ОБОСНОВАННОСТЬ ЭКСПЕРТНОЙ ОЦЕНКИ ПРОГРАММ И МЕТОДИК ИСПЫТАНИЙ

Шелестович О.А.[1], Свистунов А.В.[2]

Ключевые слова: *испытательная документация, требования к информационным системам, прослеживаемость требований, нормативные требования, эталонная проверка, достаточность проверки, верификация экспертного вывода, искусственный интеллект, большие языковые модели.*

Аннотация

Цель работы. *Определить правила обоснованной оценки ПМИ: когда предусмотренная проверка может быть признана достаточной, когда выявленный недостаток служит основанием для замечания и когда оценку следует отложить из-за неполноты исходных документов, в том числе если описание проверки распределено между несколькими документами.*

Метод исследования. *Исследование основано на декомпозиции процесса экспертной оценки ПМИ и формализации условий, связывающих применимое требование, проверяемый результат, эталонную проверку, документарное покрытие и экспертное решение. Сформулированные условия проверены на контрастных учебных случаях, моделирующих полное, распределенное и недостаточное описание проверки, а также неполноту исходных документов.*

Результаты исследования. *Сформулированы условия обоснованной оценки испытательной документации, связывающие применимое требование, самостоятельный проверяемый результат, эталонную проверку, содержание представленных документов и экспертное решение. Разграничены наличие в документации релевантного текста и достаточность предусмотренной проверки для подтверждения обязательного результата. Определены условия учета распределенного описания проверки в нескольких связанных документах и проверки предполагаемого недостатка до формирования замечания, включая поиск возможного альтернативного покрытия и повторную оценку исходного толкования требования. Показано, что неполнота исходных документов, неустановленная применимость требования или неразрешенное противоречие должны сохраняться как неопределенность и не преобразовываться в недостаток испытательной документации. Формализованные отношения используются как условия допустимости положительной оценки и проекта замечания, но не заменяют содержательного анализа применимости требования, достаточности процедуры и существенности выявленного недостатка. Предложенные положения имеют методологический характер и проиллюстрированы учебными случаями; их практическая результативность, межэкспертная согласованность и полнота выявления применимых источников требуют отдельной эмпирической проверки.*

DOI: 10.24412/1994-1404-2026-3-208-218

Введение и постановка задачи

Программа и методика испытаний связывает требования к информационной системе с действиями, по результатам которых предполагается установить их выполнение. При экспертизе необходимо выяснить, позволяют ли предусмотренные проверки получить и подтвердить обязательные результаты. Число сценариев, ссылок и форм протоколов само по себе на этот вопрос не отвечает.

Исходное требование может устанавливать отображение пользователю статуса успешной загрузки, тогда как ПМИ предусматривает лишь загрузку файла и результат «загрузка выполнена». Тематическая связь очевидна, но из нее не следует, что наблюдение пользовательского статуса включено в проверку.

Возможна и обратная ошибка: необходимое действие содержится в обязательной общей процедуре,

[1] ***Шелестович Олеся Александровна****, ведущий системный аналитик Управления проектирования и разработки информационных систем филиала публично-правовой компании «Роскадастр» Центр информационных технологий «Роскадастр-Инфотех», Москва, Россия.*

E-mail: ShelestovichOA@kadastr.ru

[2] ***Свистунов Андрей Васильевич****, кандидат технических наук, старший научный сотрудник, руководитель проектов Управления проектирования и разработки информационных систем филиала публично-правовой компании «Роскадастр» Центр информационных технологий «Роскадастр-Инфотех», Москва, Россия.*

E-mail: SvistunovAV@kadastr.ru

а способ фиксации — в связанной форме протокола. Отсутствие описания в самом сценарии тогда не доказывает недостатка. Требуется оценить совокупность документов и условия ее применения.

Для правовой информатики существенна связь оценки с нормативными и договорными основаниями. При переходе к проверяемому результату необходимо сохранять условия применимости, обязательные свойства и содержание отсылочных положений. Иначе точная ссылка может сопровождать неверное толкование требования.

Ближайшую методологическую основу составляют две работы, опубликованные в журнале «Правовая информатика». О.А. Шелестович рассматривает восстановление процессов деятельности из проверенного корпуса источников: отбор по применимости, выделение самостоятельных предписаний и экспертное подтверждение интерпретаций [1]. Для экспертизы ПМИ существенны сохранение происхождения требования и различение установленного основания и предварительной интерпретации. Единица анализа в этой работе — процесс деятельности; в настоящем исследовании — предусмотренная проверка самостоятельного результата.

А.В. Свистунов и Д.В. Митавский рассматривают сквозную прослеживаемость, связывающую предписание применимого источника с соответствующим процессом деятельности, решением о его автоматизации, требованием к информационной системе и подтверждением его выполнения. В этой модели уже разграничены процедура проверки, доказательство, документ-носитель и подтверждение соответствия; предусмотрены анализ изменений и аудит доказательного контура [2]. Эти положения принимаются как опубликованная основа. При оценке ПМИ необходимо установить, соответствует ли представленная Исполнителем испытательная документация требованиям контракта и иных применимых источников: предусмотрены ли все необходимые проверки, достаточны ли их содержание и порядок выполнения для подтверждения требуемых результатов и обосновано ли конкретное замечание при выявленном несоответствии. Наличие связанных объектов само по себе этого решения не определяет.

Смежные составляющие этой задачи рассматриваются и в других российских исследованиях. В.В. Грибова, Д.В. Жданов и Е.Е. Помников предлагают модель нормативных знаний, учитывающую контекст и зависимости положений в строительной области [3]. Эта работа дает основание рассматривать формализацию нормативного содержания и связей как уже развиваемое направление. Ее предмет, однако, отличается от оценки того, достаточно ли предусмотренных в ПМИ проверок для обоснованного экспертного заключения.

