

УДК 004.5

DOI: 10.15827/0236-235X.119.447-455

Дата подачи статьи: 20.01.17

2017. Т. 30. № 3. С. 447–455

СОВРЕМЕННЫЕ ТЕНДЕНЦИИ В АВТОМАТИЗИРОВАННОЙ ОЦЕНКЕ ЮЗАБИЛИТИ И ПОВЕДЕНЧЕСКИЕ ФАКТОРЫ В АЛГОРИТМАХ ПОИСКОВЫХ СИСТЕМ

*М.А. Бакаев, к.т.н., старший научный сотрудник, maxis81@gmail.com
(Новосибирский государственный технический университет,
просп. К. Маркса, 20, г. Новосибирск, 630073, Россия)*

В настоящее время сохраняются неопределенность относительно количественного измерения юзабилити, а также ненужное противопоставление его объективных составляющих (продуктивности и эффективности) и субъективного аспекта удовлетворенности пользователя.

В данной обзорной статье отмечается, что практическая цель заключается не столько в измерении юзабилити веб-сайта, сколько в улучшении качества его использования или его оценки со стороны соответствующих алгоритмов поисковых систем, которые все более повышают важность поведенческих факторов в ранжировании. Описаны традиционные методы оценки и проектирования юзабилити веб-сайтов и современные инструменты, вносящие в них элементы автоматизации.

В рамках автоматизированной оценки юзабилити выделены подходы, основанные на взаимодействиях, метриках и моделировании, указаны возможности и ограничения каждого из них. Кроме того, дан краткий обзор гибридных подходов, которые все в большей степени используют методы искусственного интеллекта и машинного обучения. Так, эффективным методом для оценки качества взаимодействия с веб-сайтом являются нейронные сети, применяемые, в частности, в рамках эволюционных алгоритмов проектирования для вычисления фитнес-функции для решений-кандидатов. Описаны методы автоматизированной оценки юзабилити сайтов со стороны глобальных поисковых систем (Яндекс, Гугл), приводится перечень некоторых значимых поведенческих факторов. Фактически поисковые системы рассматривают юзабилити как «черный ящик» и активно применяют методы машинного обучения для аппроксимации его оценки.

Таким образом, концептуализация и практическая оценка юзабилити со стороны поисковых систем, которые в значительной степени формируют веб-среду на сегодняшний день, постоянно эволюционируют и требуют соответствующих гибких и адаптивных подходов со стороны веб-дизайнеров и разработчиков.

Ключевые слова: *человеко-компьютерное взаимодействие, проектирование юзабилити, веб-дизайн, искусственный интеллект, поисковая оптимизация.*

В настоящее время все меньше организаций стремятся лишь создать сайт: большинство из них понимают, что важнее то, сколько потенциальных клиентов посещают сайт и с какой целью. Это, пожалуй, лучше всего отражает формула, предложенная Я. Нильсеном: $B \sim V * C * L$, где B – бизнес-эффект от сайта; V – количество уникальных посетителей; C – показатель конверсии (превращение посетителей в клиентов); L – уровень лояльности (отражает повторные визиты, средний размер заказа и т.п.) [1]. Традиционно считается, что юзабилити (здесь и далее этим термином будем обозначать эргономичность, удобство использования) веб-сайта влияет на C и L , хотя для лояльности эффект не совсем прямой и касается примерно 1/3 изменений этого показателя [2]. В наши дни, однако, юзабилити влияет и на V , поскольку поисковые системы начали учитывать так называемые поведенческие факторы в ранжировании веб-сайтов. Таким образом, неудивительно, что интерес к юзабилити растет со стороны не только дизайнеров-практиков, но и исследователей.

Понятие «юзабилити» внесено в стандарты относительно давно, например, в ISO 9241-11 (1998) оно определяется как степень, с которой продукт может быть использован определенными пользователями при определенном контексте использования для достижения определенных целей с долж-

ной эффективностью, продуктивностью и удовлетворенностью. Юзабилити – это качественное понятие, специфичное для пользователей, их задач (целей) и контекста взаимодействия. Продуктивность означает полноту и точность достижения пользователями своих целей и, как правило, измеряется следующими параметрами:

- соотношение между успешными и ошибочными действиями пользователя;
- доля используемого функционала и команд (чем выше, тем лучше);
- когнитивная нагрузка на пользователя, выражающаяся в потреблении ресурсов памяти, внимания, воображения и т.д. (чем меньше, тем лучше).

Из этих трех параметров когнитивная нагрузка является, пожалуй, наименее очевидной, и, соответственно, нет общепринятого подхода к ее измерению. Так, в [3] в ходе сравнения результатов в трех различных условиях обучения был введен индекс когнитивной нагрузки и соотнесен с относительной продуктивностью. Некоторые исследовательские работы в этой области основывались на показателях жизнедеятельности организма человека, таких как кровяное давление или движение глаз и реакция зрачков [4], хотя в рамках таких подходов трудно различить когнитивные и физические аспекты нагрузки. В других исследованиях время

рассматривается как показатель производительности и отражение когнитивной сложности задачи (см. [5]).

Эффективность (минимизация объема ресурсов для выполнения задачи) для веб-сайтов обычно понимается как скорость, с которой пользователи могут выполнить свои задачи. Соответственно, декомпозиция может быть выполнена следующим образом:

- изучаемость (легкость, с которой пользователи могут начать выполнение своих задач на незнакомом сайте, или время, необходимое для того, чтобы научиться им пользоваться; вообще пользователи редко тратят время на изучение работы с конкретным веб-сайтом, поскольку их много);
- время на выполнение задачи;
- ошибки (сколько ошибок допускает пользователь, насколько они серьезны, сколько времени необходимо для их исправления и т.д.);
- запоминаемость (насколько легко можно вернуться к результативной работе с веб-сайтом после значительного перерыва).

