УДК [006.05+006.07]:004.08

МЕТОД ОЦЕНКИ ВЛИЯНИЯ ПАРАМЕТРОВ СТАНДАРТИЗАЦИИ НА ЭФФЕКТИВНОСТЬ СОЗДАНИЯ И ПРИМЕНЕНИЯ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Гарбук С.В., канд. техн. наук, ст. науч. сотр., директор по научным проектам НИУ «Высшая школа экономики», председатель ТК 164 «Искусственный интеллект»

В статье предложены метод и математическая модель, позволяющие количественно оценивать влияние документов по стандартизации на эффективность управления процессами жизненного цикла систем искусственного интеллекта (ИИ).

В основу разработанного метода положен принцип функциональной декомпозиции жизненного цикла (ЖЦ) информационной системы на типовые процессы с выявлением факторов, оказывающих существенное влияние на реализацию этих процессов, специфичных для систем ИИ и доступных для управления с помощью соответствующих нормативно-технических документов. Показано, что для каждого из выявленных факторов может быть определен документ, устанавливающий требования по компенсации возможного негативного влияния данного фактора на ЖЦ системы ИИ, причем вся совокупность документов может быть разбита на несколько (в рассматриваемом случае – пять) типовых групп, для каждой из которых, в свою очередь, может быть обоснован унифицированный количественный показатель, определяющий эффективность управления процессами ЖЦ системы в зависимости от полноты и качества соответствующих нормативно-технических документов. Рассчитанные таким образом унифицированные показатели могут быть использованы для оценки вероятностей типовых отказов, свидетельствующих о существенных нарушениях в реализации процессов создания и применения системы ИИ.

Полученные результаты могут быть использованы для оценки эффективности нормативно-технического регулирования в области искусственного интеллекта, а также для решения обратной задачи – обоснования требований к структуре и составу документов по стандартизации, исходя из установленных требований к эффективности реализации процессов жизненного цикла систем ИИ.

Ключевые слова: искусственный интеллект, процессы жизненного цикла систем искусственного интеллекта, показатели качества систем искусственного интеллекта, функциональная надежность систем искусственного интеллекта, эффективность стандартизации, стандарты искусственного интеллекта.

ВВЕДЕНИЕ

Ожидается, что технологии искусственного интеллекта обеспечат автоматизированное решение некоторых сложных задач, которые ранее могли быть успешно решены исключительно человеком, обладающим определенными интеллектуальными способностями: задачи распознавания образов, принятия решений в непредвиденных условиях, извлечение знаний из больших данных и некоторые другие [1]. Подобные интеллектуальные технологии обработки данных все чаще находят практическое применение при проектировании и математическом моделировании информационных систем и процессов [2], при производстве и применении сложных технических объектов. Системы ИИ [3] находят применение в ходе развития сетевых технологий, обеспечивающих надежную передачу потоков данных от сенсоров к средствам обработки и хранения данных, в таких областях, как государственное управление, транспорт, здравоохранение, обеспечение безопасности [4], в том числе и информационной [5].

Вопросы необходимости стандартизации систем ИИ в условиях цифровой экономики приходится решать в реальном времени, когда существует потребность как в реализации методов переработки данных, так в разработке стандартов, связанных аппаратно-программными средствами в области ИИ [6], что определено Перспективной программной стандартизации «Искусственный интеллект» на 2021–2024 годы [7].

Оценка эффективности стандартизации применительно к конкретным отраслям экономики, социальной сферы и технологическим направлениям (т.н. «сквозным» технологиям), всегда являлась и является важным направлением управления качеством [8, 9]. Модели оценки эффективности разрабатываются с учетом специфики процессов жизненного цикла (ЖЦ) объектов стандартизации, принятых в конкретных отраслях (см., например, [10]). В настоящей работе предложены метод и математическая модель, которые могут быть использованы для оценки влияния параметров комплекса стандартов на показатели качества процессов создания и применения прикладных систем искусственного интеллекта (СИИ).

ФОРМИРОВАНИЕ МАТЕМАТИЧЕСКОЙ МОДЕЛИ

При формировании и поддержании в актуальном состоянии комплекса отраслевых документов по стандартизации возникает задача обоснования оптимальной структуры и состава документов. Общими требованиями к комплексу отраслевых нормативно-технических документов являются [11]:

1. *Полнота* – стандарты должны обеспечивать регулирование основополагающих вопросов в заданной области, а также снятие основных существующих нормативно-технических барьеров.
2. *Безизбыточность* – стандарты должны концентрироваться на специальных вопросах создания и применения технологий искусственного интеллекта в рассматриваемой отрасли, минимально затрагивая другие аспекты стандартизации, не дублируя и не вступая в противоречия с существующими нормативно-техническими и нормативными правовыми документами.
3. *Непротиворечивость* – обеспечение совместимости с национальными и международными стандартами.
4. *Востребованность* – обеспечение подтверждаемого применения стандартов для решения практически значимых задач.

С учетом вышеперечисленных требований задача оптимизации комплекса стандартов может быть сформулирована в виде задачи линейного программирования, в которой в качестве целевой функции могут быть выбраны:

а) значения показателей качества Q типовых процессов, услуг и продукции [12], подлежащих стандартизации и определяющих целенаправленную деятельность в отрасли, как максимизация результативности стандартизации (требование 1 – полноты покрытия стандартами);
б) затраты С на разработку новых и актуализацию существующих документов по стандартизации ИИ (минимизация расходов на стандартизацию – требование 2);
в) отношение значений показателей качества к ресурсам, необходимым для выполнения работ по стандартизации Q/C (максимизация эффективности стандартизации).

