

Jakość stron WWW.

Analiza śladów użytkowników.

Spis treści.

1. Dane służące do analizy wyników badań użyteczności.
2. Metryki użyteczności.
3. Zastosowanie clicktracking'u.
4. Motiontracking. Charakterystyka danych. Ścieżka wizyty użytkownika.
5. System ewaluacji zachowań użytkowników na stronach internetowych bez udziału moderatora i użytkownika:
 - a. Architektura.
 - b. Model.
 - c. Kontroler.
6. Case study: Czytodziala.pl.
 - a. Zadania dla użytkowników.
 - b. Wyniki.
7. Metody porównywania danych, uzyskanych poprzez zdalny system oraz eyetracker.

• Dyrektor operacyjny Internet Works, autor Czytodziala.pl (monitoring dostępności WWW)

1. Dane służące do analizy wyników badań użyteczności.

Po każdym przeprowadzonych badaniach i testach jakości użytkowej stron WWW uzyskujemy zbiór informacji, który wymaga odpowiedniej analizy oraz podsumowania w postaci raportu.

W zależności od użytej metod badania można uzyskać następujące dane:

- Notatka obserwatora dotycząca wykonywanych zadań przez użytkowników.
- Lista kontrolna z odpowiedziami.
- Uporządkowany zbiór danych ze skończoną liczbą pól, uzupełniany w trakcie badania w bazie danych lub arkuszu kalkulacyjnym przez prowadzącego badania lub obserwatora.
- Wywiad przed- i potestowy.
- Dane pozyskane automatycznie za pomocą sprzętu (np. Eyetrack 6000) lub oprogramowania (np. GazeTracker, CrazyEgg, ClickMapa).
 - ślad skupienia wzroku (Gaze Trail),
 - obszary patrzenia 3D (Look zone 3D),
 - analiza konturów (Contour Analysis),
 - mapa cieplna (HeatMap),
 - mapa kliknięć (ClickMap).
- Statystyczne dane:
 - średni czas wykonywania każdego zadania oraz średnia ilość kliknięć,
 - rozkład uczestników kończących zadania, uporządkowany wg pomyślności ukończenia.
- Analiza ekspercka:
 - analiza heurystyk i problemy,
 - analiza porównawcza konkurencji,
 - raport ekspercki wraz z propozycją rozwiązań problemów i uzasadnieniem uwag.

Mając powyższego dane, w całości lub częściowo, należy utworzyć dokument, nazywany audytem, lub raportem badania użyteczności.

Metody badań użyteczności z pasywnym udziałem użytkowników.

Metody z pasywnym udziałem użytkownika nie wymagają podjęcia dodatkowych działań z jego strony. Do takich metod zaliczane są badania eyetracking czy ewaluację z użytkownikiem, podczas której można pozyskać takie dane jak:

- ślady drogi (ślady myszki),
- mapę cieplną,
- mapę kliknięć,
- obszary skupienia wzroku,
- pliki z historią dokonanych akcji na stronie.

2. Metryki użyteczności.

Istnieje kilka modeli metryki użyteczności - Metrics for Usability Standards in Computing (*MUSiC*), Software Usability Measurement Inventory (*SUMI*), the Skill Acquisition Network (*SANe*), semi-Automated Interface Designer and Evaluator (*AIDE*), Diagnostic Recorder for Usability Measurement (*DRUM*).

Metryki użyteczności służą po to, by ocenić o ile bardziej użyteczny jest produkt po wprowadzonych poprawkach uzyskanych z testów użyteczności.

Najprostszym przykładem jest wyznaczenie kilku zadań do wykonania i zmierzenie takich wartości jak:

- ilość kliknięć potrzebna do osiągnięcia celu,
- czas potrzebny do osiągnięcia celu.

Zadanie/Czas	Przed poprawkami	Po przeprojektowaniu	Poprawa %
Zadanie 1	20 s	6 s	233
Zadanie 2	80 s	55 s	45
Zadanie 3	150 s	160 s	-6
Czas razem	250 s	121 s	
Arytmetycznie			107% lepiej
Geometrycznie		(biorąc pod uwagę spadek użyteczności)	63% lepiej

Po naniesieniu poprawek w produkcie, należy powtórzyć badania stosując te same czynniki w zbliżonych do natywnych warunkach.