На практике при оценке юзабилити изучаемость и запоминаемость измеряются довольно редко, а начиная с 1990-х годов Я. Нильсен ввел и обосновал подход, называемый «бюджетным юзабилити» и основанный на простых, быстрых и дешевых методах проектирования юзабилити, которые, как правило, не подразумевают количественные измерения.

Наконец, под удовлетворенностью подразумевается то, насколько приятно использовать веб-сайт. Это субъективный эмоциональный аспект, который тоже является специфичным с точки зрения пользователя, задачи и контекста (к примеру, удовлетворенность при редких визитах на сайт электронного магазина и при ежедневной работе с интранет-системой проистекают из различных факторов хотя бы с точки зрения визуального дизайна). Следует заметить, что простота понятия «приятно» во многом обманчива и этот компонент юзабилити на настоящий момент остается в некотором противопоставлении остальным, поскольку юзабилити «для использования» и юзабилити «чтобы было» существенно различаются [2]. Сравнительная важность компонентов юзабилити также не определена (например, может быть только лишь удовлетворенность, то есть то, что пользователь помнит после использования сайта или продукта, имеет значение), и, соответственно, количественное выражение юзабилити остается проблематичным.

В данной статье приведены обзор измерений и оценка юзабилити в настоящее время с использованием автоматизированных и автоматических способов. Охвачен также ряд концептуальных моментов, касающихся юзабилити, поскольку обычно считается, что улучшение какого-либо показателя (а оптимизация человеко-компьютерного взаимо-

действия – цель проектирования юзабилити) вряд ли возможно без понимания того, как он может измеряться. Кроме того, в работе описываются методы оценки юзабилити.

Традиционные методы оценки юзабилити

Практически любой процесс оценки юзабилити включает в себя сбор и обработку данных о взаимодействии с пользователем (будь оно реальным, совершаемым в ходе тестирования, или же умозрительным), их анализ и интерпретацию результатов, на основании которых вырабатываются рекомендации по улучшению юзабилити. В отличие от показателей эргономичности для интерфейсов человеко-машинных систем, которые зачастую оцениваются количественно в ходе специально поставленных экспериментов (см., например, [6] о применении в авиации), в большинстве веб-проектов применяются подходы на качественном уровне. Традиционные методы проектирования веб-юзабилити, которые начали активно использоваться еще в 1990-х годах, были направлены на повышение конверсии, снижение отскока, улучшение очевидных юзабилити-метрик и общее создание полезного, привлекательного и современного веб-дизайна. Далее описываются основные методы в этой области, классифицируемые в основном по их отношению к триаде пользователь–задача–контекст, и говорится о том, как в настоящее время осуществляется распределение работы между специалистами-людьми и компьютерами.

Явное и опосредованное наблюдение за пользователями. Наблюдение за пользователями – пожалуй, наиболее очевидный метод в инструментарии проектировщика юзабилити, поскольку он подразумевает, что реальные пользователи выполняют реальные задачи в реальном контексте, и это практически полностью соответствует определению юзабилити, за исключением возможного эффекта наблюдателя. В отсутствие непосредственного доступа к пользователям наблюдение может производиться с использованием различных существующих инструментов, которые предоставляют визуальное отображение визитов пользователей на сайт, теплокарты кликов и т.д. Например, сервис Яндекс.Вебвизор позволяет просматривать видео поведения пользователей на сайте – с траекториями движения мыши, кликами, навигацией между страницами, выбором текста, прокруткой – таким образом давая множество данных для анализа юзабилити. Существуют также более специализированные инструменты для отображения и подробного анализа теплокарт, такие как FengGui или ClickHeat.

Инструменты для логирования поведения пользователей и анализа логов серверов могут быть отнесены к этой категории, поскольку они фактиче-

ски документируют действия пользователей на сайте. Крупные сервисы веб-аналитики, такие как Яндекс.Метрика или Google.Analytics (GA), предоставляют все больше средств для оценки метрик, связанных с юзабилити, и отслеживания взаимодействий с пользователем (в GA называемых событиями) на бесплатной или условно-бесплатной основе. Так, в GA исследователями [7] был выделен набор метрик, которые наиболее интересны для проектировщиков взаимодействия в сфере электронной коммерции: average page views per visit, percentage of time spent visits, percentage of click depth visits, bounce rate, order conversion rate, average searches per visit, percent of visits using search, search results to site exits ratio, cart start rate, cart completion rate, checkout start rate, checkout completion rate и, наконец, information find conversion rate. Среди отчетов это Audience Overview (Mobile), Social Network Referrals, Goals Overview (Conversions), Path Length (to conversion), Top Events [8]. Однако все инструменты, упомянутые выше, в лучшем случае позволяют частичную автоматизацию, поскольку подразумевают участие юзабилити-экспертов в интерпретации данных, полученных в ходе наблюдения за пользователем, и в выработке рекомендаций по улучшению сайтов.

Юзабилити-тестирование – это один из наиболее эффективных в проектировании юзабилити метод [9]. Фактически тестирование проходит в лабораторных условиях, но для его проведения стремятся выбрать участников, максимально похожих на реальных пользователей, и дать им специально разработанные реалистичные задания. Успешность в значительной степени зависит от квалификации юзабилити-эксперта [10], а реалистичность контекста использования остается в большинстве случаев недостижимой.

