

Analiza wybranych metod oceny użyteczności w procesie tworzenia aplikacji internetowych

Krzysztof Nowak*, Daniel Samolej

Politechnika Lubelska, Katedra Informatyki, Nadbystrzycka 36B, 20-618 Lublin, Polska

Streszczenie. Artykuł przedstawia analizę wykorzystania testów A/B, metody KLM, heurystyk Nielsena oraz badań na użytkownikach, w procesie tworzenia autorskiej gry przeglądarkowej typu multiplayer. Badania przeprowadzono w trakcie tworzenia aplikacji, na różnych etapach jej powstawania. Celem badań było zweryfikowanie korzyści płynących z zastosowania, wspomnianych metod.

Słowa kluczowe: użyteczność; KLM; Nielsen

*Autor do korespondencji.

Adres e-mail: krzysztof376@gmail.com

Analysis of selected usability assessment methods in the process of creating web applications

Krzysztof Nowak*, Daniel Samolej

Department of Computer Science, Lublin University of Technology, Nadbystrzycka 36B, 20-618 Lublin, Poland

Abstract. The article presents the analysis of the use of A/B tests, KLM method, Nielsen heuristics and user research in the process of creating a proprietary multiplayer browser game. The research was carried out during application development, at various stages of its creation. The aim of the study was to verify the benefits of using these methods.

Keywords: usability; KLM; Nielsen

*Corresponding author.

E-mail address: krzysztof376@gmail.com

1. Wstęp

Szybki rozwój technologiczny na świecie, owocuje coraz większą liczbą powszechnie dostępnych aplikacji internetowych. Ze względu na dużą konkurencyjność tychże aplikacji, bardzo ważne stało się ich odpowiednie zaprojektowanie. Mając dużą liczbę alternatywnych rozwiązań, użytkownik może bardzo szybko porzucić daną aplikację na rzecz innej, istotnym elementem każdej aplikacji, stała się więc, jej użyteczność. Spowodowało to pojawienie się szeregu metod do jej oceny, które mogą być świetnym sposobem na poprawę aspektów czytelności i ergonomii interfejsów w tworzonych aplikacjach.

W poniższym artykule przedstawione zostaną wyniki badań dotyczących analizy wykorzystania wybranych metod oceny użyteczności w procesie tworzenia aplikacji internetowej. Metody te to: testy A/B, metoda KLM, heurystyki Nielsena oraz badania na użytkownikach. Aby sprawdzić wpływ tych metod na proces tworzenia aplikacji została stworzona gra przeglądarkowa typu multiplayer. Podczas jej tworzenia zostały wykorzystane wyżej wspomniane metody.

2. Cel badań

Celem badań było zweryfikowanie korzyści płynących z zastosowania, wspomnianych we wstępie metod, w procesie tworzenia aplikacji.

W artykule postawiono następującą tezę:

Wykorzystanie testów A/B, metody KLM, heurystyk Nielsena oraz badań na użytkownikach, w procesie tworzenia aplikacji, pozwala na znaczne zmniejszenie liczby błędów w interfejsie.

3. Przegląd literatury

Metody oceny użyteczności służą podnoszeniu jakości interfejsów graficznych. Wśród często stosowanych metod wymienia się m.in. testy A/B, Keystroke-Level Model, ocenę heurystyczną, czy badania użytkowników.

Wśród najefektywniejszych metod wymienia się badania z udziałem użytkowników. Mogą być one przeprowadzone przy niewielkich kosztach – jak wspomina Jakob Nielsen w artykule „Why You Only Need to Test with 5 Users” [1], wystarczy 5 badanych osób, aby wykryć większość problemów z interfejsem. Wnioskując on również, że przeprowadzanie kilku mniejszych badań przynosi lepsze efekty, niż przeprowadzenie jednego większego badania. Nie wymaga ono również stosowania drogiego sprzętu - wystarczy podstawowe stanowisko komputerowe [10].

