

ISSN 2500-316X (Online)

<https://doi.org/10.32362/2500-316X-2019-7-6-9-24>



УДК 004.415.53(075)

О тестировании удобства использования веб-приложений

Б.М. Басок^{1,@},
А.Н. Рожанская¹,
С.Л. Френкель²

¹МИРЭА – Российский технологический университет, Москва 119454, Россия

²Федеральный исследовательский центр «Информатика и управление», Российская академия наук, Москва 119333, Россия

@Автор для переписки, e-mail: basok@mirea.ru

В статье рассматривается задача организации тестирования удобства использования находящихся в опытной или промышленной эксплуатации веб-приложений (юзабилити-тестирование). Описываются методика юзабилити-тестирования, объединяющая разработку, отладку и выполнение тестовых сценариев, и анализ полученных данных тестирования. Разработка тестовых сценариев основывается на использовании накопленного опыта пользователей, уже работавших с этим приложением, или с другим, схожим по функциональному назначению, на применении средств веб-аналитики, которые могут предоставить сценарии поведения данных пользователей на сайте, на использовании статистических данных по посещениям конкретных страниц. Одновременно с данным подходом при построении тестов используется другой подход, основывающийся на разработке тестов, ориентированных на выявление предполагаемых в программе дефектов. Отладка и выполнение тестовых заданий осуществляется аналогично тому, как это реализуется при выполнении функционального тестирования веб-приложений с использованием средств автоматизации тестирования. Кроме того, анализ полученных при эксплуатации данных с помощью средств веб-аналитики позволяет сформировать группу тестировщиков-респондентов, возможности которых отражают возможности всей совокупности вероятных пользователей данного веб-приложения. Изложенные в работе подходы были реализованы на практике. В статье в качестве примера приводятся данные тестирования страницы приемной комиссии МИРЭА – Российского технологического университета – riem.mirea.ru. Полученные экспериментальные данные показали, что, несмотря на тестирование удобства использования программного обеспечения на этапах разработки, некоторые ошибки при эксплуатации веб-приложений остаются необнаруженными. Цена обнаружения и ликвидации

данных ошибок при этом существенно возрастает, поэтому в работе рекомендуется повышать уровень юзабилити уже на ранних этапах разработки. В частности, для оперативного предсказания уровня юзабилити желательно иметь математические средства моделирования поведения проектируемой системы и пользователя.

Ключевые слова: веб-приложение, юзабилити-тестирование, методика тестирования, средства веб-аналитики.

Для цитирования: Басок Б.М., Рожанская А.Н., Френкель С.Л. О тестировании удобства использования веб-приложений. *Российский технологический журнал*. 2019;7(6):9-24. <https://doi.org/10.32362/2500-316X-2019-7-6-9-24>

On web-applications usability testing

Boris M. Basok^{1,@},
Alena N. Rozhanskaya¹,
Sergey L. Frenkel²

¹MIREA – Russian Technological University, Moscow 119454, Russia

²Federal Research Center “Computer Science and Control,” Russian Academy of Sciences, Moscow 119333, Russia

@Corresponding author, e-mail: basok@mirea.ru

The paper discusses the task of organizing the usability testing of web applications that are in pilot or industrial operation. A usability testing technique is described that combines the development, debugging and execution of test scenarios, and analysis of the received test data. The development of test scenarios is based: on the use of the accumulated experience of users who have already worked with this application or with another, similar in functionality; on the use of web analytics, which can provide scenarios of the behavior of user data on the site; and on the use of statistical data on visits to specific pages. Simultaneously with this approach, a different approach is used in the construction of tests. It is based on the development of tests aimed at identifying defects in the program. Debugging and execution of test tasks is carried out in the same way as it is realized when performing functional testing of web applications using testing automation tools. In addition, the analysis of the data obtained during the operation using web analytics makes it possible to form a group of respondent testers, whose capabilities reflect the capabilities of the entire set of probable users of this web application. The approaches outlined in the work were put into practice. As an example, the article provides test data for the page of the admissions committee of MIREA – Russian Technological University – priem.mirea.ru. The obtained experimental data showed that, despite testing the usability of software at the development stages, some errors in the operation of web applications remain undetected. The cost of detecting and eliminating these errors increases significantly. Therefore, it is recommended in the work to increase the level of usability already in the early stages of development. In particular, for the operational prediction of the level of usability, it is desirable to have mathematical tools for modeling the behavior of the designed system and the user.

Keywords: web application, usability testing, testing methodology, web analytics.

For citation: Basok B.M., Rozhanskaya A.N., Frenkel S.L. On web applications usability testing. *Rossiiskii tekhnologicheskii zhurnal* = Russian Technological Journal. 2019;7(6):9-24 (in Russ.). <https://doi.org/10.32362/2500-316X-2019-7-6-9-24>

Введение

Важнейшей характеристикой любой интерактивной системы является удобство ее использования. Удобство использования (*usability*), определяемое в стандарте ISO/IEC 25066-2016¹ как степень, в которой система, продукт или услуга могут использоваться пользователями для достижения определенных целей с эффективностью, результативностью и удовлетворенностью в определенном контексте применения. Удобство использования (в дальнейшем – юзабилити) графического интерфейса пользователя в программных системах – это качественный признак, который определяет, насколько интерфейс пользователя удобен в эксплуатации. Другими словами, интерфейс считается удобным в том случае, когда для выполнения необходимых действий пользователю требуется достаточно малые затраты времени для чтения, изучения и анализа.

Согласно опытным данным [1], на обеспечение юзабилити тратится до десяти процентов бюджета проекта. Повышенное внимание к юзабилити, начиная с ранних стадий разработки программных продуктов, позволит на этапе их эксплуатации добиться увеличения производительности, продаж и доходов, уменьшения затрат на обучение, развитие и обслуживание.