В связи с развитием широкополосных каналов связи, способных передавать видео в реальном времени, в том числе на основе пиринговых сетей [11], появился ряд сервисов для дистанционного юзабилити-тестирования, таких как Userlytics, OpenHallway и другие. Многие из них помогают с подбором участников (которые могут оставаться в привычной для себя обстановке) и, как правило, способны снизить финансовые и временные затраты, особенно для случаев, когда целевые пользователи рассеяны по миру. Среди минусов – более строгие требования к понятности заданий, предлагаемых пользователям, а также риск набора участников, которые слишком много раз участвовали в юзабилити-тестировании и потому более не похожи на реальных пользователей [12]. Существуют специализированные программные инструменты, способные автоматизировать некоторые аспекты юзабилити-тестирования, такие как запись видео и ввод с клавиатуры и мыши, организация заданий, хранение и организация данных и т.д. (например, ПО Morage компании TechSmith способно делать

многое из перечисленного). Однако ни один из инструментов среди попавших в наше поле зрения не способен осуществлять юзабилити-тестирование без активного участия юзабилити-экспертов, и прежде всего в разработке заданий для участников, качественном анализе результатов оценки юзабилити и формулировании рекомендаций по повышению юзабилити.

Опросы и обследования. Несмотря на давно известное первое правило юзабилити, которое гласит: «Не слушайте, что говорят пользователи» (но смотрите, что они делают) [13], опросы пользователей все-таки имеют место в оценке юзабилити. Ранее наиболее распространенным подходом было просто попросить пользователя заполнить форму обратной связи на сайте (например, автоматически показывая ее, когда пользователь собирается закрыть вкладку браузера) в надежде на его добросовестность, но в настоящее время бесплатные или условно-бесплатные сервисы, такие как Usabilityhub, Usabilla, Askusers.ru и др., предлагают иным образом организовать процесс. Владелец веб-сайта или юзабилити-специалист дает ссылку на сайт или скриншот, описывает задания и вопросы, а затем, основываясь на ответах потенциальных пользователей или экспертов, получает оценку юзабилити. Этот метод обычно имеет тенденцию собирать по большей части негативные ответы, но он все же может быть полезен для определения юзабилити-проблем, хотя нет гарантии, что респонденты отражали совокупность целевых пользователей и добросовестно следовали предложенным заданиям.

Еще одна популярная и тоже «непрямая» группа методов – это обследование юзабилити, осуществляемое экспертами, а не пользователями. Юзабилити веб-сайтов проверяется на соответствие некоторым эвристическим рекомендациям, принципам или задачам – в любом случае обследующие должны адекватно представлять особенности и потребности целевых пользователей. Помимо того, что некоторые из сервисов юзабилити-опросов, которые упоминались выше, могут быть использованы для сбора мнений экспертов, существуют и специализированные инструменты для проверки соответствия рекомендациям. Они, однако, чаще направлены на проверку доступности для всех категорий пользователей, поскольку рекомендации, касающиеся этой сферы, хорошо описаны в соответствующих стандартах (W3C's Web Content Accessibility Guidelines, US Section 508 и многих других – например, для пользователей с нарушениями зрения рекомендации существуют более десятка лет [14]) и зачастую могут быть проверены прямо по программному коду веб-сайта. Автоматическая проверка рекомендаций, связанных с юзабилити структуры сайта, сценариев взаимодействия, элементов интерфейса и пр., остается гораздо менее достижимой (относительно свежий и полный обзор дается в [15]).

Следует отметить, что сбор субъективных мнений пользователей остается основным методом для измерения такого аспекта юзабилити, как удовлетворенность, который находится в некоторой изоляции от продуктивности и эффективности, обычно воспринимаемых как более объективные. И хотя активные дискуссии на тему, что же такое удовлетворенность пользователя, продолжаются (см. [16]), в Японии с 1980-х годов началось использование вполне прагматичного подхода в этой области – так называемой эмоциональной (кансей) инженерии, которая, можно сказать, рассматривает удовлетворенность пользователя от взаимодействия с продуктом как «черный ящик» [17]. Этот метод позволяет идентифицировать успешный веб-сайт или другой артефакт через набор субъективных оценок по нескольким эмоциональным шкалам, а затем выработать проектные решения для нового продукта с целью произвести желаемое впечатление на пользователей. Применение эмоциональной инженерии в веб-инженерии, однако, пока не носит системный характер [18], и случаи широко используемой на практике автоматизации этого метода неизвестны.

Автоматизация оценки юзабилити

Основные типы подходов к автоматизации оценки юзабилити. Среди преимуществ автоматической оценки юзабилити веб-сайтов в основном отмечают следующие [19]: снижение затрат и требований к квалификации экспертов, более полный и последовательный охват, возможность прогнозировать потери от юзабилити-проблем, оперативно пробовать различные версии дизайна и т.д. Соответствующие подходы были заложены изначально еще в начале 2000-х годов (см., например, [20]) и в настоящее время могут быть разделены на три большие группы.

Прежде всего это *оценка юзабилити, основанная на взаимодействии*. Можно сказать, это наиболее прямолинейный подход, поскольку он использует данные, возникающие в результате взаимодействия (реального или постановочного) с сайтом, юзабилити которого оценивается. В частности, в ходе анализа могут рассматриваться движение курсора мыши или характеристики ввода с клавиатуры, сравниваться лог реального взаимодействия с некоторым описанным идеалом или изучаться потенциальные юзабилити-угрозы [21]. Однако отмечается, что при помощи таких технологий затруднительно получить достоверные количественные значения, учитывая разнообразие контекстов использования [2]. Фундаментальным недостатком подхода является также то, что для проведения оценки требуется достаточный объем данных о реальных взаимодействиях. Предоставление клиен-

там потенциально низкокачественного веб-интерфейса влечет за собой риск для репутации или даже финансовые потери.

Соответственно, *оценка юзабилити, основанная на метриках*, пытается определить и квантифицировать характеристики веб-сайтов, отражающие их юзабилити – от соответствия рекомендациям по дизайну [22] и визуальной сложности веб-интерфейсов [5] до высокоуровневых факторов дизайна, например объема содержимого страницы или соотношения объемов графики и текста [23]. Полагаем, что уровень юзабилити веб-сайта (особенно, если он низкий) действительно возможно предсказать, основываясь на некоторых факторах, и современные инструменты автоматизации в этой области показывают неплохую корреляцию с результатами оценщиков-людей [15]. Однако этот подход имеет пока не разрешенную проблему с учетом влияния различных категорий пользователей и контекстов использования на сравнительную значимость метрик.