Kolejną często stosowaną metodą jest Keystroke-Level Model. Jest ona stosowana do oszacowania czasu wykonania zadania przy pomocy klawiatury i myszki [6]. Stosowanie jej polega na stworzenie dwóch wersji tego samego rozwiązania,

a następnie rozbić ich na elementarne akcje, takie jak wciśnięcie klawisza, czy przesunięcie kursora myszki, w celu wyliczenia czasu wykonania każdego z nich [7]. Rozwiązanie z krótszym czasem wykonania uznawane jest jako lepsze. Jak wspomniano w artykułach „Using the Keystroke-Level Model to Estimate Execution Times” [2] oraz „The Keystroke-Level Model for User Performance Time with Interactive Systems” [4], rzeczywiste czasy wykonywanych akcji mogą się jednak różnić pomiędzy użytkownikami, w zależności od stopnia zaawansowania.

Wśród innych często stosowanych metod jest ocena heurystyczna. Polega ona na ocenie systemu przez ekspertów na podstawie opracowanych zasad. Jakob Nielsen w jednym ze swoich artykułów, zatytułowanym „How to Conduct a Heuristic Evaluation” [3] opisuje, że optymalnym rozwiązaniem jest analiza interfejsu przez czterech ekspertów, gdyż pozwala to na wykrycie większości problemów. Jednak taką analizę może przeprowadzić nawet jedna osoba, co pozwoli na znalezienie około 35% błędów [9].

Kolejną często wykorzystywaną metodą są testy A/B. Ich stosowanie polega na przygotowaniu dwóch wariantów produktu lub prototypu, które będą się nieznacznie różnić [5][8]. Następnie są one prezentowane użytkownikom i na podstawie ich opinii wybierane są najlepsze rozwiązania.

4. Procedura badawcza

W celu potwierdzenia postawionej we wstępie hipotezy przeprowadzony został szereg badań na różnych etapach tworzenia aplikacji. Ze względu na istotną kolejność ich zastosowania podczas jej tworzenia, cały proces został omówiony poniżej.

1. Przeprowadzenie testów A/B dotyczących układu interfejsu aplikacji

W zaprojektowanej aplikacji można wyróżnić dwa główne widoki. Jest to widok strony przed zalogowaniem oraz widok strony po zalogowaniu. Dla każdego widoku zostały stworzone po dwa prototypy, które następnie zostały zaprezentowane 5. różnym osobom. Uczestnicy mieli za zadanie wybrać opcję, która ich zdaniem jest lepsza.

2. Stworzenie prototypu aplikacji

Dla wybranej, w poprzednim etapie, wersji układu został stworzony prototyp przedstawiający funkcjonalności aplikacji.

3. Wykorzystanie metody KLM

Po zbudowaniu prototypu gry, zostały zaprojektowane różne rozwiązania dla wybranych funkcjonalności. Następnym krokiem było wybranie zaprojektowanych rozwiązań na podstawie wyników otrzymanych po zastosowaniu metody KLM.

4. Implementacja gry

W tym miejscu przystąpiono do implementacji gry, zgodnie z projektem oraz rozwiązaniami wypracowanymi w poprzednich etapach.

5. Wykorzystanie heurystyk Nielsena

Mając finalną wersję aplikacji wykorzystano heurystyki Nielsena w celu wykrycia potencjalnych błędów. Istotnym elementem na tym etapie był brak wprowadzenia poprawek dla błędów wskazywanych przez heurystyki Nielsena. Fakt ten został podyktowany chęcią ustalenia czy błędy wykryte tą metodą zostaną zauważone podczas badań na użytkownikach.

6. Badania na użytkownikach

W celu sprawdzenia ogólnego odbioru aplikacji, wykrycia ewentualnych błędów oraz zweryfikowania rozwiązań wypracowanych przy zastosowaniu poprzednich metod, zostały przeprowadzone badania na użytkownikach. Była to grupa 5 osób, która po raz pierwszy miała styczność z aplikacją wykorzystywaną do testów. Uczestnicy zostali poinformowani o typie aplikacji z jaką mają do czynienia, jednak nie zostały im wyjaśnione zasady rozgrywki. Zostało to podyktowane chęcią sprawdzenia, czy interfejs jest odpowiednio przystosowany dla nowych użytkowników. Wszyscy uczestnicy badania mieli za zadanie wykonać szereg czynności, według podanego poniżej scenariusza.