В исследовании И.М. Малиновского и соавторов обсуждаются управление требованиями, атрибуты источников, версии, согласование и матрицы соответствия в авиадвигателестроении [4]. Для рассматриваемой задачи существенно, что требование и подтверждающая информация могут быть связаны и организованы систематически. Наличие такой организации не снимает вопроса о содержательной достаточности конкретной проверки. Сравнение функций систем управления требованиями также не является проверкой эффективности предлагаемого здесь метода.

Ближе к предмету испытательной документации находятся работы, связывающие выявление требований с последующим тестированием и прослеживаемостью. Е.А. Герлиц, Д.С. Кильдишев и А.В. Хорошилов рассматривают выявление функциональных требований из документации программного интерфейса для задач функционального тестирования [5]. Б.А. Позин и Г.Н. Циперман рассматривают прослеживаемость требований как основу проектирования функционально-логической архитектуры программной системы [6]. Для настоящей задачи существенно сохранение связи между исходным требованием и объектами последующей проверки; сама прослеживаемость, однако, не определяет достаточность конкретной процедуры ПМИ.

К.А. Шишин, И.В. Муравьёв и Е.К. Куликов исследуют автоматическую генерацию модульных тестов с учетом контекста Java-приложений [7]. Эта работа подтверждает возможность автоматизации отдельных операций формирования тестов, но решает иную задачу: тест создается для программной реализации, тогда как в настоящем исследовании оценивается уже представленная испытательная документация относительно независимо определенного состава необходимых проверок.

O. Kosenkov и соавторы в систематическом обзоре требований к регуляторному соответствию программных систем рассматривают роль инженерии требований, связанные проблемы и практики, а также ее взаимодействие с другими процессами жизненного цикла [8]. Для экспертизы ПМИ это подтверждает необходимость связывать применимые внешние требования с объектами разработки и проверки, но не дает самостоятельного критерия достаточности испытательной процедуры.

A. Murugesan и соавторы исследуют автоматизированный семантический анализ assurance cases — структурированных аргументов, связывающих утверждения и свидетельства [9]. Для настоящей работы существенна сама идея проверки не только наличия отдельных элементов, но и связей между ними. Вместе с тем корректность исходного толкования применимого требования остается отдельной содержательной задачей.

R. Kamoi и соавторы в критическом обзоре самокоррекции языковых моделей показывают зависимость результата от типа обратной связи и отмечают более надежные результаты при наличии внешней обратной связи [10]. Это существенно при проверке предполагаемого недостатка с применением языковых моделей: повторный запрос к той же модели не должен трактоваться как независимое подтверждение исходного вывода.

В совокупности эти исследования охватывают отдельные составляющие рассматриваемой задачи — выявление и прослеживаемость требований, регуляторное со-

Таблица 1

Исследовательские направления и задача обоснования оценки ПМИ

Направление и источник	Результат, существенный для постановки задачи	Что необходимо дополнительно определить для оценки ПМИ
Установление требований по источникам [1, 3, 5]	Отбор применимого содержания, выделение предписаний и функциональных требований, учет контекстных связей	Как из установленного требования определить самостоятельный проверяемый результат и необходимую проверку
Прослеживаемость и управление требованиями [2, 4, 6]	Связи между источниками, требованиями, решениями и подтверждением; учет версий и изменений	Когда совокупность связанных фрагментов ПМИ содержательно достаточна для проверки
Автоматизация формирования тестов [7]	Генерация тестов с учетом программного контекста	Как отделить создание новой проверки от экспертизы уже представленной ПМИ
Инженерия регуляторных требований [8]	Систематизация проблем и практик обеспечения соответствия программных систем внешним требованиям	Как применимое требование используется как основание оценки конкретной испытательной процедуры
Семантический анализ аргументации [9]	Проверка структуры утверждений, связей и свидетельств в assurance cases	Как обосновать переход от найденных фрагментов к экспертному выводу о достаточности или недостатке
Самокоррекция языковых моделей [10]	Зависимость самопроверки от типа и надежности обратной связи	Как проверять предполагаемый недостаток без подмены независимой проверки повторным ответом модели

ответствие, формирование тестов, проверку аргументации и машинную самопроверку. Предмет настоящей работы уже: условия, при которых содержание ПМИ позволяет обоснованно признать проверку достаточной, подтвердить конкретный недостаток либо сохранить нерешенный вопрос.

Сопоставление этих направлений с задачей настоящего исследования представлено в таблице 1.

Прослеживаемость, формализация нормативных знаний, автоматизация тестирования и машинная поддержка анализа не рассматриваются в настоящей работе как новые сами по себе. Рассмотренные исследования показывают существование самостоятельных решений для этих задач [1—10]. Предлагаемое решение сосредоточено на условиях перехода от применимого требования к оценке уже предусмотренной проверки ПМИ и к обоснованию конкретного экспертного решения. Отсутствие полного аналога в отобранных работах не используется как доказательство мирового приоритета.

Исследовательский вопрос состоит в том, какие условия позволяют различить три возможных результата документарной экспертизы ПМИ: признание предусмотренной проверки достаточной, подтверждение существенного недостатка и сохранение неопределенности, препятствующей обоснованной оценке. Для решения этой задачи рассматривается последовательная связь между применимым требованием, самостоятельным проверяемым результатом, эталонной проверкой, содержанием представленной Исполнителем документации и экспертным решением. Условия перехода между этими элементами формализуются и раскрываются на учебных случаях, включающих локальное и распределенное описание проверки, недостаточность предусмотренной процедуры и отсутствие документа, необходимого для установления содержания требования.

Предлагаемое решение связывает три группы условий: установление необходимой проверки на основе применимых требований, оценку ее обеспечения совокупностью положений ПМИ и связанных документов, а также проверку предполагаемого недостатка путем поиска возможного альтернативного покрытия до формирования замечания. По отношению к рассмотренным исследованиям это позволяет конкретизировать задачу именно для экспертизы испытательной документации: необходимый состав проверок определяется до сопоставления с документацией Исполнителя, а отрицательная оценка требует проверки возможного альтернативного покрытия.