В рамках *оценки юзабилити, основанной на моделировании*, преодоление описанной выше проблемы представляется возможным, поскольку методы искусственного интеллекта (например симуляция), применяемые для оценки юзабилити без реального взаимодействия, могут учитывать модель задач, как и модели контекста использования (пользователя, платформы и среды), которые обычно и создаются в рамках модели ориентированной парадигмы разработки интерфейсов [24]. Например, система MASP+MeMo [25] моделирует взаимодействие с учетом контекста использования и задач (хотя модель пользователя не выделяется, будучи включенной в контекст), в то время как в [26] были предложены настраиваемые модели. В целом подход выглядит многообещающим, особенно с учетом современного бурного развития методов искусственного интеллекта. Однако следует заметить, что вычислительная сложность моделей может оказаться слишком высокой (если только веб-интерфейсы не создаются для относительно простых задач или особых контекстов использования), а адаптация моделей к меняющимся тенденциям в проектировании юзабилити слишком затратной.

Гибридные подходы к оценке юзабилити на основе методов искусственного интеллекта. Описанные выше основные подходы к автоматизированной оценке юзабилити можно обобщить так, как показано в таблице. Из нее видно, что никакой отдельный подход или инструмент не является универсальным решением (отсутствие общепринятого количественного выражения юзабилити, безусловно, способствует такой ситуации) и их комбинация часто оказывается более выигрышной, в том числе с точки зрения экономической эффективности.

Обобщение автоматизированных подходов к оценке юзабилити

Generalization of automated approaches to assessing usability

Подход	Пользователи, веб-сайты, взаимодействие	Ограничения
Основанный на взаимодействии	Реальные пользователи или их представители. Реальные веб-сайты или их прототипы (динамические). Реальные взаимодействия (отраженные в логах)	В покрытии – в основном «собственные» веб-сайты. В затратах времени – требуется накопить достаточное количество взаимодействий. В плохом сервисе – потенциальные потери от низкого качества
Основанный на метриках	Без пользователей. Реальные веб-сайты (как правило, статический анализ). Без взаимодействий (анализ кода или дизайна)	В обобщении – метрики и их сравнительные веса зависят от контекста использования. В специалистах – могут требоваться эксперты или пользователи-аннотаторы
Основанный на моделировании	Модель пользователей или их симуляция. Модель, прототип или реальный сайт. Симулированное взаимодействие, анализ данных других взаимодействий	В ресурсах – требуются собранные данные, актуальные модели, вычислительные ресурсы. В знаниях – модели должны быть созданы и обучены

Гибридные подходы к оценке юзабилити зачастую основываются на методах искусственного интеллекта и машинного обучения, что может включать в себя обработку данных о взаимодействии в реальном времени, основанные на моделях эволюционные подходы [27], юзабилити-модели, обученные на данных взаимодействий или опросов [2], и т.п. Результаты могут варьироваться от только лишь полученной оценки качества взаимодействия до определения потенциальных юзабилити-проблем [28], указания на релевантные рекомендации по дизайну [29] и даже до регенерации интерфейса на лету [30]. Особо можно отметить новый и довольно многообещающий инструмент под названием WaPPU [2], который совмещает анализ взаимодействий, сбор субъективных юзабилити-метрик и обучаемые юзабилити-модели. Этот гибкий подход позволяет как осуществлять количественную оценку юзабилити, так и предсказывать статистически значимые отличия между юзабилити различных версий пользовательских интерфейсов; правда, для этого требуется значительное количество взаимодействий, а интеграция с системами веб-аналитики пока не предусмотрена.

Поскольку известно, что проектирование веб-юзабилити – это итеративный процесс, в котором стратегия «часто проверяй, часто изменяй» является одной из лучших, неудивительно, что в последнее время в этой области расширяется применение эволюционных алгоритмов, основанных на методах искусственного интеллекта [31]. Преимущества эволюционных алгоритмов по сравнению, например, с более широко используемым методом A/B тестирования в реальной среде включают в себя потенциально более быстрые итерации, меньшую необходимость вовлечения дизайнера-человека при создании решений-кандидатов, а также возможность избежать работы реальных пользователей с неудачными версиями интерфейсов. Проблемы с применением эволюционных алгоритмов в проектировании взаимодействия в основном свя-

заны с формулировкой и вычислением фитнес-функции: поскольку оно должно производиться на каждом шаге, его вычислительная сложность критична для общей производительности алгоритма. Поэтому для вычисления фитнес-функции в эволюционных алгоритмах все чаще применяются нейронные сети, фактически симулирующие особенности и предпочтения пользователей (см., например, [32, 33]), что позволяет осуществлять выбор кандидатов на основе аппроксимации фитнес-функции – только приемлемые дизайны показываются пользователям и участвуют в сборе данных в ходе реальных взаимодействий. Нейронные сети рассматривают юзабилити как «черный ящик» и могут адаптироваться к меняющимся требованиям или среде [34], например, к эволюционирующему пониманию юзабилити со стороны поисковых систем.

