техническим отчётом[10]) больших данных. При этом указано, что использование методов ИИ для решения задач информационной безопасности имеет следующие особенности:

- концепция «больших данных», неразрывно связанная с областью ИИ, приводит к размыванию границ между традиционно рассматриваемыми по отдельности процессами сбора, хранения и предоставления доступа к данным. Обеспечение ИБ усложняется в условиях появления таких новых угроз, как расширенная постоянная угроза, распределенный отказ в обслуживании (DDoS), интеллектуальный анализ данных и нарушение конфиденциальности на основе машинного обучения;
- информационная безопасность должна быть обеспечена в распределённой среде обработки и хранения данных, что предъявляет соответствующие требования к средствам ИБ, в том числе в части их масштабируемости;
- методы машинного обучения могут быть использованы в средствах ИБ для поиска персональных данных в распределённых и неструктурированных хранилищах информации и предотвращения, таким образом, несанкционированной утечки этих данных.

6 ISO/IEC 24027:2020 Information technology - Artificial Intelligence (AI) – Bias in AI systems and AI-aided decision making.

7 ISO/IEC DTR 24030 - Information Technology -- Artificial Intelligence – Use cases.

8 ISO/IEC CD 20547-4 Information technology – Big Data Reference Architecture – Part 4: Security and Privacy.

9 ISO/IEC 27000:2018 Information technology — Security techniques — Information security management systems — Overview and vocabulary.

10 ISO/IEC TR 27550:2019 Information technology — Security techniques — Privacy engineering for system life cycle processes.

В дорожной карте стандартов в области больших данных[11] в качестве одного из трёх наиболее ярких примеров, иллюстрирующих прикладное значение интеллектуальных методов обработки больших объёмов эмпирических данных, приведена задача в ИБ, при решении которой защищающаяся сторона получит преимущество перед злоумышленником.

В Европейском Союзе основное внимание уделяется формированию единых этически сбалансированных правил использования технологий ИИ. В мае 2018 года принят Общий регламент защиты персональных данных ЕС (General Data Protection Regulation, GDPR) – документ Европейского Союза, в соответствии с которым устанавливаются унифицированные правила защиты персональных данных европейцев[12]. Документ GDPR существенно способствует развитию технологий ИИ, так как определяет правовую базу в области использования персональных данных при обучении систем ИИ. При этом документ является универсальным, в явном виде технологии искусственного интеллекта в нём не упоминаются. Что касается вопросов информационной безопасности, то документ GDPR определяет организационные меры в области защиты информации (персональных данных), требования к средствам ИБ в нём также не рассматриваются.

Появления технических стандартов, устанавливающих требования к использованию алгоритмов ИИ в системах информационной безопасности, следует ожидать также в Европейской организации по стандартизации ETSI, на официальном сайте которой данный аспект указывается в качестве одного из перспективных направлений стандартизации[13]. Однако по состоянию на конец 2020 года организацией ETSI разработан и опубликован единственный обзор[14], в котором изучаются вопросы обеспечения информационной безопасности интеллектуальных систем.

В работе [8] отмечается, что для эффективного совершенствования средств ИБ на основе использования интеллектуальных алгоритмов национальными регуляторами должна создаваться нормативная база, необходимая для динамического тестирования, валидации и сертификации средств ИИ для систем ИБ. На международном уровне необходимо вырабатывать общие нормы в области исследований и разработок и формировать разумные ограничения на распространение знаний и возможностей в этой технологической области. Подчёркивается также, что ускорение процессов создания нормативной базы в области ИИ достигается также благодаря использованию высокоуровневых принципов, например, принципов доверенного ИИ (trustworthy AI), разработанных ОЭСР.

Таким образом, проведённый анализ показал, что разработка стандартов, устанавливающих требования в области применения технологий ИИ в системах информационной безопасности, находится на самой начальной стадии. Это обуславливает особую актуальность работ по обоснованию рационального перечня задач нормативно-технического регулирования, который, с одной стороны, учитывал бы необходимость преодоления приведённых выше барьеров, а с другой – не создавал искусственных препятствий на пути развития и прикладного использования технологий ИИ.

Задачи нормативно-технического регулирования процессов жизненного цикла интеллектуальных систем информационной безопасности

Выявление особенностей и разработка предложений по нормативно-техническому регулированию вопросов создания и применения ИСИБ могут быть выполнены на основе декомпозиции процессов жизненного цикла (ЖЦ) систем в соответствии с национальным стандартом ГОСТ Р 57193-2016[15] (табл.2). Подобный подход был апробирован и хорошо зарекомендовал себя, например, при решении задач управления качеством и рисками в жизненном цикле сложных систем [9]. Отметим, что последовательность, в которой процессы приведены в данном стандарте, не подразумевает предписывающего порядка их использования. Последовательность использования процессов вводится при определении модели ЖЦ конкретной системы ИБ с учетом воздействия множественных факторов, включая социальные, торговые, организационные и технические, каждый из которых может меняться в период жизни системы.

Задачи нормативно-технического регулирования, приведённые в табл.2, могут быть сгруппированы в следующие функциональные группы:

1) обоснование требований и измерение существенных функциональных характеристик и характеристик безопасности ИСИБ (2.5.а, 3.4.б, 3.7.а, 3.7.б, 3.8, 4.1.б, 4.2.а, 4.6, 4.9, 4.11);

2) оценка функциональных возможностей и уровня безопасности квалифицированного человека-оператора, осуществляющего решение заданной прикладной задачи ИБ в ручном режиме (2.4.а, 4.1.а);

3) формализация предусмотренных условий эксплуатации ИСИБ (4.2.б, 4.2.в);

4) обеспечение конфиденциальности данных при создании и применении ИСИБ (3.4.а, 3.6.д, 3.6.е, 4.12.б, 4.14);

5) унификация и повышение доступности наборов данных, необходимых для создания и оценки соответствия ИСИБ (3.1.а, 3.6.а, 3.6.б, 3.6.в, 3.6.г);

6) управление дообучением ИСИБ на стадии эксплуатации, тиражирование разработанных программно-алгоритмических решений на смежные задачи ИБ (2.5.б, 2.6.а, 4.12.а).

11 ISO/IEC CD 20547-5 Information technology – Big Data Reference Architecture – Part 5: Standards Roadmap.

12 Regulation (EU) 2016/679 of the European Parlament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation, GDPR)

13 Сайт Европейской организации по стандартизации ETSI. https://www.etsi.org/technologies/securing-artificial-intelligence. Дата обращения 12.02.2021.