При этом для различных целевых функций ограничения накладываются на:

а) максимальные расходы на стандартизацию $C \le C_{max}$; максимальное расхождение комплекса стандартов с существующей нормативной базой в области ИТ $\Delta \le \Delta_{max}$ (требование 3); минимальные значения прогнозируемого уровня востребованности документов $D \ge D_{min}$, под которыми могут пониматься, например, количество заинтересованных потребителей стандартов, количество вариантов прикладного использования стандартов, частота использования стандартов и другие показатели (требование 4);
б) минимальные значения показателей качества объектов стандартизации $Q \ge Q_{min}$; максимальное расхождение комплекса стандартов $\Delta \le \Delta_{max}$; минимальные значения показателей востребованности $D \ge D_{min}$;
в) максимальное расхождение комплекса стандартов $\Delta \le \Delta_{max}$; минимальные значения показателей востребованности $D \ge D_{min}$.

Оптимизации во всех перечисленных выше случаях подлежат такие параметры комплекса стандартов S, как перечень и параметры объектов стандартизации, количество и вид документов по стандартизации. Таким образом, для наиболее распространенного случая синтеза оптимального комплекса отраслевых стандартов, обеспечивающего максимизацию качества объектов стандартизации при наличии ресурсных и иных ограничений, задача оптимизации будет иметь вид (случай целевой функции «а»):

$$S_{opt} = argmax\,[Q(S) | C \le C_{max}, \Delta \le \Delta_{max}, D \ge D_{min}]. \quad (1)$$

При реализации такого подхода прямой задачей является нахождение зависимости значений показателей качества объектов стандартизации от параметров комплекса стандартов $Q(S)$.

Под экономическим эффектом стандартизации [8], понимают выраженную в денежных или натуральных показателях экономию живого и овеществленного труда в общественном производстве в результате внедрения стандарта с учетом необходимых затрат. При этом предусматривается возможность оценки экономического эффекта:

- на всех стадиях ЖЦ продукции, включая проектирование, производство и эксплуатацию;
- для различных видов стандартизации (единичный стандарт, комплексные программы стандартизации);
- для различных объектов стандартизации (продукция, предприятие, отрасль в целом);
- для различных видов стандартов (типы и основные параметры продукции; технические требования, правила

эксплуатации и ремонта; методы контроля и правила приемки; типовые технологические процессы и др.).

В существующих методиках оценки показателей качества $Q(S)$ реализуемых процессов и создаваемой продукции от параметров применяемого комплекса стандартов не рассматриваются. Считается, что значения этих показателей определяются либо набором внешних факторов, либо оценивается эффективность по результатам стандартизации [8, 10], что обеспечивает высокую универсальность подхода, но ограничивает его прямое применение для оценки эффективности стандартизации в конкретной отрасли.

При разработке модели учитывалось, что СИИ обладают следующими особенностями, выделяющими их среди других систем обработки данных [1, 7]:

- разработка СИИ предполагает обязательный этап обучения на прецедентах (обучающих наборах данных);
- интеллектуальные алгоритмы обработки информации СИИ могут принципиально не обладать свойствами интерпретируемости, объяснимости процесса вычислений и получаемых результатов;
- при аппаратно-программной реализации СИИ используются, как правило, специальные программные компоненты (программные библиотеки для машинного обучения и др.) и вычислительные средства (векторные, тензорные, нейроморфные процессоры и др.);
- значительная часть СИИ рассчитана на автоматизацию естественных интеллектуальных способностей человека;
- в СИИ предусматривается совершенствование (дообучение) алгоритмов на стадии применения системы;
- обработка данных в СИИ может приводить к росту уровня конфиденциальности обрабатываемых данных.

Для обеспечения адекватности модели ЖЦ СИИ использовался подход, ранее предложенный для анализа особенностей технического регулирования вопросов создания и применения интеллектуальных технологий в различных прикладных системах ИИ: системах информационной безопасности [5] и космических системах дистанционного зондирования Земли [13].

В основу этого подхода была положена функциональная декомпозиция жизненного цикла в соответствии с национальным стандартом ГОСТ Р 57193–2016 [14], обеспечивающая полное покрытие процессов ЖЦ с учетом многообразия вариантов реализации конкретных информационных систем. Универсальность такой декомпозиции достигается за счет того, что в стандарте [14] не подразумевается предписывающий порядок использования процессов ЖЦ и не учитывается взаимозависимость процессов. В результате подход, продемонстрированный в работах [5, 13] на различных задачах ИИ, позволил выявить исчерпывающий перечень факторов F_Σ, влияющих на эффективность реализации процессов ЖЦ СИИ.

Перечень F_Σ может быть разделен на функционально однородные группы F_1-F_9, типовое распределение которых по этапам ЖЦ СИИ представлено в табл. 1. Значения в ячейках таблицы соответствуют количеству факторов определенной группы на этапах ЖЦ, специфичных для СИИ: внешнее проектирование и выбор типовых решений (R), обучение (L1), тестирование при вводе в эксплуатацию (T1), эксплуатация (U), пробное дообучение на вновь поступающих данных (L2), повторное тестирование после пробного дообучения (T2) и модификация системы при успешном дообучении (M).

Отметим, что полученный перечень F_Σ ориентирован на решение задач анализа эффективности нормативно-технического регулирования процессов создания и применения СИИ, поэтому некоторые общие факторы качества процессов ЖЦ информационных систем, актуальные также и для СИИ, оставлены без рассмотрения:

- не учитывались требования, не учитывающие особенности используемых алгоритмов обработки данных (авторы статьи исходят из того, что в СИИ используются плохо интерпретируемые алгоритмы машинного обучения);
- в соответствии с общими принципами оценки экономической эффективности стандартизации [8] не рассматривались факторы, связанные полнотой и качеством основополагающих документов по стандартизации (термины и определения, классификации, обозначения);
- преимущественно рассматривались вопросы, связанные с управлением данными и разработкой алгоритмов ИИ, а факторам аппаратной реализации систем ИИ уделено меньше внимания, что не снижает, однако, общности предлагаемого методического подхода;
- не рассматривались факторы квалификации персонала, осуществляющего разработку и применение алгоритмов и систем ИИ, так как данные вопросы не относятся к области нормативно-технического регулирования.