Zadanie 1

- Zarejestruj się
- Dodaj opis oraz zdjęcie do swojego profilu w grze
- Zmień hasło do swojego konta

Zadanie 2

- Sprawdź co się dzieje w budynku: Centrum
- Rozbuduj wybrany przez siebie budynek
- Stwórz 10 mieczników
- Zmień nazwę swojej osady
- Sprawdź ogólny stan posiadania jednostki typu miecznik (we wszystkich osadach łącznie)
- Zaatakuj osadę innego gracza
- Odczytaj raport z przeprowadzonego ataku
- Zobacz co się dzieje w twojej drugiej osadzie

Zadanie 3

- Sprawdź profil swojego sojuszu
- Ustal kto jest założycielem sojuszu oraz kto ma uprawnienia administratora
- Opuść sojusz

Zadanie 4

- Zmień opis oraz zdjęcie swojego sojuszu
- Nadaj uprawnienia administratora wybranemu członkowi który takich uprawnień nie posiada
- Zaproś gracza, który jest 2-gi w rankingu ogólnym, do swojego sojuszu
- Przyjmij jednego z graczy, który wysłał prośbę o dołączenie do twojego sojuszu
- Rozwiąż sojusz

Zadania 2 i 3 zostały przeprowadzone na specjalnie przygotowanym koncie, na którym: gracz posiadał 2 lokacje,

należał do sojuszu jako zwykły członek, posiadał w dzienniku kilka archiwalnych raportów z bitew. Zadanie 4 zostało przeprowadzone na koncie podobnym do tego wykorzystanego przy zadaniach 2 i 3 z tą różnicą, że na tym koncie gracz miał uprawnienia administratora w sojuszu do którego należał.

7. Implementacja odrzuconych układów interfejsu

Na tym etapie zostały zaimplementowane układy interfejsu, odrzucone w testach A/B, w celu przeprowadzenia weryfikacji o której mowa w następnym punkcie.

8. Weryfikacja wyboru układu w wersji prototypowej na finalnej wersji aplikacji

Badanie to polegało na zaprezentowaniu układów aplikacji, opisanych w punkcie 1, w wersji finalnej użytkownikom, którzy brali udział w wyborze wersji prototypowej. Etap ten został przeprowadzony w celu potwierdzenia wyboru układów w wersji prototypowej.

9. Implementacja alternatywnych rozwiązań funkcjonalności

Na tym etapie zaimplementowano rozwiązania odrzucone po wykorzystaniu metody KLM.

10. Porównanie odrzuconych i zaimplementowanych rozwiązań

Etap ten został przeprowadzony w celu potwierdzenia „trafności” rozwiązań wybranych metodą KLM. Badania polegały na prezentacji obu wersji rozwiązań, grupie użytkowników, którzy brali udział w badaniach w punkcie 6.

5. Wyniki badań

5.1. Testy A/B

Wersje układów zostały oznaczone odpowiednio A i B. Zestawienie wyborów użytkowników dotyczące układu interfejsu gry, przedstawia tabela 1 oraz tabela 2.

Tabela 1. Wybory układu przed zalogowaniem

	Użytkownik					Podsumowanie
	1	2	3	4	5	
Prototyp	B	B	B	A	B	B
Wersja finalna	A	A	A	A	B	A

Tabela 2. Wybory układu gry (po zalogowaniu)

	Użytkownik					Podsumowanie
	1	2	3	4	5	
Prototyp	B	A	A	B	B	B
Wersja finalna	B	A	B	B	B	B

5.2. Metoda KLM

Pierwsze i drugie rozwiązanie dla wybranych funkcjonalności zostały oznaczone odpowiednio A i B. Wyniki otrzymane metodą KLM, przedstawia tabela 3. Wybory użytkowników, otrzymane podczas weryfikacji metody KLM przedstawia tabela 4. Zestawienie tych wyborów z wynikami metody KLM oraz rozwiązań zaimplementowanych na jej podstawie, prezentuje tabela 5.