14 ETSI GR SAI 004 V1.1.1 (2020-12). Securing Artificial Intelligence (SAI); Problem Statement. Group Report.

15 ГОСТ Р 57193-2016. Системная и программная инженерия. Процессы жизненного цикла систем (ISO/IEC/IEEE 15288:2015, NEQ).

Таблица 2

Особенности процессов жизненного цикла ИСИБ

Группа процессов ЖЦ	Процесс ЖЦ	Задачи нормативно-технического регулирования при реализации процесса ЖЦ для ИСИБ
1 Процессы соглашения	1.1 Приобретение	-
	1.2 Поставка	-
2 Процессы организационного обеспечения проекта	2.1 Управление моделью ЖЦ	-
	2.2 Управление инфраструктурой	-
	2.3 Управление портфелем	-
	2.4 Управление человеческими ресурсами	а) В том случае, если ИСИБ предназначена для автоматизации интеллектуальной деятельности, обычно выполняемой человеком вручную, для оценки качества ИСИБ может понадобиться референтная группа квалифицированных операторов [6, 7, 12]
	2.5 Управление качеством	а) Для типовых прикладных задач ИБ должны быть разработаны перечни существенных функциональных характеристик ИСИБ и метрики, обеспечивающие оценку качества как степени соответствия значений функциональных характеристик установленным требованиям. Подробнее вопросы оценки качества систем ИИ рассмотрены в [10] б) Если при эксплуатации ИСИБ предусмотрено дообучение системы, то должны быть установлены процедуры периодического контроля качества на этапе функционирования
	2.6 Управление знаниями	а) Должны быть разработаны требования к процедурам повторного использования знаний, полученных в ходе обучения ИСИБ, в том числе, с использованием процедур «переноса обучения» (transfer learning)
3 Процессы технического управления	3.1 Планирование проекта	а) Процесс планирования проекта в обязательном порядке должен предусматривать задачи по формированию обучающих, тестовых и иных наборов данных, специфичных для конкретной прикладной задачи ИБ
	3.2 Оценка и контроль проекта	-
	3.3 Управление решениями	-
	3.4 Управление рисками	а) На этапе обучения и функционирования ИСИБ в обязательном порядке должны учитываться риски, связанные с возможным неконтролируемым повышением уровня конфиденциальности обрабатываемых данных
		б) На этапах верификации, валидации и функционирования в обязательном порядке должны учитываться риски, связанные с уязвимостью интеллектуальных алгоритмов к воздействию специфических атак на исходные данные, включая состязательные (adversarial) атаки
	3.5 Управление конфигурацией	-

Группа процессов ЖЦ	Процесс ЖЦ	Задачи нормативно-технического регулирования при реализации процесса ЖЦ для ИСИБ
	3.6 Управление информацией	а) Должны быть предусмотрены требования к специфическим процедурам обработки информации, связанным с разметкой данных, необходимых для обучения, дообучения и тестирования ИСИБ (далее – НД ИСИБ)
		б) Должны быть разработаны специальные требования к качеству НД ИСИБ
		в) Должны быть разработаны требования к процедурам расширения (аугментации) НД ИСИБ
		г) Должны быть разработаны унифицированные форматы представления НД ИСИБ. Данная задача имеет место для любых информационных систем, но в случае систем ИИ приобретает особую актуальность вследствие размывания границ между стадиями сбора, хранения и предоставления доступа к данным
		д) Должны быть разработаны требования к обеспечению конфиденциальности НД ИСИБ, исключающих использование этих данных для недобросовестной конкуренции или использование злоумышленниками сведений об этих данных для повышения эффективности реализации угроз ИБ в отношении ИСИБ
		е) Должны быть разработаны требования к методам и средствам гарантированной деклассификации НД ИСИБ, то есть такого преобразования данных, при котором уровень их конфиденциальности необратимо становится достаточно низким и появляется возможность предоставления этих данных широкому кругу разработчиков ИСИБ и другим заинтересованным лицам без рисков нарушения конфиденциальности по 3.6.д
	3.7 Измерения	а) Должны быть разработаны требования к унифицированным процедурам измерения функциональных характеристик ИСИБ, основанным на тестировании систем на представительных наборах данных. Подробнее принципы измерения функциональных характеристик систем ИИ рассмотрены в [10] в рамках предложенного направления «интеллометрия»
		б) Если некорректная работа ИСИБ может привести к непосредственным угрозам безопасности третьих лиц, то должны быть предусмотрены унифицированные процедуры измерения (оценки) вероятности реализации таких угроз
	3.8 Гарантии качества	а) Унифицированные процедуры оценки функциональных характеристик (3.7.а), характеристик безопасности (3.7.б) и качества (2.5) ИСИБ должны обеспечивать получение соответствующих оценок с заданными точностью и достоверностью в условиях действующих рисков (3.4) и применительно к планируемым условиям эксплуатации ИСИБ по 4.2
4 Технические процессы	4.1 Анализ бизнеса или назначения	а) Если ИСИБ предназначена для автоматизации интеллектуальной деятельности человека, то должны быть предусмотрены процедуры формализации соответствующей интеллектуальной прикладной задачи и определение функциональных возможностей 3.7.а и уровня безопасности 3.7.б референтной группы квалифицированных операторов (2.4.а)

Группа процессов ЖЦ	Процесс ЖЦ	Задачи нормативно-технического регулирования при реализации процесса ЖЦ для ИСИБ
		б) Для ИСИБ должны быть сформулированы функциональные требования (3.7.а) и требования по безопасности для третьих лиц (3.7.б). При этом могут учитываться возможности квалифицированных операторов 4.1.а по решению заданной прикладной задачи ИБ в ручном режиме
	4.2 Определение потребностей и требований заинтересованной стороны	а) Для типовых прикладных задач ИБ должны быть разработаны перечни существенных функциональных характеристик ИСИБ (см. также 2.5.а)
		б) Для предусмотренных условий эксплуатации ИСИБ должны быть определены перечни внешних (не зависящих от ИСИБ) факторов, существенным образом влияющих на функциональные характеристики и характеристики ИСИБ (перечень существенных условий эксплуатации)
		в) Для предусмотренных условий эксплуатации должны быть определены диапазоны допустимых значений существенных условий эксплуатации (4.2.б), при которых должна сохраняться возможность применения ИСИБ по назначению (область применения системы, domain) с гарантией по 3.8
	4.3 Определение системных требований	-
	4.4 Определение архитектуры	-
	4.5 Определение проекта	-
	4.6 Системный анализ	а) Должны быть предусмотрены процедуры, обеспечивающие выбор рационального компромисса между объяснимостью (понятностью, transparency, explainability) алгоритма работы ИСИБ и качеством системы (по 2.4 и 2.5)
	4.7 Реализация	-
	4.8 Комплексирование	-
	4.9 Верификация	а) Должны быть разработаны требования к унифицированным методикам измерения существенных функциональных характеристик по 3.7.а и 4.2.а и методикам оценки рисков для третьих лиц по 3.7.б
		б) При необходимости должны быть разработаны фрагменты тестовых наборов данных, на которых должно осуществляться измерение характеристик 4.9.а, и/или описания тестовых ситуаций, в которых должно осуществляться измерение характеристик 4.9.а. Фрагменты тестовых наборов данных и описания тестовых ситуаций должны обеспечивать формирование тестовых НД, обладающих представительностью в условиях эксплуатации 4.2
		в) Должна быть обеспечена конфиденциальность тестовых наборов данных, используемых для оценки соответствия систем ИИ предъявляемым требованиям, исключающая снижение достоверности получаемых оценок, вызванное переобучением ИСИБ
	4.10 Передача	-