Полученный перечень факторов F_Σ дает общее представление о направлениях стандартизации в области СИИ, но не позволяет полноценно анализировать зависимость $Q(S)$ показателей качества СИИ Q от параметров комплекса стандартов ИИ S, так как не учитывает вид стандартов, устанавливающих требования по компенсации негативного влияния соответствующих факторов, структуру взаимосвязей факторов

Таблица 1

Типовое распределение факторов, определяющих качество процессов ЖЦ и принадлежащих различным функциональным группам, по этапам ЖЦ СИИ

ГРУППА ФАКТОРОВ		КОЛИЧЕСТВО ФАКТОРОВ ПО ЭТАПАМ СИИ				
№	СОДЕРЖАНИЕ ГРУППЫ ФАКТОРОВ	R	L1, L2	T1, T2	U	M
F_1	Точность измерения функциональных характеристик и характеристик безопасности СИИ	8	12	11	9	1
F_2	Качество оценки функциональных возможностей квалифицированного человека-оператора, осуществляющего решение интеллектуальной задачи в ручном режиме	1	1	1	2	0
F_3	Полнота выявленных существенных условий эксплуатации СИИ	1	2	0	2	0
F_4	Наличие доверенных унифицированных аппаратно-программных средств, преимущественно – на отечественных компонентах	0	3	2	0	0
F_5	Возможности по масштабированию и тиражированию алгоритмов ИИ на смежные прикладные интеллектуальные задачи	2	2	0	0	0
F_6	Уровень унификации и качество наборов данных, используемых при создании и оценке качества СИИ. Надежность деклассификации данных для обеспечения безопасного доступа заинтересованных разработчиков	3	4	5	3	0
F_7	Уровень конфиденциальности данных при создании и применении СИИ	0	0	3	5	5
F_8	Уровень интерпретируемости и верифицируемости результатов работы СИИ эксплуатирующим персоналом	0	1	1	1	0
F_9	Эффективность дообучения СИИ на стадии эксплуатации	1	1	1	6	1
Всего (F_Σ)		16	26	24	28	7

и особенности их влияния на определенные показатели качества ЖЦ СИИ.

ФАКТОРЫ КАЧЕСТВА ЖИЗНЕННОГО ЦИКЛА СИИ

Для восстановления зависимости $Q(S)$ была разработана модель типового ЖЦ СИИ, направленная на преодоление этих недостатков (рис. 1). В соответствии с предложенной моделью, на первом этапе (R) формируется список функциональных характеристик и характеристик безопасности, значимых для решения конкретной прикладной задачи; устанавливаются значения (допустимые диапазоны значений) для этих характеристик; выбираются типовые программно-алгоритмические решения, ранее хорошо зарекомендовавшие себя при решении соответствующих задач; формируется список факторов внешней среды, существенным образом влияющих на сложность решения прикладной задачи ИИ (перечень существенных факторов эксплуатации); задаются допустимые диапазоны изменения значений этих факторов (предусмотренные условия эксплуатации) [1].

На этапе первоначального обучения (L1) подготавливаются обучающие наборы данных (НД), которые затем используются для синтеза алгоритма обработки данных, рассчитанного на решение конкретной прикладной задачи ИИ. Этап L1 повторяется до тех пор, пока результаты тестирования на этапе T1 не подтвердят соответствие функциональных характеристик системы требованиям, установленным на этапе R. Тестирование СИИ осуществляется на специальных НД, причем точность и достоверность получаемых оценок функциональных характеристик и характеристик безопасности системы ИИ определяется, прежде всего, представительностью этих тестовых НД. При отрицательных результатах тестирования осуществляется повторное обучение на доработанных (расширенных) обучающих НД. Возможности по расширению обучающих НД могут быть исчерпаны, а требования к системе ИИ не достигнуты. В этом случае осуществляется возврат на этап R с целью смягчения требований к характеристикам и/или условиям эксплуатации системы.

При выполнении требований к характеристикам СИИ на этапе тестирования T1 осуществляется переход к эксплуатации системы (этап U), которая может сопровождаться появлением новых информативных обучающих НД как сформированных самой системой ИИ, так и полученных