Tabela 3. Wyniki metody KLM

Funkcjonalność	Rozwiązanie	Czas [s]
1	A	11,46
	B	11,46
2	A	12,62
	B	15,12
3	A	17,84
	B	12,84
4	A	5,00
	B	5,00

Tabela 4. Wybory użytkowników

Funkcjonalność	Użytkownik				
	1	2	3	4	5
1	B	B	B	B	B
2	B	B	A	A	A
3	B	A	A	A/B	B
4	B	B	B	B	B

Tabela 5. Wyniki badań weryfikujących metodą KLM

Funkcjonalność	Użytkownicy	KLM	Implementacja
1	B	A/B	A
2	A	A	A
3	A/B	B	B
4	B	A/B	A

5.3. Heurystyki Nielsen

Wyniki analizy aplikacji przeprowadzone tą metodą, przedstawia tabela 6. Podczas badań na użytkownikach, żaden z nich nie zwrócił uwagi na problemy prezentowane w tabeli 6.

Tabela 6. Wyniki oceny heurystycznej

Lp.	Problem	Naruszane heurystyki
1.	Brak informacji o liczbie możliwych do utworzenia jednostek	5. Zapobiegaj błędom
2.	Zbędna informacja o niedostępnych jednostkach	8. Dbaj o estetykę i umiar
3.	Wymiana surowca na taki sam rodzaj	5. Zapobiegaj błędom
4.	Brak okna dialogowego z potwierdzeniem usunięcia członka sojuszu	5. Zapobiegaj błędom
5.	W całej aplikacji brakuje informacji, o miejscu w którym znajdują się użytkownik	1. Pokazuj status systemu

5.4. Badania na użytkownikach

Wyniki badań na użytkownikach przedstawia tabela 7 oraz tabela 8.

Tabela 7. Zestawienie pomyłek popełnianych przez użytkowników

Zadania	Użytkownicy				
	1	2	3	4	5
1	a				
	b				
	c				
2	a				
	b				
	c				
	d				
	e				
	f				
	g				
	h				
3	a				
	b				
	c				
4	a				
	b				
	c				
	d				
	e				

Legenda

	Brak pomyłek
	Wystąpiły pomyłki, ale zadanie zostało ukończone
	Liczba pomyłek, która zaowocowała pominięciem zadania

Tabela 8. Czasy wykonywania zadań przez użytkowników

Zadanie	Czas [s] wykonania zadania dla poszczególnych użytkowników			
	1	2	3	4
1	93	145	150	80
2	374	350	433	254
3	115	46	84	48
4	125	190	125	170

Podczas badania zanotowano szereg spostrzeżeń, których podsumowanie prezentuje się następująco.

Spostrzeżenia badającego:

- Użytkownicy mieli wyraźny problem z przełączeniem się do odpowiedniego widoku w grze. Głównie w zadaniu 2, w podpunktach: a, c, d, e, f i g, oraz w zadaniu 4c. Opisana sytuacja bardzo często powtarzała się dla kilku użytkowników.
- Problemy z odnalezieniem funkcjonalności w danym widoku, występowały sporadycznie. Pojawiły się one w zadaniach: 1a, 2e, 4a. Były to jednak jednorazowe przypadki.
- Wśród użytkowników często pojawiał się problem ze skojarzeniem danej strony w grze z funkcjonalnością wymaganą do ukończenia zadania. Dotyczy to głównie zadania 2d, gdzie problem powtórzył się aż u 3 użytkowników. Podobna sytuacja występowała w zadaniu 2f, gdzie opisany problem wystąpił u 2 badanych osób. Pozostałe przypadki wystąpienia

problemu, to sytuacje dotyczące tylko jednej z badanych osób.

Spostrzeżenia użytkowników

Wszystkie poniżej przedstawione sugestie użytkowników to spostrzeżenia unikalne – nie powtarzające się wśród pozostałych badanych:

- W miejscu wpisywania hasła, mogłaby istnieć opcja jego chwilowego podglądu. Pozwoliłoby to szybką korektę błędnie wpisanego hasła, zamiast konieczności jego ponownego wpisywania.
- Ikony oznaczające poszczególne uprawnienia członków sojuszu mogłyby być większe.
- Opcja pozwalająca na przełączenie się do rankingu graczy mogłaby być w menu głównym.
- W widoku budynku, mógłby istnieć przycisk, przełączający do widoku w którym można rozbudowywać budynki.
- Przycisk otwierający opcję edytowania uprawnień członków sojuszu mógłby być większy oraz jego podświetlenie po najechaniu na niego powinno być wyraźniejsze.