На уровень юзабилити программных систем, как следует из его определения, влияют эффективность, результативность и удовлетворенность, поэтому его оценка, в конечном счете, сводится к оценке данных свойств анализируемой системы. Существует множество методов оценки юзабилити, но основной из них и самый полезный – это метод тестирования. Благодаря юзабилити-тестированию удастся с помощью эксперимента измерить основные характеристики взаимодействия пользователя с системой.

В данной работе рассматривается задача организации юзабилити-тестирования веб-приложений, находящихся в опытной или в промышленной эксплуатации, описывается методика тестирования, основывающаяся на классических подходах к тестированию программных систем [1, 2]. В основу данной методики положены разработка и отладка тестовых сценариев, ориентированных на выявление ошибок пользователей и используемых ими программ. Предлагаемая методика одновременно использует совокупность различных методов юзабилити-тестирования с учетом их достоинств и недостатков, в том числе анкетирование и опрос пользователей. В данной методике широко используются средства автоматизации тестирования и веб-аналитики. В работе приводятся некоторые экспериментальные данные применения разработанных авторами подходов.

1. Основные понятия и определения

Прежде чем рассматривать вопросы тестирования удобства использования веб-приложений, приведем несколько определений и понятий с учетом стандарта ISO/IEC 25066-2016:

Под эффективностью (*effectiveness*) понимается связь между достигнутым результатом и использованными ресурсами. В качестве показателей эффективности юзабилити можно использовать:

- скорость работы пользователя;
- скорость обучения;
- количество человеческих ошибок.

Под результативностью (*efficiency*) понимается степень реализации запланированной деятельности и достижения запланированных результатов. К показателям результативности юзабилити можно отнести:

¹ Международный стандарт ISO/IEC 25066:2016. Systems and software engineering – Systems and software Quality Requirements and Evaluation (SQuaRE) – Common Industry Format (CIF) for Usability – Evaluation Report.

- соотношение выполненных и невыполненных заданий;
- уровень нагрузки на пользователя. Этот показатель сложно измерить непосредственно, но косвенно он проявляется в показателях эффективности, например, в зависимости скорости работы от интенсивности нагрузки.

Под удовлетворенностью юзабилити понимается отсутствие дискомфорта и положительное отношение к использованию продукции.

Дефект удобства использования (*usability defect*) – это свойства продукта, которые приводят к несоответствию между намерениями и/или действиями пользователя, а также системными свойствами и поведением.

Типичные дефекты удобства использования включают:

- дополнительные шаги, не требуемые как часть выполнения задачи;
- наличие информации, вводящей в заблуждение пользователя;
- недостаточную информацию о пользовательском интерфейсе;
- неспецифицированные в документации неожиданные системные ответы;
- ограничения в навигации;
- неэффективные механизмы восстановления ошибок использования;
- физические характеристики пользовательского интерфейса, которые не подходят для физических характеристик пользователя.

Юзабилити-тестирование (в дальнейшем ЮТ) – это набор методов и инструментов, позволяющих обнаружить приведенные выше дефекты и измерить отмеченные характеристики взаимодействия пользователя с продуктом (в нашем случае – с программным продуктом, веб-приложением) для оценки уровня пользовательских свойств. Цель ЮТ – обнаружить и объяснить поведение человека в сложных системах при специфических обстоятельствах, «неэффективное» с той или иной точки зрения, а также свойства продукта, которые приводят к несоответствию между намерениями и/или действиями пользователя, а также системными свойствами и поведением.

Таким образом, в соответствии с вышесказанным ЮТ содержит:

- а) проверки для выявления дефектов удобства использования и потенциальных проблем удобства использования, включая:
 - отклонения объекта оценки от определенных критериев, таких как требования пользователей, руководящие принципы проектирования или установленные договоренности;
 - потенциальные проблемы удобства использования при попытке выполнить одну или несколько задач с целью оценки;
- б) наблюдение за пользователями, включая:
 - наблюдение за поведением пользователей для определения фактических результатов использования;
 - измерение производительности и реакции пользователя (например, время, затрачиваемое на выполнение задачи, количество ошибок использования, проводимость кожи или расширение зрачка);
- в) опросы пользователей, включая:
 - выявление проблем, мнений и впечатлений от пользователей (качественные опросы пользователей);
 - измерение уровня удовлетворенности или восприятия пользователя, например, оценочные значения шкалы для удовлетворения или для субъективно воспринимаемой эффективности или результативности (количественные опросы пользователей);

– другие данные, сообщенные пользователем (например, данные, полученные от человека в сочетании с данными наблюдений).

ЮТ выполняется как на этапе разработки программного продукта (промежуточное тестирование), так и во время его эксплуатации, по требованию пользователей (полное тестирование) [3].

В данной работе рассматривается задача полного ЮТ программного продукта, находящегося в опытной или промышленной эксплуатации.

2. Методы юзабилити-тестирования и постановка задачи

Среди известных методов, используемых в процедурах полного ЮТ, следует выделить [4]: метод экспертной оценки, метод сегментации пользовательской аудитории, метод оценки производительности, анализ рабочих заданий и метод, известный как «Мысли вслух». Коротко рассмотрим эти методы.

Экспертная оценка

Данный метод предполагает привлечение к тестированию экспертов в области юзабилити и профессиональных дизайнеров интерфейсов. В ходе исследования продукта они выявляют имеющиеся, на их взгляд, проблемы. Эксперт способен определять основные проблемы благодаря знанию стандартов интерфейсов, как официальных, так и неофициальных (например, корпоративных). Также опыт работы позволяет находить проблемы, которые ранее встречались при тестировании других продуктов. Качество экспертной оценки можно улучшить, увеличив количество экспертов.

Однако данный метод обладает своими недостатками. Основной – субъективность оценки. Также экспертная оценка зависит от опыта эксперта, что может привести либо к низкому качеству тестирования, либо к высокой оплате услуг эксперта.