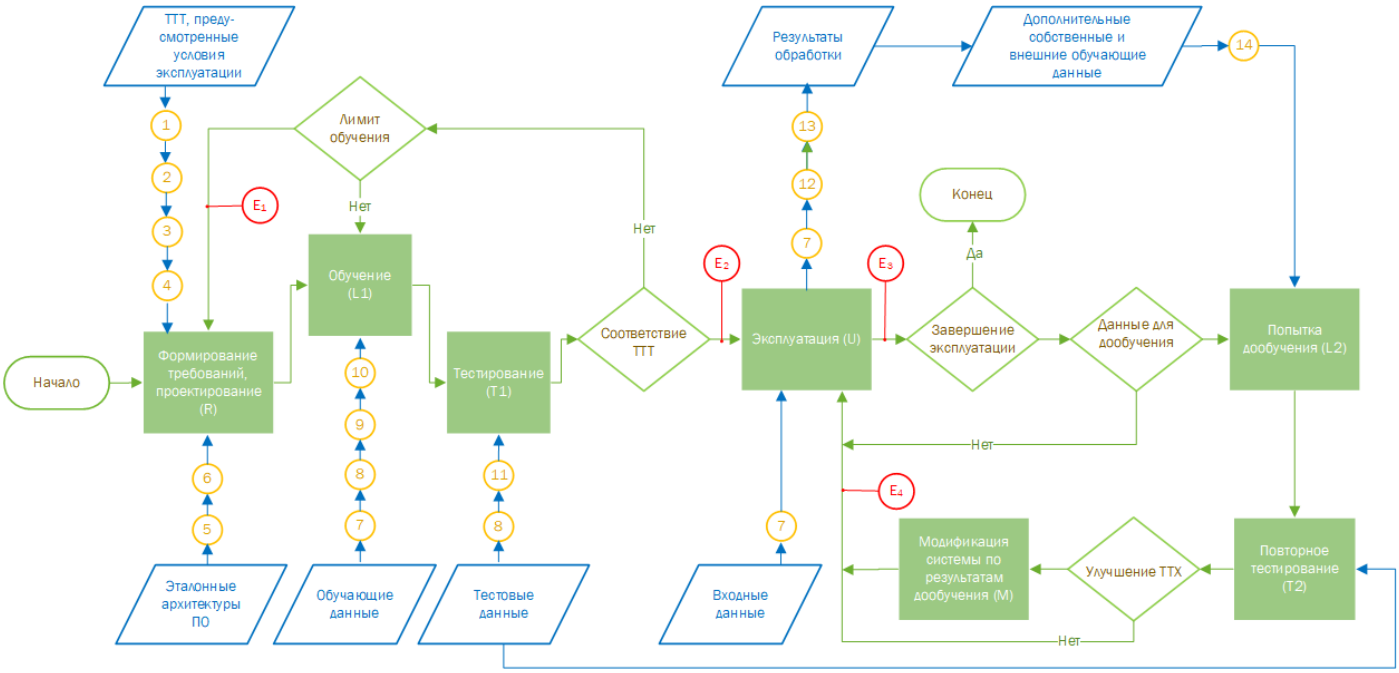

Рис. 1. Модель жизненного цикла СИИ

Цифрами обозначены номера требований, существенным образом влияющие на показатели качества процессов создания и применения СИИ; E_1, E_2, E_3, E_4 – точки локализации типовых отказов, определяющих значения показателей качества СИИ

из внешних источников. Эти данные используются для дообучения алгоритма обработки данных, выведенного из оперативного контура эксплуатации, и при подтверждении улучшения функциональных характеристик и характеристик безопасности этого алгоритма на этапе повторного тестирования (этап T2) дообученный алгоритм используется для модификации эксплуатируемой системы ИИ (этап M). Модификация системы может осуществляться также в ответ на необходимость эффективной защиты информации, уровень конфиденциальности которой существенно вырос в процесс ее накопления и обобщения.

В модели на рис.1 с учетом множества факторов F_Σ сформулированы требования к процессам ЖЦ СИИ, причем каждой группе факторов поставлена в соответствие одна или несколько групп требований, однородных по видам документов по стандартизации, устанавливающих соответствующие требования:

$$F_i \rightarrow \{(s_1^i, v_1^i), \dots (s_{K_i}^i, v_{K_i}^i)\},$$

где i – номер фактора, i =1…9; s_j^i –j-я группа требований, обеспечивающая учет -го фактора, $s_j^i \in \{S_1, \dots, S_{14}\}$; v_j^i – вид стандарта, устанавливающего требования группы s_j^i, $v_j^i \in \{V_1, \dots, V_5\}$; K_i – количество групп требований, обеспечивающих учет i-го фактора.

Описание сформированных таким образом групп требований к СИИ, оказывающих наибольшее влияние на качество процессов создания и применения систем, представлено в (табл. 2).

По видам стандартов требования к процессам ЖЦ СИИ были объединены в следующие пять групп:

- V_1 – стандарты, устанавливающие требования к процессам внешнего проектирования (обоснование тактико-технических требований и предусмотренных условий эксплуатации СИИ, обеспечение информационного сопряжения СИИ по входным и выходным данным);
- V_2 – квалиметрические стандарты, учитывающие особенности оценки функциональных характеристик и характеристик безопасности информационных систем на основе плохо интерпретируемых алгоритмов (т.н. «интеллометрические» стандарты [15]);
- V_3 – стандарты, устанавливающие единые подходы к оценке функциональных возможностей (компетенций) человека-оператора при решении типовых прикладных задач ИИ;
- V_4 – стандарты в области унификации данных и программного обеспечения;
- V_5 – стандарты в области защиты информации (преимущественно – обеспечения конфиденциальности).

В рамках предложенной модели могут быть оценены вероятности возникновения частных видов отказов, свиде-