DOI:10.21681/2311-3456-2021-3-68-83

Группа процессов ЖЦ	Процесс ЖЦ	Задачи нормативно-технического регулирования при реализации процесса ЖЦ для ИСИБ
	4.11 Валидация	а) Должны быть разработаны унифицированные методики подтверждения возможности успешного решения соответствующих прикладных задач ИБ в условиях эксплуатации 4.2
	4.12 Функционирование	а) На стадии функционирования должны быть установлены требования к процедурам дообучения ИСИБ, учитывающие требования по управлению качеством 2.5.б
		б) Должны быть предусмотрены процедуры непрерывного контроля уровня конфиденциальности обрабатываемых данных, исключающие реализацию рисков в соответствии с 3.4.а
	4.13 Сопровождение	-
	4.14 Изъятие и списание	а) Должны быть установлены требования к процедурам оценки уровня конфиденциальности данных, накопленных (сформированных) в процессе функционирования системы. Данные процедуры должны исключать нарушение конфиденциальности информации при изъятии и списании ИСИБ

Отметим, что первые четыре задачи ориентированы преимущественно на устранение нормативного барьера, связанного с отсутствием гарантий качества и безопасности работы ИСИБ, а две последних – на преодоление сложностей создания ИСИБ, связанных с ограничениями на доступ к данным и низким уровнем унификации ИСИБ. В последующих разделах будут рассмотрены варианты решения этих задач, предусматривающие корректировку существующих и разработку новых нормативно-технических документов.

Обоснование требований и измерение функциональных характеристик ИСИБ

Целью данной группы задач нормативно-технического регулирования является обеспечение метрологического единства при измерении функциональных характеристик ИСИБ, предназначенных для решения типовых задач ИБ (табл.1). Как уже было отмечено выше, особенностью интеллектуальных систем ИБ, как и любых других интеллектуальных систем обработки данных, является необходимость использования представительных тестовых наборов исходных данных при оценке функциональных характеристик таких систем. Проведённый анализ терминологических стандартов в области информационной безопасности[16] свидетельствует о том, что данная особенность ИСИБ в действующих нормативно-технических документах не учитывается и соответствующие стандарты должны быть дополнены рядом терминов, характеризующих используемые при проектировании, разработке и тестировании ИСИБ наборы данных (рис.1):

- базовый демонстрационный набор данных – образцовый размеченный НД и/или описание способов формирования такого НД (включая, при необходимости, описание тестовых сценариев, в которых необходимо осуществлять испытания системы ИИ), включённые в состав технического стандарта, устанавливающего унифицированные требования к проведению испытаний прикладных систем ИИ определённой функциональности. Подобные НД и/или правила их формирования являются опциональными и характеризуют наиболее общие особенности данной прикладной интеллектуальной задачи, в рассматриваемом случае – в области ИБ;
- дополнительный демонстрационный набор данных – дополнительный размеченный НД, предоставляемый заказчиком системы ИИ при формировании уточнённых требований к системе с учётом специфики конкретной решаемой прикладной интеллектуальной задачи;
- демонстрационный набор данных – совокупность базового и дополнительного демонстрационных НД;
- обучающий набор данных – набор данных, формируемый разработчиком системы ИИ на основе демонстрационного НД и необходимый для создания прикладной технологии ИИ. При формировании обучающего НД разработчиком широко применяются технологии аугментации данных, позволяющие повысить качество создаваемой системы ИИ;
- тестовый набор данных – НД, формируемый в испытательной лаборатории на основе демонстрационного НД и необходимый для проведения сертификации, испытаний или аттестации системы ИИ. Особенностью тестового НД является то, что этот набор не должен быть известен разработчику системы ИИ (в противном случае испытания окажутся неэффективными), но при этом разработчик и другие заин-

16 1)ГОСТ Р 50922-2006 Защита информации. Основные термины и определения.
2)ГОСТ Р 53114-2008 Защита информации. Обеспечение информационной безопасности в организации. Основные термины и определения.