РЕШЕНИЕ ПОСТАВЛЕННОЙ ЗАДАЧИ

Предмет оценки и применимость исходных оснований

Предмет метода — достаточность программы и методики испытаний вместе со связанными сценариями, общими процедурами, приложениями и формами протоколов. Далее этот комплект обозначается как ПМИ. Оценивается предусмотренная проверка самостоятельного результата, а не документ или число ссылок.

Область оценки задается объектом, этапом работ, видом испытаний, требованиями и редакциями документов. Положительный вывод относится к предусмотренной проверке. Он не подтверждает проведение испытания или соответствие системы требованию.

Недостаток ПМИ также не доказывает отсутствия функции в системе.

Исходным основанием может выступать положение договора, технического задания, применимого нормативного правового, нормативно-технического или методического документа. Для его использования требуется установить не только реквизиты, но и конкретное содержание, относящееся к рассматриваемому объекту и условиям. Упоминание документа в общем перечне не заменяет такого установления. Вместе с тем содержание требования может быть распределено: один документ задает обязанность, другой определяет обязательный состав результата или порядок его оценки. Исключение второго документа способно изменить смысл первого.

Отбор источников по редакции, применимости и роли в исследуемой области опирается на подход О.А. Шелестович [1]. Критерии выделения предписаний используются с учетом иной задачи: самостоятельность процесса и самостоятельность проверяемого результата не тождественны.

При восстановлении требования необходимо сохранять условия, ограничения, исключения и содержание документов, к которым оно отсылает. Современные исследования regulatory requirements engineering рассматривают выявление и анализ внешних требований как самостоятельную часть обеспечения соответствия программных систем [8]. Для каждого включенного положения необходимо объяснить, какой существенный элемент оно добавляет. Такая полнота относится к установленной области экспертизы и не означает получения всех существующих источников по теме.

Например, требование предоставить результат загрузки «в порядке, установленном приложением» не позволяет до получения приложения заключить, что обязательным является именно отображение статуса на экране пользователя. В равной мере отсутствие этого приложения не позволяет исключить соответствующую обязанность. До выяснения содержания основания возникает нерешенный вопрос. Замечание об отсутствии проверки экранного статуса в такой ситуации было бы преждевременным: обязательность самого результата еще не установлена.

Применимость источника необходимо отличать от отнесения проверки к конкретному виду контроля. Документ, определяющий обязательный состав результата, может быть необходим для понимания требования, даже если отдельная проверка одного его свойства не относится к рассматриваемой опытной эксплуатации. Сначала восстанавливается полное содержание требования, затем определяется, какие его результаты должны подтверждаться в установленной области испытаний. Сужение области контроля не должно незаметно обеднять исходное требование.

Договорная область экспертизы и юридическая применимость требований также не тождественны. Настоящий метод не устанавливает универсального порядка определения юридической силы документов и не предполагает, что договорный перечень исчерпывает все обязательные требования. Потенциально применимое основание вне согласованного предмета работы должно быть отмечено и рассмотрено отдельно. Вывод о достаточности ПМИ ограничивается той областью, для которой применимость и содержание оснований действительно установлены.

Связь с основанием должна позволять восстановить, почему результат обязателен для данного объекта и какие его свойства проверяются. Эта же связь обосновывает замечание: эксперт указывает, какое требование нельзя надлежащим образом оценить по представленной документации.

Независимый эталон как основание сравнения

Метод включает две стадии. На первой стадии по требованиям контракта и применимым положениям иных документов, уточняющих их содержание, формируется необходимый состав проверяемых результатов и условий их подтверждения. На второй стадии этот состав сопоставляется с ПМИ Исполнителя. Проверяемая документация не должна определять обязательный состав результатов первой стадии.

Результат первой стадии — внутренний эталон проверок. Это основание сравнения, а не образец оформления для Исполнителя. Эталон не задает структуру документов и способ реализации системы, если они не установлены применимым требованием.

Независимость означает формирование эталона из оснований до сопоставления с ПМИ. Она не гарантирует безошибочности толкования. Разделение заданий одного эксперта или программного средства не доказывает независимости ошибок; соблюдение установленного порядка должно быть проверяемо по материалам работы.

Самостоятельный результат допускает отдельный вывод о соответствии. Загрузка данных и отображение пользователю ее статуса связаны одной операцией, но подтверждают разные свойства: совпадение загруженных записей с исходными не доказывает наблюдения статуса.

Условие, параметр или критерий не становятся отдельной проверкой автоматически. Условие «после успешной загрузки» сохраняется при оценке отображения. Разделять проверки следует там, где нужны самостоятельные выводы, а объединять — когда единая процедура и свидетельства позволяют принять общее решение без потери смысла.

Эталонная проверка определяет основание, результат, исходные условия и данные, действия, критерий и предполагаемый способ фиксации. Обязательные варианты и исключения сохраняются в объеме полного требования.

Состав и подробность описания зависят от способа проверки. Для анализа наличия раздела документа и наблюдения поведения системы нужны разные действия. Элемент существен, если без него нельзя

обосновать проведение проверки или оценку результата. Общая формулировка «работает корректно» не заменяет признака соответствия.

Различение процедуры, свидетельства, его документального носителя и решения о соответствии следует модели Свистунова и Митавского [2]. Здесь оно применяется к планируемому испытанию: оценка ПМИ касается возможности получить свидетельство, а не признания уже полученного доказательства.

В эталоне указывается предполагаемый вид свидетельства, а не полученный результат. Протокол может связывать операцию, наблюдение и заключение. Скриншот допустим как один из способов фиксации, но его обязательность нельзя выводить из предпочтения эксперта.

До сопоставления проверяются основания эталона, представление обязательных результатов, допустимость объединений и отсутствие добавленных свойств. Редакция фиксируется. Новое существенное основание требует явного пересмотра эталона и затронутых оценок; скрытая подгонка под ПМИ недопустима.

Условия достаточности и допустимости экспертного вывода

Вторая стадия начинается с поиска фрагментов, относящихся к эталонной проверке. Найденная связь — кандидат на покрытие. Для вывода о достаточности необходимо проверить содержание действий, критериев и фиксации результата.