Оценка поведенческих факторов поисковыми системами. Как упоминалось ранее, крупные поисковые системы имеют собственные алгоритмы оценки юзабилити веб-сайтов, и оно сейчас является значимым фактором, влияющим на ранжирование в результатах поиска. С учетом постоянно идущей битвы против черной и серой поисковых оптимизаций алгоритмы полностью не разглашаются и подвергаются постоянным модификациям. Поэтому специалисты в этой области постоянно ищут малейшие намеки (например, цитаты из выступлений представителей Гугла или Яндекса), чтобы оптимизировать свои сайты под них. С этой точки зрения поисковые системы все более ответственны за формирование практического определения юзабилити, пока оно остается объектом теоретических дискуссий в исследовательском сообществе. Хотя компании время от времени патентуют свои технологии для оценки юзабилити и перечни учитываемых факторов (так, Yahoo! получила патент «Количественный анализ загроможденности веб-страницы и ее влияния на субъективные впечатления» еще в 2006 г.), они публикуются с суще-

ственной задержкой и таким образом не имеют особой практической значимости. Например, заявка на патент Гугл «Модификация ранжирования в результатах поиска в зависимости от неявной реакции пользователей и модель ошибки представления» (US 8938463 B1) была подана в 2007 г., а дата его публикации – 2015 г. В этой ситуации специалисты по оптимизации работают над реальными веб-проектами исходя из собственного представления о значимых юзабилити-факторах, или же подобно сервисам типа WooRank или GetGoodRank стараются непосредственно конвертировать в деньги свои знания (адекватность которых, однако, проверить трудно).

Фактически алгоритмы поисковых систем соответствуют гибридным подходам к оценке юзабилити, хотя в основном полагаются на измерения в рамках взаимодействий (что понятно, поскольку они имеют доступ к огромным объемам данных о поведении пользователей), и оптимизация под поисковые системы в настоящий момент скорее направлена на такие поведенческие факторы. Хотя строгое обоснование и точные количественные оценки отсутствуют, широко признается, что перечисленные далее юзабилити-факторы, основывающиеся на взаимодействии, важны для ранжирования [35].

- Высокий *естественный показатель кликабельности* (CTR) для ключевых слов, на который влияет контент веб-сайта, отображаемый в результатах поиска.

- Низкий *показатель отскоков*: веб-страницы, с которых пользователи быстро уходят, вероятно, оказались им не очень полезны (хотя это необязательно означает, что у этих страниц низкое юзабилити).

- Высокая *посещаемость*: если веб-сайт популярен, особенно с высоким количеством повторных посещений, это может свидетельствовать о его юзабилити.

- Большое *время пребывания на сайте*: если посетители проводят на нем долгое время, это считается свидетельством в пользу его высокого юзабилити, хотя вообще-то обычно показатель юзабилити «время выполнения задачи» стремятся минимизировать, а не максимизировать.

- Высокая *активность пользователей*: если пользователи активно комментируют содержимое сайта, значит, они смогли как оценить его полезность, так и успешно осуществить взаимодействие с сайтом, вводя данные и отправляя их для публикации.

В настоящее время есть ряд свидетельств, что поисковые системы в своей оценке юзабилити начинают прибегать и к моделированию, например, патент Гугл (US 8903690 B2) «Связь между инструментами тестирования графических пользовательских интерфейсов и моделированием производительности человека для оценки юзабилити»,

опубликованный в 2014 г. Конечно же, поисковые системы хотели бы иметь и возможность оценивать сайты на основе их программного кода, и уже в 2006 г. в упомянутом ранее патенте Yahoo! было предложено строить линейную регрессию с использованием метрик контента (текста), ссылок, дизайна и т.п. Системно элементы оценки юзабилити, основанной на метриках, были впервые введены Яндексом в районе 2009 г., когда были запущены алгоритмы семейства MatrixNet, позволившие этой системе на тот момент получить преимущество перед Гугл, запустившим свой алгоритм машинного обучения Panda лишь в 2011 г. [36]. В таких алгоритмах позитивные оценки юзабилити, полученные от людей-оценщиков, сопоставляются очень большому количеству возможных метрик с постоянным пересчетом их сравнительной значимости – так что, однозначный ответ на вопрос о поведенческих факторах, наиболее важных для ранжирования сайтов в результатах поиска, при всем желании не могут дать даже представители поисковых систем.

Таким образом, следует сделать вывод, что даже крупные поисковые системы не определяют юзабилити в явном виде, а скорее рассматривают его как «черный ящик» (что соответствует подходу, используемому в нейронных сетях) и аппроксимируют оценку, поскольку для точной оценки юзабилити требуется способность полного понимания текстовой информации, которой не достигли еще даже поисковые системы. Следовательно, концептуализация и практическая оценка юзабилити со стороны поисковых систем, которые в значительной степени формируют веб-среду на сегодняшний день, постоянно эволюционируют и требуют соответствующих гибких и адаптивных подходов со стороны веб-дизайнеров и инженеров.

Заключение

Экономический эффект от проектирования юзабилити остается высоким, однако явно требуются дальнейшие инновации в этой области, поскольку все очевидные и легко устранимые ошибки в основном уже были исправлены [1]. Существуют инструменты, которые совмещают несколько подходов к оценке юзабилити, стремясь обеспечить более полную оценку и тесную связь с проектированием юзабилити, то есть не только измерение и идентификацию юзабилити-проблем, но и помощь в их исправлении, например, посредством связи с релевантными рекомендациями или шаблонами проектирования.

Результаты данного обзора позволяют предположить, что барьеры на пути к более полной и эффективной автоматизации оценки юзабилити включают в себя сохраняющиеся неясности в количественном выражении юзабилити в сочетании с меняющимися стандартами веб-индустрии и ожи-

даниями пользователей. На практике в настоящее время поисковые системы осуществляют оценку юзабилити на основе комбинации подходов, основанных на взаимодействии, метриках и моделях. Новейшие методы в этой области также включают в себя эволюционные алгоритмы для проектирования взаимодействия (как, например, было предложено в [31]), которые, в частности, могут работать с меняющейся фитнес-функцией, адаптирующейся к постоянно меняющемуся концептуальному представлению об юзабилити.

Исследование выполнено при поддержке РФФИ в рамках научного проекта № 16-37-60060 мол_а_дк.

