

DOI: 10.15514/ISPRAS-2023-35(6)-1



Валидация требований и ее влияние на качество при использовании программного обеспечения: тематическое исследование

¹ Л. Канчари, ORCID: 0000-0003-2382-7724 <lcanchari@continental.edu.pe>

² П. Ангелери, ORCID: 0000-0002-0719-7235 <paula.angeleri@ub.edu.ar>

^{1,3} А. Давила, ORCID: 0000-0003-2455-9768 <abraham.davila@puclp.edu.pe>

¹ Перу, Уанкайо, Континентальный университет.

² Аргентина, Буэнос-Айрес, Университет Бельграно.

³ Перу, Лима, Папский католический университет Перу.

Аннотация. Исследование посвящено изучению и анализу взаимосвязи между повышением качества требований к программному обеспечению и качеством используемого программного продукта. Анализ проводился при разработке реальных требований к программному продукту и вычислении метрик двух программных продуктов в соответствии со стандартом ISO/IEC 25010. Результаты показывают, что мероприятия по проверке качества, проводимые на этапе разработки требований к программному обеспечению, положительно влияют на качество реализуемых программных продуктов. В отношении изученного программного обеспечения можно сказать, что повышение качества требований способствовало повышению качества использования программных продуктов.

Ключевые слова: требования к программному обеспечению; проверка качества требований; качество при использовании программного обеспечения; стандарт ISO/IEC 25010.

Для цитирования: Канчари Л., Ангелери П., Давила А. Валидация требований и ее влияние на качество при использовании программного обеспечения: тематическое исследование. Труды ИСП РАН, том 35, вып. 6, 2023 г., стр. 7–28. DOI: 10.15514/ISPRAS–2023–35(6)–1.

Благодарности: Представленное исследование поддерживалось Папским католическим университетом Перу, группой исследований и разработок в области программной инженерии GIDIS.

Requirements Validation and Its Impact in Software Quality in Use: A Case Study

¹ L. Canchari, ORCID: 0000-0003-2382-7724 <lcanchari@continental.edu.pe>

² P. Angeleri, ORCID: 0000-0002-0719-7235 <paula.angeleri@ub.edu.ar>

^{1,3} A. Dávila, ORCID: 0000-0003-2455-9768 <abraham.davila@puclp.edu.pe>

¹ Universidad Continental, Huancayo, Perú.

² Universidad de Belgrano, Buenos Aires, Argentina.

³ Pontificia Universidad Católica del Perú, Lima, Perú.

Abstract. In this study, the relationship between the improvement of software requirement quality and the software product quality in use was explored and analyzed. Analysis was based on the design of software product quality-in-use and the measure of metrics from ISO/IEC 25010 standard in two software products. The results show that the validation activities introduced in the software requirements stage have a positive relationship with the quality in use of the software products analyzed. In the software studied, it can be said that the improvement of the quality of the requirements has contributed to the improvement of the quality in use of software products.

Keywords: software requirements; requirement quality validation; software product quality-in-use; ISO/IEC 25010 standard.

For citation: Canchari L., Angeleri P., Dávila A. Requirements Validation and its impact in Software Quality in Use: A case study. *Trudy ISP RAN/Proc. ISP RAS*, vol. 35, issue 6, 2023. pp. 7-28 (in Russian). DOI: 10.15514/ISPRAS-2023-35(6)-1.

Acknowledgements. This research was supported by Software Engineering Research & Development Group (GIDIS) – Pontifical Catholic University of Peru.

1. Введение

Стандарт ISO/IEC 12207 был установлен в правовой базе Перу [1] в качестве обязательного. Подразделениями информационных технологий (Information Technology Unit, ITU) принятие стандарта ISO/IEC 29110-5-1-2 рассматривалось как возможный прогрессивный способ внедрения стандарта ISO/IEC 12207 в рамках общего развития технологий. Стандарт ISO/IEC 29110-5-1-2 – это руководство для небольших организаций по разработке программного обеспечения, которое в явном виде включает задачи верификации и валидации [2]. В частности, он определяет валидацию требований к программному обеспечению в качестве первичной задачи [2], которая положительным образом влияет на качество программного обеспечения [3].

Кроме того, некоторые исследования установили взаимосвязь между требованиями к программному обеспечению и успешности его разработки [4-7]. Другие авторы указывали, что анализ требований является одним из наиболее важных этапов [3, 5, 7-9]. В работе [10] указывается, что проблемы с требованиями вызвали кризис программного обеспечения. Таким образом, понятно, что действия, направленные на повышение качества требований, способствуют повышению продуктивности проекта и качества передаваемого пользователям разработанного программного продукта. Авторы также указывают, что существует заинтересованность исследователей в изучении качества при использовании программного продукта [4, 11-13].

В нашей предыдущей статье [14] характеристики, результаты и преимущества валидации требований к программному обеспечению в контексте общественного учреждения ITU были представлены в контексте некоторых характеристик качества и характеристик проекта. Настоящее исследование тоже базировалось на характеристиках и результатах работы [14]. Как видно из [14], нам удалось улучшить как процесс поддержания качества программного обеспечения, так и качество готового программного продукта. Чтобы это проверить, качество программного продукта измерялось с помощью метрик эффективности проекта и характеристик программного продукта.

В этой статье, беря за основу качество при использовании, авторы в качестве основной цели выбрали задачу определения преимуществ применения процедур валидации требований к программному обеспечению. По этой причине при оценке качества при использовании программного обеспечения, основанной на стандарте ISO/IEC 25010 и связанных с ним материалах, используется тот же контекст, что и в предыдущей работе [14], в которой исследовалось программное обеспечение информационных систем для общественных учреждений в Лиме. Остальная часть статьи структурирована следующим образом: в разделе 2 представлены принципы обеспечения “качества при использовании”, основанные на стандартах серии ISO/IEC 25000, в разделе 3 описан подход к проведению тематического исследования, в разделе 4 представлены результаты этого исследования, а в разделе 5 – итоговое обсуждение и будущие работы.

2. Основные концепции

В этом разделе представлены некоторые основные концепции и модель качества при использовании, введенные стандартом ISO/IEC 25000, а также кратко излагается содержание некоторых работ, тем или иным образом связанных с нашим исследованием.

2.1 Модель “качества при использовании” программного обеспечения в ISO

Серия стандартов ISO/IEC 25000 (Требования и оценка качества систем и программных средств) представляет собой набор технических документов, разработанных в отношении качества программного продукта в рамках проекта SQuaRE [15-18]. Эти документы объединяют все предыдущие стандарты и предлагают новые подходы. Стандарты ISO/IEC 25022 [15] и ISO/IEC 25010 [19] определяют качество при использовании как степень, в которой конкретные пользователи могут использовать продукт или систему для удовлетворения своих потребностей и достижения конкретных целей с эффективностью, результативностью, удовлетворением и свободой от риска в конкретных контекстах использования [15]; как это видно на рис. 1. В процессе определения модели качества программного продукта [20] набор характеристик, подхарактеристик и показателей выбирается из базовой модели (например, внутренней или внешней модели ISO/IEC 25010) и, в некоторых случаях, для всего программного обеспечения или его части устанавливаются их эталонные значения.

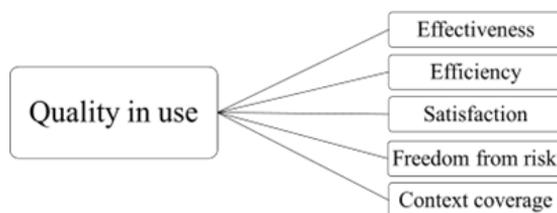


Рис 1. Модель качества при использовании ISO/IEC 25010, адаптированная из [19] (Эффективность, Результативность, Удовлетворенность, Свобода от рисков, Контекстный охват)
Fig 1. Quality in use model ISO/IEC 25010 adapted from [19]

2.2 Показатели качества при использовании

Показатель – это определенный метод измерения, связанный со шкалой оценок [20-21]. Измерение – это набор операций, целью которых является определение некоторого численного значения [15]. Качество при использовании может быть измерено на основе предложений стандарта ISO 25010, которые создают основу для более комплексного подхода к спецификации и измерению качества [22-23]. В табл. 1 представлены показатели качества при использовании в соответствии с ISO/IEC 25022.

Согласно стандарту ISO/IEC 25022 [15], существует множество факторов, влияющих на качество при использовании программного продукта. Измерение или оценка этих факторов представляет собой сложный процесс [24]. В целом, программная инженерия предлагает и методы, и инструменты, помогающие снизить сложность разработок. Например, иногда применяются такие две методики: методика “Размышление вслух” и “Опрос удобства использования системы после изучения” (Post-Study System Usability Questionnaire, PSSUQ). По методике исследования “Размышление вслух” участники должны произносить вслух любые слова, которые приходят им на ум, пока они работают над заданием [25]. Согласно [26], участники озвучивают свои мысли по мере прохождения через программный интерфейс. Методика PSSUQ, согласно [27], обладает следующими характеристиками:

- (i) это стандартизированный вопросник из 16 пунктов, который после выполнения задания используется для измерения удовлетворенности пользователей веб-сайтом, программным обеспечением, системой или продуктом,
- (ii) пункты написаны в прошедшем времени для обозначения недавно выполненных заданий,
- (iii) для присвоения баллов по каждому пункту используется 7-уровневая шкала, которая варьируется от 1, что означает “полностью согласен”, до 7, что означает “полностью не согласен”,
- (iv) по трем группам – полезность системы (пункты 1-6), качество информации (пункты 7-12), качество интерфейса (пункты 13-15) – вычисляется среднее значение, отдельный 16-й пункт используется для общей оценки.