Сегментация пользовательской аудитории

Сегментирование целевой аудитории – это разделение аудитории на группы, где они объединены по признаку схожих потребностей (запросов). В рамках этого метода вся пользовательская аудитория делится на различные группы по определенным похожим характеристикам их представителей. Это необходимо для понимания целей пользователей каждой группы. Для формирования репрезентативной выборки необходимо создать «портрет» представителей каждой группы целевой аудитории, персонифицировать их. В результате должна быть получена выборка, отражающая потребности всех возможных пользователей, для которых был разработан продукт. В работе [5] показано, что достаточно использовать выборку из 5–8 респондентов. При этом при использовании выборки из пяти респондентов находят в среднем 85% проблем, а при использовании выборки из восьми респондентов фиксируется почти 95% проблем. В работах [6–8] рекомендуется увеличить размер выборки до двенадцати. По мнению авторов, таким образом можно обеспечить более полное представительство различных групп целевой аудитории и тем самым сделать ЮТ более эффективным.

Оценка производительности

Данный метод в большей степени направлен на получение количественных данных. Пользователи (выбранные респонденты) выполняют определенный набор заданий, в ходе которых замеряются показатели затраченного времени, количество выполненных задач, совершенных действий, совершенных ошибок. Важно отметить, что общение между организатором и респондентом должно быть сведено к минимуму, без вмешательства

в процесс работы пользователя. Для получения данных об удовлетворенности от работы с системой рекомендуется провести небольшую беседу или анкетирование после выполнения заданий. Также можно попытаться сделать выводы при наблюдении за пользователем в ходе тестирования по косвенным признакам и реакциям.

Особо важной частью этого метода является анализ полученной информации. Количественные данные могут помочь определить области, в которых сосредоточены основные недостатки интерфейса. Стоит учитывать, что найденная ошибка может быть вызвана не проблемами продукта, а невнимательностью данного респондента. Кроме того, необходимо учитывать модель поведения пользователя. Дело в том, что пользователь может выбрать сознательно при исполнении заданий не оптимальный по времени путь. Это может быть связано с его личными качествами или недостаточной подготовкой для работы с данным интерфейсом. Из этого следует необходимость подробного исследования ошибок для определения вызвавших их причин.

Анализ рабочих заданий

Данный метод – это набор заданий в форме анкетирования или интервью для получения представления о том, как люди в настоящий момент выполняют конкретные задания.

«Мысли вслух»

Метод заключается в том, что респондент проговаривает все возникающие у него мысли в течение работы с тестируемым продуктом. Благодаря этому можно понять, как именно пользователь воспринимает продукт и чем могут быть обусловлены его действия. В ходе выполнения задания допускается помощь наблюдателя с использованием наводящих вопросов, но не прямых подсказок касательно интерфейса. Данный метод может показывать высокую эффективность даже с небольшим количеством респондентов, однако он не направлен на получение точных количественных данных. Причина в том, что потребность в озвучивании своих мыслей снижает скорость работы. Также важно отметить, что такой подход может повлиять на пользователей, побуждая их действовать логично и обдуманно, завышая результаты выполнения заданий.

В целом, анализ рассмотренных методов показал, что, несмотря на ряд исследований и достижений, в данной области отсутствует универсальная методика оценки юзабилити программных приложений, применение которой позволит инженеру-тестировщику достаточно эффективно организовать ЮТ. В данной работе предлагается такая методика, основывающаяся, с одной стороны, на сочетании рассмотренных выше методов ЮТ, а с другой стороны, использующая стратегию классического тестирования программных средств.

3. Основные положения методики юзабилити-тестирования

В соответствии с классическим подходом к тестированию программных продуктов рассматриваемая методика ЮТ состоит из следующих разделов:

- определение целей исследования – объектов тестирования;
- создание команды тестирования и распределение обязанностей;
- разработка и отладка системы тестов;
- создание среды тестирования;
- выполнение тестов;
- документирование результатов тестирования;
- анализ полученных данных и выработка рекомендаций по повышению качества.

Рассмотрим каждый из данных разделов.

Цели исследования формируются согласно поведению посетителей сайта. Определение целей начинается с определения основных пользовательских задач сайта. Последние представляют собой задачи, которые ставят перед собой пользователи в ходе применения продукта. Для определения этих задач можно использовать инструменты веб-аналитики, например, сервис Яндекс.Метрика [9, 10], предназначенный для сбора статистики посещений, с включенным в него инструментальным средством вебвизор, одна из полезных функций которого – запись всех действий пользователя на сайте [11].

Из полученного списка задач далее выбираются ключевые цели для проведения тестирования. В процессе выбора следует руководствоваться двумя принципами:

- задачи должны быть реальными, обычными для пользователя. Делать задания слишком сложными не имеет смысла – не стоит искать проблему, с которой никто не сталкивался.
- протестировать абсолютно все пользовательские задачи представляется маловероятным. Тестировать стоит только важные задачи, а именно, популярные задачи и те, при выполнении которых предположительно возможны ошибки.

Результатом выполнения данного этапа является подлежащий тестированию список целей.

В соответствии с классическим подходом к тестированию программ основными участниками процесса тестирования являются разработчики тестов и исполнители-тестировщики. При ЮТ функции разработчика тестовых сценариев выполняет юзабилити-специалист, который выбирает подходящие сценарии для определенных целей тестирования и составляет для них описания, которые будут переданы в качестве заданий тестировщикам. Кроме того, этот специалист должен уметь пользоваться средствами автоматизации тестирования и веб-аналитики.