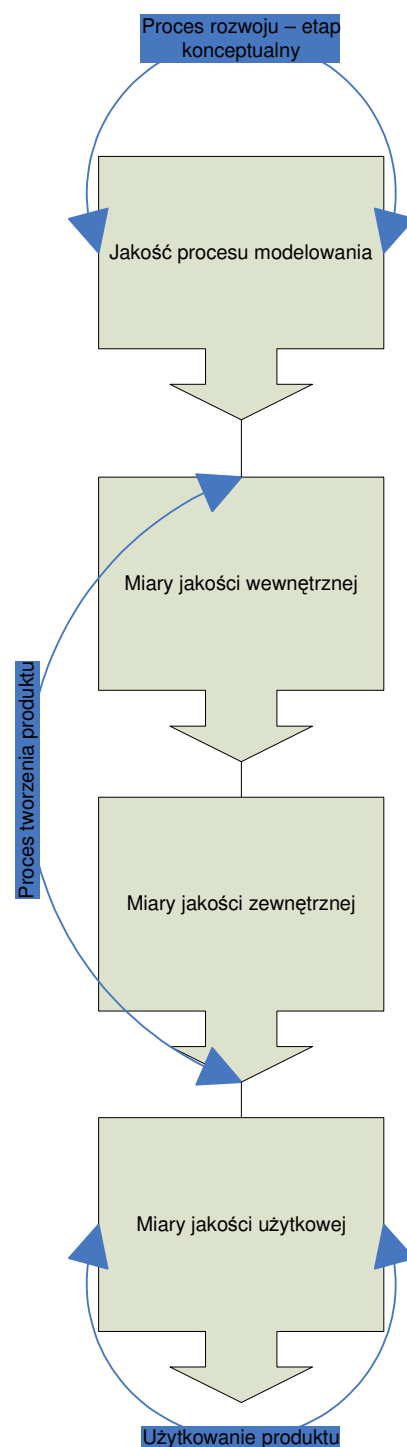
Z kolei w metryce jakości użytkowej ISO 9126 zawarta jest charakterystyka pozwalająca wymiennie ocenić produkt, co jest szczególnie ważne by móc porównać wskaźniki przed i po wprowadzonych zmianach.

Pierwsze pojęcie to **efektywność**, którą cechują trzy miary:

- skuteczność zadaniowa,
- kompletność zadaniowa,
- częstotliwość błędów.

Produktywność jest ściśle związana z użytkownikiem i zadaniami, które on wykonuje. Występują pojęcia:

- czasu wykonania zadania, czyli jak długo trwało zakończenie zadania.
- skuteczności zadaniowej, czyli jak skuteczni są użytkownicy, ile zadań w jednostce czasu potrafią wykonać poprawnie.
- produktywność ekonomiczna – jak korzystny kosztowo jest użytkownik, czyli skuteczność zadaniowa / koszt zadania.



Bezpieczeństwo produktu względem użytkownika, rozumiane jest jako wpływ produktu na postawę i zdrowie.

Strata finansowa jest miarą, która określa skalę możliwości poniesienia strat z powodu niewystarczającej użyteczności produktu.

Przy pomocy miar **zadowolenia** można zmierzyć:

- jak bardzo jest usatysfakcjonowany użytkownik (np. ankieta potestowa, rozmowa z ekspertem),
- na ile z poszczególnych cech produktu jest zadowolony użytkownik?

Na podstawie powyższych, oraz innych modeli metryki użyteczności, można stworzyć jeden skonsolidowany, hierarchiczny model mierzenia użyteczności. Taki model nosi nazwę Quality In Use Integrated Measurement (*QUIM*). W modelu QUIM zawartych jest 10 składników: wydajność, efektywność, produktywność, zadowolenie, zdolność uczenia się, bezpieczeństwo, zaufanie, dostępność, uniwersalność i używalność – każdy odpowiada za odpowiedni faset użyteczności, który jest definiowany w istniejących standardach i modelach. Następnie tych 10 czynników, jest rozłożone na 26 mniejszych, mierzalnych kryteriów, które ostatecznie składają się z 127 konkretnych metryk.