Табл. 1. Показатели качества при использовании в соответствии с ISO/IEC 25022 [15]

Table 1. Quality in use metrics according to ISO/IEC 25022 [15]

Показатель	ID	Имя
Показатели результативности	Ef-1-G	Выполненные задачи
	Ef-2-S	Достигнутые цели
	Ef-3-G	Ошибки в задаче
	Ef-4-G	Задачи с ошибками
	Ef-5-G	Интенсивность ошибок задачи
Показатели эффективности	Ey-1-G	Время выполнения задачи
	Ey-2-S	Эффективность затраченного времени
	Ey-3-S	Экономическая эффективность
	Ey-4-S	Коэффициент полезного времени
	Ey-5-S	Ненужные действия
	Ey-6-S	Последствия усталости
Показатели удовлетворенности	SUs-1-G	Общая удовлетворенность
	SUs-2-G	Удовлетворенность функциями
	SUs-3-G	Дискреционное использование
	SUs-4-G	Использование функций
	SUs-5-G	Доля жалующихся пользователей
	SUs-6-G	Доля жалоб пользователей на ту или иную функцию
Показатели доверия	STr-1-G	Доверие пользователей
Показатели удовольствия	SPI-1-G	Удовольствие пользователя
Показатели эргономического комфорта	SCo-1-G	Физический комфорт
Показатели свободы риска	Rec-1-G	Рентабельность инвестиций (ROI)
	Rec-2-G	Время окупаемости инвестиций
	Rec-3-G	Эффективность бизнеса
	Rec-4-G	Преимущества инвестиций в ИТ
	Rec-5-G	Обслуживание клиентов
	Rec-6-G	Количество посетителей сайта, ставших клиентами
	Rec-7-G	Выручка от каждого клиента
	Rec-8-G	Ошибки с экономическими последствиями
Показатели по снижению рисков для здоровья и безопасности труда	RHe-1-G	Частота возникновения проблем со здоровьем пользователей
	RHe-2-G	Воздействие на здоровье и безопасность пользователей
	RHe-3-G	Безопасность людей, пострадавших от использования системы
Показатели по снижению экологических рисков	REn-1-G	Экологическое воздействие
Показатели полноты контекста	CCm-1-G	Полнота контекста
Показатели гибкости	CFI-1-S	Использование гибких связей
	CFI-2-S	Гибкость продукта
	CFI-3-S	Независимость от квалификации

2.3 Обзор литературы

В научной литературе имеются описания исследований, которые помогают понять комплексность качества программного обеспечения, предлагая модели, атрибуты и показатели. Однако существует лишь несколько исследований, в которых анализируется влияние качества при использовании как следствия совершенствования процесса валидации требований. Эти тематически близкие работы рассматриваются далее.

В работе [28] авторы приходят к выводу, что валидация требований к программному обеспечению имела большое значение с точки зрения снижения стоимости проекта.

В работе [29] указаны:

- (i) важность интеграции характеристик качества продукта в процессах их производства,
- (ii) замечание о том, что организации проводят сертификацию своих процессов без учета влияния этих процессов на характеристики качества продукции, и,
- (iii) подтверждение того, что качество продукта зависит от качества процесса его получения.

На основании проведенного исследования делается вывод о том, что качество взаимодействия влияет на решения пользователей о внедрении той или иной системы.

3. Разработка тематического исследования

В ходе проведенной нами работы было проведено тематическое исследование (case study, CS) с учетом рекомендаций, предложенных в работах [30-31]. В табл. 2 показана адаптация этапов CS. Необходимые теоретические обоснования (P1-3) были представлены в разделе 2.

Табл. 2. Фазы и шаги схемы тематического исследования, адаптированные из [30] и [31]

Table 2. Adaptation of the case study phases from [30] and [31]

Фаза	Функция
P1 Разработка тематического исследования	(P1-1) Цель исследования
	(P1-2) Что вы изучаете
	(P1-3) Требуемая теория
	(P1-4) Вопросы исследования
	(P1-5) Методы сбора
	(P1-6) Отбор данных
P2 Определение сбора данных	(P2-1) Определение метода сбора данных
E1 Сбор данных	(E1-1) Сбор данных
E2 Анализ данных	(E2-1) Анализ данных
	(E2-2) Интерпретация данных
R1 Отчетность	(R1-1) Представление результатов

3.1 Объекты исследования

Информационные системы, рассмотренные в данном тематическом исследовании, являются частью нашего предыдущего исследования [15] общественных учреждений ИТУ. В ходе ранее проведенного исследования было установлено, что качество требований к программному обеспечению для одной из систем превосходит качество требований к другой системе. Кроме того, было доказано, что качество требований напрямую связано со стоимостью, графиком работ и качеством процесса создания программного обеспечения информационной системы. В нашем новом исследовании мы стремимся определить взаимосвязь между повышением качества требований и качеством при использовании программного обеспечения, а также определить, как эти понятия преобразуются в преимущества для пользователя.

В настоящем исследовании из-за ограничений ИТУ использовались только две информационные системы из тех шести, что были представлены в [15]. Для обеспечения преемственности к предыдущему исследованию [15] были выбраны системы SI02 и SI03. Для

этих систем были определены участвовавшие пользователи, контекст для каждой из систем, протоколы оценки, методы получения данных и показатели для проведения измерений качества. Систему SI02 используют 40 должностных лиц ИТУ, которые отвечают за функционирование процесса государственных закупок. Эта система представляет собой настольное приложение, основанное на клиент-серверной архитектуре и реляционной базе данных. При разработке системы методы проверки требований не применялись. Система SI03 используется 20-ю должностными лицами ИТУ, отвечающими за набор и мониторинг сотрудников для консультационной службы. SI03 – это веб-приложение, и, как и SI02, оно было разработано в клиент-серверной архитектуре для доступа к реляционной базе данных. При разработке этого программного обеспечения были применены методы проверки требований.

3.2 Цели тематического исследования (P1-1)

Основная цель этого исследования (P1-1) состоит в том, чтобы выявить преимущества внедрения валидации требований к программному обеспечению с точки зрения качества при использовании, то есть качества программного обеспечения в специфическом контексте пользователя.

Контекст исследования имеет следующие характеристики:

- (i) программные продукты эксплуатируются более 12 месяцев,
- (ii) участники используют программное обеспечение в качестве основного инструмента для своей рабочей деятельности, и
- (iii) системное окружение, то есть компьютеры, оборудование рабочих мест, графики работы и даже отвлекающие факторы, оставлены такими же, как в реальной жизни.

Для оценки качества при использовании была подготовлена копия приложений, связанная с копией базы данных, а на стороне пользователя были настроены ярлыки для этих приложений.

3.3 Что является объектом изучения в этом эмпирическом исследовании (P1-2)?

В ходе этого исследования определялись и анализировались показатели качества при использовании программного обеспечения информационных систем в ИТУ. Объектами анализа для данного исследования являлись два программных продукта. В одном случае проверка требований к программному обеспечению включалась в состав работ по улучшению процесса, а в другом – не включалась. Программы, выбранные для исследования (см. табл. 3), входили в состав проектов, анализировавшихся в предыдущем исследовании [14]. В проекте SI03 с самого начала проводилась проверка требований к программному обеспечению (улучшение процесса разработки программного обеспечения), в то время как в проекте SI02 использовались обычные практики, существовавшие ранее.

В ходе исследования измерялось качество при использовании программных систем, затем полученные результаты сравнивались между собой. Сравнение проводилось с учетом показателей эффективности, результативности и удовлетворенности атрибутом и без учета оценки безопасности и охвата контекстов использования. Такое решение обосновывалось следующими аргументами:

- (i) показатели экономического риска не имеют отношения к делу, поскольку обе программные системы не ставили задач по достижению экономической выгоды,
- (ii) одна из программ обрабатывает конфиденциальную информацию о людях, а другая – нет, сравнивать показатели риска в работе с людьми нецелесообразно,

- (iii) считается, что улучшения, внесенные в процесс валидации требований, не влияют на окружающую среду, по этой причине экологический риск не оценивается, и
- (iv) наличие пользователей было ограничением исследования, что вынудило нас проводить оценку в едином (реальном) контексте, поэтому охват контекста не оценивался.

Табл. 3. Список проектов исследования

Table 3. List of projects studies

Проект	Описание	Группа
SI02	Программная система управления процессом закупок	G-Cop: Контрольная группа
SI03	Программная система мониторинга сотрудников консалтинговой службы	G-Exp: Экспериментальная группа

3.4 Вопросы исследования (P1-4)

В этом исследовании был задан следующий вопрос: “Какие характеристики качества при использовании улучшились в результате принятия мер по проверке требований к программному обеспечению?” Этот вопрос разбит на следующие гипотезы:

- Внедрение методов проверки требований к программному обеспечению позволяет получить программное обеспечение, которое работает более эффективно в специфическом контексте применения.
- Внедрение методов проверки требований к программному обеспечению позволяет получить программное обеспечение, которое работает более результативно в специфическом контексте использования.
- Внедрение методов проверки требований к программному обеспечению позволяет получить программное обеспечение, удовлетворяющее потребностям пользователя в специфическом контексте использования.