Выбор тестировщиков при ЮТ имеет свои особенности. Это должны быть представители типичных для пользователей групп (сегментов целевой аудитории), определяемых с помощью средств веб-аналитики. При этом на выбор тестировщиков также должны влиять:

- опыт работы с продуктом. Рекомендуется набирать половину респондентов с опытом работы для определения проблем, связанных с переучиванием, а половину – без опыта работы, для определения скорости обучения;
- уровень компьютерной грамотности. Рекомендуется соблюдать такую пропорцию: три четверти – с уровнем, совпадающим с опытом целевой аудитории, у одной четвертой уровень ниже, благодаря чему можно определить большее количество проблем;
- возраст тестировщиков. Рекомендуется три четверти тестировщиков выбирать возраста целевой аудитории, а оставшуюся четверть – вне пределов этого возраста.

Кроме перечисленных критериев, на выбор тестировщиков должна влиять их эмоциональная открытость, благодаря которой при опросах и оценке удовлетворенности можно получить необходимую для анализа информацию.

Как уже упоминалось в предыдущем разделе, количество тестировщиков можно ограничить примерно десятком человек, которые смогут обнаружить до 95% имеющихся проблем.

Важнейшим этапом ЮТ является разработка тестовых сценариев, поскольку от качества тестов зависит, в конечном счете, оценка удобства использования программного продукта. Тестовые сценарии должны соответствовать выбранным ранее целевым задачам. При их успешном выполнении должна быть получена конкретная, заранее известная требуемая информация (сообщения, рисунки, таблицы, результаты вычислений и т. д.). При этом следует отметить, что тестовые задания при ЮТ существенно отличаются от

классических тестовых заданий. Дело в том, что при классическом тестировании тестовые задания реализуются строго по шагам, указанным в сценарии, в то время как при ЮТ в сценарии приводятся общая постановка задачи, ее конечная цель и действия, необходимые для начала тестирования. В остальном тестирующему предоставлена полная свобода действий.

Для разработки тестовых сценариев можно использовать накопленный опыт пользователей, уже работавших с этим приложением или с другим, схожим по функциональному назначению. Для этого можно провести устный опрос этих пользователей, выяснить, что они ждут от данного приложения, на какие его возможности рассчитывают в первую очередь. Результат опроса – разработка тестовых сценариев, демонстрирующих эти возможности.

Кроме того, можно воспользоваться средствами веб-аналитики, которые могут предоставить сценарии поведения данных пользователей на сайте. Например, вебвизор может продемонстрировать все действия, совершенные пользователем за конкретный сеанс, из чего можно сделать вывод о том, какие цели он преследовал на сайте. Также в составлении сценариев может помочь статистика по посещениям конкретных страниц.

Однако не стоит всецело полагаться на данные, полученные из сервиса. Статистика показывает успешно достигнутые цели, но отображает и то, чего пользователи искали, но не нашли. Поэтому в данном случае разработчик тестов выбирает подходящие сценарии для определенных целей тестирования и составляет для них описания, которые будут поставлены в качестве заданий перед тестирующими.

Третий подход к составлению тестовых сценариев – это самостоятельная работа разработчика тестов на основе прилагаемой документации и накопленного опыта работы с программным продуктом.

Созданные вышеописанными способами тестовые сценарии объединяются в совокупность тестов проверки удобства использования. Затем происходит минимизация их количества путем сокращения схожих по своему назначению сценариев. После определения совокупности тестов все указанные в них задания выполняются разработчиком тестов.

Тестовые процедуры тщательно контролируются, выполняются неоднократно, ошибки исправляются, а полученные таким образом данные, например, такие, как время прохождения теста, количество совершенных переходов и т. п., можно считать эталонными при оценке действий тестируемых.

Затем осуществляется окончательная проверка сценариев с использованием небольшой группы тестируемых, например, состоящей из одного-двух представителей от каждой группы целевой аудитории. Главная цель этого тестирования – убедиться, что тестируемые правильно понимают поставленные задачи, что они не вызывают никаких трудностей. Низкая оценка сценария может служить основанием для его переработки. При этом на данном этапе осуществляется экспресс-оценка удобства использования анализируемого веб-приложения. Эта проверка является некоторым аналогом смюк-тестирования при классическом тестировании программ [1, 2].

По завершении окончательной проверки все тестовые сценарии вновь выполняются разработчиком тестов. При этом все действия запоминаются путем их автоматической записи с помощью Selenium IDE – инструмента для автоматизации действий в браузере в режиме Записи / Воспроизведения (Recorder / Player) [12]. Сценарии сохраняются в формате HTML. Таким образом, получаются копии тестовых сценариев, которые можно в любой момент запускать в автоматическом режиме.

Полученный автотест может в дальнейшем использоваться для повторной проверки сценария в случае каких-либо проблем при проведении тестирования. Например, если у тестирующего появились трудности во время выполнения задания и он не может достичь цели. В этом случае путем запуска нужного автотеста можно убедиться, что цель либо достижима, либо сайт был изменен, а цель находится в другом разделе сайта или уже недоступна.

При выполнении тестов тестирующие последовательно проходят на сайте подготовленные сценарии. Можно разделить все задания между участниками так, чтобы часть выполняла их согласно методу «Мысли вслух», озвучивая все свои рассуждения в ходе выполнения заданий, а другая часть – без рассуждений, с простым замером количественных данных с использованием метода «Оценка производительности». Очевидно, что первая группа должна быть существенно меньше второй, поскольку, как указывалось выше, в первом случае требуются при выполнении дополнительные затраты времени на проговаривание мыслей, а действия тестирующего становятся более разумными. При этом использование метода «Мысли вслух» позволит лучше понять, как воспринимается программный продукт пользователем, а оценка времени выполнения тестов определяется путем сравнения данных, полученных с использованием метода «Оценки производительности».

Для оценки уровня удовлетворенности и для получения дополнительных данных по завершении выполнения заданий тестирующие проходят ряд опросов или анкетирование.