Таблица 2

Взаимосвязь факторов качества F_i и требований S_j к процессам ЖЦ СИИ

F_i	ГРУППА ТРЕБОВАНИЙ			
	s_j^i	СОДЕРЖАНИЕ	ОБЪЕКТ	ВИД
F_1	S_1	Полнота набора существенных функциональных характеристик и характеристик безопасности системы ИИ и достаточность этого набора для оценки возможности применения систем ИИ по назначению	Тактико-технические требования (ТТТ), предусмотренные условия эксплуатации системы ИИ	V_1
F_2	S_2	Обоснованность критериальных значений и метрик для существенных характеристик (функциональных и безопасности) систем ИИ с учетом мирового уровня техники, требований заинтересованных лиц, характеристик источников информации и потребителей результатов обработки информации в СИИ или из других соображений		V_2
F_3	S_3	Качество оценки возможностей квалифицированного человека-оператора при решении соответствующих прикладных задач ИИ вручную		V_3
F_3	S_4	Полнота набора факторов внешней среды, учитываемых при оценке возможности применения системы ИИ по назначению, (существенных факторов эксплуатации) и достаточность этого набора для описания предусмотренных условий эксплуатации систем ИИ		V_1
F_4	S_5	Уровень унификации эталонных архитектур систем ИИ, наличие эталонных архитектур для решения типовых прикладных задач ИИ	Эталонные архитектуры ПО, используемого для разработки алгоритмов ИИ	V_4
F_5	S_6	Возможности по тиражированию типовых эталонных архитектур на смежные задачи ИИ		V_4
F_6	S_7	Уровень унификации форматов представления данных, используемых при создании системы ИИ и в процессе ее эксплуатации	Все данные	V_4
F_6	S_8	Уровень репрезентативности (полноты, несмещенности, точности и достоверности) обучающих и тестовых НД	Обучающие и тестовые наборы данных	V_2
F_6	S_9	Уровень возможностей по гарантированной деклассификации (включая анонимизацию) данных, обеспечивающей эффективный доступ к обучающим НД заинтересованных разработчиков систем ИИ	Обучающие наборы данных	V_5
F_7	S_{10}	Уровень конфиденциальности НД. Степень соответствия этого уровня требованиям по предотвращению возможности использования злоумышленниками сведений о НД для повышения эффективности реализации угроз ИБ в отношении системы ИИ		V_5
F_7	S_{11}	Уровень конфиденциальности тестовых НД, используемых для оценки соответствия систем ИИ предъявляемым требованиям. Степень соответствия этого уровня требованиям к достоверности тестирования (предотвращение переобучения алгоритмов ИИ)	Тестовые наборы данных	V_5
F_7	S_{12}	Качество автоматизированных процедур оценки уровня конфиденциальности данных, обрабатываемых в системе ИИ в процессе эксплуатации, с учетом возможности возрастания уровня конфиденциальности при накоплении и агрегировании данных	Результаты обработки	V_5
F_8	S_{13}	Качество интерпретации результатов работы системы ИИ, с учетом уровня объяснимости (понятности) используемых интеллектуальных алгоритмов обработки данных		V_1
F_9	S_{14}	Достоверность оценок информативности дополнительных обучающих данных (собственных, сформированных в процессе эксплуатации системы, и полученных из внешних источников) и целесообразности использования этих данных для дообучения системы ИИ в процессе эксплуатации		V_2

тельствующих о существенных нарушениях в реализации процессов создания и применения системы ИИ:

1) заданные потребителем тактико-технические требования (ТТТ) не достигнуты при разработке системы (вероятность отказа E_1, рис.1);
2) время достижения ТТТ при разработке системы превысило заданный лимит t_0 (E_2);
3) тактико-технические характеристики системы в процессе эксплуатации не соответствуют подтвержденным на этапе тестирования, то есть частота выхода значений функциональных характеристик за пределы

установленного диапазона превысила пороговое значение (E_3);
4) конфиденциальность обрабатываемых данных нарушается в процессе применения и утилизации системы ИИ (E_4).

Интегральная оценка качества процессов разработки и применения системы ИИ в этом случае может быть выполнена с помощью выражения:

$$Q = Q_{max} \frac{\sum_{i=1}^{4} a_i(1-E_i)}{\sum_{i=1}^{4} a_i}, \quad (2)$$

где a_i – коэффициент значимости i-го процесса ЖЦ системы ИИ; Q_{max} – максимальное значение показателя качества, определяемое информационными возможностями сенсоров, поставляющих информацию для системы ИИ, а также возможностями метасистемы (например, объекта управления), получающей информацию от системы [15].

Отметим, что перечень отказов E_1-E_4 не является исчерпывающим и при необходимости может быть дополнен для обеспечения сбалансированности интегральной оценки качества (2).

В соответствии с разработанным методом оценки влияния нормативно-технических документов на эффективность реализации процессов ЖЦ СИИ, значения вероятностей частных видов отказов могут быть вычислены с учетом структуры взаимосвязи факторов, влияющих на соответствующий отказ, и видов стандартов, устанавливающих требования по компенсации негативного влияния того или иного фактора. Так, вероятность недостижения СИИ заданных технических характеристик может быть рассчитана по формуле:

$$E_1 = 1 - P_R P_{L1} \quad (3)$$

где P_R – вероятность успешной реализации этапа формирования требований и проектирования СИИ; P_{L1} – вероятность успешной реализации этапа обучения СИИ.

Вероятность отказа, связанного с чрезмерной продолжительностью создания СИИ, имеет вид:

$$E_2 = 1 - P_R P_{L1}, P(t \le t_0) \quad (4)$$

где $P(t \le t_0)$ – вероятность того, что продолжительность создания СИИ не превышает заданный порог t_0.

Учитывая локализацию и вид требований к процессам ЖЦ СИИ (рис. 1), а также предполагая события, связанные с выполнением этих требований, статистически независимыми, вероятности, входящие в состав выражений (3) и (4), могут быть оценены по формулам:

$$P_R = V_1(S_1, S_2, S_4)\, V_3(S_3)\, V_4(S_5, S_6), \quad (5)$$

$$P_{L1} = V_2(S_8)\, V_4(S_7)\, V_5(S_9, S_{10}) \quad (6)$$

где $V_k(S_{J_1}, \ldots, S_{J_k})$ – коэффициент, зависящий от наличия и результативности применения стандартов k-го типа для обеспечения выполнения требований $S_{J_1}, \ldots, S_{J_k}$ (табл. 2) – унифицированные показатели эффективности нормативно-технического регулирования процессов ЖЦ СИИ.

Функциональная декомпозиция требований к процессам ЖЦ СИИ и их распределение по типовым группам используемых стандартов (табл. 2) позволяет сформулировать оценочные зависимости коэффициентов V_1-V_5, исходя из особенностей влияния частных факторов F_i на эффективность создания и применения СИИ. При этом предполагается, что коэффициенты могут принимать значения в диапазоне [0,1], причем минимальное значение коэффициента означает практическую невозможность успешной реализации соответствующего этапа ЖЦ, учитывая мультипликативный характер зависимостей (5) – (6). На практике это означает, что зависимости (5) – (6) позволяют получить нижние (пессимистические) оценки вероятностей успешной реализации соответствующих этапов ЖЦ.