Показатели результативности оценивают точность и полноту, с которыми пользователи достигают конкретных целей. Эти показатели не учитывают, как достигаются цели, а только степень, в которой они достигнуты [15]. Показатели эффективности оценивают затраченные ресурсы в зависимости от точности и полноты, с которыми пользователи достигают поставленных целей. Для этого:

- (i) в качестве основного ресурса для измерения эффективности используется время,
- (ii) эффективность пользователя сравнивается с эффективностью эксперта, и
- (iii) при выполнении задач устанавливается лимит времени [15].

Показатели удовлетворенности оценивают степень удовлетворения потребностей пользователей при использовании продукта или системы в специфическом контексте использования [15].

3.5 Методы сбора (P1-5)

Часть первичных данных была получена с помощью методики "Размышление вслух", для чего в протокол оценки программного обеспечения был включен специально разработанный набор задач. Применение этой методики было связано с выполнением следующих заданий:

- (i) идентифицировать участников,
- (ii) выдать участникам запланированные для них задания, и
- (iii) ожидать, что, выполняя задания, пользователи будут вслух выражать свои мысли.

Для протоколирования времени выполнения каждой задачи, допущенных ошибок, ненужных действий и других данных использовались специальные программы, записывавшие состояние экрана. Каждая запись воспроизводилась, и из полученных записей извлекались

следующие данные: количество достигнутых целей, количество допущенных ошибок, количество ненужных действий, количество запросов о помощи, использованное время (минуты), время запроса о помощи (минуты), время устранения ошибок (минуты) и количество жалоб.

Для получения дополнительных данных был использован вопросник PSSUQ, причем вопросы участникам задавались сразу же после того, как они завершали выполнение протокольных заданий. При этом для получения нужных данных использовался следующий инструментарий: протокол оценки, разработанный для каждой программной системы (см. приложение А), и анкета PSSUQ, в которую для получения дополнительной информации для измерения удовлетворенности были добавлены два пункта (см. приложение В). В протокол включались сведения о цели задания, об основных индикаторах, наборы задач, действий и данных, предлагаемых для выполнения задач. Тестовые задачи, включенные в протокол, были предложены и заранее оценены опытным пользователем-экспертом, который устанавливал максимальное время, нужное для выполнения каждой задачи. Время, необходимое для выполнения всех задач в каждом протоколе, оценивалось в 40 минут.

Выполнение протокола было осуществлено после согласования с лицом, ответственным за прикладную область, при этом учитывались три аспекта: во-первых, обеспечение того, чтобы контекст использования был очень похож на реальный, во-вторых, предложение пользователям озвучивать свои мысли во время выполнения задач и, в-третьих, запись действий и комментариев пользователей во время выполнения протокола. Участники выполняли протокольные задания на своих рабочих местах и своих компьютерах в свое рабочее время по расписанию. Каждому участнику были предоставлены наушники, была активирована аудио- и видеозапись.

Число пользователей, выполнивших задачи протокола оценки программной системы SI02, составило 10 человек, то есть 83% от общего числа пользователей. Для системы SI03 также было 10 пользователей, что составляло 90% от общего числа пользователей. Собрать всех пользователей для исследования оказалось невозможно, но количество участников представляет собой значительную выборку. Опыт исполнителей работы [15] показал, что надежные результаты могут быть получены при выборке всего из восьми участников. Все участники пользовались программным обеспечением более пяти месяцев и могут быть квалифицированы как продвинутые пользователи, но не как эксперты. В данном исследовании пользователь – это человек, который работает с программным обеспечением как с основным инструментом своей рабочей деятельности, взаимодействует с ним каждый день и посвящает ему более половины своего рабочего времени.

Численные значения, собранные, обработанные и проанализированные в данном исследовании, представляют собой итоги измерения результативности, эффективности и удовлетворенности пользователей качеством при использовании информационных систем. Смысл показателей задавался на основе стандартов ISO/IEC 25022 [16]:

- (i) показатели результативности оценивают точность и полноту, с которыми пользователи достигают своих специфических целей,
- (ii) показатели эффективности оценивают затраченные ресурсы в зависимости от точности и полноты, с которыми пользователи достигают целей, и
- (iii) показатели удовлетворенности оценивают степень удовлетворения потребностей пользователей при использовании программного продукта в конкретном контексте.

В исследовании была использована только часть показателей результативности, эффективности и удовлетворенности. Кроме того, в качестве метода наблюдения и сбора первичных данных для показателей, предложенных стандартом ISO/IEC 25022, использовалась методика “Размышления вслух”. После завершения исследования для оценки

удобства использования системы в качестве инструмента сбора данных, дополняющих показатели и характеристики удовлетворенности, был использован вопросник PSSUQ.

3.6 Отбор данных (P1-6)

Показатели, предложенные ISO/IEC 25022, требуют первичного набора данных, эти данные были извлечены из собранных видеороликов. Для каждого задания подсчитывались следующие первичные данные: достигнутые цели (OA), количество ошибок (NE), ненужные действия (UA), запросы на помощь (AR), использованное время (TU), время запроса помощи (HRT), время восстановления после ошибки (ERT), количество жалоб (NC). На основе этих данных были рассчитаны показатели результативности и эффективности. С другой стороны, показатели удовлетворенности были рассчитаны с использованием оценок факторов и пунктов PSSUQ.

Для измерения качества при использовании систем SI02 и SI03 применялись показатели ISO/IEC 25022, представленные в табл. 4. Результативность измеряли четырьмя показателями, эффективность – пятью, удовлетворенность – тоже пятью. Кроме того, были использованы три фактора из вопросника PSSUQ. Показатель “Sus-1-G Общая удовлетворенность” измерялся на основе интегральной оценки PSSUQ, а “Sus-2-G Удовлетворенность функциями”, “STI-1-G Доверие пользователей”, “SPI-1-G Удовольствие пользователей”, “Sco-1-G Физический комфорт” на основе баллов, полученных по пунктам 17, 18, 13 и 4 PSSUQ.

4. Результаты

В этом разделе представлены результаты и анализ валидности. Здесь же приведен анализ показателей результативности, эффективности и удовлетворенности. Валидность включает в себя внутреннюю, внешнюю и конструктивную.

4.1 Показатели результативности

В табл. 5 приведены данные по показателям результативности программного обеспечения системы SI02. Значения показаны с детализацией по всем участникам (P01, P02, ...), они были рассчитаны с использованием измерительных функций стандарта ISO/IEC 25022. Чтобы получить обобщенное значение, представляющее показатели на уровне всей системы, рассчитывалось среднее значение. Табл. 6 показывает данные, аналогичные данным табл. 5, но табл. 6 относится к системе SI03.

Значение показателя “Ef-1-G Выполненные задачи” для SI03 больше, чем значение для SI02, что указывает на то, что участники выполнили в системе SI03 больше заданий, чем в системе SI02. Значение показателя “Ef-2-S Достигнутые цели” для SI03 немного выше значения для SI02, это позволяет сделать вывод, что участники достигли того же количества целей с помощью системы SI03, что и с системой SI02. Значение показателя “Ef-3-G Ошибки в задаче” для SI03 меньше значения для SI02, это указывает на то, что участники допустили меньше ошибок в заданиях для SI03, чем в заданиях для SI02. Значение показателя “Ef-4-G Задачи с ошибками” для SI03 меньше значения для SI02, что указывает на то, что в системе SI03 без ошибок было выполнено больше задач, чем в системе SI02.

С использованием данных, приведенных в табл. 5 и табл. 6, были подготовлены диаграммы (см. рис. 2), которые показывают распределение значений показателей результативности для систем SI02 и SI03.

На рис. 2.а по показателю “Ef-1-G Выполненные задачи” видно, что участники, пользовавшиеся системой SI03, выполнили больше задач, чем участники, пользовавшиеся системой SI02. Этот вывод основан на положительном смещении, которое наблюдается для SI03, по сравнению с отрицательным смещением, наблюдаемым для SI02.