Один из наиболее простых, но не самых надежных способов оценки удовлетворенности – анкета по словам [7, 8]. Тестирующему предоставляется список случайно подобранных прилагательных, половина которых носит позитивную оценку, а другая – негативную, например, «эффективный», «нечеткий», «неудобный», «бесполезный», «последовательный», «качественный» и т. д. Все слова располагаются в случайном порядке. Участник должен выбрать те слова, которые, на его взгляд, описывают продукт. По завершении подсчитывается разница между числом негативных и позитивных слов.

Еще один способ – формальная анкета. В отличие от предыдущего, это менее универсальный способ, поскольку анкета требует адаптации под конкретный проект. Она содержит несколько вопросов или положений. Например, анкета может включать следующие положения: «Я доволен скоростью своей работы», «Во время выполнения заданий я чувствовал себя вполне уверенно», «В любой момент времени я понимаю, что должен сделать дальше». Ответы пользователя представляются в виде оценки по пятибалльной шкале от –2 до +2 баллов [7, 8].

В завершение в форме опроса или в форме беседы можно получить развернутый ответ на интересующие вопросы. Например, на следующие:

- Какие задания были для Вас наиболее сложными?
- Что бы Вы хотели поменять в сайте?
- Какие Вы можете предложить улучшения?
- Какие решения, наоборот, кажутся Вам удачными?

Не стоит забывать, что при проведении тестирования не исключается возможность обнаружения проблем, не найденных ранее, что послужит причиной возвращения к этапу разработки тестовых сценариев для исправления. Среди таких проблем могут быть, например, некорректное понимание заданий большинством участников, изменение сайта.

После завершения выполнения тестовых заданий выполняется анализ полученных результатов. Итогом данного этапа будут выводы или гипотезы об уровне удобства исполь-

зования тестируемой версии сайта. Формируются они следующим образом. Количественные данные, такие как количество совершенных действий, успешность выполнения заданий, количество ошибок, сравниваются с эталонным значением по каждому заданию, выполненному как с использованием метода «Оценка производительности», так и с использованием метода «Мысли вслух». Серьезная разница у большинства участников может говорить о наличии дефектов удобства использования при выполнении конкретных заданий.

Оценка времени выполнения заданий путем сравнения с эталонными данными осуществляется только для тех тестировщиков, которые реализовали тестовые сценарии методом «Оценка производительности».

Уровень удовлетворенности определяется согласно выбранной анкете. Обычно в качестве эталонного результата берется максимальное значение, которое можно получить в анкете. Низкий показатель уровня удовлетворенности конкретного тестировщика может быть основанием более подробного изучения качественных данных, полученных от него. Однако при этом не стоит забывать, что могут проявляться и какие-то личные качества, не зависящие от объекта тестирования.

Анализ качественных данных, всех видео- и аудиоматериалов, а также опросов требуют иного подхода. Проблемы, с которыми сталкиваются тестировщики в этом случае, отображаются наглядно. Однако они могут быть индивидуальными, например, из-за личной невнимательности. Идентификация такой ситуации возможна с помощью вебвизора – возможно, кто-то из реальных посетителей также столкнулся с этой проблемой. Явно обозначит дефект юзабилити и повторяющаяся среди тестировщиков ошибка.

При опросе, релевантном как опыту тестировщиков, так и рассматриваемым дефектам, респонденты также могут указать на то, что они считают неудобным, или на то, что, по их мнению, можно улучшить.

Все найденные проблемы необходимо зафиксировать и присвоить им приоритеты. Описание проблемы должно быть точным и лаконичным, мало отличающимся от описания дефектов при классическом функциональном тестировании программ [13]. В этом описании должны быть указаны: краткое содержание проблемы, тестовые задания, при выполнении которых эта проблема была обнаружена, среда, в которой осуществлялась тестирование. Для воспроизведения данной проблемы разработчиками веб-приложения, возможно, стоит добавить перечень пройденных шагов от начала выполнения тестового задания до шага, на котором возникает рассматриваемая проблема, а при необходимости приложить соответствующие аудио- и видеоматериалы. Оптимальный перечень приоритетов можно описать так:

- критическая ошибка – мешает выполнению задачи;
- серьезная ошибка – вызывает сложности у большинства пользователей;
- незначительная ошибка – вызывает раздражение, но особых затруднений не создает.

Используя результаты анализа, юзабилити-специалист должен предложить рекомендации по улучшению удобства использования веб-приложений и передать их разработчикам сайта.

4. Особенности реализации и некоторые результаты

Предложенная методика была использована при оценке удобства использования priem.migea.ru – страницы приемной комиссии МИРЭА – Российского технологического университета (РТУ МИРЭА). Основную функцию сайта можно определить, как информационную. В большинстве случаев он посещается с целью найти нужную информацию.

Сайт используется для поиска информации о подходящих программах и направлениях обучения, о порядке подачи документов в Приемную комиссию. Поэтому цель исследования – узнать, насколько легко пользователю найти нужную ему информацию.

Согласно данным, полученным с помощью Яндекс.Метрики, удалось определить четыре основные группы пользователей-посетителей данного сайта. Эти группы объединяют в своем составе порядка 90% всех пользователей сайта. Первую группу образуют посетители моложе 18 лет (абитуриенты), вторую и третью – пользователи в возрасте 18–24 лет (студенты РТУ МИРЭА и студенты других вузов), четвертую группу образуют пользователи в возрасте 35–54 лет (родители абитуриентов).

Уровень компьютерной грамотности каждой группы подразумевается как средний или выше среднего, то есть предполагается, что работа респондента с компьютером является частью обычной деятельности в течение двух или более лет. Подразумевается, что опыт работы с сайтом есть у группы студентов из РТУ МИРЭА. В состав остальных групп включены респонденты, которые уже пользовались сайтами других университетов. Также для исследования юзабилити в каждую группу было включено хотя бы по одному представителю, кто не имел опыта работы с подобными сайтами вообще.

Исходя из поставленных целей и сформированной целевой аудитории, к тестированию было привлечено 12 пользователей – по три представителя от каждой группы, среди которых были респонденты с указанными выше особенностями.