Литература

- Nielsen J. Usability ROI declining, but still strong. 2008. URL: <https://www.nngroup.com/articles/usability-roi-declining-but-still-strong/> (дата обращения: 20.01.2017).
- Speicher M., Both A., and Gaedke M. Ensuring web interface quality through usability-based split testing. LNCS, 2014, vol. 8541, pp. 93–110.
- Paas F.G. and Van Merriënboe J.J. The efficiency of instructional conditions: an approach to combine mental effort and performance measures. Jour. of the Human Factors and Ergonomics Society, 1993, vol. 35, no. 4, pp. 737–743.
- Buettner R. Cognitive workload of humans using artificial intelligence systems: towards objective measurement applying eye-tracking technology. LNCS, 2013, vol. 8077, pp. 37–48.
- Bakaev M. and Avdeenko T. A quantitative measure for information transfer in human-machine control systems. Proc. IEEE Int. Siberian Conf. on Control and Communications (SIBCON), 2015, pp. 1–4.
- Карпов А.А., Ронжин А.Л., Усов В.М. Инструментальная методика тестирования интерактивного бесконтактного человеко-машинного взаимодействия при использовании шлема виртуальной реальности // Пилотируемые полеты в космос. № 3. 2015. С. 43–53.
- Hasan L., Morris A. and Proberts S. Using Google Analytics to evaluate the usability of e-commerce sites. LNCS, 2009, vol. 5619, pp. 697–706.
- Cardello J. Five essential analytics reports for UX strategists. 2014. URL: <https://www.nngroup.com/articles/analytics-reports-ux-strategists/> (дата обращения: 20.01.2017).
- Nielsen J. The most important usability activity. 2012. URL: <https://www.nngroup.com/articles/the-most-important-usability-activity/> (дата обращения: 20.01.2017).
- Nielsen J. Anybody can do usability. 2009. URL: <https://www.nngroup.com/articles/anybody-can-do-usability/> (дата обращения: 20.01.2017).
- Савельев А.И. Алгоритмы обработки данных в контролируемых аккаунтах системы видеоконференцсвязи // Информационно-управляющие системы. 2016. Т. 82. № 3. С. 906–913.
- Schade A. Remote usability tests: moderated and unmoderated. 2013. URL: <https://www.nngroup.com/articles/remote-usability-tests/> (дата обращения: 20.01.2017).
- Nielsen J. First rule of usability? Don't listen to users. 2001. URL: <https://www.nngroup.com/articles/first-rule-of-usability-dont-listen-to-users/> (дата обращения: 20.01.2017).
- Leporini B. and Paternò F. Increasing usability when interacting through screen readers. Universal access in the information society, 2004, vol. 3, no. 1, pp. 57–70.
- Dingli A. and Cassar S. An intelligent framework for web-site usability. Advances in Human-Computer Interaction, 2014, vol. 5, pp. 1–13.
- Speicher M., Both A., and Gaedke M. SOS: does your search engine results page (SERP) need help? Proc. 33rd Annual ACM Conf. on Human Factors in Computing Systems, 2015, pp. 1005–1014.
- Schütte S.T.W., Eklund J., Axelsson J.R.C., Nagamachi M. Concepts, methods and tools in Kansei engineering. Theoretical Issues in Ergonomics Science, 2004, vol. 5, no. 3, pp. 214–231.
- Bakaev M., Gaedke M., Khvorostov V., Heil S. Extending Kansei engineering for requirements consideration in web interaction design. LNCS, 2016, vol. 9671, pp. 513–518.
- Dingli A. and Mifsud J. Useful: A framework to mainstream web site usability through automated evaluation. IHCI, 2011, vol. 2, no. 1, pp. 10–30.
- Vanderdonckt J., Beirekdar A., Noirhomme-Fraiture M. Automated evaluation of web usability and accessibility by guideline review. LNCS, 2004, vol. 3140, pp. 17–30.
- Grigera J., Garrido A., and Rivero J.M. A tool for detecting bad usability smells in an automatic way. LNCS, 2014, vol. 8541, pp. 490–493.
- Gay G. and Qi Li C. AChecker: open, interactive, customizable, web accessibility checking. Proc. 2010 Intern. Cross Disciplinary Conf. on Web Accessibility (W4A'10), 2010, p. 23.
- Lin Y.C., Yeh C.H., and Wei C.C. How will the use of graphics affect visual aesthetics? A user-centered approach for web page design. Int. Jour. of Human-Computer Studies, 2013, vol. 71, no. 3, pp. 217–227.
- Kieras D. Model-based evaluation. In Jacko, J.A. & Sears, A. (eds) The human-computer interaction handbook. NJ: Lawrence Erlbaum Associates, 2003, pp. 1139–1151.
- Feuerstack S., et al. Automated usability evaluation during model-based interactive system development. LNCS, 2008, vol. 5247, pp. 134–141.
- Ladry J., Palanque P., Barboni E., and Navarre D. Model-based usability evaluation and analysis of interactive techniques. Proc. 5th Intern. Workshop on Model Driven Development of Advanced User Interfaces at ACM CHI, 2010, pp. 21–24.
- Pleuss A., Wollny S., and Botterweck G. Model-driven development and evolution of customized user interfaces. Proc. 5th ACM SIGCHI Symposium on Engineering Interactive Computing Systems, 2013, pp. 13–22.
- Grigera J., Garrido A., Rivero J.M., Rossi G. Automatic detection of usability smells in web applications. Intern. Jour. of Human-Computer Studies, 97, 2017, pp. 129–148.
- Viswanathan S. and Peters J.C. Automating UI guidelines verification by leveraging pattern based UI and model based development. ACM CHI'10 EA on Human Factors in Computing Systems, 2010, pp. 4733–4742.
- Barrera-León L., et al. TUKUCHIY: a dynamic user interface generator to improve usability. Intern. Jour. of Web Information Systems, 2016, vol. 12, iss. 2, pp. 150–176.
- Bakaev M., Gaedke M. Application of Evolutionary Algorithms in Interaction Design. Proc. IEEE EIConRusNW, 2016, pp. 125–130.
- Qu Q.X. Kansei knowledge extraction based on evolutionary genetic algorithm: an application to e-commerce web appearance design. Theoretical Issues in Ergonomics Science, 2015, vol. 16 (3), pp. 299–313.
- Guo F., et al. Optimization design of a webpage based on Kansei engineering. Human Factors and Ergonomics in Manufacturing & Service Industries, 2016, vol. 26, no. 1, pp. 110–126.
- Yao X. Evolving artificial neural networks. Proc. IEEE, 1999, vol. 87, no. 9, pp. 1423–1447.
- Dean B. Google's 200 Ranking Factors: The Complete List. 2016. URL: <http://backlinko.com/google-ranking-factors> (дата обращения: 20.01.2017).
- Siotos M. A panda wandering in the Siberian steppe... 2011. URL: <http://www.freshegg.co.uk/blog/technical-seo/search-algorithms/panda-wondering-siberian-steppe> (дата обращения: 20.01.2017).