Кроме того, значения S_1-S_{14}, характеризующие полноту выполнения соответствующей группы требований, считаются бинарными величинами: $S_j = 1$, если требования полностью выполняются, и $S_j = 0$ – в противном случае. В выражения для расчета коэффициентов V_k переменные S_j входят в виде аддитивных или мультипликативных составляющих в зависимости от критичности наличия стандартов соответствующей группы требований для предотвращения того или иного отказа. В результате для коэффициентов V_k с учетом табл. 2 могут быть получены следующие выражения:

$$V_1 = \beta_{m1} S_1 (\beta_{01} + \sum_{j \in \{2,4,13\}} \beta_j S_j), \quad (7)$$

$$V_2 = \beta_{m2} (\beta_{02} + \sum_{j \in \{8,14\}} \beta_j S_j), \quad (8)$$

$$V_3 = \beta_{m3} (\beta_{03} + S_3), \quad (9)$$

$$V_4 = \beta_{m4} (\beta_{04} + \sum_{j \in \{5,6,7\}} \beta_j S_j), \quad (10)$$

$$V_5 = \beta_{m5} S_{11} (\beta_{05} + \sum_{j \in \{9,10,12\}} \beta_j S_j), \quad (11)$$

где $0 < \beta_{0k} \ll 1$ – значение коэффициента V_k, соответствующее случаю отсутствия всех необязательных документов по стандартизации (аддитивные составляющие) и присутствия обязательных (мультипликативные составляющие); $0 \ll \beta_{mk} < 1$ – максимальное значение коэффициента V_k, соответствующее случаю наличия всех документов по стандартизации; β_j – весовые коэффициенты аддитивных составляющих, выбираемые экспертным путем с учетом требования нормировки (сумма β_j для каждого коэффициента V_k должна быть равна $1 - \beta_{0k}$).

Выражения (7) – (11) могут быть использованы для расчета вероятностей (5) – (6), причем при неполном наборе аргументов для того или иного коэффициента V_k в расчетных формулах (7) – (11) отсутствующие аргументы S_j принимаются равными 1.

Для оценки вероятности $P(t \le t_o)$ в формуле (4) в табл. 2 могут быть выделены требования, существенным образом влияющие на продолжительность создания СИИ на этапах R и $L1$. В частности, на время проектирования СИИ t_R оказывает существенное влияние уровень унификации ПО S_5 и наличие нормативно-технических документов, обеспечивающих перенос предобученных алгоритмов ИИ на другие прикладные задачи S_6:

$$t_R = t_{mR} + (1 - S_6)\,[t_{05}\,(1 - S_5) + t_{06}], \qquad (12)$$

где t_{mR} – минимальное время проектирования при высоком уровне унификации $(S_5 = 1)$ и нормативно-технической поддержке переноса алгоритмов ИИ на смежные прикладные задачи $(S_6 = 1)$; t_{05} – прирост времени проектирования из-за низкого уровня унификации $(S_5 = 0)$; t_{06} – прирост времени проектирования СИИ из-за невозможности переноса алгоритмов ИИ в условиях отсутствия необходимых нормативных документов $(S_6 = 0)$.

К возрастанию времени обучения системы t_{L1} приводят необоснованное завышение тактико-технических требований к СИИ (отсутствие документов, устанавливающих требования к критериальным значениям ТТХ для типовых прикладных задач ИИ – S_2, S_3), завышение количества и вариативности существенных условий эксплуатации СИИ (S_4), а также недостаточные возможности по доступу разработчиков к представительным наборам обучающих данных (S_7, S_8, S_9). Все вышеперечисленные требования непосредственно влияют на размер необходимого обучающего НД и на время обучения, причем для оценки времени t_{L1} могут быть приняты следующие допущения:

- при необоснованном завышении ТТТ время формирования НД и время обучения возрастают полиномиально;
- завышение количества существенных условий эксплуатации приводит к экспоненциальному росту t_{L1};
- сокращение возможностей по использованию ранее созданных НД и необходимость формировать новые – к линейному росту t_{L1}.

В итоге для времени t_{L1} может быть использована следующая верхняя оценка:

$$t_{L1} \le t_{mL1} + t_{0L1}\Big[(1 - S_2 S_3)\beta_{23}{}^2 + \\ + (1 - S_4) N_4^{\beta_4} + \sum_{j \in \{7,8,9\}} (1 - S_j)\beta_j\Big], \qquad (13)$$

где t_{mL1} – минимальное время, необходимое для обучения СИИ; t_{0L1} – нормирующий коэффициент, связывающий безразмерные составляющие от каждого негативного фактора с приростом времени обучения СИИ; β_{23} – коэффициент завышения ТТТ к СИИ (характерные значения: 1,1 – 2,0, зависимость времени обучения от ТТТ в данном примере принята квадратичной); N_4 – среднее количество значений, которое принимает каждый параметр, характеризующий существенные условия эксплуатации СИИ (характерные значения: 10–1000); β_4 – избыточно учтенное количество параметров, существенно влияющих на качество работы СИИ в реальных условиях эксплуатации; β_j – весовые коэффициенты аддитивных составляющих, связанных с влиянием нормативной базы на возможности по использованию ранее созданных НД.

Учитывая, что $t = t_R + t_{L1}$, выражения (12, 13) могут быть использованы для оценки вероятности выполнения требования по предельной продолжительности создания СИИ $P(t \le t_0)$ в (4). Из (13) видно также, что наибольший вклад в значение времени обучения t_{L1} вносит составляющая $N_4^{\beta_4}$, связанная с избыточным завышением размерности описания внешней среды, в которой функционирует СИИ (количеством и вариативностью существенных факторов эксплуатации СИИ). Следовательно, стандартизация перечней таких факторов для типовых прикладных задач является одной из приоритетных задач нормативно-технического регулирования ИИ.