Возможность представления одного объекта в нескольких документах уже предусмотрена в модели прослеживаемости [2]. Для экспертизы требуется дополнительно оценить обязательность их совместного применения и содержательную согласованность.

Документарное покрытие — обеспечение эталонной проверки положениями ПМИ и связанных документов. Оно может быть локальным или распределенным: это характеристика размещения содержания, а не его достаточности.

Для распределенного покрытия должны выполняться дополнительные условия связи. Фрагменты относятся к рассматриваемой операции и совместимым редакциям документов; их совместное применение обязательно в заданном контексте; между ними нет противоречия, влияющего на вывод; недостающее действие или критерий не восстанавливаются догадкой. Общая форма протокола сама по себе не доказывает обязательность наблюдения, если сценарий и общая процедура такого действия не предусматривают.

Формализацию этих условий удобно ограничить несколькими обозначениями. Пусть q — эталонная проверка, $E(q)$ — множество ее существенных элементов, D — фиксированный доступный комплект испытательной документации, а $F \subseteq D$ — совокупность рассматриваемых фрагментов. Обозначим через $B(q)$ установленность и достаточность основания проверки; через $A(F,q)$ — обязательность и согласованность применения фрагментов к этой проверке; через $C(F,e)$ — содержательное обеспечение элемента e данными фрагментами. Наконец, U означает наличие критической неразрешенности в основаниях соответствующего вывода.

Эти отношения требуют содержательного обоснования. Например, отношение $C(F,e)$ для наблюдения статуса должно опираться на предусмотренное действие и доступный оценке результат, а не только на наличие в тексте слова «статус». В отличие от инструментального анализа формальной аргументации [9], здесь не реализован автоматический вывод из модели. Формулы ниже описывают условия допуска проекта оценки; они не являются алгоритмом проверки смысла и не доказывают полноты системы решений.

Условие допуска положительной документарной оценки записывается следующим образом:

$$\text{Sufficient}(q,D) \Leftarrow B(q) \wedge \exists F \subseteq D \left[A(F,q) \wedge \bigwedge_{e \in E(q)} C(F,e) \right] \wedge \neg U.$$

В этой записи ¬U означает установленное отсутствие блокирующей неопределенности. Неизвестное значение необходимого условия не разрешает вывод. Формальная запись оставляет открытыми проверку полноты $E(q)$ и учет противоречащих обязательных положений вне выбранного F; эти вопросы нельзя считать решенными одним наличием подтверждающих фрагментов.

Отрицательная оценка не получается простым отрицанием положительной. Ненайденное покрытие может означать дефицит, неполный поиск, недоступность документа или неясность основания. Причину требуется установить.

Пусть $G(g,q,D)$ означает установленный существенный дефицит g, препятствующий проверке результата q по комплекту D, а $K(g,q,D)$ — выполненное рассмотрение альтернативного покрытия, не устранившее данный дефицит. Тогда условие формирования проекта замечания имеет вид:

$$\text{DraftRemark}(g,q,D) \Leftarrow B(q) \wedge G(g,q,D) \wedge K(g,q,D) \wedge \neg U.$$

Для G указываются конкретный дефицит и его существенность; для K — рассмотренные альтернативы и причины, по которым они не устранили дефицит. Вывод ограничен установленным корпусом. Конечный поиск не доказывает отсутствия неизвестных документов.

У Шелестович предусмотрен второй эксперт для спорных случаев и сохранение неподтвержденного статуса до согласования [1]. Здесь предполагаемый недостаток подвергается дополнительной проверке: необходимо установить, не устраняется ли он другими положениями испытательной документации либо иной допустимой интерпретацией исходного требования. Само участие второго проверяющего еще не доказывает