**MODERN TRENDS IN AUTOMATED USABILITY EVALUATION AND BEHAVIORAL FACTORS
IN SEARCH ENGINE ALGORITHMS**

M.A. Bakaev¹, Ph.D. (Engineering), Senior Researcher, maxis81@gmail.com

¹Novosibirsk State Technical University, K. Marks Ave. 20, Novosibirsk, 630073, Russian Federation

Abstract. Nowadays, there are persisting ambiguities in usability conceptualization for quantitative measurement and a certain gap between its more objective components (effectiveness and efficiency) and a fairly subjective aspect of user satisfaction.

The paper mentions that the practical goal is not necessarily measuring website's usability, but rather improving its quality-in-use or assessment by respective algorithms of major search engines, which increase the importance of "behavioral factors" in their rankings. The paper describes traditional methods in usability evaluation and design, as well as up-to-date tools that introduce certain automation into them.

The authors identify approaches within automated usability evaluation and summarize possibilities and limitations of each one. The approaches are based on interactions, metrics and modeling. Furthermore, there is an overview of hybrid approaches that use more and more AI and machine learning methods. Therefore, an effective method for website interaction quality evaluation is neural networks, which are applied within evolutionary web design algorithms to calculate fitness function for candidate solutions. The final part of the paper describes how global search engines (such as Yandex and Google) perform automated website usability evaluation and provides a list of the most important behavioral factors. In fact, they mostly consider usability as a "black box" and actively apply machine-learning methods to approximate its calculation.

Thus, usability conceptualization and practical assessment by search engines, which nowadays shape the web to a great extent, is constantly evolving and, in our opinion, calls for respectively flexible approaches from web designers and engineers.

Keywords: human-computer interaction, usability engineering, web design, artificial intelligence, search engines optimization.

Acknowledgements. The reported study has been funded by RFBR within the research project no. 16-37-60060 mol_a_dk.

References

1. Nielsen J. *Usability ROI Declining, But Still Strong*. 2008. Available at: <https://www.nngroup.com/articles/usability-roi-declining-but-still-strong/> (accessed January 20, 2017).
2. Speicher M., Both A., Gaedke M. Ensuring web interface quality through usability-based split testing. *LNCS*. 2014, vol. 8541 (Web Engineering), pp. 93–110.
3. Paas F.G., Van Merriënboe J.J. The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Jour. of the Human Factors and Ergonomics Society*. 1993, vol. 35 (4), pp. 737–743.
4. Buettner R. Cognitive Workload of Humans Using Artificial Intelligence Systems: Towards Objective Measurement Applying Eye-Tracking Technology. *LNCS*. 2013, vol. 8077 (KI 2013: Advances in Artificial Intelligence), pp. 37–48.
5. Bakaev M., Avdeenko T. A quantitative measure for information transfer in human-machine control systems. *IEEE Int. Siberian Conf. on Control and Communications (SIBCON)*. 2015, pp. 1–4.
6. Karpov A.A., Ronzhin A.L., Usov V.M. Instrumental Methods of Testing an Interactive Contactless Human-Machine Interaction When Using a Helmet-Mounted Display. *Pilotiruemye polety v kosmos [Manned Spaceflight]*. 2015, no. 3 (16), pp. 43–53.
7. Hasan L., Morris A., Proberts S. Using Google Analytics to evaluate the usability of e-commerce sites. *LNCS*. 2009, vol. 5619 (Human centered design), Springer Berlin Heidelberg Publ., pp. 697–706.
8. Cardello J. *Five Essential Analytics Reports for UX Strategists*. 2014. Available at: <https://www.nngroup.com/articles/analytics-reports-ux-strategists/> (accessed January 20, 2017).
9. Nielsen J. *The Most Important Usability Activity*. 2012. Available at: <https://www.nngroup.com/articles/the-most-important-usability-activity/> (accessed January 20, 2017).
10. Nielsen J. *Anybody Can Do Usability*. 2009. Available at: <https://www.nngroup.com/articles/anybody-can-do-usability/> (accessed January 20, 2017).
11. Savelev A.I. Algorithms of Data Processing in Supervised Accounts of a Videoconferencing System. *Informatsionno-upravliaiushchie sistemy [Information and Control Systems]*. 2016, vol. 82 (3), pp. 906–913 (in Russ.).
12. Schade A. *Remote Usability Tests: Moderated and Unmoderated*. 2013. Available at: <https://www.nngroup.com/articles/remote-usability-tests/> (accessed January 20, 2017).
13. Nielsen J. *First Rule of Usability? Don't Listen to Users*. 2001. Available at: <https://www.nngroup.com/articles/remote-usability-tests/> (accessed January 20, 2017).
14. Leporini B., Paternò F. Increasing usability when interacting through screen readers. *Universal access in the information society*. 2004, vol. 3 (1), pp. 57–70.
15. Dingli A., Cassar S. An intelligent framework for website usability. *Advances in Human-Computer Interaction*. 2014, vol. 5, pp. 1–13.
16. Speicher M., Both A., Gaedke M. SOS: Does Your Search Engine Results Page (SERP) Need Help? *Proc. 33rd Annual ACM Conf. on Human Factors in Computing Systems*. 2015, pp. 1005–1014.