Табл. 4. Показатели, взятые из стандартов ISO/IEC 250XX и вопросника PSSUQ и выбранные для измерения качества при использовании

Table 4. Metrics selected to measure the quality in use, taken from ISO/IEC 250XX and PSSUQ

Id	Название	Функция	Источник
Ef-1-G	Выполненные задачи	$X = A / B$ A = Количество выполненных уникальных задач B = Общее количество выполнявшихся уникальных задач	ISO/IEC
Ef-2-S	Достигнутые цели	$\{X = 1 \sum A_i X \geq 0\}$ A_i = Пропорциональное значение каждой пропущенной или неправильной цели в выходных данных задачи (максимальное значение = 1)	ISO/IEC
Ef-3-G	Ошибки в задаче	$X = A$ A = Количество ошибок, допущенных пользователем во время выполнения задачи	ISO/IEC
Ef-4-G	Задачи с ошибками	$X = A / B$ A = Количество задач с ошибками B = Общее количество задач	ISO/IEC
Ey-1-G	Время выполнения задачи	$X = T$ T = Время выполнения задачи	ISO/IEC
Ey-2-S	Эффективность затраченного времени	$X = A / T$ A = Количество достигнутых целей T = Время	ISO/IEC
Ey-3-S	Экономическая эффективность	$X = A / B$ A = Общая стоимость выполнения задачи B = Количество достигнутых целей	ISO/IEC
Ey-4-S	Коэффициент полезного времени	$X = T_a / T_b$ где T_a = Продуктивное время, то есть – время, затраченное на выполнение задачи – время, затраченное на получение помощи – время, затраченное на восстановление после ошибок – время, затраченное на безрезультатный поиск T_b = Время, затраченное на выполнение задачи	ISO/IEC
Ey-5-S	Ненужные действия	$X = A/B$ A = Количество действий, которые фактически не были необходимы для выполнения задачи B = Количество действий, выполненных пользователем	ISO/IEC
SUs-1-G	Общая удовлетворенность	$X = \sum (A_i)$ A_i = Ответ на вопрос	ISO/IEC
SUs-2-G	Удовлетворенность функциями	$X = \sum (A_i)$ A_i = Ответ на вопрос, связанный с определенной характеристикой	ISO/IEC
STr-1-G	Доверие пользователя	$X = A$ A = Значение психометрической шкалы доверия	ISO/IEC
SPI-1-G	Удовольствие пользователя	$X = A$ A = Значение психометрической шкалы удовольствия	ISO/IEC
SCo-1-G	Физический комфорт	$X = A$ A = Значение психометрической шкалы комфорта	ISO/IEC
SysU	Полезность системы (SysUse)	$X = A$ A = Значение психометрической шкалы полезности	PSSUQ
InfoQ	Качество информации (InfoQual)	$X = A$ A = Значение психометрической шкалы качества информации	PSSUQ
InterQ	Качество интерфейса (InterQual)	$X = A$ A = Значение психометрической шкалы качества интерфейса	PSSUQ

На рис. 2.b по показателю “Ef-2-S Достигнутые цели”, видно, что участники, которые использовали систему SI03, достигли в среднем такого же количества целей, что и участники, которые использовали систему SI02. Несмотря на то, что существует положительное смещение для SI03 и отрицательное смещение для SI02, разница в средних значениях очень мала, что не позволяет установить статистические различия.

На рис. 2.c по показателю “Ef-3-G ошибки в задаче”, замечено, что участники, которые использовали ПО SI03, допустили меньше ошибок, чем участники, которые использовали ПО SI02. Это верно, потому что SI03 имеет отрицательное смещение по отношению к асимметричному и концентрированному распределению SI02.

На рис. 2.d по показателю “Ef-4-G задачи с ошибками”, видно, что при использовании SI03 задач с ошибками меньше, чем при использовании SI02. Среднее значение, отрицательное

смещение SI03 по отношению к симметричному распределению SI02 и разброс данных позволяют нам заключить, что показатель SI03 “Ef-4-G задачи с ошибками” лучше, чем показатель SI02.

Табл. 5. Показатели результативности для SI02

Table 5. Software effectiveness metrics for SI02

PU	Ef-1-G	Ef-2-G	Ef-3-G	Ef-4-G
P01	0.81	0.81	7	0.50
P02	0.39	0.42	15	0.83
P03	0.48	0.52	5	0.50
P04	0.61	0.65	6	0.50
P05	0.81	0.84	6	0.50
P06	0.81	0.84	6	0.83
P07	0.81	0.81	7	0.67
P08	0.58	0.65	5	0.67
P09	0.52	0.58	7	0.83
P10	0.84	0.90	8	0.67
Среднее значение	0.66	0.70	7.20	0.65

Табл. 6. Показатели результативности для SI03

Table 6. Software effectiveness metrics for SI03

PU	Ef-1-G	Ef-2-G	Ef-3-G	Ef-4-G
P01	0.63	0.63	0	0.00
P02	0.75	0.75	7	0.57
P03	0.95	0.95	3	0.29
P04	0.90	0.90	2	0.14
P05	0.53	0.53	6	0.43
P06	1.00	1.00	0	0.00
P07	0.63	0.63	4	0.43
P08	0.45	0.45	6	0.43
P09	0.65	0.65	6	0.57
P10	0.65	0.65	9	0.57
Среднее значение	0.71	0.71	4.30	0.34

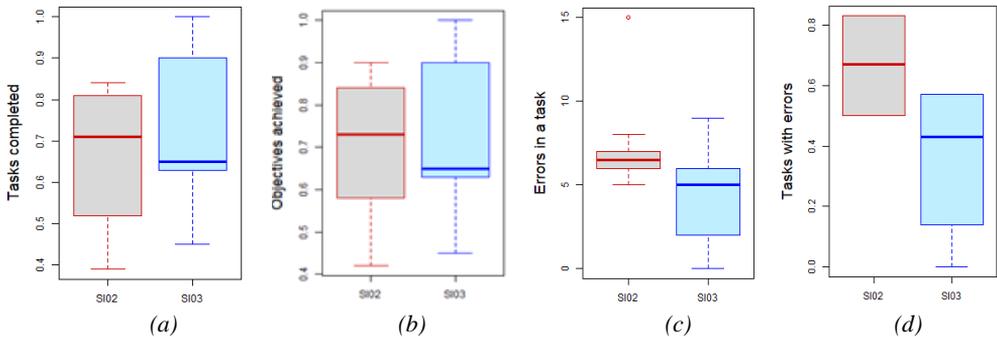


Рис 2. Сравнение показателей результативности для SI02 и SI03: (a) “Ef-1-G Выполненные задачи”, (b) “Ef-2-S Достигнутые цели”, (c) “Ef-3-G Ошибки в задаче”, (d) “Ef-4-G Задачи с ошибками”

Fig 2. Effectiveness metrics comparison for SI02 and SI03: (a) “Ef-1-G Tasks completed”, (b) “Ef-2-S Objectives achieved”, (c) “Ef-3-G Errors in a task”, (d) “Ef-4-G Tasks with errors”

Анализ данных в табл. 5 и табл. 6 и оценка диаграмм рис. 2 позволяют нам сделать вывод, что показатель “Ef-2-S Достигнутые цели” не показывает существенной разницы между SI03 и SI02; в то время как остальные три показателя для SI03 лучше, чем для SI02. Это позволяет сделать вывод, что программная система SI03 обладает большей результативностью при использовании, чем система SI02.

Процент улучшения был рассчитан путем вычисления по формуле

$$(avg02 - avg03) / avg02 \times 100$$

где $avg02$ – среднее значение показателей результативности системы SI02, которое вычислялось до проведения работ по ее улучшению, а $avg03$ – среднее значение показателей системы SI03, полученных после улучшений. Исходя из табл. 5 и 6, улучшение особенно хорошо заметно на примере показателей по количеству ошибок в одной задаче (на 40%) и по общему числу ошибок в задачах (на 47%).

4.2 Показатели эффективности

В табл. 7 и табл. 8 приведены данные показателей эффективности программных систем SI02 и SI03 соответственно. Данные представлены по каждому из участников, они были рассчитаны с использованием измерительных функций, рекомендованных стандартом ISO/IEC 25022. Чтобы получить значение, представляющее показатель на обобщенном уровне системы, также рассчитывалось среднее значение по всем участникам.

Табл. 7. Показатели эффективности для системы SI02

6Table 7. Efficiency metrics for software SI02

PU	Ey1G	Ey2S	Ey3S	Ey4S	Ey5S
P01	49	0.51	1.96	0.88	0.22
P02	41	0.30	3.42	0.83	0.33
P03	47	0.32	3.13	0.94	0.23
P04	47	0.40	2.47	0.94	0.28
P05	45	0.56	1.80	0.80	0.22
P06	46	0.54	1.84	1.00	0.20
P07	55	0.45	2.20	0.96	0.22
P08	49	0.36	2.72	0.92	0.35
P09	51	0.31	3.19	0.94	0.26
P10	53	0.49	2.04	0.89	0.25
Среднее значение	48	0.42	2.48	0.91	0.26

Табл. 8. Показатели эффективности для системы SI03

Table 8. Efficiency metrics for software SI03

PU	Ey1G	Ey2S	Ey3S	Ey4S	Ey5S
P01	48	0.52	1.92	0.96	0.04
P02	49	0.61	1.63	0.82	0.10
P03	64	0.59	1.68	0.94	0.07
P04	41	0.88	1.14	0.93	0.05
P05	53	0.40	2.52	0.83	0.22
P06	47	0.85	1.18	1.00	0.00
P07	52	0.48	2.08	0.96	0.14
P08	40	0.45	2.22	0.90	0.26
P09	40	0.65	1.54	0.93	0.15
P10	40	0.65	1.54	0.83	0.20
Среднее значение	47	0.61	1.75	0.91	0.12

Значение показателя “Ey-1-G Время выполнения задачи” для SI03 немного меньше значения для SI02, это указывает на то, что участники, которые использовали SI03, выполнили задания за более короткое время, чем участники, использовавшие SI02. Значение показателя “Ey-2-S Эффективность затраченного времени” для SI03 выше, чем значение для SI02, это указывает на то, что за время, затраченное на выполнение задач, пользователи программного обеспечения SI03 достигли большего количества целей. Значение показателя “Ey-3-S Экономическая эффективность” для SI02 имеет более высокое значение, чем SI03, это указывает на то, что участникам потребовалось больше времени для выполнения задач SI02. Значение показателя “Ey-4-S Коэффициент полезного времени” для SI02 и SI03 равны и имеют значение 0,91, это указывает на то, что большинство действий были эффективными и способствовали достижению целей задачи. Значение показателя “Ey-5-S Ненужные действия” для SI03 меньше значения для SI02, это указывает на то, что участники выполнили меньше ненужных действий с системой SI03, чем с системой SI02.