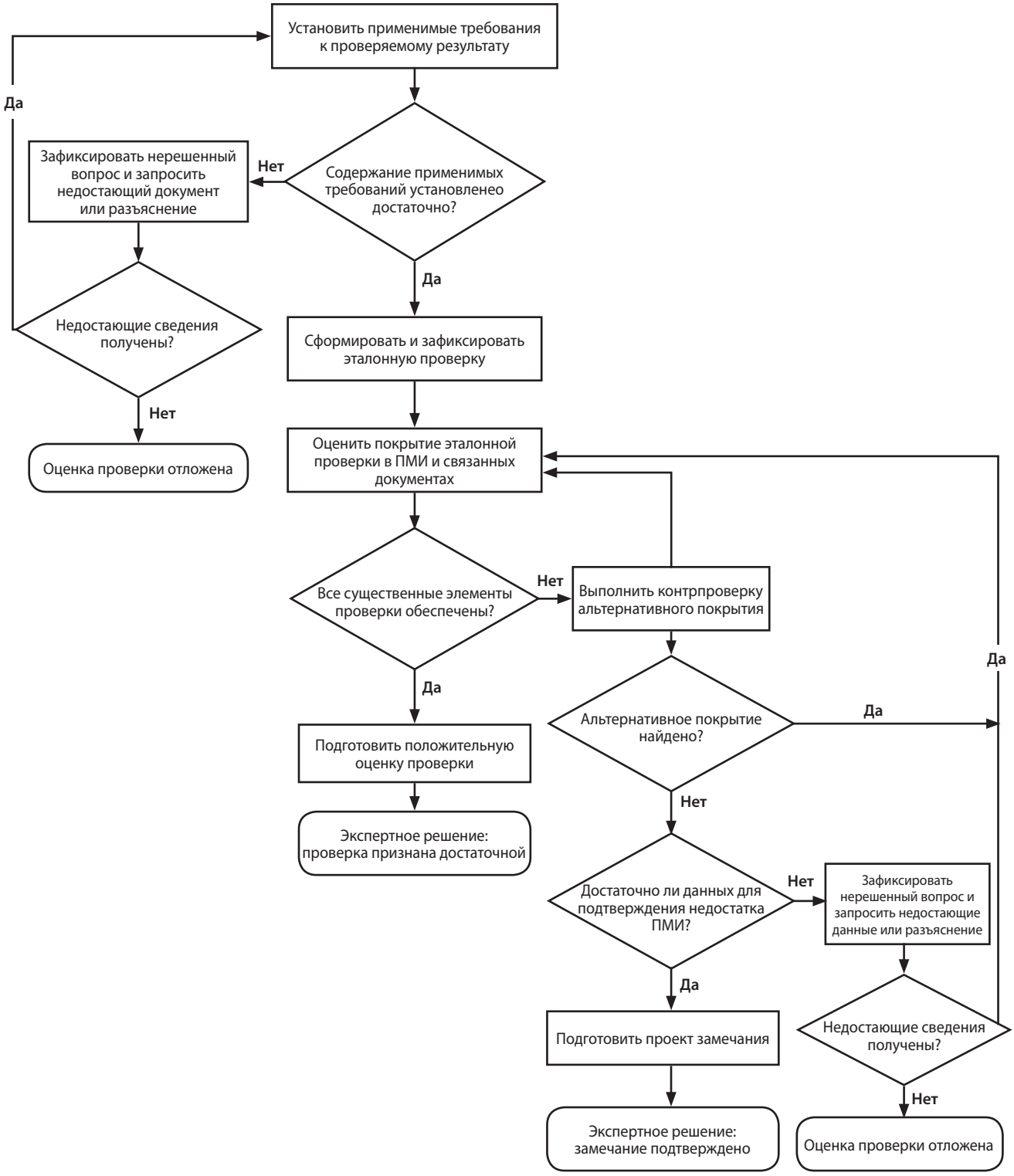

Рис. 1. Условия перехода от основания проверки к экспертному решению

Таблица 2

Условия и границы трех исходов оценки

Исход	Необходимые основания	Допустимое действие	Недопустимая подмена
Достаточность	Обоснованная проверка; согласованное обязательное покрытие всех существенных элементов; разрешенные основания вывода	Подготовить положительную оценку данной проверки для экспертного подтверждения	Признать достаточность по тематическому совпадению или заполнению граф
Подтвержденная недостаточность	Существенный дефицит; его связь с обязательным результатом; рассмотренные альтернативы; отсутствие блокирующей неопределенности	Сформулировать предметный проект замечания	Приравнять ненайденный фрагмент к отсутствующей проверке без поиска возможного альтернативного покрытия
Неразрешенность	Указанная неопределенность, препятствующая содержательному решению	Зафиксировать зависимую область и запросить недостающее основание либо решение	Выдать недостаток данных за недостаток ПМИ или за подтвержденную достаточность

независимости ошибок. Рассматриваются другие сценарии, общие процедуры, приложения, формы фиксации и распределенные связи. Одновременно проверяется, не добавил ли эксперт отсутствующее в основании требование. Повторения того же поискового запроса недостаточно: проверка предполагаемого недостатка должна завершаться обоснованием того, почему рассмотренное альтернативное покрытие его не устраняет.

Неразрешенность означает, что недостаток основания, противоречие или отсутствие данных мешают обосновать содержательный вывод. Она требует получения документа, уточнения применимости или разрешения спорной интерпретации. Зависимая область остается открытой, даже если независимые проверки завершены. Доказанная неприменимость требования отличается от неизвестной применимости; в приведенных формулах и схеме это различие отдельно не формализовано.

Общая логика переходов представлена на схеме (рис. 1).

Если не предусмотрена проверка одного обязательного пользовательского результата, отсутствие действия наблюдения, критерия и фиксации может описываться одним замечанием. Это не исключает разделения независимых дефицитов. Основанием группировки служит содержание необходимого исправления, а не число отсутствующих граф.

Состав обоснования систематизирован в таблице 2.

Эксперт проверяет основание, существенность дефицита, альтернативное покрытие и формулировку. Его аналитический вывод оформляется и утверждается в порядке, установленном для конкретной экспертизы. Метод сам по себе не определяет полномочий на направление замечания или приемку результата работ.

Для поиска и подготовки вариантов толкования могут использоваться программные средства, в том числе языковые модели. Содержательная оценка их предложений остается необходимой: правильность структуры записи не подтверждает правильности толкования. Исследования самокоррекции языковых моделей показывают существенную зависимость результата от типа и надежности обратной связи [10]. Поэтому повторная генерация не заменяет проверку исходного источника и независимую содержательную оценку.

Сквозная учебная демонстрация

Рассмотрим учебный пример. Применимое требование устанавливает: «После успешной загрузки набора данных система должна отображать пользователю статус успешного завершения операции». Задача экспертизы состоит в том, чтобы установить, позволяет ли предусмотренная в ПМИ проверка подтвердить именно этот результат, а не только факт успешной загрузки. Поэтому эталонная проверка должна предусматривать успешное выполнение загрузки, наблюдение пользователем соответствующего статуса, однозначный критерий его связи с завершенной операцией и фиксацию результата наблюдения.

Первый вариант: проверка результата не предусмотрена. В ПМИ предусмотрены загрузка файла и подтверждение ее успешности путем сопоставления загруженных данных с исходным набором. Форма протокола позволяет зафиксировать выполнение загрузки и результат сопоставления. Однако ни в сценарии, ни в общей процедуре, ни в форме протокола не предусмотрены наблюдение пользовательского статуса, критерий его отображения и фиксация этого результата. Таким образом, документация позволяет подтвердить успешность загрузки, но не обеспечивает проверку предусмотренного требованием пользовательского результата.

Проверка связанных документов не выявляет положения, которое устраняло бы этот недостаток. Поэтому может быть подготовлен следующий проект замечания:

В ПМИ предусмотрены выполнение загрузки и подтверждение ее успешности сопоставлением данных,

однако не предусмотрена проверка отображения пользователю статуса успешного завершения операции. Необходимо предусмотреть проверку указанного результата, однозначный критерий его оценки и фиксацию результата проверки.

Такое замечание относится к испытательной документации. Оно не означает, что соответствующая функция отсутствует в системе, и не предписывает конкретный способ ее реализации.