В рамках предлагаемого метода, можно получить также следующее выражение для вероятности отказа, связанного с несоответствием реальных ТТХ СИИ характеристикам, полученным в процессе тестирования:

$$E_3 = 1 - V_2\,(S_8, S_{14})\,V_5\,(S_{10}, S_{11}), \qquad (14)$$

где V_2 и V_5 – коэффициенты, определяющие эффективность нормативно-технического регулирования, если в качестве показателя эффективности принята вероятность отказа E_3 (вычисляются по формулам (8) и (11), соответственно).

Вероятность отказа, связанного с нарушением конфиденциальности обрабатываемых данных в процессе эксплуатации СИИ, может быть оценена с использованием выражения:

$$E_4 = 1 - V_5\,(S_{10}, S_{12}). \qquad (15)$$

Отметим, что формула (15) учитывает исключительно факторы обеспечения конфиденциальности данных, специфичные для систем ИИ и связанные с процессами обучения и дообучения систем. Общие вопросы обеспечения конфиденциальности учитываются при расчете коэффициента β_{m5} по формуле (11).

ЗАКЛЮЧЕНИЕ

Предложенная математическая модель жизненного цикла системы ИИ отражает структуру взаимовлияния факторов, определяющих эффективность реализации процессов жизненного цикла, специфичных для систем ИИ, и используется для обоснования выражений для расчета унифицированных показателей нормативно-технического регулирования и вероятностей типовых отказов.

Частные показатели, рассчитанные с использованием выражений (3), (4), (14) и (15), могут быть использованы для оценки интегрального показателя (2), что, в свою очередь, обеспечивает возможность решения как прямой задачи оценки эффективности нормативно-технического регулирования процессов жизненного цикла систем ИИ, так и обратной задачи – обоснования требований к структуре и составу документов по стандартизации, исходя из установленных требований к эффективности реализации процессов жизненного цикла.

Список использованных источников и литературы

1. Гарбук С.В., Губинский А.М. Искусственный интеллект в ведущих странах мира: стратегии развития и военное применение. – М.: Знание, 2019. 590 с.
2. Бурый А.С., Шевкунов М.А. Суррогатное моделирование распределенных информационных систем по большим данным // Информационно-экономические аспекты стандартизации и технического регулирования. 2019. № 5 (51). С. 43–50.
3. Куприков Н.М., Башкирова Е.А. Стандартизация в сфере искусственного интеллекта // Информационно-экономические аспекты стандартизации и технического регулирования. 2021. № 5 (63). С. 8–13.
4. Бурый А.С., Усцелемов В.Н. Онтологический подход к формированию когнитивных моделей оценки кибербезопасности // Информационно-экономические аспекты стандартизации и технического регулирования. 2020. № 3 (55). С. 77–84.
5. Гарбук С.В. Задачи нормативно-технического регулирования интеллектуальных систем информационной безопасности // Вопросы кибербезопасности. 2021. № 3(43). С. 68–83. DOI: 10.21681/2311-3456-2021-3-68-83
6. Гарбук С. Интеллектуальные технологии вместо человека: оценка соответствия // Открытые системы. СУБД. 2018. № 2. С. 20.
7. Перспективная программа стандартизации в области приоритетного направления «Искусственный интеллект» на 2021–2024 годы. Утверждена 22.12.2020. – URL: https://www.tc164.ru/Национальная-стандартизация (дата обращения 15.04.2022).
8. Оценка экономической эффективности мероприятий по повышению качества продукции и услуг: учебное пособие / М.В. Галушко, С.В. Горбачев. – Оренбург: Оренбургский гос. ун-т, 2019. 102 с.
9. ГОСТ Р ИСО 9000–2015. Системы менеджмента качества. Основные положения. Словарь [Текст]. – Введ. 2015–11–01. – М.: Стандартинформ, 2015.
10. Черных Е.В., Иванова Г.Н. Стандартизация в устойчивом развитии предприятий. – М.: Издательство «Научный консультант», 2020. 178 с.
11. Гарбук С.В., Шалаев А.П. Перспективная структура национальных стандартов в области искусственного интеллекта // Стандарты и качество. 2021. № 10. С. 26–33. DOI: 10.35400/0038-9692-2021-10-26-33
12. ГОСТ 1.1–2002 Межгосударственная система стандартизации. Термины и определения [Текст]. – Введ. 2003–07–01. – М.: ИПК Изд-во стандартов, 2003.
13. Гарбук С.В. Задачи нормативно-технического регулирования интеллектуальных систем обработки данных дистанционного зондирования Земли // Современные проблемы дистанционного зондирования Земли из космоса. 2022. Т. 19. № 1. С. 107–122. DOI: 10.21046/2070-7401-2022-19-1-107-122
14. ГОСТ Р 57193–2016. Системная и программная инженерия. Процессы жизненного цикла систем [Текст]. – Введ. 2017–11–01. – М.: Стандартинформ, 2016.
15. Garbuk S.V. Intellimetry as a Way to Ensure AI Trustworthiness // Proceedings – 2018 International Conference on Artificial Intelligence: Applications and Innovations, IC-AIAI 2018, Nicosia, 31 октября – 02 ноября 2018 года. – Nicosia, 2019. P. 27–30. DOI: 10.1109/IC-AIAI.2018.8674447

METHOD FOR ASSESSING THE IMPACT OF STANDARDIZATION PARAMETERS ON THE EFFECTIVENESS OF THE CREATION AND APPLICATION OF ARTIFICIAL INTELLIGENCE SYSTEMS

Garbuk S.V., Candidate of Technical Sciences, Senior Researcher, Director of Research Projects, Higher School of Economics

This paper suggests a method and mathematical model that allow to quantify the effect of standardization documents on the performance in managing the life cycle processes for artificial intelligence (AI) systems.