На основе данных, приведенных в табл. 7 и табл. 8 были подготовлены диаграммы (см. рис. 3), которые представляют распределение значений показателей эффективности SI02 и SI03.

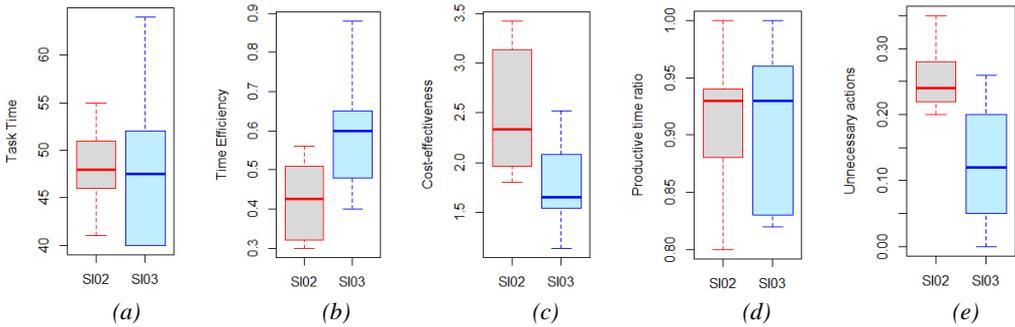


Рис 3. Сравнение показателей эффективности для SI02 и SI03.

(a) “Еу-1-Г Время выполнения задачи”, (b) “Еу-2-С Эффективность затраченного времени”,
(c) “Еу-3-С Экономическая эффективность”, (d) “Еу-4-С Коэффициент полезного времени”,
(e) “Еу-5-Е Ненужные действия”

Fig 3. Efficiency metrics comparison for SI02 and SI03. (a) “Еу-1-Г Task time”,
(b) “Еу-2-С Time efficiency”, (c) “Еу-3-С Cost-effectiveness”, (d) “Еу-4-С Productive time ratio”,
(e) “Еу-5-С Unnecessary actions”

На рис. 3.а показатель “Еу-1-Г Время реализации задачи” указывает на то, что система SI03 для выполнения участниками своих задач требует меньше времени, чем система SI02, это подтверждается отрицательным смещением SI03, симметричным распределением SI02 и средними значениями. Однако из-за высокой изменчивости показателей SI03 и небольшого отличия от медианного значения это утверждение не является решающим.

На рис. 3.б показатель “Еу-2-С Эффективность затраченного времени” показывает, что система SI03 имеет лучшую эффективность по времени по сравнению с системой SI02, низкая дисперсия распределения его данных и медиана, превышающая SI02, подтверждают это утверждение.

На рис. 3.с по показателю “Еу-3-С Экономическая эффективность”, система SI03 имеет более хорошую оценку экономической эффективности, чем система SI02, более концентрированное распределение данных и более низкое среднее значение, что приводит к выводу, что SI03 имеет более высокую оценку эффективности, чем SI02.

Рис. 3.д иллюстрирует показатели “Еу-4-С Коэффициент полезного времени”. Этот коэффициент для системы SI02 статистически эквивалентен аналогичному показателю для системы SI03, он смещен в отрицательную сторону по распределению данных систем, средние значения для двух систем тоже эквивалентны, что подтверждает сделанное утверждение.

Рис. 3.е демонстрирует показатели “Еу-5-С Ненужные действия”. Из него видно, что доля ненужных действий в системе SI03 меньше, чем эта доля в системе SI02. Среднее значение, достигнутое в системе SI03, существенно меньше, чем среднее, полученное в системе SI02, симметричное распределение в системе SI03, положительное смещение в системе SI02 и концентрированная дисперсия в обоих случаях позволяют нам заключить, что показатель “Еу-5-С Ненужные действия” для системы SI03 лучше, чем тот же показатель для системы SI02.

Анализ данных в табл. 7 и табл. 8, а также оценка диаграмм рис. 3 подтверждают, что два показателя эквивалентны или не показывают существенных различий между системами SI03 и SI02. В то же время по трем показателям значения для SI03 лучше, чем для SI02. Это позволяет сделать вывод, что SI03 обладает большей эффективностью в использовании, чем SI02.

Процент улучшения, согласно табл. 7 и 8, значительно повышен для показателя “Эффективность затраченного времени” (45%) и снижен для показателей “Рентабельность” (29%) и “Ненужные действия” (53%).

4.3 Показатели удовлетворенности

В табл. 9 и 10 приведены значения показателей удовлетворенности (см. табл. 4) для программных продуктов SI02 и SI03. Значения были рассчитаны с использованием критериев PSSUQ, а для получения значения, представляющего совокупный показатель по всем участникам, было рассчитано среднее значение. Эти же данные показаны на рис. 4.

Табл. 9. Показатели удовлетворенности для системы SI02

Table 9. Satisfaction metrics for software SI02

PU.	SUs-1-G	SUs-2-G	STr-1-G	SPI-1-G	SCo-1-G	SysU	InfoQ	InterQ
P01	3.25	6.00	2.00	4.00	3.00	2.83	3.50	4.00
P02	5.06	5.00	6.00	7.00	5.00	5.17	4.67	5.67
P03	5.25	4.00	5.00	5.00	5.00	5.00	4.83	6.00
P04	4.94	4.00	7.00	5.00	4.00	5.50	4.00	5.33
P05	3.75	3.00	3.00	5.00	2.00	2.50	5.17	3.67
P06	3.00	2.00	3.00	5.00	3.00	2.17	3.17	4.33
P07	3.19	5.00	2.00	3.00	3.00	2.50	3.17	4.33
P08	3.31	3.00	4.00	1.00	2.00	3.33	3.50	2.33
P09	3.81	5.00	5.00	4.00	4.00	4.00	3.83	3.67
P10	2.63	6.00	1.00	3.00	1.00	2.17	2.67	3.67
Среднее значение	3.82	4.30	3.80	4.20	3.20	3.52	3.85	4.30

Табл. 10. Показатели удовлетворенности для системы SI03

Table 10. Satisfaction metrics for software SI03

PU.	SUs-1-G	SUs-2-G	STr-1-G	SPI-1-G	SCo-1-G	SysU	InfoQ	InterQ
P01	3.13	4.00	4.00	4.00	3.00	3.17	3.00	3.33
P02	3.50	4.00	4.00	2.00	4.00	3.83	3.67	2.33
P03	2.25	5.00	2.00	2.00	2.00	1.67	2.50	3.00
P04	3.00	6.00	2.00	2.00	3.00	2.50	3.50	2.67
P05	3.25	3.00	3.00	3.00	3.00	3.17	3.50	3.00
P06	2.88	5.00	4.00	3.00	2.00	2.50	3.00	3.33
P07	1.63	2.00	1.00	1.00	1.00	1.50	1.83	1.33
P08	2.31	3.00	2.00	2.00	2.00	1.83	2.50	2.67
P09	4.25	5.00	5.00	4.00	3.00	3.00	5.33	4.67
P10	2.56	3.00	2.00	3.00	2.00	2.33	2.50	3.00
Среднее значение	2.88	4.00	2.90	2.60	2.50	2.55	3.13	2.93

На рис. 4 для сравнения показателей удовлетворенности использованы гистограммы. Оцениваемые системы обозначены полосами разных цветов, красный цвет выбран для системы SI02, синий для SI03. На рис. 4 показано, что показатели удовлетворенности SI03 лучше, чем эти показатели для SI02, так как согласно критериям PSSUQ, более низкие баллы показывают более полное удовлетворение программной системой.

Рис. 5.а показывает данные по показателю “SUs-1-G Общая удовлетворенность”, по нему видно, что система SI03 позволяет пользователям достичь большего удовлетворения от работы по сравнению с системой SI02. Более низкое среднее значение, отрицательное смещение данных и меньшая дисперсия позволяет сделать заключение, что система SI03 полнее удовлетворяет пользователей, чем система SI02. Более того, можно сделать вывод, что разница между системами статистически значима, поскольку на рис. 5.а видно, что продолжение одной из медиан не пересекает прямоугольник другой.

Рис. 5.б построен на основании данных по показателю “SUs-2-G Удовлетворенность функциями”, этот показатель измеряется с помощью пункта 17 вопросника и связан с восприятием пользователем полноты функциональных возможностей. Рис. 5.б показывает, что пользователи считают функциональные возможности обеих систем SI02 и SI03 неполными, это видно из того, что полученные оценки в основном сосредоточены между значениями 3 (“частично согласен”) и 5 (“частично не согласен”). Близость значений дисперсии, разброс оценок и их среднее значение позволяют нам считать, что по показателю “SUs-2-G Удовлетворенность функциями” системы SI02 и SI03 эквивалентны.