17. Schütte S.T.W., Eklund J., Axelsson J.R.C., Nagamachi M. Concepts, methods and tools in Kansei Engineering. *Theoretical Issues in Ergonomics Science*. 2004, vol. 5 (3), pp. 214–231.
18. Bakaev M., Gaedke M., Khvorostov V., Heil S. Extending Kansei Engineering for Requirements Consideration in Web Interaction Design. *LNCS*. 2016, vol. 9671, pp. 513–518.
19. Dingli A., Mifsud J. Useful: A framework to mainstream web site usability through automated evaluation. *IJHCI*. 2011, vol. 2 (1), pp. 10–30.
20. Vanderdonck J., Beirekdar A., Noirhomme-Fraiture M. Automated evaluation of web usability and accessibility by guideline review. *LNCS*. 2004, vol. 3140 (Web Engineering), pp. 17–30.
21. Grigera J., Garrido A., Rivero J.M. A tool for detecting bad usability smells in an automatic way. *LNCS*. 2014, vol. 8541 (Web Engineering), pp. 490–493.
22. Gay G., Qi Li C. AChecker: open, interactive, customizable, web accessibility checking. *Proc. 2010 Int. Cross Disciplinary Conf. on Web Accessibility (WAA'10)*. 2010, no. 23, p. 23.
23. Lin Y.C., Yeh C.H., Wei C.C. How will the use of graphics affect visual aesthetics? A user-centered approach for web page design. *Int. Journ. of Human-Computer Studies*. 2013, vol. 71 (3), pp. 217–227.
24. Kieras D. Model-based evaluation. *The human-computer interaction handbook*. In Jacko, J.A. & Sears, A. (Eds.) NJ, Lawrence Erlbaum Associates Publ., 2003, pp. 1139–1151.
25. Feuerstack S., Blumendorf M., Kern M., Kruppa M., Quade M., Runge M., Albayrak S. Automated usability evaluation during model-based interactive system development. *LNCS*. 2008, vol. 5247 (Engineering Interactive Systems), pp. 134–141.
26. Ladry J., Palanque P., Barboni E., Navarre D. Model-based usability evaluation and analysis of interactive techniques. *Proc. 5th Int. Workshop on Model Driven Development of Advanced User Interfaces at ACM CHI*. 2010, pp. 21–24.
27. Pleuss A., Wollny S., Botterweck G. Model-driven development and evolution of customized user interfaces. *Proc. 5th ACM SIGCHI Symp. on Engineering Interactive Computing Systems*. 2013, pp. 13–22.
28. Grigera J., Garrido A., Rivero J.M., Rossi G. Automatic detection of usability smells in web applications. *Int. Jour. of Human-Computer Studies*. 2017, vol. 97, pp. 129–148.
29. Viswanathan S., Peters J.C. Automating UI guidelines verification by leveraging pattern based UI and model based development. *ACM CHI'10 EA on Human Factors in Computing Systems*. 2010, pp. 4733–4742.
30. Barrera-León L., Mejia-Molina N., Carrillo-Ramos A., Flórez-Valencia L., Pavlich-Mariscal J.A. TUKUCHIY: A Dynamic User Interface Generator to Improve Usability. *Int. Jour. of Web Information Systems*. 2016, vol. 12, iss. 2, pp. 150–176.
31. Bakaev M., Gaedke M. Application of Evolutionary Algorithms in Interaction Design. *Proc. IEEE EIConRusNW*. 2016, pp. 125–130.
32. Qu Q.X. Kansei knowledge extraction based on evolutionary genetic algorithm: an application to e-commerce web appearance design. *Theoretical Issues in Ergonomics Science*. 2015, vol. 16 (3), pp. 299–313.
33. Guo F., Liu W.L., Cao Y., Liu F.T., Li M.L. Optimization design of a webpage based on Kansei Engineering. *Human Factors and Ergonomics in Manufacturing & Service Industries*. 2016, vol. 26 (1), pp. 110–126.
34. Yao X. Evolving artificial neural networks. *Proc. IEEE*. 1999, vol. 87, no. 9, pp. 1423–1447.
35. Dean B. *Google's 200 Ranking Factors: The Complete List*. Available at: <http://backlinko.com/google-ranking-factors> (accessed January 20, 2017).
36. Siotos M. *A Panda Wandering in the Siberian Steppe...* *The Fresh Egg Blog*. 2011. Available at: <http://www.freshegg.co.uk/blog/technical-seo/search-algorithms/panda-wondering-siberian-steppe> (accessed January 20, 2017).

Примеры библиографического описания статьи

1. Бакаев М.А. Современные тенденции в автоматизированной оценке юзабилити и поведенческие факторы в алгоритмах поисковых систем // Программные продукты и системы. 2017. Т. 30. № 3. С. 447–455. DOI: 10.15827/0236-235X.119.447-455.
2. Bakaev M.A. Modern trends in automated usability evaluation and behavioral factors in search engine algorithms. *Programmnye produkty i sistemy* [Software & Systems]. 2017, vol. 30, no. 3, pp. 447–455 (in Russ.). DOI: 10.15827/0236-235X.119.447-455.