Для составления актуальных сценариев поведения пользователей на сайтах университетов был проведен небольшой опрос среди целевой аудитории. Для составления сценариев также использовались статистические данные просмотров страниц сайта, полученные с использованием возможностей веб-аналитики. При этом, на основе изучения данных веб-аналитики, до начала ЮТ были выявлены потенциальные дефекты. В частности, по данным карты кликов сайта priem.mirea.ru можно предположить, что пользователям не очевидно назначение ссылки «Гид по специальностям» как способа выбора подходящей специальности. В то время как большой популярностью пользуются ссылки «Программы обучения, число мест».

На основании полученных таким образом данных были сформированы следующие задания:

1. Найти направление(-я) бакалавриата с подходящими предметами, проходными баллами и ценой обучения;
2. Найти подробную информацию о направлении магистратуры;
3. Найти подробную информацию о направлении аспирантуры;
4. Найти подробную информацию о программах среднего профессионального образования;
5. Найти перечень необходимых документов для поступления в магистратуру;
6. Найти расписание приема документов;
7. Найти свидетельство государственной аккредитации университета;
8. Найти информацию о предоставлении общежития;
9. Найти информацию о порядке зачисления на программы бакалавриата;
10. Найти, где задать вопрос приемной комиссии;
11. Найти дату ближайшей встречи по вопросам поступления;
12. Найти информацию о порядке перевода в МИРЭА из других вузов.

Каждый тестирующий (табл. 1) получил по 2 задания, при выполнении которых он озвучил свои мысли согласно методу «Мысли вслух». Еще 4 задания выполнялись согласно методу «Оценка производительности». Задания у каждого тестирующего не повторялись. Не было повторений при распределении заданий по методам в рамках группы.

Таблица 1. Распределение заданий между тестирующими

Группа	Метод «Мысли вслух»	Метод «Оценка производительности»
Абитуриенты	1, 7	2, 6, 8, 11
	6, 10	3, 5, 9, 12
	2, 8	1, 4, 7, 10
Студенты других вузов	5, 11	1, 4, 7, 10
	3, 9	2, 6, 8, 12
	4, 12	3, 5, 9, 11
Студенты РТУ МИРЭА	1, 5	2, 6, 8, 11
	3, 7	4, 5, 9, 12
	9, 11	1, 3, 7, 10
Родители	2, 6	1, 7, 10, 11
	4, 8	2, 6, 9, 12
	10, 12	3, 4, 5, 8

В соответствии с предлагаемой методикой было проведено предварительное (пилотное) тестирование. Для его реализации было выделено два тестирущика. В ходе тестирования с первым респондентом были скорректированы формулировки некоторых задач на более конкретные. В остальном же оба респондента не столкнулись с какими-либо трудностями.

Далее все задания были многократно выполнены вручную разработчиком тестов – юзобилити-специалистом. Полученные таким образом данные (время прохождения каждого теста, количество выполненных шагов, маршруты переходов по ссылкам и т. д.) будут использоваться как эталонные и рекомендуемые при выполнении заданий тестирующими. Кроме того, с помощью Selenium IDE были записаны все действия юзобилити-специалиста и получены таким образом 12 автоматических тестов, которые всегда можно запустить для анализа действий тестирущиков и проверки отсутствия изменений в содержимом рассматриваемого сайта.

После того, как каждый тестирущик в индивидуальном порядке выполнил все ответственные ему задания, были получены следующие количественные результаты, приведенные в табл. 2 и 3.

Таблица 2. Время выполнения заданий, мин:с

№ задания	Метод «Мысли вслух»		Метод «Оценка производительности»			
	1	2	3	4	5	6
1	3:31	1:07	2:54	1:05	5:21	1:24
2	2:56	1:49	1:12	0:39	4:51	3:12
3	1:44	2:41	3:16	0:46	0:54	1:50
4	0:57	1:35	4:51	1:29	0:37	0:58
5	5:23	0:33	0:16	0:17	0:15	0:28
6	0:51	0:28	0:34	0:51	1:27	1:03
7	1:32	0:58	1:40	0:35	2:08	0:41
8	2:08	1:23	0:31	0:47	2:03	0:25
9	0:48	0:38	0:37	1:03	0:10	0:41
10	0:42	0:20	0:56	0:11	0:39	0:49
11	0:46	3:37	5:04	0:22	0:38	1:40
12	1:07	0:31	0:19	0:50	0:35	0:27

Таблица 3. Количество ошибок при выполнении заданий

№ задания	Метод «Мысли вслух»		Метод «Оценка производительности»			
	2	1	2	1	3	1
1	2	1	2	1	3	1
2	3	1	1	1	3	2
3	1	1	2	1	0	2
4	1	0	2	3	0	0
5	2	0	0	0	0	0
6	1	0	0	1	0	1
7	1	1	1	2	1	1
8	0	0	0	0	0	0
9	0	0	0	0	0	0
10	0	0	0	0	0	0
11	0	2	3	0	0	3
12	0	0	0	0	0	0

Больше всего ошибок было допущено при выполнении заданий с 1 по 4. Данные задания образуют группу, связанную с получением информации о направлениях обучения. Стабильное совершение ошибок при выполнении заданий, а также существенные затраты времени при выполнении этих заданий респондентами в рамках метода «Оценка производительности» говорят о наличии дефекта юзабилити.

Для получения качественной информации об уровне удовлетворенности при работе с сайтом был проведен небольшой опрос, с помощью которого можно узнать сформированное отношение к сайту, наиболее сложные или, наоборот, простые задания, какие решения в дизайне сайта являются удачными, получить предложения по улучшению сайта. Респонденты проходили опрос по завершении основной части тестирования. Опрос включал следующие вопросы:

- Какие задания были наиболее сложными?
- Что бы хотели поменять в сайте?
- Какие можете предложить улучшения?
- Какие решения, наоборот, кажутся удачными?