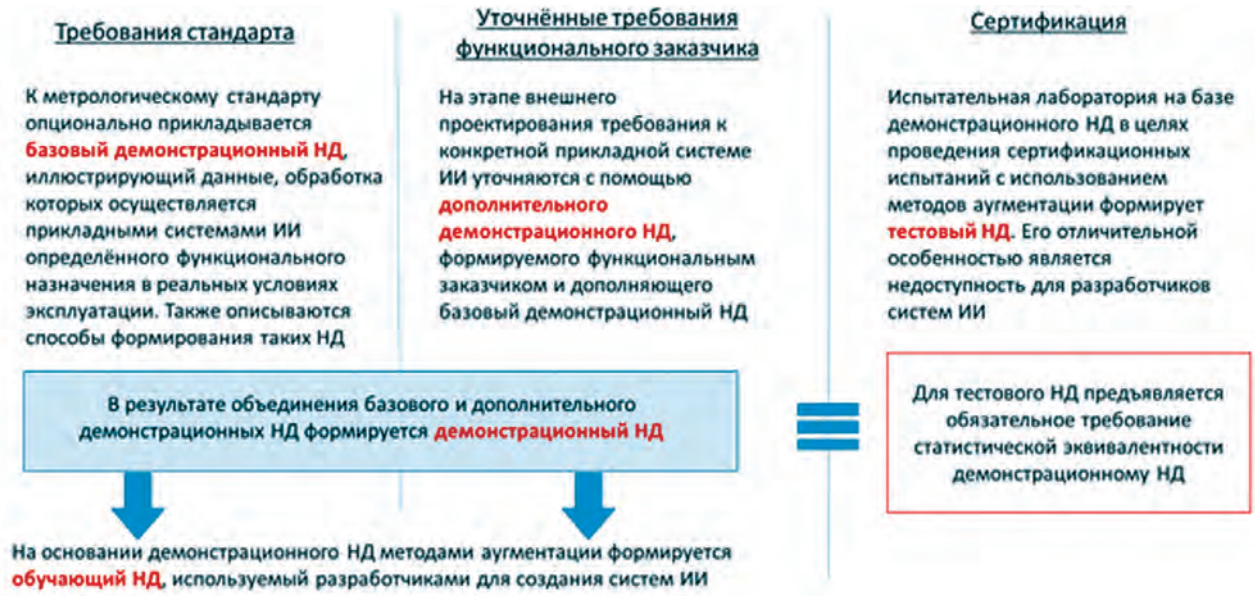

Рис.1. Виды наборов данных, используемых при создании систем ИИ

тересованные стороны должны иметь возможность проверять корректность формирования тестового набора, то есть принадлежность демонстрационного и тестового наборов данных одной генеральной совокупности (в противном случае результаты испытаний будут являться непредставительными). Достижение компромисса в выполнении этих конфликтующих требований (конфиденциальность тестовых НД и прозрачность процедуры сертификации) предполагает корректировку руководящих документов, регламентирующих работу органов по сертификации средств ИБ.

Описанная совокупность наборов данных позволяет реализовать предложенный в [10] метрологический подход к оценке соответствия систем ИИ установленным требованиям, поименованный как «интеллометрия». В соответствии с данным подходом на начальном этапе предполагается формирование перечня характеристик систем ИИ, значимых для решаемой прикладной задачи $\boldsymbol{G}$. Совокупность значений таких характеристик для системы $\boldsymbol{S}$ может быть представлена в виде вектора функциональных характеристик (ВФХ) $F(\boldsymbol{S})=\{f_1, f_2, \ldots f_N\}$, компоненты которого f_i определены на шкалах значений, выбранных, исходя из физического смысла соответствующих характеристик (числовая шкала, шкала категорий, отношений и др.).

Далее для каждой компоненты ВФХ f_i. предполагается определение способа измерения значений соответствующей характеристики системы ИИ, причём, учитывая особенности интеллектуальных методов обработки данных, измерение значений должно осуществляться на представительной выборке тестовых исходных данных $T \in \Gamma$, где Γ – множество возможных значений исходных данных для задачи $\boldsymbol{G}$, учитывающее предусмотренные условия эксплуатации системы ИИ. При этом $T = T_1 \bigcup T_2 \bigcup \ldots \bigcup T_N$, где T_i – выборка исходных данных, представительная для измерения характеристики f_i.

В N-мерном пространстве функциональных характеристик обосновывается метрика $M^{ab} \in R^1$ как правило сравнения функциональных возможностей систем $\boldsymbol{S^a}$ и $\boldsymbol{S^b}$, заданных векторами F^a и F^b:

$$M^{ab} = M\left(F^a, F^b\right) = \sum_i^N v_i \mu_i\left(f_i^a, f_i^b\right),$$

где v_i – коэффициент, характеризующий важность i-й характеристики f_i в контексте решаемой задачи ИБ; $\mu_i\left(f_i^a, f_i^b\right) \in R^1$ – частная метрика для i-й характеристики, определённая на соответствующей шкале значений.

$M^{ab} = -M^{ba}$, причём $M^{ab} = 0$, если функциональные возможности систем $\boldsymbol{S^a}$ и $\boldsymbol{S^b}$ совпадают, $M^{ab} > 0$, если функциональные возможности системы $\boldsymbol{S^a}$ превышают возможности системы $\boldsymbol{S^b}$.

В этом случае представительность тестовой выборки T означает, что её объём и вариативность позволяют с заданной вероятностью γ_T утверждать, что модуль метрики для оценок ВФХ системы ИИ $\boldsymbol{S^a}$, полученных на тестовой выборке F_T^a и на выборке, соответствующей реальным условиям эксплуатации, $\widehat{F_\Gamma^a}$ не будет превышать некоторую заданную величину ε_T:

$$P\left(\left|M\left(\widehat{F_T^a}, \widehat{F_\Gamma^a}\right)\right| \le \varepsilon_T\right) = \gamma_T.$$

Требования к тестовой выборке существенным образом зависят от сложности решаемой интеллектуальной задачи ИБ. При оценке сложности задачи должны учитываться:

- информационные свойства источника данных, зависящие от многообразия возможных состояний контролируемых объектов и контекстов, а также пропускной способности сенсоров (объективная оценка сложности задачи);
- возможности метасистемы (потребителя) по усвоению результатов обработки данных в системе ИИ (субъективная оценка сложности);

DOI:10.21681/2311-3456-2021-3-68-83

- функциональные возможности другой (референтной) информационной системы, с приемлемым качеством решающей заданную задачу ***G***. Например, для интеллектуальных систем, предназначенных для замены человека при решении рутинных интеллектуальных задач, необходимо учитывать компетенции квалифицированного человека-оператора, обеспечивающего решение задачи с заданным качеством (референтная оценка качества);
- сложность формализованного, в том числе вербального, описания интеллектуальной задачи (семантическая оценка сложности);
- другие существенные факторы.

Применительно к задачам информационной безопасности полнота тестового набора данных может обеспечиваться путём:

- использования наиболее полных, своевременно актуализируемых БД, содержащих сведения о выявленных угрозах ИБ, антивирусных БД и т. п., в том числе, подготовленных разными разработчиками и исследователями;
- полиморфной модификации программного кода, реализующего угрозы ИБ (вариант аугментации набора данных применительно к области информационной безопасности). Возможные варианты таких преобразований рассмотрены, например, в [11];
- использования мутационных и порождающих фаззеров, обеспечивающих оптимизированный перебор возможных информационных атак на интеллектуальную систему ИБ;
- использования интеллектуальных симуляторов для моделирования поведения злоумышленников, в частности, при реализации угроз ИБ, связанных с применением методов социальной инженерии;
- использования других подходов.

Далее для системы ИИ должен быть обоснован функционал качества $Q^a = Q\left(M^{ar}\right)$, значение которого (показатель качества системы) определяет степень соответствия характеристик F^a системы Sa характеристикам F^r некоторой эталонной (референтной) системы $\boldsymbol{S^r}$, а также объективную полезность системы $\boldsymbol{S^a}$ для решения интеллектуальной задачи ***G***.