Второй вариант: проверка описана в нескольких связанных документах. В сценарии установлена обязательность общей процедуры, которая наряду с подтверждением успешности загрузки требует проверить отображение пользователю статуса этой операции. Форма протокола предусматривает фиксацию наблюдаемого статуса и результата проверки. В этом случае отсутствие полного описания непосредственно в сценарии не является недостатком: необходимые элементы проверки распределены между связанными документами и должны оцениваться совместно.

Однако наличие всех необходимых элементов еще не означает, что процедура достаточна. Если наблюдение статуса начинается только после сопоставления загруженных данных, кратковременно отображаемый статус может быть пропущен. Исходное требование не устанавливает обязанность сохранять его до окончания сопоставления, поэтому такое свойство нельзя добавлять к требованию при экспертизе. До уточнения порядка наблюдения положительная оценка этой проверки преждевременна.

Третий вариант: содержание требования установлено не полностью. Предположим, что исходный документ требует предоставлять результат загрузки «в порядке и составе, установленных приложением», однако само приложение отсутствует. В такой ситуации нельзя заранее считать обязательным именно отображение пользовательского статуса: содержание требования еще не восстановлено полностью. Детально описанная процедура загрузки в ПМИ этого пробела не устраняет. Сначала необходимо получить и проверить отсутствующий документ, установить полное содержание требования и лишь затем формировать соответствующую эталонную проверку и оценивать ее наличие в ПМИ.

Три варианта показывают разные основания экспертного решения. В первом случае после проверки возможного альтернативного покрытия подтверждается конкретный недостаток ПМИ. Во втором необходимое содержание найдено в совокупности документов, но остается вопрос о достаточности предусмотренной процедуры. В третьем имеющихся исходных документов недостаточно даже для установления того, что именно должна подтверждать проверка. Следовательно, отсутствие текста в одном месте документации само по себе не доказывает недостатка, а подробность ПМИ не компенсирует неполноту требования, с которым она сопоставляется.

Ограничения обоснования и дальнейшая проверка

Предложенный подход позволяет проследить экспертную оценку ПМИ от применимого требования до конкретной проверки и сформированного замечания. Однако обоснованность такого вывода зависит от того, насколько корректно определен исходный корпус требований. Прослеживаемость найденных положений подтверждает происхождение использованных требований, но сама по себе не доказывает, что выявлены все документы, способные повлиять на содержание проверки. Поэтому границы рассмотренного корпуса, недоступные документы и неразрешенные отсылки должны сохраняться в материалах экспертизы, а выводы — относиться к той области, для которой применимость и содержание требований действительно установлены.

Вторая группа ограничений связана с содержательной интерпретацией. Не все решения могут быть получены из формальных связей между документами. В частности, эксперт должен определить, образует ли положение самостоятельный проверяемый результат, какие элементы необходимы для его подтверждения и обеспечивает ли совокупность фрагментов ПМИ достаточную проверку. Формализация делает основания таких решений явными и позволяет локализовать разногласия, но не подменяет профессионального анализа содержания требований и испытательной документации.

Отдельное ограничение связано с независимостью стадий анализа. Формирование эталонной проверки до изучения соответствующего содержания ПМИ снижает риск подгонки необходимой проверки под представленную документацию. Аналогично, отрицательная или частичная оценка требует дополнительной проверки: поиска возможного альтернативного покрытия и проверки исходного толкования требования. Вместе с тем организационное разделение стадий еще не гарантирует независимости ошибок: одинаковые исходные предпосылки или одна и та же ошибочная интерпретация могут воспроизводиться на нескольких этапах анализа. В самой статье поэтому различается разделение процедур и доказанная независимость их результатов.

Технологии искусственного интеллекта в этой схеме целесообразно рассматривать как средство поддержки отдельных аналитических операций: выявления ссылок и потенциально применимых источников, поиска относящихся к эталонной проверке фрагментов ПМИ, подготовки вариантов интерпретации и поиска возможного альтернативного покрытия. Их использование особенно полезно при работе с объемным и распределенным документарным корпусом. При этом машинно найденное соответствие не является доказательством достаточности проверки, а предложенная интерпретация — установленным содержанием требования. Содержательная оценка применимости

источника, достаточности предусмотренной проверки и обоснованности замечания остается предметом экспертного решения. В предлагаемом подходе технологии искусственного интеллекта рассматриваются как средство поддержки отдельных аналитических операций; оценка применимости требований, достаточности проверки и обоснованности замечания остается предметом экспертного решения.

Дальнейшая проверка подхода может быть направлена не только на итоговые замечания, но и на устойчивость отдельных стадий: полноту формирования корпуса требований, согласованность выделения самостоятельных результатов, воспроизводимость оценки распределенного покрытия и результативность проверки предполагаемых недостатков. На реальных обезличенных комплектах ПМИ с независимой экспертной разметкой можно также отдельно исследовать, как применение ИИ влияет на поиск релевантных фрагментов, выявление альтернативного покрытия, число требующих ручного разрешения случаев и общую трудоемкость экспертизы.

Результаты и выводы

В работе сформулированы условия обоснованной документарной оценки программ и методик испытаний. Исходной точкой такой оценки являются требования контракта и применимые положения иных источников, определяющие обязательные результаты и условия их проверки. На их основе до сопоставления с документацией Исполнителя формируется эталонная проверка, после чего устанавливается, предусмотрена ли соответствующая проверка в ПМИ и позволяет ли она однозначно подтвердить требуемый результат. Тем самым содержание представленной ПМИ не используется для определения того, что именно должно быть проверено.

Основным объектом оценки выступает не отдельная ссылка или фрагмент текста, а предусмотренная проверка самостоятельного результата. Наличие относящегося к требованию текста еще не означает достаточности проверки. Необходимые действия, критерии и способы фиксации могут быть распределены между сценарием, общей процедурой, приложением и формой протокола; поэтому их следует оценивать в совокупности с учетом обязательности совместного применения и согласованности содержания. Положительная оценка допустима, когда эта совокупность позволяет выполнить необходимую проверку и однозначно оценить ее результат. При этом такая оценка относится к испытательной документации и сама по себе не подтверждает фактическое выполнение требования информационной системой.