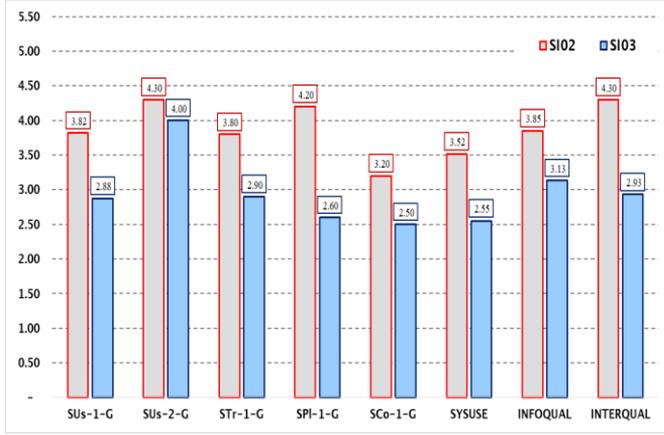


Рис 4. Сравнение показателей удовлетворенности, гистограмма средних значений
Fig 4. Satisfaction metrics comparison, bar chart of averages

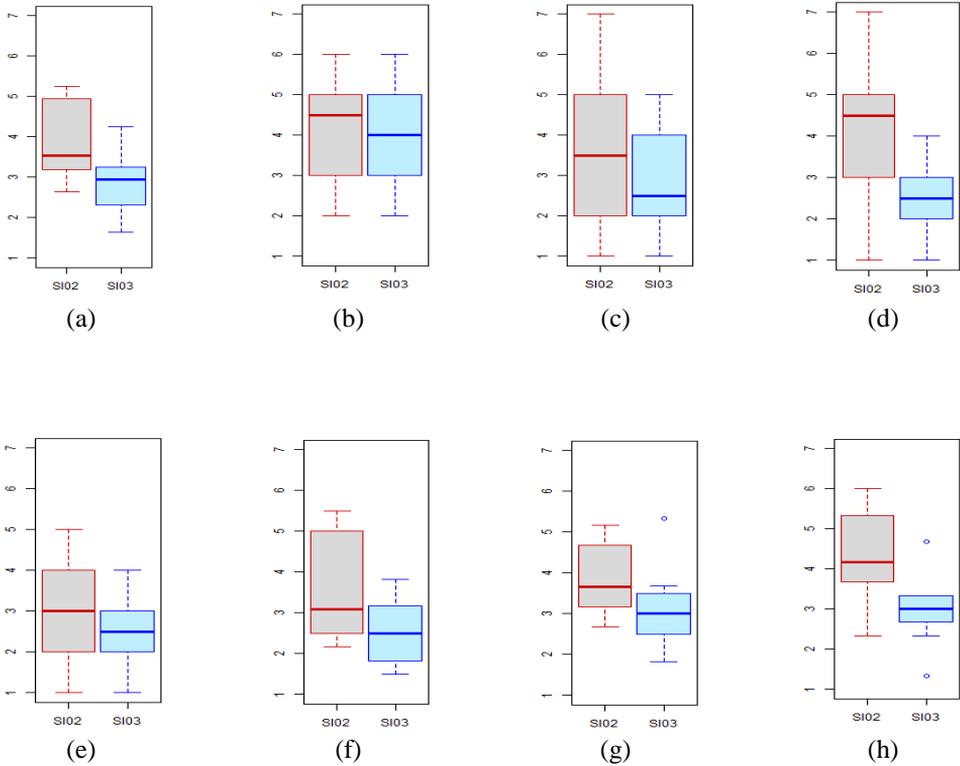


Рис 5. Сравнение показателей удовлетворенности для SI02 и SI03,
(a) “SUs-1-G Overall satisfaction”, (b) “SUs-2-G Satisfaction with features”, (c) “STr-1-G User trust”,
(d) “SPI-1-G User pleasure”, (e) “SCo-1-G Physical comfort”, (f) “SysUse System Usefulness”,
(g) “InfoQual Information Quality”, (h) “InterQual Interface Quality”
Fig 5. Satisfaction metrics comparison for SI02 and SI03,
(a) “SUs-1-G Overall satisfaction”, (b) “SUs-2-G Satisfaction with features”, (c) “STr-1-G User trust”,
(d) “SPI-1-G User pleasure”, (e) “SCo-1-G Physical comfort”, (f) “SysUse System Usefulness”,
(g) “InfoQual Information Quality”, (h) “InterQual Interface Quality”

Рис. 5.с демонстрирует показатель “STr-1-G Доверие пользователей”, этот показатель измеряется с помощью пункта 18 вопросника, здесь видно, что пользователи больше доверяют системе SI03, чем системе SI02. Подтверждение этого утверждения находим в том, что дисперсия и разброс оценок при более низких абсолютных значениях тоже низки.

Рис. 5.d иллюстрирует показатель “SP1-1-G Удовольствие пользователя”, этот показатель измеряется с помощью пункта 13 вопросника. Видно, что пользователи воспринимают систему SI03 как более приятную, чем систему SI02. Подтверждение этого утверждения также находим в том, что дисперсия и разброс оценок при более низких абсолютных значениях низки. Кроме того, продолжение медианы значений для одной из систем не пересекает прямоугольник значений для другой системы, что делает разницу оценок статистически значимой, а вывод убедительным.

Рис. 5.e составлен по данным для показателя “SCo-1-G Физический комфорт”, этот показатель измеряется с помощью пункта 4 вопросника, из него видно, что пользователи считают, что система SI03 обеспечивает более высокий уровень комфорта, чем система SI02. Подтверждается это утверждение тем, оценки по этому показателю имеют меньший разброс, а также меньшую дисперсию при более низких абсолютных значениях.

Рис. 5.f составлен по данным для показателя “SysUse Полезность системы”. Этот рисунок показывает, что пользователи считают систему SI03 более полезной, чем систему SI02. Полученные оценки имеют меньший разброс и меньшую дисперсию при меньших абсолютных значениях, что подтверждает это утверждение.

На рисунке 5.g показан показатель “InfoQual Качество информации”. Легко видеть, что по мнению пользователей система SI03 предлагает информацию более высокого качества, чем система SI02. Можно видеть не очень большой разброс значений, меньшую дисперсию при небольших абсолютных значениях, симметричное распределение оценок системы SI03 по сравнению с положительно смещенным распределением оценок SI02. Продолжение линии медианы одной из систем не пересекает поле значений другой системы, что делает разницу показателей статистически значимой и вывод убедительным.

На рис. 5.h демонстрируется диаграмма для показателя “InterQual Качество интерфейса”. По ней видно, что пользователи считают, что у системы SI03 более хороший пользовательский интерфейс, чем у системы SI02. Здесь видны небольшой разброс значений, небольшое значение дисперсии при небольших абсолютных значениях показателя, симметричное распределение значений системы SI03 по сравнению с положительным смещением распределения значений для системы SI02. Продолжение линии медианы одной из систем не пересекает поле значений другой системы, что делает разницу показателей статистически значимой и вывод убедительным.

Анализ содержания табл. 9 и табл. 10, а также рис. 4 и рис. 5 показывают, что из восьми показателей, измеряющих удовлетворенность использованием, семь благоприятны для SI03, а по одному показателю система SI03 эквивалентна системе SI02. Это позволяет сделать вывод, что качество при использовании системы SI03 выше, чем качество при использовании системы SI02.

4.4 Угрозы валидности

Чтобы гарантировать валидность проведенного исследования, была проведена валидация требований к качеству программ и собранной информации, использовавшейся для расчета показателей. Валидация проводилась путем получения экспертных оценок с использованием критериев ясности, объективности, организованности, возможности, актуальности, достаточности, согласованности и непротиворечивости. Была применена методология, предложенная Эрнандесом-Ньетто [32], где в качестве измерения использовался коэффициент валидности контента (Content Validity Coefficient, CVC). Коэффициент CVC оценивает степень согласия экспертов относительно каждого из пунктов и инструмента в целом. Был

получен коэффициент 0,8719, который квалифицирует приборы и конструкцию как валидные, сама процедура валидации была описана в [33].

Такие понятия, как качество при использовании, результативность при использовании и удовлетворенность при использовании, широко изучаются разработчиками программного обеспечения и специалистами по использованию программ. Эти концепции были пересмотрены, согласованы и стандартизированы в рамках серии стандартов ISO/IEC 25000.

- Внутренняя валидность. Это исследование сосредоточено на измерении качества программного обеспечения с использованием модели, предложенной ISO/IEC 25010. Для обеспечения валидности измерений были использованы: показатели и рекомендации из ISO/IEC 25022, стандартизированный вопросник PSSUQ и методика тестирования "Размышления вслух".

Чтобы обеспечить валидность исследования, контекст, в котором проводится оценка программного обеспечения, в точности совпадает с реальным: физическая среда, компьютерное оборудование, графики работы, пользователи и даже мешающие элементы были такими же, как в реальной обстановке.