$Q^a \in \left[0; Q_{max}\right]$, причём $Q^a = 1$ при $M^{ar} = 0$; $Q^a = 0$ при фактическом отсутствии у системы $\boldsymbol{S^a}$ функциональных возможностей, значимых для решения интеллектуальной задачи ***G***; $Q^a = Q_{max}$, если дальнейшее улучшение функциональных характеристик F^a не приводит к повышению качества решения задачи ***G***, но вызывает дополнительные ресурсные затраты и, соответственно, приводит к снижению эффективности решения задачи ***G***. Предельное значение Q_{max} определяется ограниченными информационными возможностями сенсоров, поставляющих информацию для системы $\boldsymbol{S^a}$, а также ограниченными возможностями метасистемы (например, объекта управления), получающей информацию от $\boldsymbol{S^a}$. Примерный вид функционала $Q\left(M^{ar}\right)$ показан на рис.2.

Для ИСИБ, предназначенных для автоматизации деятельности человека, в качестве характеристик F^r референтной системы S^r могут быть приняты возможности квалифицированного оператора, с приемлемым качеством решающего прикладную интеллектуальную задачу ***G***. Оценка F^r должна осуществляться с привлечением группы экспертов, причём размеры этой группы при заданной тестовой выборке T могут быть определены с использованием коэффициента конкордации, характеризующего степень согласованности мнений экспертов [12].

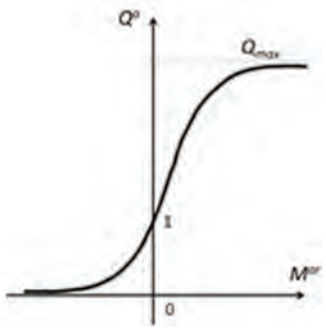

Рис.2. Примерный вид функционала качества системы ИИ

Завершающим этапом «интеллометрической» процедуры подтверждения соответствия системы ИИ установленным требованиям является выбор критериального значения функционала качества $Q^* \in]0; Q_{max}]$. Для определённого критериального значения Q^* решение о соответствии ИСИБ $\boldsymbol{S^a}$ установленным требованиям принимается при выполнении условия: $Q^a \geq Q^*$.

До настоящего времени в нормативных документах в области качества информационных систем [4] требования к способам измерения и сравнения функциональных характеристик, функционалам качества ИСИБ, требования к тестовым НД, включая требования в области полноты, не разработаны или носят формальный характер. Так, например, в стандарте по оценке качества программных средств ГОСТ 28195-89[17] одним из оценочных элементов фактора «надёжность ПС» упоминается «Наличие тестов для проверки значений входных данных», однако способы формирования и требования к полноте таких тестов не регламентируются. В ГОСТ Р 52447-2005[18] приведены номенклатура показателей качества и характеризуемые ими свойства СЗИ от несанкционированного доступа, а также номенклатура показателей качества и характеризуемые ими свойства программных СЗИ. При этом в составе показателей качества также отсутствуют понятия, характеризующие полноту обучающего НД, использованного при создании ИСИБ.

Данное обстоятельство существенно ограничивает возможности по применению таких систем. Так, например, одно из перспективных направлений применения технологий ИИ в средствах криптографической защи-

17 ГОСТ 28195-89 Оценка качества программных средств. Общие положения.

18 ГОСТ Р 52447-2005 Защита информации. Техника защиты информации. Номенклатура показателей качества.

ты информации охватывает методы преобразования биометрических параметров человека в уникальный криптографический ключ с использованием нейросетевых алгоритмов. В разработанном национальном стандарте[19] описывается подход к обучению нейронной сети, позволяющей формировать ключ длиной 30 бит на основе биометрических данных. В настоящее время авторами ГОСТ Р 52633.5-2011 в рамках проекта ГОСТ Р 52633.5-XX «Защита информации. Техника защиты информации. Автоматическое обучение сетей квадратичных нейронов с многоуровневым квантованием биометрических данных» предлагается усовершенствованный подход, обеспечивающий формирование ключа длиной 60 бит. Один из проблемных вопросов, связанных с применением предлагаемого метода формирования криптографических ключей, заключается в отсутствии убедительных доказательств влияния уровня конфиденциальности параметров нейронной сети на эффективность атаки на криптографический ключ. Получение таких доказательств предполагает тестирование нейросетевого алгоритма формирования ключа на представительном наборе биометрических данных с моделированием атак на ключ в условиях полной или частичной доступности злоумышленнику параметров нейронной сети.

Оценка функциональных возможностей квалифицированного человека-оператора при решении задач безопасности в ручном режиме

В соответствии с ГОСТ Р ИСО/МЭК 9126-93[20] основными методами оценки качества средств ИСИБ являются регистрационный и экспертный, что объясняется присущей системам ИИ непрозрачностью алгоритмов обработки данных. Регистрационный метод основан на получении информации во время испытаний или функционирования программных средств (ПС), когда регистрируются и подсчитываются определённые события, например, время и число сбоев и отказов СЗИ и др.

Определение значений показателей качества ПС экспертным методом осуществляется группой экспертов-специалистов, компетентных в решении данной задачи, на базе их опыта и интуиции (например, при выявлении с использованием методов ИИ обфусцированного исходного кода). При этом возникает задача сопоставления измеренных функциональных характеристик ИСИБ с возможностями экспертов при решении той или иной прикладной задачи информационной безопасности.

Действующая нормативная база не устанавливает единых требований по измерению функциональных возможностей экспертов, что затрудняет принятие решений о замене человека-оператора на ИСИБ при решении ответственных задач ИБ. Некоторые подходы к сопоставлению функциональных характеристик прикладных интеллектуальных систем и квалифицированного человека-оператора рассмотрены в [6, 7, 12].

Формализация предусмотренных условий эксплуатации ИСИБ

Необходимо отметить, что анализ полноты тестовых НД и, соответственно, представительности оценок функциональных характеристик ИСИБ возможен лишь при наложении определённых ограничений на условия применения системы. В англоязычных документах, в частности, в проекте основополагающего стандарта, разрабатываемого рамках подкомитета ISO/IEC JTC 1 SC 42[21], для обозначения предусмотренных условий эксплуатации системы ИИ используется термин «domain».

Формализация предусмотренных условий применения предполагает выявление внешних факторов, существенным образом влияющих на сложность решаемой прикладной интеллектуальной задачи ИБ, и определение диапазонов значений, которые могут принимать эти факторы в реальных условиях эксплуатации ИСИБ. В действующих нормативных документах перечни таких существенных условий эксплуатации для типовых задач ИБ не определены. Требования к унифицированным способам подтверждения полноты учёта существенных условий эксплуатации при создании и тестировании ИСИБ также не разработаны.