Отрицательная оценка требует дополнительного обоснования. Если часть необходимой проверки первоначально не найдена, прежде чем формировать замечание, рассматриваются другие связанные документы, возможное распределенное описание и правильность исходного толкования требования. Только после исключения альтернативного покрытия установленный недостаток может служить основанием для проекта замечания. Если же содержание требования не может быть достоверно определено из-за отсутствующего документа, неясной применимости источника или неразрешенного противоречия, формируется нерешенный вопрос, а зависимая оценка откладывается. Поэтому в методе различаются три результата: достаточность предусмотренной проверки, подтвержденная недостаточность и невозможность обоснованной оценки до получения недостающих сведений.

Учебная демонстрация показывает различие этих ситуаций на одном проверяемом результате. Подтверждение успешной загрузки данных не заменяет проверки отображения пользователю статуса завершения операции. Вместе с тем отсутствие такого действия непосредственно в сценарии не является недостатком, если оно предусмотрено обязательной общей процедурой и надлежащим образом фиксируется. Даже при наличии распределенного описания необходимо оценивать выполнимость всей процедуры: порядок действий может не позволять наблюдать требуемый результат. Если же обязательное содержание требования определяется отсутствующим приложением, сопоставление ПМИ с предполагаемой проверкой до получения этого документа было бы необоснованным.

Формальная запись используется в работе для фиксации условий, при которых допустим переход к положительной оценке или проекту замечания. Она не заменяет содержательной оценки применимости требования, самостоятельности проверяемого результата, достаточности предусмотренной процедуры и согласованности связанных документов. Такой способ представления делает явными основания экспертного решения и позволяет отделить установленный недостаток ПМИ от неполноты исходных данных или поисковой процедуры.

Технологии искусственного интеллекта могут поддерживать отдельные операции этого процесса: поиск ссылок и потенциально применимых источников, выявление относящихся к эталонной проверке фрагментов документации, подготовку вариантов интерпретации и поиск возможного альтернативного покрытия. Полученный машиной результат рассматривается как аналитический материал, а не как самостоятельное экспертное решение. Оценка применимости требований, достаточности проверки и обоснованности замечания сохраняется за экспертом.

Предложенный подход обеспечивает прослеживаемую связь между применимым требованием, проверяемым результатом, необходимой проверкой, содержанием ПМИ и экспертным решением. Такая связь позволяет установить не только результат оценки, но и основания, по которым предусмотренная проверка признана достаточной, выявленный недостаток подтвержден либо принятие решения отложено до получения недостающих сведений.

Информация о финансировании.
Исследование выполнено в инициативном порядке.

Литература

1. Шелестович О.А. Метод идентификации процессов деятельности при создании информационных систем // Правовая информатика. 2026. № 2. С. 114—126. DOI: 10.24412/1994-1404-2026-2-114-126.
2. Свистунов А.В., Митавский Д.В. Нормативно-ориентированная архитектура сквозной прослеживаемости требований к информационным системам // Правовая информатика. 2026. № 2. С. 346—369. DOI: 10.24412/1994-1404-2026-2-346-369.
3. Грибова В.В., Жданов Д.В., Помников Е.Е. Онтологическое моделирование нормативных знаний в строительной отрасли // Онтология проектирования. 2026. Т. 16. № 2 (60). С. 226—239. DOI: 10.18287/2223-9537-2026-16-2-226-239.
4. Малиновский И.М., Стародумов А.В., Мягков К.А., Юсипов Б.Х. Актуальность внедрения системы управления требованиями в производственный процесс авиадвигателестроительной отрасли // Вестник Самарского университета. Аэрокосмическая техника, технологии и машиностроение. 2024. Т. 23. № 2. С. 167—177. DOI: 10.18287/2541-7533-2024-23-2-167-177.
5. Герлиц Е.А., Кильдишев Д.С., Хорошилов А.В. Выявление функциональных требований в документации программного интерфейса приложения для функционального тестирования // Труды Института системного программирования РАН. 2022. Т. 34. Вып. 1. С. 7—22. DOI: 10.15514/ISPRAS-2022-34(1)-1.
6. Позин Б.А., Циперман Г.Н. Прослеживаемость требований как основа проектирования функционально-логической архитектуры программной системы // Труды Института системного программирования РАН. 2022. Т. 34. Вып. 1. С. 23—34. DOI: 10.15514/ISPRAS-2022-34(1)-2.
7. Шишин К.А., Муравьёв И.В., Куликов Е.К. Об автоматической генерации модульных тестов для Java-приложений, использующих фреймворк Spring // Труды Института системного программирования РАН. 2024. Т. 36. Вып. 2. С. 59—72. DOI: 10.15514/ISPRAS-2024-36(2)-5.
8. Kosenkov O., Elahidoost P., Gorschek T., Fischbach J., Mendez D., Unterkalmsteiner M., Fucci D., Mohanani R. Systematic Mapping Study on Requirements Engineering for Regulatory Compliance of Software Systems // Information and Software Technology. 2025. Vol. 178. Art. 107622. DOI: 10.1016/j.infsof.2024.107622.
9. Murugesan A., Wong I., Arias J., Stroud R., Varadarajan S., Salazar E., Gupta G., Bloomfield R., Rushby J. Automating Semantic Analysis of System Assurance Cases Using Goal-Directed ASP // Theory and Practice of Logic Programming. 2024. Vol. 24. No. 4. P. 805—824. DOI: 10.1017/S1471068424000425.
10. Kamoi R., Zhang Y., Zhang N., Han J., Zhang R. When Can LLMs Actually Correct Their Own Mistakes? A Critical Survey of Self-Correction of LLMs // Transactions of the Association for Computational Linguistics. 2024. Vol. 12. P. 1417—1440. DOI: 10.1162/tacl_a_00713.