- Внешняя валидность. Исследование ограничено сравнением двух программных продуктов – информационных систем, поддерживающих жизненный цикл разработки программ в общественном учреждении, и результаты не поддаются обобщению. Он представлен научному сообществу разработчиков программного обеспечения в качестве эмпирического исследования, в котором устанавливается взаимосвязь между качеством требований и качеством при использовании программного продукта. Утверждения являются результатами описательного статистического анализа.

5. Завершающее обсуждение и будущая работа

В этом исследовании оценивалось и сравнивалось качество при использовании двух программных продуктов, которые участвовали в другом, ранее проведенном исследовании. Анализ показателей продемонстрировал, что программное обеспечение системы SI03 обладает лучшим качеством при использовании, чем SI02. В исследовании [14] было определено качество требований к программному обеспечению (SRQI), и система SI02 получила оценку SRQI, равную 0,5054, а система SI03 получила значение SRQI, равное 0,8495. Разница была объяснена улучшением процесса проверки требований в SI03. Мы считаем, что на разницу в качестве при использовании также повлияло внедрение методов проверки на этапе формирования требований к программам, аналогичных тем, что описаны в работе [34], подтверждающих, что качество процесса влияет на качество продукта, как указано в работе [29].

Анализ результатов показывает, что внедрение методов валидации на этапе разработки требований оказывает положительное влияние на качество в использовании. Результаты выражаются в терминах:

- (i) точной и добросовестной работы с ними для достижения их целей,
- (ii) оптимизации используемых ресурсов или повышения их производительности, и
- (iii) удовлетворенности работой пользователя в отношении используемого программного обеспечения.

Все это вместе повлияло на то, что пользователи оценили систему SI03 лучше, чем систему SI02, что соответствует результатам исследования [35]. Подводя итог, можно утверждать, что валидация требований позволяет получить программное обеспечение с лучшим качеством при использовании.

Результативность, эффективность и удовлетворенность моделью качества при использовании ISO/IEC 25010 – это показатели, важные для сравнения качества при

использовании двух программных продуктов, это утверждение согласуется с выводами работ [36-37, 22].

Применение методологий "Размышления вслух" и PSSUQ подтверждает правильность исследования, и, в частности, PSSUQ позволяет избежать повторного возникновения вопросов и беспорядка в их постановке, решая одну из проблем измерения удовлетворенности, поднятых в работе [38].

При выборе показателей удовлетворенности на данном этапе была исключена группа показателей из стандарта ISO/IEC 25022, которые измеряют полезность; в качестве будущей работы предлагается измерять эти показатели путем автоматического сбора данных и статистического моделирования для анализа, предложенного в работе [39].

Наконец, в этом исследовании доступ к информации о конкретных примерах был ограничен. Было невозможно разработать экспериментальную модель, которая объясняла бы с помощью проверки гипотез взаимосвязь между валидацией требований к программному обеспечению и качеством при использовании.

Приложения А и В на испанском языке оригинала доступны по ссылке: https://drive.google.com/open?id=169Z4QhsO4nVxoRaDj_kEndoGS7JdyBey.

Список литературы / References

- [1]. PCM: RM. N° 041-2017-PCM. Uso Obligatorio de la Norma Técnica Peruana NTP-ISO/IEC 12207:2016, Peru (2017).
- [2]. ISO/IEC: ISO/IEC TR 29110-1:2016 Systems and software engineering — Lifecycle profiles for Very Small Entities (VSEs) — Part 1: Overview, Geneva (2016).
- [3]. Demirel, S.T., Das, R.: Software requirement analysis: Research challenges and technical approaches. In: 6th International Symposium on Digital Forensic and Security, ISDFS 2018 - Proceeding. pp. 1–6. Institute of Electrical and Electronics Engineers Inc. (2018). <https://doi.org/10.1109/ISDFS.2018.8355322>.
- [4]. Atoum, I., Baklizi, M.K., Alsmadi, I., Otoom, A.A., Alhersh, T., Ababneh, J., Almalki, J., Alshahrani, S.M.: Challenges of Software Requirements Quality Assurance and Validation: A Systematic Literature Review, (2021). <https://doi.org/10.1109/ACCESS.2021.3117989>.
- [5]. Anuar, U., Ahmad, S., Emran, N.A.: A simplified systematic literature review: Improving Software Requirements Specification quality with boilerplates. In: 2015 9th Malaysian Software Engineering Conference, MySEC 2015. pp. 99–105. Institute of Electrical and Electronics Engineers Inc. (2015). <https://doi.org/10.1109/MySEC.2015.7475203>.
- [6]. García-Mireles, G.A.: Addressing product quality characteristics using the ISO/IEC 29110. In: Mejia J., Munoz M., Rocha Á., C.-M.J. (ed.) Trends and Applications in Software Engineering. Advances in Intelligent Systems and Computing. pp. 25–34. Springer, Cham, Sinaloa (2016). https://doi.org/https://doi.org/10.1007/978-3-319-26285-7_3.
- [7]. Hussain, A., Mkpojiogu, E.O.C., Kamal, F.M.: The Role of Requirements in the Success or Failure of Software Projects. In: International Soft Science Conference (ISSC) in International Review of Management and Marketing. pp. 306–311. EconjournalsLimonluk Mah. 24117 sokSahinRezidance, A6 33100 Yenisehir, Langkawi Island, Kedah, Malaysia (2016).
- [8]. Niazi, M., Mahmood, S., Alshayeb, M., Qureshi, A.M., Faisal, K., Cerpa, N.: Toward Successful Project Management in Global Software Development. *International Journal of Project Management*. 34, 1553–1567 (2016). <https://doi.org/https://doi.org/10.1016/j.ijproman.2016.08.008>.
- [9]. Bhardwaj, M., Rana, A.: Key Software Metrics and its Impact on each other for Software Development Projects. *International Journal of Electrical and Computer Engineering (IJECE)*. 6, 242–248 (2016). <https://doi.org/https://doi.org/10.11591/ijece.v6i1.8247>.
- [10]. Duran Toro, A.: Un Entorno Metodológico de Ingeniería de Requisitos para Sistemas de Información, <https://idus.us.es/handle/11441/15365>, (2000).
- [11]. Mishra, D., Abdalhamid, S.: Software Quality Issues in SCRUM: A Systematic Mapping. *Journal of Universal Computer Science*. 24, 1690–1716 (2018).
- [12]. Atoum, I.: A novel framework for measuring software quality-in-use based on semantic similarity and sentiment analysis of software reviews. *Journal of King Saud University - Computer and Information Sciences*. 32, 113–125 (2020). <https://doi.org/10.1016/j.jksuci.2018.04.012>.

- [13]. Salomón, S., Duque, R., Montaña, J.L., Tenés, L.: Towards automatic evaluation of the Quality-in-Use in context-aware software systems. *Journal of Ambient Intelligence and Humanized Computing*. (2022). <https://doi.org/10.1007/s12652-021-03693-w>.
- [14]. Canchari, L., Dávila, A.: Requirements Validation in the Information Systems Software Development: An Empirical Evaluation of Its Benefits for a Public Institution in Lima. In: Mejia J., Muñoz M., Rocha Á., A.C.-M.J. (ed.) *Trends and Applications in Software Engineering*. CIMPS 2019. *Advances in Intelligent Systems and Computing*. pp. 23–35. Springer Nature Switzerland AG 2020, Guanajuato, Mexico (2020). https://doi.org/https://doi.org/10.1007/978-3-030-33547-2_3.
- [15]. ISO/IEC: ISO/IEC 25022:2016 Systems and software engineering — Systems and software quality requirements and evaluation (SQuaRE) — Measurement of quality in use. Springer US, Geneva (2016).
- [16]. Kim, S.-H., Kim, W.-J.: Evaluation of Software Quality-in-use Attributes Based on Analysis Network Process. *Cluster Computing*. 22, 2101–2114 (2019). <https://doi.org/https://doi.org/10.1007/s10586-018-2309-6>.
- [17]. ISO/IEC: ISO/IEC 25001:2014 Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Planning and management, Geneva (2014).
- [18]. Marín, B., Condori-Fernández, N., Pastor, O.: Calidad en Modelos Conceptuales: Un Análisis Multidimensional de Modelos Cuantitativos Basados en la ISO 9126. In: VIII Conferencia Anual de la Asociación Española de Métricas de Sistemas Informáticos. In *Revista de Procesos y Métricas (RPM) - AEMES*. pp. 153–167 (2007).
- [19]. ISO/IEC: ISO/IEC 25010:2011 Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models, Geneva (2011).
- [20]. Fenton, N., Nell, M.: Software Metrics: Roadmap. In: *Conference on the Future of Software Engineering - In 22nd International Conference on Software Engineering - ICSE*. pp. 357–370, Limerick Ireland (2000). <https://doi.org/https://doi.org/10.1145/336512.336588>.
- [21]. Estayno, M., Dapozo, G., Cuenca Pletsch, L., Greiner, C.: Modelos y Métricas para Evaluar Calidad de Software. In: XI Workshop de Investigadores en Ciencias de la Computación. pp. 382–388, San Juan, Argentina (2009).
- [22]. Bevan, N.: Los Nuevos Modelos de ISO para la Calidad y la Calidad en Uso del Software. In: Calero, C., Moraga, A., and Piattini, M. (eds.) *Calidad del Producto y Proceso Software*. pp. 55–78. RA-MA, Madrid - España (2012).
- [23]. Kurosu, M.: Usability, Quality in Use and the Model of Quality Characteristics. In: *Human-Computer Interaction: Design and Evaluation - HCI*. In *Lecture Notes in Computer Science*. pp. 227–237. Springer International Publishing Switzerland 2015, Los Angeles (2015). https://doi.org/10.1007/978-3-319-20901-2_21.
- [24]. Atoum, I., Bong, C.H., Kulathuramaiyer, N.: Towards Resolving Software Quality-in-Use Measurement Challenges. *Journal of Emerging Trends in Computing and Information Sciences*. 5, 877–885 (2014).
- [25]. Charters, E.: The Use of Think-aloud Methods in Qualitative Research An Introduction to Think-aloud Methods. *Brock Education Journal*. 12, 68–82 (2003). <https://doi.org/https://doi.org/10.26522/brocked.V12I2.38>.
- [26]. Nielsen, J.: Thinking Aloud: The #1 Usability Tool, <https://www.nngroup.com/articles/thinking-aloud-the-1-usability-tool/>, last accessed 2021/05/30.
- [27]. Sauro, J.: 10 Things to Know About the Post Study System Usability Questionnaire, <https://measuringu.com/pssuq/>, last accessed 2019/03/03.
- [28]. Allasi, D., Dávila, A.: Financial impact on the adoption of software validation tasks in the analysis phase: A business case. In: Mejia J., Muñoz M., Rocha Á., Quiñonez Y., C.-M.J. (ed.) *Trends and Applications in Software Engineering*. CIMPS 2017. *Advances in Intelligent Systems and Computing*. pp. 106–116. Springer International Publishing AG 2018, Zacatecas, Mexico (2018). https://doi.org/https://doi.org/10.1007/978-3-319-69341-5_10.
- [29]. Pardo, C., García, F., Pino, F., Piattini, M.: Producto y Proceso: Una Relación Compleja en la Ingeniería de Software. *El Hombre y la Máquina*. 67–72 (2013).
- [30]. Runeson, P., Höst, M.: Guidelines for Conducting and Reporting Case Study Research in Software Engineering. *Empirical Software Engineering*. 14, 131–164 (2009). <https://doi.org/https://doi.org/10.1007/s10664-008-9102-8>.
- [31]. Genero Bocco, M., Cruz-Lemus, J.A., PiattiniVelthuis, M.G.: Métodos de Investigación en Ingeniería de Software. Ra-Ma, Madrid, España (2013).