The suggested method is based on the principle of functional decomposing the life cycle (LC) of an information system into standard processes in line with the standard ISO/IEC/IEEE 15288:2015 Systems and software engineering – System life cycle processes, with subsequent identification of factors that have a significant impact on the implementation of these processes which are specific to AI systems and available for management using relevant regulatory and technical documents.

We have shown that, for each of the factor identified, we can define a standard that sets requirements for mitigating the possible negative impact of this factor on the LC of an AI system, wherein the entire package of documents can be divided into several groups of typical standards, for each of which we can justify a unified quantitative indicator determining the performance of the AI system's LC process management depending on the completeness and quality of the relevant regulatory and technical documents. The paper proposes the following five groups of typical standards: requirements for external design processes; qualimetric standards that take into account the specifics of assessing the functional parameters and safety parameters of AI systems; standards that describe uniform approaches to assessing the functional capabilities (competencies) of a human operator when solving typical applied AI tasks; standards in the field of data and software unification; information security standards.

The unified indicators calculated in this way can be used to assess the probabilities of typical failures indicating significant violations in the processes of creating and using an AI system. As examples, the paper provides mathematical expressions for quantifying the probability of the following typical failures: the requirements specified by the customer were not achieved when developing the system; the time to achieve the requirements during the system development exceeded the set limit; the system characteristics during its operation differ significantly from the declared ones; the confidentiality of the data processed is impaired during the use and disposal of the AI system.

The results obtained can be used to assess the performance of technical regulations in the field of artificial intelligence, as well as to solve the inverse problem of justifying the requirements for the structure and composition of standardization documents based on the existing requirements for the performance of AI systems' life cycle processes.

Keywords: artificial intelligence, life cycle processes of artificial intelligence systems, quality indicators of artificial intelligence systems, functional reliability of artificial intelligence systems, standardization efficiency, artificial intelligence standards.

References

1. Garbuk S.V., Gubinskij A.M. Iskusstvennyj intellekt v vedushchih stranah mira: strategii razvitiya i voennoe primenenie. Moscow, Znanie Publ., 2020, 590 p.
2. Buryi A.S., Shevkunov M.A. Surrogatnoe modelirovanie raspredelennyh informacionnyh sistem po bol'shim dannym. Informacionno-ekonomicheskie aspekty standartizacii i tekhnicheskogo regulirovaniya, 2019, no 5 (51), pp. 43–50.
3. Kuprikov N.M., Bashkirova E.A. Standartizaciya v sfere iskusstvennogo intellekta. Informacionno-ekonomicheskie aspekty standartizacii i tekhnicheskogo regulirovaniya, 2021, no 5 (63), pp. 8–13.

4. Buryi A.S., Uscelemov V.N. Ontologicheskij podhod k formirovaniyu kognitivnyh modelej ocenki kiberbezopasnosti. Informacionno-ekonomicheskie aspekty standartizacii i tekhnicheskogo regulirovaniya, 2020, no 3(55), pp. 77–84.
5. Garbuk S.V. Zadachi normativno-tekhnicheskogo regulirovaniya intellektual'nyh sistem informacionnoj bezopasnosti. Voprosy kiberbezopasnosti, 2021, no. 3(43), pp. 68–83. DOI: 10.21681/2311-3456-2021-3-68-83
6. Garbuk S. Intellektual'nye tekhnologii vmesto cheloveka: ocenka sootvetstviya. Otkrytye sistemy. SUBD, 2018, no 2, p. 20.
7. Perspektivnaya programma standartizacii v oblasti prioritetnogo napravleniya «Iskusstvennyj intellekt» na 2021–2024. Utverzhdena 22 dekabrya 2020. Available at: https://www.tc164.ru/Nacional'naya-standartizaciya (accessed 14 March 2022).
8. Galushko M.V., Gorbachev S.V. Ocenka ekonomicheskoj effektivnosti meropriyatij po povysheniyu kachestva produkcii i uslug: uchebnoe posobie. Orenburg, Orenburg State University Publ., 2019, 102 p.
9. GOST R ISO 9000–2015 Quality management systems. Fundamentals and vocabulary. Moscow, Standartinform Publ., 2015. (In Russian)
10. Chernyh E.V., Ivanova G.N. Standartizaciya v ustojchivom razvitii predpriyatij. Moscow, Izdatel'stvo "Nauchnyj konsul'tant" Publ., 2020, 178 p.
11. Garbuk S.V., Shalaev A.P. Perspektivnaya struktura nacional'nyh standartov v oblasti iskusstvennogo intellekta. Standarty i kachestvo, 2021, no. 10, pp. 26–33.
12. GOST 1.1–2002 Interstate System for Standardization. Terms and definitions. Moscow, Izdatel'stvo standartov Publ., 2003. (In Russian)
13. Garbuk S.V. Zadachi normativno-tekhnicheskogo regulirovaniya intellektual'nyh sistem obrabotki dannyh distancionnogo zondirovaniya Zemli. Sovremennye problemy distancionnogo zondirovaniya Zemli iz kosmosa, 2022, vol. 19, no. 1, pp. 107–122.
14. GOST R 57193–2016 Systems and software engineering. System life cycle processes. Moscow, Standartinform Publ., 2016. (In Russian)
15. Garbuk S.V. Intellimetry as a Way to Ensure AI Trustworthiness. Proceedings – 2018 International Conference on Artificial Intelligence: Applications and Innovations, IC-AIAI 2018, Nicosia, October 31 – November 02, 2018. Nicosia, 2019, pp. 27–30. DOI: 10.1109/IC-AIAI.2018.8674447