Также была использована формальная анкета со следующими вопросами:

- Мне нравится внешний вид сайта – Нет/Да
- Во время заданий я не столкнулся с существенными трудностями – Нет/Да
- Во время выполнения заданий я не ошибался – Нет/Да
- Я доволен своей скоростью работы – Нет/Да
- Всякий раз, когда я ошибался, я с легкостью замечал и исправлял ошибку – Нет/Да

Кроме того, каждый пункт приведенной в предыдущем разделе анкеты для определения уровня удовлетворенности оценивался, как указывалось выше, по пятибалльной шкале от –2 до +2 (максимальный средний уровень удовлетворенности соответственно равен 10).

После выполнения всех тестовых заданий, опроса и анкетирования были получены следующие результаты:

- При тестировании было выявлено 6 серьезных и 3 незначительных дефекта, критических дефектов не обнаружено.
- Среди найденных серьезных дефектов можно назвать, например, неудачное расположение ссылки «Гид по специальностям». Только двое респондентов с самого начала

воспользовались ссылкой на главной странице. Остальные находили ссылку на него внизу страниц «Программы обучения, число мест». Большинство, даже пролистав страницу до конца, не заметили ссылку. Двое из респондентов заметили ее только на странице программ обучения среднего профессионального образования, из-за короткой таблицы направлений. Как отметили респонденты, назначение «Гида» непонятно, пока его не загрузишь. Таким образом, тестирование подтвердило гипотезу о редком использовании данной ссылки после анализа сайта с помощью средств веб-аналитики.

- Среди незначительных ошибок следует упомянуть ошибку, связанную с выполнением задания 11 («Найти дату ближайшей встречи по вопросам поступления»). Для трех из шести выполнявших это задание респондентов была непонятна сама формулировка задания и его назначение. Также все отметили, что долго не обращали внимания на нужную для выполнения данного задания кнопку, поскольку видели первое слово «Вебинары» и поэтому отбрасывали данный вариант и продолжали ее искать.

- Средний уровень удовлетворенности в ходе работы с сайтом на основании формальной анкеты равен 6 из 10 возможных баллов. Среди всех 12 тестировщиков подсчет уровня удовлетворенности не был ниже 0, что говорит о том, что в большей степени отношение к сайту у тестировщиков положительное. Большинство отметило, что данный сайт – один из самых понятных, которые доводилось посещать, среди сайтов других университетов.

Заключение

В статье рассмотрена методика полного юзабилити-тестирования веб-приложений, находящихся в опытной или промышленной эксплуатации. Данная методика использует возможности методов классического тестирования и ЮТ программных средств, а также возможности средств автоматизации тестирования и веб-аналитики. Предлагаемая методика была успешно реализована и использовалась при оценке качества страницы приемной комиссии priem.mirea.ru сайта РТУ МИРЭА.

Полученные экспериментальные данные показали, что, несмотря на тестирование удобства использования программного обеспечения на этапах разработки, некоторые ошибки остаются необнаруженными. Это говорит о том, что необходимо повышать уровень юзабилити уже на начальных этапах разработки. В частности, для оперативного предсказания уровня юзабилити желательно иметь математические средства моделирования поведения проектируемой системы и пользователя. Поскольку поведение пользователей весьма вариативно (как в виду различных личностных характеристик, так и разного уровня осведомленности в конкретной предметной области), разумно использовать вероятностно-статистические модели. Можно использовать модель цепи Маркова в запросах к серверу, в которой состояния представляют веб-страницы, а переходы между состояниями представляют активности работы с гипертекстом. Схема перемещения посетителя на веб-сайте может рассматриваться как последовательность переходов из одного состояния в другое с добавлением дополнительного состояния, чтобы обозначить случай, когда посетитель находится за пределами веб-сайта.

Разработка таких моделей, на наш взгляд, является перспективным направлением в области юзабилити-тестирования веб-приложений.

Благодарности

Работа выполнена при частичной финансовой поддержке РФФИ (проекты № 18-07-00669 и 18-29-03100).

Литература:

1. Басок Б.М., Красовский В.Е. Тестирование программного обеспечения. М.: МИРЭА, 2010. 120 с. ISBN 978-5-7339-08 07-6
2. Куликов С.А. Тестирование программного обеспечения. Базовый курс [Электронный ресурс]. – URL: http://svyatoslav.biz/software_testing_book/ (дата обращения: 20.08.2019).
3. Нильсен Я., Лоранжер Х. Web-дизайн: удобство использования Web-сайтов: пер. с англ. / ред. В.С. Иващенко. М.: Вильямс, 2009. 368 с. ISBN 978-5-8459-1222-0
4. Сергеев С.Ф. Методы тестирования и оптимизации интерфейсов информационных систем. СПб.: НИУ ИТМО, 2013. 117 с.
5. Nielsen J., Landauer T.K. A mathematical model of the finding of usability problems // *Proceed. of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*. Amsterdam, the Netherlands, April 24-29, 1993. P. 206-213. <https://doi.org/10.1145/169059.169166>
6. Руководство по юзабилити-тестированию – опыт Mail.Ru Group. [Электронный ресурс]. – URL: <https://vc.ru/flood/19132-ux-testing-mrg> (дата обращения 20.08.2019).
7. Rubin J., Chisnell D. Handbook of usability testing: How to plan, design, and conduct. Second ed. New York: John Wiley & Sons, 2008. P. 15–53. ISBN 978-0-470-18548-3
8. Головач В.В. Юзабилити-тестирование по дешевке. [Электронный ресурс]. – URL: <https://medium.com/usethics-doc/%D1%8E%D0%B7%D0%B0%D0%B1%D0%B8%D0%BB%D0%B8%D1%82%D0%B8-%D1%82%D0%B5%D1%81%D1%82%D0%B8%D1%80%D0%BE%D0%B2%D0%B0%D0%BD%D0%B8%D0%B5-%D0%BF%D0%BE-%D0%B4%D0%B5%D1%88%D0%B5%D0%B2%D0%BA%D0%B5-2e853250960f> (дата обращения 20.08.2019).
9. Google Analytics vs Яндекс.Метрика. [Электронный ресурс]. – URL: <https://blog.sibirix.ru/2016/07/06/ga-vs-yam/> (дата обращения 20.08.2019).
10. Яндекс.Метрика. [Электронный ресурс]. – URL: <http://znet.ru/instrumenty/yandeks-metrika/> (дата обращения 20.08.2019).
11. Вебвизор – анализ поведения пользователей на сайте. [Электронный ресурс]. – URL: <https://webliberty.ru/vebvisor-sayt-glazami-polzovatelya/> (дата обращения 20.08.19).
12. Selenium / WebDriver. Автоматизация веб-приложений через браузер. [Электронный ресурс]. – URL: <https://www.selenium2.ru/docs/selenium-ide> (дата обращения 20.08.2019).
13. Басок Б.М., Головин С.А. Тестирование готового к использованию программного продукта. М.: МИРЭА – Российский технологический университет, 2019. 96 с. ISBN 978-5-7339-1469-5