Устранению данного нормативного пробела будет способствовать корректировка национального стандарта ГОСТ 28195-89 путём дополнения в группу показателей универсальности программных средств (гибкость, мобильность и модифицируемость) комплексного показателя «пригодность», характеризующего, в том числе, уровень соответствия обучающего и тестового НД планируемым условиям эксплуатации ИСИБ.

В стандарте ГОСТ Р ИСО/МЭК 9126-93 отмечается, что такие атрибуты программного обеспечения, как эффективность, надёжность, практичность и сопровождаемость должны определяться в конкретных условиях эксплуатации, а возможность использования ПО в различных условиях характеризуется такими атрибутами, как изменяемость и адаптируемость. Однако в данном документе также отсутствуют правила формализации условий эксплуатации ПО и атрибуты, в явном виде характеризующие уровень соответствия наборов данных, используемых при создании и тестировании систем, этим условиям.

Обеспечение конфиденциальности данных при создании и применении ИСИБ

Обучающие и тестовые НД для интеллектуальных систем ИБ могут содержать конфиденциальную информацию, несанкционированное распространение которой способно привести к следующим негативным последствиям:

- наборы данных могут быть использованы злоумышленниками для повышения эффективности реализации угроз ИБ в дальнейшем;

19 ГОСТ Р 52633.5-2011 Защита информации. Техника защиты информации. Автоматическое обучение нейросетевых преобразователей биометрия-код доступа.

20 ГОСТ Р ИСО/МЭК 9126-93 Информационная технология. Оценка программной продукции. Характеристики качества и руководства по их применению.

21 ISO/IEC 22989:2021(X) Information Technology — Artificial Intelligence — Artificial Intelligence Concepts and Terminology.

- наборы данных могут быть использованы злоумышленниками для оценки эффективности (в том числе – уровней бескомпроматности и безуликовости) ранее реализованных угроз ИБ;
- может быть нанесён репутационный и экономический ущерб организациям, являющимся разработчиками, владельцами и эксплуатантами автоматизированных систем, сведения о реализованных угрозах ИБ в отношении которых содержатся в соответствующих наборах данных.

Особенностью интеллектуальных систем ИБ является возможность бесконтрольного повышения уровня конфиденциальности данных, накапливаемых в системе в процессе её функционирования и используемых, в частности, для дообучения системы. Так, при получении из публичных источников отрывочных данных об уязвимостях информационных систем и способах реализации компьютерных атак, в процессе обобщения и разметки (сопровождения метаданными) этих данных в определённый момент могут сформироваться сведения, содержащие коммерческую или иную тайну.

При такой постановке задачи априорные сведения об уровнях конфиденциальности отдельных информационных объектов, агрегированных в определённой комбинации, отсутствуют или считается, что объекты не являются конфиденциальными, независимо от их обобщения, что не позволяет использовать мандатную модель управления безопасностью Low-Watermark (LWM) и иные модели конечных состояний [13]. Для оценки уровня конфиденциальности агрегированных данных может быть использован риск-ориентированный подход [14, 15], основанный на определении величины потенциального ущерба, который может наступить при разглашении тех или иных сведений.

Соответственно для ИСИБ должны быть установлены унифицированные процедуры автоматизированного контроля уровня конфиденциальности накапливаемых данных с уведомлением оператора системы о критическом повышении этого уровня и необходимости принятия дополнительных организационно-технических мер по защите обрабатываемой информации. Отметим, что в рамках существующей парадигмы нормативно-технического регулирования в области ИБ такие возможности не предусмотрены, что сдерживает применение технологий ИИ не только в области информационной безопасности, но и при решении других прикладных задач искусственного интеллекта.

Унификация и повышение доступности наборов данных, необходимых для создания и оценки соответствия ИСИБ

Наличие наборов данных, содержащих описание представительной совокупности инцидентов информационной безопасности, является обязательным условием создания ИСИБ. Важным инструментом формирования таких НД являются международные и национальные базы данных (БД), содержащие сведения о выявленных уязвимостях ПО и способах реализации компьютерных атак. Так, например, БД может содержать описание вариантов реализации полиморфной атаки, примеры вредоносного кода и т.п. При этом необходимо, чтобы содержащаяся в этих БД информация удовлетворяла двум конфликтующим требованиям: во-первых, была достаточно полной для разработки (обучения) и тестирования ИСИБ; во-вторых, не приводила к нарушению конфиденциальности чувствительной информации.

В Российской Федерации накопление сведений об уязвимостях информационных систем осуществляется в соответствии с приказом ФСТЭК России[22] и специальным Регламентом[23], определяющим порядок взаимодействия ФАУ «ГНИИИ ПТЗИ ФСТЭК России», обеспечивающего функционирование банка данных угроз безопасности информации, с разработчиками и производителями программного обеспечения и программно-аппаратных средств, с организациями и специалистами, которые выявляют (обнаруживают) уязвимости программного обеспечения и программно-аппаратных средств, при включении информации об уязвимостях программного обеспечения и программно-аппаратных средств в банк данных угроз безопасности информации ФСТЭК России. Упомянутые Положение и Регламент создавались без учёта требований, связанных с необходимостью обучения и тестирования интеллектуальных (в частности – нейросетевых) алгоритмов, должны быть пересмотрены и в случае необходимости – скорректированы.

На международном уровне описание уязвимостей осуществляется с использованием формата CVE (Common Vulnerabilities and Exposures), который в настоящее время практически является общепринятым стандартом. Описания выявленных уязвимостей в формате CVE представлены, например, на сайте базы данных уязвимостей NVD (National Vulnerability Database), поддерживаемой Национальным институтом технологий и стандартов США (NIST).

Обобщённые сведения о базах данных уязвимостей представлены в табл.3. Отметим, что зарубежные информационные ресурсы, как и отечественный банк данных угроз безопасности информации, создавались без учёта возможности их использования для обучения и тестирования алгоритмов ИИ. Это обстоятельство может потребовать пересмотра нормативно-технических документов, устанавливающих требования к порядку ведения этих БД.

Для предотвращения возможности использования сведений, содержащихся в тестовых и обучающих НД, в целях повышения эффективности реализации угроз ИБ и недобросовестной конкуренции режим конфиденциальности в той или иной степени должен распространяться на:

- способы эксплуатации выявленных угроз ИБ;
- наименование, географическое месторасположение, ведомственную и отраслевую принадлежность организаций, в автоматизированных системах которых были выявлены уязвимости ИБ;
- размер ущерба, причинённого организациям в результате реализации угроз ИБ;
- при необходимости – другие сведения.