- [32]. Pedrosa, I., Suárez-Álvarez, J., García-Cueto, E.: Evidencias sobre la Validez de Contenido: Avances Teóricos y Métodos para su Estimación. *Acción Psicológica*. 10, 3–18 (2014). <https://doi.org/https://doi.org/10.5944/ap.10.2.11820>.
- [33]. Canchari, L.: La validación de requisitos de software como base del éxito de los proyectos de sistemas informáticos desarrollados e implementados en la Comisión Nacional para el Desarrollo y Vida sin Drogas-DEVIDA. (2018).
- [34]. Pérez-Verdejo, J.M., Sánchez-García, J., Ocharán-Hernández, J.O., Mezura-Montes, E., Cortés-Verdín, K.: Requirements and GitHub Issues: An Automated Approach for Quality Requirements Classification. *Programming and Computer Software*. 47, (2021). <https://doi.org/10.1134/S0361768821080193>.
- [35]. González-Sánchez, J.-L., Montero-Simarro, F., Gutiérrez-Vela, F.-L.: Evolución del Concepto de Usabilidad como Indicador de Calidad del Software. *El Profesional de la Información*. 21, 529–536 (2012). <https://doi.org/https://doi.org/10.3145/epi.2012.sep.13>.
- [36]. Covella, G., Olsina, L.: Medición y Evaluación de Calidad en Uso: Un Caso de Estudio para una Aplicación E-Learning. In: Castro, J., Cernuzzi, L., and Gordillo, S.E. (eds.) IX Conferencia Iberoamericana de Software Engineering - CibSE. pp. 317–330. , La Plata, Argentina (2006).
- [37]. Sierra González, J.C.: Métodos de Evaluación de Usabilidad para Sistemas de Información Web: Una Revisión. In: Conferencia Colombiana en Gestión de Sistemas de Información y de TIC - GSTIC. pp. 1–13. Universidad Nacional de Colombia, Manizales (2013).
- [38]. Hornbæk, K.: Current Practice in Measuring Usability: Challenges to Usability Studies and Research. *International Journal of Human Computer Studies*. 64, 79–102 (2006). <https://doi.org/10.1016/j.ijhcs.2005.06.002>.
- [39]. Dávila Nicanor, L., MejíaAlvarez, P.: Evaluación de la Calidad de Software en Sistemas de Información en Internet. In: Congreso de IngenieríaEléctrica, CInvestAv-IPN. pp. 1–11. , Zacatenco (2003).

Информация об авторах / Information about authors

Luis CANCHARI, Universidad Continental, Huancayo, Perú - Master's degree in Systems Engineering, Computer Engineer, Specialist in Government and Digital Transformation, Software Project Manager. Computer Engineer with a master's degree in Systems Engineering. Experienced in planning, directing, and executing information technology projects in the public sector. Responsible for implementing technological tools to optimize management of hospital centers in ESSALUD (Junín Network and INCOR). Part of a group of IT professionals who monitor servers and services of the SUNAT systems. Academic and researcher in software quality and the application of ICTs in educational innovation. Education degree holder, with experience in both basic and university teaching.

Луис КАНЧАРИ, Универсидад Континенталь, Хуанкайо, Перу - Магистр системной инженерии, инженер информатики, специалист по управлению и цифровому преобразованию, управляющий проектами программного обеспечения. Инженер-информатик с магистерской степенью в системной инженерии. Опыт в планировании, руководстве и выполнении проектов в области информационных технологий в государственном секторе. Ответственный за внедрение технологических инструментов для оптимизации управления больничными центрами в ESSALUD (Red Junín e INCOR). Был частью команды информатиков, отслеживающих серверы и службы систем SUNAT. Академический исследователь в области качества программного обеспечения и применения ИКТ в инновационном образовании. Преподавал в основной и высшей школе.

Paula Maria ANGELERI, Universidad de Belgrano, Buenos Aires, Argentina. Paula Maria Angeleri currently works at the Information Technology, University of Belgrano - Universidad de Belgrano, Argentina (UB). Paula does research in Teaching Methods, Software Engineering and Information Systems (Business Informatics). She is the research director of MyFEPS project (2010-current) MyFEPS': Methodologies and Framework for software product evaluation (Metodologías y Framework para la Evaluación de Productos Software). Angeleri is an expert in IT standardization, and she participates actively in ISO/IEC JTC1/SC7 Software and Systems Engineering Subcommittee. She is an expert in the field of Software Quality Standards (SW processes,

management systems, software products, IT services and information security, among others standards)

Паула Мария АНГЕЛЕРИ Университет Бельграно, Буэнос-Айрес, Аргентина. Паула Мария Ангелери настоящее время работает в Университете Бельграно, Буэнос-Айрес, Аргентина. Паула занимается исследованиями в области методов преподавания, программной инженерии и информационных систем (бизнес-информатики). Она является руководителем исследовательского проекта "МайФЭПС" (с 2010 года и по настоящее время) "Методологии и Фреймворк для оценки программного обеспечения". Ангелери является экспертом в стандартизации информационных технологий и активно участвует в ИСО/МЭК JTC1/SC7 Подкомитете по программной и системной инженерии. Она является экспертом в области стандартов качества программного обеспечения (процессы ПО, системы управления, программные продукты, ИТ-услуги и информационная безопасность, среди прочих стандартов).

Abraham Eliseo DÁVILA RAMÓN, Pontificia, Universidad Católica del Perú, Lima, Perú. Abraham Dávila is a researcher and professor at the Pontifical Catholic University of Peru (PUCP) since 2000. He is a principal research of the ProCalProSer Project (2013-2016 Phase I and 2017-2019 Phase II) and he is a founding member of the research group in software engineering (GIDIS-PUCP). Master's degree in Computer Science from PUCP and bachelor's degree in science with a major in Mechanical Engineering. He is member of the ISO/IEC working group of ISO/IEC 29110 standards. Their main areas of interest are computer quality (at the level of software process, products and service management) and education in software engineering.

Авраам Элисео ДАВИЛА РАМОН, Папский католический университет Перу, Лима, Перу. Авраам Давила — исследователь и профессор Папского католического университета Перу (PUCP) с 2000 года. Он является главным научным сотрудником проекта ProCalProSer (Фаза I 2013–2016 гг. и Фаза II 2017–2019 гг.), а также одним из основателей исследовательской группы в области разработки программного обеспечения (GIDIS-PUCP). Степень магистра компьютерных наук PUCP и степень бакалавра наук по специальности «Инженер-механик». Является членом рабочей группы ISO/IEC по стандартам ISO/IEC 29110. Основные области интересов — качество компьютеров (на уровне управления программными процессами, продуктами и услугами) и образование в области разработки программного обеспечения.