6. Analiza wyników

6.1. Testy A/B

Przy wyborze układu strony przed zalogowaniem w wersji prototypowej, zdecydowanie wygrała opcja B (4 na 5 użytkowników) jednak przy wyborze w wersji finalnej sytuacja się odwróciła. Widząc gotową wersję 4 na 5 użytkowników wybrałoby opcję A. Ciekawy jest również fakt, że tylko 2 na 5 użytkowników podtrzymało swój wybór prototypu na wersji finalnej. Warto również dodać, że były to dwie różne wersje układu.

Przy wyborze prototypu układu gry 3 na 5 użytkowników wybrało opcję B. Wybór tej wersji potwierdził się podczas wyboru wersji finalnej, gdyż wszyscy Ci trzech użytkownicy podtrzymali swój wybór. Dodatkowo widok finalnego układu w wersji B przekonał jednego z użytkowników, który w wersji prototypowej wybrał opcję A. Tylko jeden z użytkowników pozostał przy wersji A.

6.2. Metoda KLM

Wybory użytkowników, zebrane podczas przeprowadzonego badania weryfikującego wykorzystanie metody KLM, tylko w połowie pokrywają się z tymi wybranymi dzięki metodzie KLM. Istotne jest to, że niezgodność wyboru użytkowników oraz metody KLM dotyczy rozwiązań, które według metody KLM mają taki sam czas i wybór odpowiedniego rozwiązania został dokonany przez twórców aplikacji. W niektórych przypadkach użytkownicy byli bardziej przychylni wersji która według metody KLM jest dłuższa.

Użytkownicy mieli możliwość wyboru obu rozwiązań, jeśli te odpowiadałyby im w równym stopniu. Podczas badań

tylko raz padło takie stwierdzenie. Zdecydowana większość opowiadała się po stronie jednego z rozwiązań.

6.3. Heurystyki Nielsena

Najczęściej pojawiającym się problemem, wykrytym tą metodą, jest brak zapobiegania błędom w niektórych miejscach aplikacji. Jest to istotny problem, gdyż pomimo braku jego wykrycia podczas badań na użytkownikach, może on w znaczący sposób wpływać na frustrację graczy w momencie jego wystąpienia.

6.4. Badania na użytkownikach

Badania pozwoliły na wykrycie szeregu problemów z którymi nowi użytkownicy mogą mieć problemy. Pozwoliły one również na odkrycie kilku nowych elementów, które mogłyby usprawnić interfejs aplikacji. Pomimo błędów popełnianych przez badane osoby, nie zaobserwowano żadnego elementu w interfejsie, który uniemożliwiłby wykonanie któregoś z zadań wszystkim użytkownikom.

Podczas badań w części dotyczącej strony gry, najistotniejszy okazał się problem z przełączaniem do odpowiedniego widoku w grze, który zawierał funkcjonalności wymagane do ukończenia zadania. Wśród użytkowników pojawiały się również problemy ze skojarzeniem danego widoku jako miejsca zawierającego daną funkcjonalność. Problemy te w dużym stopniu były spowodowane pierwszym zetknięciem z aplikacją oraz z nieznanymi mechanikami rozgrywki. Po zakończeniu badania, użytkownikom umożliwiono swobodne zapoznanie się z grą i testowanie zastosowanych w niej mechanik. Po takim zapoznaniu większość użytkowników stwierdziła, że spora część mechanik i rozwiązań w niej zastosowanych jest spójna i logiczna, a popełniane wcześniej błędy wynikają z braku doświadczenia z mechaniką rozgrywki. Wszyscy stwierdzili, że gdyby mieli możliwość przejścia samouczka dotyczącego rozgrywki, zadania postawione im w trakcie badań byłyby znacznie prostsze.

Podczas badań na użytkownikach, w części dotyczącej strony przed zalogowaniem, nie stwierdzono istotnych problemów. Aspekt ten może wynikać z faktu iż każda osoba biorąca udział w badaniach miała już styczność z tym elementem aplikacji internetowych.