SECTION:
REGULATORY INFORMATICS

VALIDITY OF EXPERT ASSESSMENT OF TEST PROGRAMS AND PROCEDURES

Shelestovich O.A.[3], Svistunov A.V.[4]

Keywords: *test documentation, information system requirements, requirements traceability, regulatory requirements, reference test, test sufficiency, verification of expert conclusions, artificial intelligence, large language models.*

[3] ***Olesya A. Shelestovich****, Lead Systems Analyst, Information Systems Design and Development Department, Roskadastr-Infotech Information Technology Center, branch of the Public Law Company Roskadastr, Moscow, Russia.*
E-mail: ShelestovichOA@kadastr.ru

[4] ***Andrey V. Svistunov****, Ph.D. (Technology), Senior Researcher, Project Manager, Information Systems Design and Development Department, Roskadastr-Infotech Information Technology Center, branch of the Public Law Company Roskadastr, Moscow, Russia.*
E-mail: SvistunovAV@kadastr.ru

Abstract

Purpose. *To define the rules for a well-founded assessment of test programs and procedures (TPPs): when a specified test can be considered sufficient, when an identified deficiency provides grounds for an expert comment, and when the assessment should be deferred because the source documentation is incomplete, including cases where the description of a test is distributed across several documents.*

Research method. *The study is based on decomposing the expert assessment of TPPs and formalizing the conditions linking an applicable requirement, a verifiable result, a reference test, documentary coverage, and an expert decision. The proposed conditions are examined using contrasting instructional cases that model complete, distributed, and insufficient descriptions of a test, as well as incomplete source documentation.*

Results. *The conditions for a well-founded assessment of test documentation have been formulated by linking an applicable requirement, a distinct verifiable result, a reference test, the content of the submitted documents, and the expert decision. A distinction is made between the presence of relevant text in the documentation and the sufficiency of the specified test for confirming the required result. Conditions are defined for assessing a test whose description is distributed across several related documents and for verifying a presumed deficiency before an expert comment is formulated, including the search for possible alternative coverage and reconsideration of the initial interpretation of the requirement. It is shown that incomplete source documentation, uncertain applicability of a requirement, or an unresolved contradiction should remain sources of uncertainty rather than be treated as deficiencies in the test documentation. The formalized relations are used as conditions governing the admissibility of a positive assessment and a draft expert comment, but they do not replace substantive analysis of requirement applicability, procedural sufficiency, or the materiality of an identified deficiency. The proposed provisions are methodological in nature and are illustrated through instructional cases; their practical effectiveness, inter-expert agreement, and the completeness of identifying applicable sources require separate empirical evaluation.*

References

1. Shelestovich O.A. Metod identifikatsii protsessov deiatel'nosti pri sozdanii informatsionnykh sistem // Pravovaia informatika. 2026. No. 2. S. 114—126. DOI: 10.24412/1994-1404-2026-2-114-126.
2. Svistunov A.V., Mitavskii D.V. Normativno-orientirovannaia arkhitektura skvoznoi proslezhivaemosti trebovanii k informatsionnym sistemam // Pravovaia informatika. 2026. No. 2. S. 346—369. DOI: 10.24412/1994-1404-2026-2-346-369.
3. Gribova V.V., Zhdanov D.V., Pomnikov E.E. Ontologicheskoe modelirovanie normativnykh znanii v stroitel'noi otrasli // Ontologiia proektirovaniia. 2026. T. 16. No. 2 (60). S. 226—239. DOI: 10.18287/2223-9537-2026-16-2-226-239.
4. Malinovskii I.M., Starodumov A.V., Miagkov K.A., Iusipov B.Kh. Aktual'nost' vnedreniia sistemy upravleniia trebovaniiami v proizvodstvennyi protsess aviadvigatelestroitel'noi otrasli // Vestnik Samarskogo universiteta. Aerokosmicheskaia tekhnika, tekhnologii i mashinostroenie. 2024. T. 23. No. 2. S. 167—177. DOI: 10.18287/2541-7533-2024-23-2-167-177.
5. Gerlits E.A., Kil'dishev D.S., Khoroshilov A.V. Vyiavlenie funktsional'nykh trebovanii v dokumentatsii programmnogo interfeisa prilozheniia dlia funktsional'nogo testirovaniia // Trudy Instituta sistemnogo programmirovaniia RAN. 2022. T. 34. Vyp. 1. S. 7—22. DOI: 10.15514/ISPRAS-2022-34(1)-1.
6. Pozin B.A., Tsiperman G.N. Proslezhivaemost' trebovanii kak osnova proektirovaniia funktsional'no-logicheskoi arkhitektury programmnoi sistemy // Trudy Instituta sistemnogo programmirovaniia RAN. 2022. T. 34. Vyp. 1. S. 23—34. DOI: 10.15514/ISPRAS-2022-34(1)-2.
7. Shishin K.A., Murav'ev I.V., Kulikov E.K. Ob avtomaticheskoi generatsii modul'nykh testov dlia Java-prilozhenii, ispol'zuiushchikh freimvork Spring // Trudy Instituta sistemnogo programmirovaniia RAN. 2024. T. 36. Vyp. 2. S. 59—72. DOI: 10.15514/ISPRAS-2024-36(2)-5.
8. Kosenkov O., Elahidoost P., Gorschek T., Fischbach J., Mendez D., Unterkalmsteiner M., Fucci D., Mohanani R. Systematic Mapping Study on Requirements Engineering for Regulatory Compliance of Software Systems // Information and Software Technology. 2025. Vol. 178. Art. 107622. DOI: 10.1016/j.infsof.2024.107622.
9. Murugesan A., Wong I., Arias J., Stroud R., Varadarajan S., Salazar E., Gupta G., Bloomfield R., Rushby J. Automating Semantic Analysis of System Assurance Cases Using Goal-Directed ASP // Theory and Practice of Logic Programming. 2024. Vol. 24. No. 4. P. 805—824. DOI: 10.1017/S1471068424000425.
10. Kamoi R., Zhang Y., Zhang N., Han J., Zhang R. When Can LLMs Actually Correct Their Own Mistakes? A Critical Survey of Self-Correction of LLMs // Transactions of the Association for Computational Linguistics. 2024. Vol. 12. P. 1417—1440. DOI: 10.1162/tacl_a_00713.