22 Положение о банке данных угроз безопасности информации. Утверждено приказом ФСТЭК России от 16 февраля 2015 г. № 9 (зарегистрирован Минюстом России 17 апреля 2015 г., рег. № 36901).

23 Методический документ ФСТЭК России. Регламент включения информации об уязвимостях программного обеспечения и программно-аппаратных средств в банк данных угроз безопасности информации ФСТЭК России. 2015.

Таблица 3

Базы данных о выявленных уязвимостях информационных систем

Наименование БД	Ссылка
Банк данных угроз безопасности информации ФСТЭК России	http://www.bdu.fstec.ru
Национальная база уязвимостей США, NVD (CVE)	http://nvd.nist.gov
	http://cve.mitre.org
База уязвимостей компании Secunia	http://secunia.com
База уязвимостей от IBM ISS (X -Force)	http://xforce.iss.net
База уязвимостей компании Security Focus	http://www.securityfocus.com
Открытая база уязвимостей, OSV DB «Open Source Vulnerability Database»	http://osvdb.org
База эксплойтов	http://www.exploit-db.com
Metasploit, проект, посвященный созданию средств тестирования на проникновение (эксплойтов)	http://www.metasploit.com
Сайт компании Digital Bond, занимающейся безопасностью промышленных систем	http://www.digitalbond.com
Группа реагирования на инциденты в области промышленных систем ICS -CERT «Industrial Control Systems Cyber Emergency Response Team»	http://ics-cert.us-cert.gov
Сайт исследователя Luigi Auriemma	http://aluigi.org
	http://aluigi.altervista.org
Информационный портал, посвященный безопасности SCADA-систем	http://scadahacker.com
Информационный портал фирмы Positive Technologies	http://www.securitylab.ru
Архив форумов (почтовых рассылок, «Mailing List»), посвященных ИБ	http://seclists.org
Информационный портал, посвященный вопросам ИБ, содержащий базу уязвимостей	http://www.securelist.com

Выполнение требований к обеспечению конфиденциальности тестовых и обучающих наборов данных может быть обеспечено за счёт подготовки соответствующего нормативного правового документа, определяющего организационные мероприятия по обеспечению конфиденциальности, и разработки нормативно-технических документов, устанавливающих требования к средствам гарантированной деклассификации (калька от англ. “declassification”, рассекречивание, понижение уровня конфиденциальности до приемлемого) данных.

Управление дообучением ИСИБ на стадии эксплуатации

Возможность совершенствования (дообучения) в процессе эксплуатации является важной особенностью интеллектуальных систем обработки данных. В то же время, данная особенность практически не учтена в действующих нормативно-технических документах, что существенно затрудняет использование дообучения в процессе практического применения ИСИБ. Устранению данного недостатка будет способствовать решение следующих задач:

- включение в состав показателей качества СЗИ, предусмотренных ГОСТ Р 52447-2005[24], показателей, характеризующих возможность дообучения средств ИБ на стадии эксплуатации;
- определение в ГОСТ Р ИСО/МЭК 9126-93[25] требований к процедуре повторной оценки качества ПС при дообучении системы ИИ на стадии эксплуатации.

Выводы

Таким образом, в работе были сформулированы основные задачи нормативно-технического регулирования в области интеллектуальных систем информационной безопасности. Полнота данного перечня задач обеспечивается комплексностью рассмотрения процессов, реализуемых на стадиях жизненного цикла подобных систем, а отсутствие избыточности – направленностью задач нормативно-технического регулирования на преодоление конкретных нормативных барьеров, препятствующих созданию и применению интеллектуальных систем ИБ. Разработка и применение соответствующих нормативно-технических документов обеспечит повышение эффективности решения задач информационной безопасности за счёт использования в перспективных системах ИБ технологий искусственного интеллекта.

24 ГОСТ Р 52447-2005 Защита информации. Техника защиты информации. Номенклатура показателей качества.

25 ГОСТ Р ИСО/МЭК 9126-93 Информационная технология. Оценка программной продукции. Характеристики качества и руководства по их применению.

Литература

1. Гарбук С.В., Губинский А.М. Искусственный интеллект в ведущих странах мира: стратегии развития и применение в сфере обороны и безопасности. М.: «Знание». 2020. 590 с.
2. Червяков Н.И., Евдокимов А.А., Галушкин А.И. и др. Применение искусственных нейронных сетей и системы остаточных классов в криптографии. – М.: ФИЗМАТЛИТ, 2012. – 280 с.
3. Гарбук С.В., Гриняев С.Н., Правиков Д.Ю., Полянский А.В. Обеспечение информационной безопасности АСУ ТП с использованием метода предиктивной защиты // Вопросы кибербезопасности. 2019. №2(30). С.38-45. DOI: 10.21681/2311-3456-2019-2-38-45.
4. Соловьёв В.С., Язов Ю.К. Информационное обеспечение деятельности по технической защите информации. Вопросы кибербезопасности. 2021. №1(41). С.69-79. DOI: 10.21681/2311-3456-2021-1-69-79.
5. Гарбук С.В. Интеллектуальные автоматизированные средства тематической обработки информации в системах безопасности // Искусственный интеллект и принятие решений. 2017. №1. С.95-104.
6. Haier, Richard J. The neuroscience of intelligence / Richard J. Haier, University of California, Irvine. New York, NY: Cambridge University Press, 2017.
7. Francois Chollet – On the Measure of Intelligence. arXiv:1911.01547v2 [cs.AI] 25 Nov 2019.
8. Matteo E. Bonfanti, Kevin Kohler. Artificial Intelligence for Cybersecurity/CSS Analyses in Security Policy. №. 265, June 2020.
9. Kostogryzov A., Korolev V. Probabilistic Methods for Cognitive Solving of Some Problems in Artificial Intelligence Systems. 2019. DOI: 10.5772/intechopen.89168
10. Garbuk S.V. Intellimetry as a way to ensure AI trustworthiness//The Proceedings of the 2018 International Conference on Artificial Intelligence Applications and Innovations (IC-AIAI). Limassol, Cyprus, 6-10.10.2018. C.27-30. DOI 10.1109/IC-AIAI.2018.00012.
11. Кузьменко А.Н. Технология полиморфной модификации программного кода для повышения надежности функционирования программного обеспечения в недоверенной среде // Вопросы безопасности. 2018. № 4. С.64-77.
12. Гарбук С.В., Бакеев Р.Н. Конкурентная оценка качества технологий интеллектуальной обработки данных // Проблемы управления. 2017, №6. С.50-62.
13. Михайлов Ю. Б. Научно-методические основы обеспечения безопасности защищаемых объектов. Горячая линия-Телеком. 2015. 322 с. ISBN: 978-5-9912-0485-9
14. Парьев С.Е., Правиков Д.И., Карантаев В.Г. Особенности применения риск-ориентированного подхода для обеспечения кибербезопасности промышленных объектов // Безопасность информационных технологий. Том 27, №4(2020). DOI: http://dx.doi.org/10.26583/bit.2020.4.04.
15. Ritsuko Kawasaki (Aiba), Takeshi Hiromatsu – Proposal of a Model Supporting Decision-Making on Information Security Risk Treatment // World Academy of Science, Engineering and Technology. International Journal of Economics and Management Engineering Vol:8, No:4, 2014. C.583-589.