7. Wnioski

Badania weryfikacyjne na użytkownikach pokazały, że wybór układu interfejsu na wersji prototypowej nie zawsze pokrywa się z wyborem, dokonany przez użytkowników, na wersji finalnej. Kolejnym istotnym elementem, który wykazały przeprowadzone badania jest fakt, iż niektóre z rozwiązań wypracowanych za pomocą metody KLM, nie pokrywają się w tej kwestii z preferencjami użytkowników. Heurystyki Nielsena pozwoliły na wykrycie kilku problemów, które zawierała stworzona aplikacja, jednak wykryte problemy pozostały niezauważane przez użytkowników. Badania na użytkownikach potwierdziły iż

niektóre z rozwiązań, zaimplementowane na podstawie wyników otrzymanych metodą KLM, mogą być nietrafione.

Nawiązując do tezy badawczej należy stwierdzić, że wykorzystanie wymienionych w niej metod oceny użyteczności pozwoliłoby na znaczne ograniczenie liczby błędów w interfejsie. Metody te są w stanie znacznie usprawnić proces tworzenia aplikacji oraz niwelują liczbę popełnianych błędów. Ważne jest jednak odpowiednie ich wykorzystanie. Istotne jest również stosowanie kilku wymienionych metod jednocześnie.

W testach A/B dotyczących wyboru układu strony w wersji prototypowej istotne jest jak najdokładniejsze zaprezentowanie wizji finalnej wersji tego układu. Metoda KLM pozwala na odpowiedni wybór rozwiązania tylko wtedy gdy scenariusze badanego rozwiązania znacznie się różnią. W przypadku gdy oba scenariusze są podobne lub wręcz identyczne, najlepiej jest zaprezentować dane rozwiązanie kilku użytkownikom, aby móc wybrać to które ich zdaniem jest najlepsze. Heurystyki Nielsena pozwalają ocenić spójność całego systemu oraz wykryć jego niedopracowane elementy. Metoda ta pozwala też na wykrycie elementów, których brak, może być niezauważony przez użytkowników, a których istnienie mogłoby im znacznie uprościć posługiwanie się aplikacją.

Z przeprowadzonych badań można wywnioskować, że najlepszą metodą oceny użyteczności są badania na użytkownikach, gdyż to one pozwalają na ostateczną weryfikację zastosowanych rozwiązań. To właśnie dla użytkownika jest tworzony interfejs aplikacji i to on będzie się nim posługiwał, nie powinien więc zaskakiwać fakt, że to właśnie badania na użytkownikach pozwalają na najlepszą ocenę użyteczności.

Literatura

- [1] Why You Only Need to Test with 5 Users, <https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/>, [29.10.2019]
- [2] Using the Keystroke-Level Model to Estimate Execution Times, <ftp://www.infomus.org/pub/PPM/Publications/paper-L18-KeystrokeLevelModel.pdf>, [29.10.2019]
- [3] How to Conduct a Heuristic Evaluation, <http://www.ingenieriasimple.com/usabilidad/HeuristicEvaluation.pdf>, [29.10.2019]
- [4] The Keystroke-Level Model for User Performance Time with Interactive Systems, https://www.researchgate.net/profile/Stuart_Card/publication/220426122_The_Keystroke-Level_Model_for_User_Performance_Time_with_Interactive_Systems/links/09e4150e317b6ce55400000/The-Keystroke-Level-Model-for-User-Performance-Time-with-Interactive-Systems.pdf, [29.10.2019]
- [5] On the Complexity of A/B Testing, <http://proceedings.mlr.press/v35/kaufmann14.pdf>, [29.10.2019]
- [6] Kieras D., Using the Keystroke-Level Model to Estimate Execution Times, University of Michigan, 2001
- [7] Sauro J., Estimating productivity: Composite operators for keystroke level modeling, Berlin Heidelberg: Springer-Verlag, 2009.

- [8] An ultimate guide to A/B testing on prototypes, <https://uxplanet.org/an-ultimate-guide-to-a-b-testing-on-pre-live-apps-4bd57679e8cc>, [29.10.2019]
- [9] Ocena heurystyczna, <http://hci.pjwstk.edu.pl/index.php?page=ocena-heurystyczna>, [27.10.2019]
- [10] Wybrane metody oceny użyteczności stron i aplikacji internetowych, http://krainabiznesu.pl/wp-content/uploads/sites/2/2014/04/White_Paper-Wybrane_metody_oceny_uzytecznozci_stron_i_aplikacji_internetowych.pdf, [27.10.2019]