References:

1. Basok B.M., Krasovsky V.E. Software testing. Moscow: MIREA Publ., 2010. 120 p. (in Russ). ISBN 978-5-7339-08 07-6
2. Kulikov S.A. Software testing. Basic course. [Electronic resource] – URL: http://svyatoslav.biz/software_testing_book/ (accessed August 20, 2019) (in Russ).
3. Nielsen J. Prioritizing Web Usability. Berkeley, Calif.: New Riders, 2006. [Nielsen J., Laurent H. Web design: Usability of Web sites: transl. from Engl. Moscow: Villyams Publ., 2009. 368 p. (in Russ.)]
4. Sergeev S.F. Testing methods and optimization of information systems interfaces. Saint Petersburg: ITMO Publ., 2013. 115 p. (in Russ.).
5. Nielsen J., Landauer T.K. A mathematical model of the finding of usability problems. *Proceed. of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*. Amsterdam, the Netherlands, April 24-29, 1993. P. 206-213. <https://doi.org/10.1145/169059.169166>
6. Usability Testing Guide – Mail.Ru Group Experience. [Electronic resource]. – URL: <https://vc.ru/flood/19132-ux-testing-mrg> (August 20, 2019) (in Russ).
7. Rubin J., Chisnell D. Handbook of usability testing: how to plan, design, and conduct. Second ed. New York: John Wiley & Sons, 2008. P. 15-53. ISBN 978-0-470-18548-3
8. Golovach V.V. Usability testing at a low price. [Electronic resource]. – URL: <https://medium.com/usethics-doc/%D1%8E%D0%B7%D0%B0%D0%B1%D0%B8%D0%BB%D0%B8%D1%82%D0%B8-%D1%82%D0%B5%D1%81%D1%82%D0%B8%D1%80%D0%BE%D0%B2%D0%B0%D0%BD%D0%B8%D0%B5-%D0%BF%D0%BE-%D0%B4%D0%B5%D1%88%D0%B5%D0%B2%D0%BA%D0%B5-2e853250960f> (accessed August 20, 2019) (in Russ).
9. Google Analytics vs Yandex.Metrika. [Electronic resource]. – URL: <https://blog.sibirix.ru/2016/07/06/ga-vs-yam/> (accessed August 20, 2019) (in Russ).
10. Yandex.Metrika. [Electronic resource]. – URL: <http://znet.ru/instrumenty/yandeks-metrika/> (accessed August 20, 2019) (in Russ).
11. Webvisor – analysis of user behavior on the site. [Electronic resource]. – URL: <https://webliberty.ru/vebvisor-sayt-glazami-polzovatelya/> (accessed August 20, 2019) (in Russ).
12. Selenium / WebDriver. Automate web applications through a browser. [Electronic resource]. – URL: <https://www.selenium2.ru/docs/selenium-ide> (accessed August 20, 2019).
13. Basok B.M., Golovin S.A. Testing a ready-to-use software product. Moscow: RTU MIREA Publ., 2019. 96 p. (in Russ). ISBN 978-5-7339-1469-5

Об авторах:

Басок Борис Моисеевич, кандидат технических наук, доцент кафедры математического обеспечения и стандартизации информационных технологий Института информационных технологий ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). Scopus Author ID 6602074818, ResearcherID X-6744-2019, <https://orcid.org/0000-0002-8935-4248>.

Рожанская Алена Николаевна, бакалавр кафедры математического обеспечения и стандартизации информационных технологий Института информационных технологий ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78).

Френкель Сергей Лазаревич, кандидат технических наук, старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук (ФИЦ ИУ РАН) (119333, Россия, Москва, ул. Вавилова, д.44-2). Scopus Author ID 36865996500, <https://orcid.org/0000-0002-2150-7236>.

About the authors:

Boris M. Basok, Cand. of Sci. (Engineering), Associate Professor of the Chair of Mathematical Software and Standardization of Information Technology, Institute of Information Technology, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow 119454, Russia). Scopus Author ID 6602074818, Researcher ID: X-6744-2019, <https://orcid.org/0000-0002-8935-4248>.

Alena N. Rozhanskaya, Bachelor of Computer Science of the Chair of Mathematical Software and Standardization of Information Technology, Institute of Information Technology, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow 119454, Russia).

Sergey L. Frenkel, Cand. of Sci. (Engineering), Senior Researcher of FRC “Computer Sciences and Control” Russian Academy of Sciences (44-2, Vavilova ul., Moscow 119333, Russia). Scopus Author ID 36865996500, <https://orcid.org/0000-0002-2150-7236>.