TASKS OF TECHNICAL REGULATION OF INTELLIGENT INFORMATION SECURITY SYSTEMS

Garbuk S.V.[26]

Research aim. *Improving the efficiency of solving information security tasks by eliminating standard technical barriers that prevent the application of artificial intelligence technologies in advanced information security systems.*

Research method. *The article applies the method of functional decomposition of intelligent tasks of information security, based on the analogy of artificial and natural intelligence. With respect to the proposed functional structure, the intelligent information security system is decomposed according to the processes of its life cycle with the specific tasks of technical regulation identification, that is specific to each of the processes, and the subsequent aggregation of tasks into groups corresponding to the main areas of standardization of such systems is performed.*

Results obtained. *The research presents a structured list of information security tasks, the solution quality of which can be improved with the use of artificial intelligence technologies. It is shown that the main standard technical barriers to the effective creation and application of intelligent information security systems are associated with the shortcomings of metrological support for intelligent systems, also with the peculiarities of ensuring the confidentiality of information processed in such systems. The analysis of the current state of work on the preparation of national and international standards governing the creation and application of intelligent information security systems is carried out, and it is indicated that the work in this direction is of an initial, staged nature. The list of specific standardization tasks aimed at overcoming the identified standard technical barriers in the implementation of individual processes of*

26 Sergey Garbuk, Ph.D. (tech.), Senior research fellow, Director of research projects, National Research University «Higher School of Economics», chairman of TC 164 «Artificial intelligence», Moscow, Russia. E-mail: garbuk@list.ru.

the intelligent systems life cycle is justified. Specific tasks are grouped by the main standardization areas, for each of which the proposals for the adjustment of existing and the development of new standard technical documents in the field of artificial intelligence and information security are prepared.

Keywords: *artificial intelligence, applied tasks of artificial intelligence, intelligent tasks of information security, system life cycle, evaluation of the functional characteristics of intelligent systems, intellometry, quality of intelligent systems, information security of intelligent systems.*

References

1. Garbuk S.V., Gubinskii` A.M. Iskusstvenny`i` intellekt v vedushchikh stranakh mira: strategii razvitiia i primenenie v sfere oborony` i bezopasnosti. M.: «Znanie». 2020. 590 s.
2. Cherviakov N.I., Evdokimov A.A., Galushkin A.I. i dr. Primenenie iskusstvenny`kh nei`ronny`kh setei` i sistemy` ostatochny`kh classov v kriptografii. – M.: FIZMATLIT, 2012. – 280 s.
3. Garbuk S.V., Greeniaev S.N., Pravikov D.Iu., Polianskii` A.V. Obespechenie informatcionnoi` bezopasnosti ASU TP s ispol`zovaniem metoda prediktivnoi` zashchity` // Voprosy` kiberbezopasnosti. 2019. №2(30). S.38-45. DOI: 10.21681/2311-3456-2019-2-38-45.
4. Solov`yov V.S., Iazov Iu.K. Informatcionnoe obespechenie deiatel`nosti po tekhnicheskoi` zashchite informatcii. Voprosy` kiberbezopasnosti. 2021. №1(41). S.69-79. DOI: 10.21681/2311-3456-2021-1-69-79.
5. Garbuk S.V. Intellektual`ny`e avtomatizirovanny`e sredstva tematicheskoi` obrabotki informatcii v sistemakh bezopasnosti // Iskusstvenny`i` intellekt i priniatie reshenii`. 2017. №1. S.95-104.
6. Haier, Richard J. The neuroscience of intelligence / Richard J. Haier, University of California, Irvine. New York, NY: Cambridge University Press, 2017.
7. Francois Chollet – On the Measure of Intelligence. arXiv:1911.01547v2 [cs.AI] 25 Nov 2019.
8. Matteo E. Bonfanti, Kevin Kohler. Artificial Intelligence for Cybersecurity/CSS Analyses in Security Policy. №. 265, June 2020.
9. Kostogryzov A., Korolev V. Probabilistic Methods for Cognitive Solving of Some Problems in Artificial Intelligence Systems. 2019. DOI: 10.5772/intechopen.89168
10. Garbuk S.V. Intellimetry as a way to ensure AI trustworthiness//The Proceedings of the 2018 International Conference on Artificial Intelligence Applications and Innovations (IC-AIAI). Limassol, Cyprus, 6-10.10.2018. S.27-30. DOI 10.1109/IC-AIAI.2018.00012.
11. Kuz`menko A.N. Tekhnologiia polimorfnoi` modifikatcii programmnogo koda dlia povy`sheniia nadezhnosti funktcionirovaniia programmnogo obespecheniia v nedoverennoi` srede // Voprosy` bezopasnosti. 2018. № 4. S.64-77.
12. Garbuk S.V., Bakeev R.N. Konkurentnaia ocenka kachestva tekhnologii` intellektual`noi` obrabotki danny`kh // Problemy` upravleniia. 2017, №6. S.50-62.
13. Mihai`lov Iu. B. Nauchno-metodicheskie osnovy` obespecheniia bezopasnosti zashchishchaemy`kh ob``ektov. Goriachaia liniia-Telekom. 2015. 322 s. ISBN: 978-5-9912-0485-9
14. Par`ev S.E., Pravikov D.I., Karantaev V.G. Osobennosti primeneniia risk-orientirovannogo podhoda dlia obespecheniia kiberbezopasnosti promy`shlenny`kh ob``ektov // Bezopasnost` informatcionny`kh tekhnologii`. Tom 27, №4(2020). DOI: http://dx.doi.org/10.26583/bit.2020.4.04.
15. Ritsuko Kawasaki (Aiba), Takeshi Hiromatsu – Proposal of a Model Supporting Decision-Making on Information Security Risk Treatment // World Academy of Science, Engineering and Technology. International Journal of Economics and Management Engineering Vol:8, No:4, 2014. C.583-589.