3. Zastosowanie clicktrackingu.

Zanim wyznaczę definicję clicktrackingu, należy wspomnieć o narzędziach typu *actiontracker*, które wykorzystuje się podczas ewaluacji z użytkownikiem, zarówno pasywnej jak i aktywnej. Podczas odwiedzania poszczególnych stron, przeglądarka odnotowuje kolejne wykonywane przez użytkownika akcje, zarówno między stronami jak i na samej stronie. Do takich akcji można zaliczyć:

- wejście na stronę,
- przełączenie między stronami (zakładkami),
- przejście z jednej strony na drugą w ramach jednego okna,
- wpisanie danych formularzowych na stronie,
- kliknięcie przycisku/linku na stronie,
- zaznaczenie tekstu, kopiowanie, wklejanie,
- edycja, jeśli na taką pozwala strona (najczęściej technologia AJAX) przeciąganie, opuszczanie elementów strony.

Clicktracking jest metodą wykorzystywaną do śledzenia aktywności użytkowników na stronach WWW, lecz nie samych akcji. Używane jest także, jako narzędzie przy testach A/B (*Split A/B testing*). Podejście testowania A/B pozwala porównać dwie alternatywy do siebie, zweryfikować funkcjonalność dwóch konkurencyjnych prototypów systemu.

Clicktracking pozwala stworzyć mapę relacji użytkowników-link na stronie WWW. Można odwzorować liczbę kliknięć w każdy link, czy też dowolny inny element na stronie. Dzięki temu wykrywana jest niepopularna treść oraz miejsca, cieszące się największym zainteresowaniem.

Dane, które uzyskuje się za pomocą clicktrackingu, to:

- czas kliknięcia
- pozycja kliknięcia
- liczba kliknięć w dany element
- parametry użytkownika, który kliknął (przeglądarka, rozdzielczość, lokalizacja za pomocą GeoIP, miejsce, z którego nastąpiło kliknięcie)

Clicktracking jest bardzo przydatny w e-commerce, pomaga rozwiązywać problem tzw. product placement, czyli umieszczaniu produktów na stronie internetowej.

Możliwe sposoby prezentacji danych to:

- mapa kliknięć
- mapa cieplna

Mapa kliknięć jest dwuwymiarowym obrazem, składającym się z dwóch warstw – badanego systemu, oraz nałożonej powłoki z informacją o kliknięciach.

Mapa cieplna, inaczej termiczna jest natomiast graficzną reprezentacją punktów skupionych w pewne obszary, odległe od siebie o zdefiniowaną stałą.

Aby powstał obszar, w promieniu R (określonym w systemie generowania mapy), musi znaleźć się X kliknięć.

Możliwa jest reprezentacja graficzna najbardziej gorących miejsc na badanej stronie WWW. Wszystkie kliknięcia w danym obszarze są prezentowane jako obrazy termiczne, gdzie miejsca o najwyższym współczynniku kliknięć przybierają czerwony kolor.

Istnieje kilka znanych systemów, oferujących kolekcję statystyk dot. strony oraz prezentowania we wspomnianych formach graficznych. Do najbardziej popularnych serwisów zagranicznych można zaliczyć Clickdensity, CrazyEgg oraz ClickHeat, zaś w Polsce system statystyk stat24.com (ClickMapa) oraz SiteDoctor.

4. Motiontracking. Charakterystyka danych.

Motiontracking – to sposób na „nagrywanie” ruchów i akcji wykonywanych przez użytkownika na stronie WWW w czasie rzeczywistym. Podczas wizyt na stronie, ruchy i kliknięcia użytkownika są notowane do zdalnej bazy danych lub plików log. Zgodnie z określonym niżej modelem, każdorazowo zapamiętywana jest czwórka parametrów. Owe „nagrywanie” cechuje logowanie w bardzo krótkich odstępach czasu ($< 100\text{ms}$) ruchów wykonywanych przy pomocy myszki czy innego narzędzia wskazującego.

Odbywa się to trybie dyskretnie ciągłym, dla każdego użytkownika z osobna, gdzie ostatecznie powstaje zbiór danych zawierający takie informacje jak:

- pozycja kursora
- czas pozycji kursora w wyznaczonym miejscu na stronie
- długość przestoju kursora w jednym miejscu
- rodzaj akcji wykonanej podczas przestoju

Normalizując to do czwórki danych $MT = \{P, T, L, A\}$, otrzymamy następujące składowe:

- $P = (x, y)$, jednostką dla x, y jest 1px (1 piksel).
- $T = 00:00:00, \text{h:m:s}$.
- $L = 000 \text{ ms}$ – wyrażone w milisekundach
- A należy do zbioru $\{ \text{NULL}, \text{START}, \text{STOP}, \text{CLICK}, \text{FIX}, \text{CLOSE} \}$
 - NULL – brak podjętej akcji
 - START – naciśnięcie przycisku myszki rozpoczynające akcję (np. kliknięcie w element, przeciągnięcie, wywołanie akcji skryptu wykonywanego po stronie użytkownika)
 - STOP – ukończenie akcji
 - CLICK – kliknięcie w element
 - FIX – miejsce fiksacji, stan w który przechodzi A po minięciu określonego czasu
 - CLOSE – akcja zamknięcia strony lub opuszczenia jej na rzecz innej

Reprezentacją graficzną śladu jest dwuwymiarową krzywą powstałą poprzez aproksymację kolejnych punktów typu MT.

Na podstawie tych danych można utworzyć unikalną ścieżkę wizyty użytkownika na danej stronie internetowej. Metoda poniekąd jest pokrewna z mechanizmem tworzenia ścieżek odwiedzin stron na podstawie bazy danych przebiegów czasów. Zagadnienie jest z dziedziny eksploracji danych.

Analiza przebiegów czasowych obejmuje metody analizy przebiegów czasowych w celu znalezienia: trendów, podobieństw, anomalii oraz cykli.

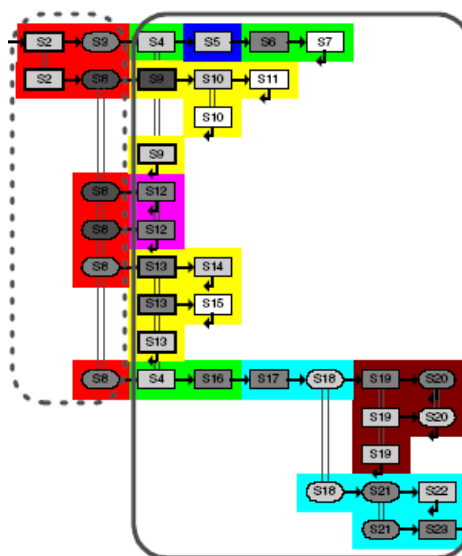
Metodę tę można zastosować również do wykrycia podobieństw między użytkownikami i ich zachowaniem na stronie, by poznać najpopularniejsze przebiegi wizyt na stronie i w następstwie skorygować prototyp strony zgodnie z ukierunkowaniem właśnie użytkowników.

Ścieżka wizyty użytkownika.

Przy metodzie *motiontracking* pomocny jest graf zachowań sieciowych użytkowników (*eng. Web behavior graph*), rysunek obok.

Ciepłe kolory w środku oznaczają poprawną ścieżkę, o odpowiednim kierunku, kierującą użytkownika do właściwej informacji, zimne – oddalenie się od wykonania zadania czy postawionego celu. Za pomocą tego grafu badając stronę internetową można wykryć obszary, które wymagają poprawy.

W przypadku „kreślenia” ścieżki użytkownika w systemie webowym, nie wyznaczamy punktu, który oznacza spełnienie celu, lecz badamy wszystkie ścieżki użytkowników. Określana na wstępie jest idealna ścieżka od pierwszego punktu do ostatniego, która stanowi wzorzec dla pozostały śladów, które są do niej porównywane. Po przekroczeniu ustalonego progu zgodności, można założyć, iż użytkownik osiągnął założony cel lub nie.



5. System ewaluacji zachowań użytkowników na stronach internetowych bez udziału moderatora i użytkownika.

Przedstawiam opis sposobu działania aplikacji webowej typu *tracker*, wdrażanej pod nazwą *obitracks.com*. Za pomocą niej można uzyskać mapę cieplną (*heatmap*) oraz mapę kliknięć (*clickmap*). *ObiTracks.com* wykorzystuje sposób wizualizacji oparty o rozwiązanie *ClickHeat* firmy *Labsmedia*. System jest wykonany w technologii PHP 5.2, w oparciu o bazę danych MySQL 5 oraz biblioteki Javascript logujące zdarzenia. Aby rozpocząć kolekcjonowanie danych, należy umieścić kod Javascript przed znacznikiem `</body>` w sekcji strony, która ma być badana. Przykład:

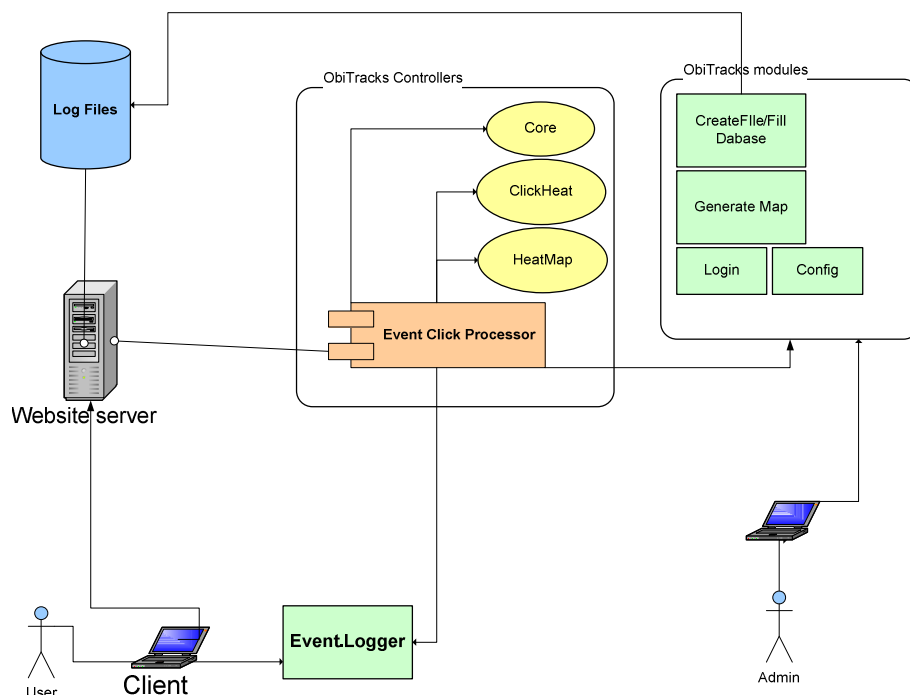
```
<script type="text/javascript" src="http://czytodziala.pl/svetovid/js/clickheat.js"></script><script type="text/javascript"><!--
clickHeatSite = 'czytodziala.pl';clickHeatGroup = document.title;clickHeatServer =
'http://czytodziala.pl/svetovid/click.php';initClickHeat(); //--> </script></body>
```

Jak widać, zasada *Event.logger* jest podobna do tej, którą stosuje *Google Analytics*, przy kolekcjonowaniu danych statystycznych.

Dane są kolekcjonowane do plików tekstowych:

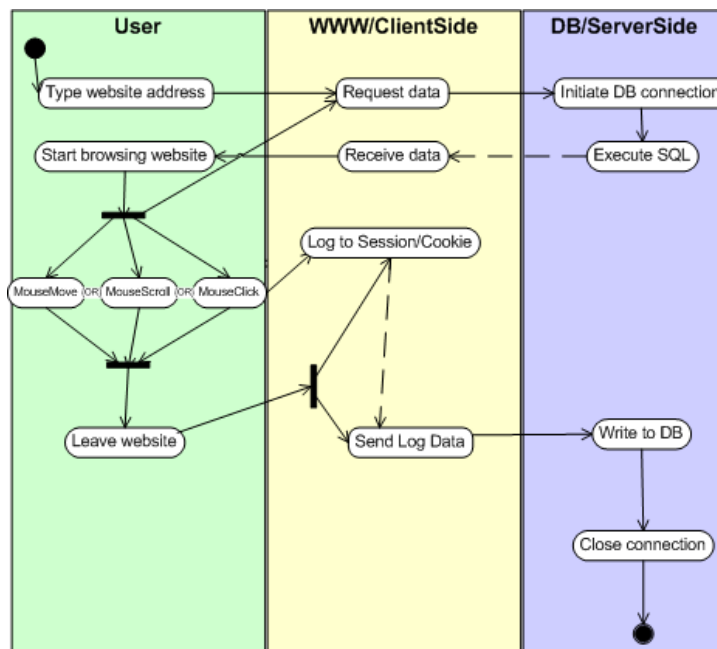
XIY|szerokość_przeglądarki|liczba_klikniec|czas_klikniecia|numer_IP_uzytkownika

Aplikacje cechuje następująca architektura:



W zależności od badanej strony, należy ustawić dla Event.loggera informację o tym, jaki układ ona posiada. Może to być strona wycentrowana o zmiennym rozmiarze, stałym lewym menu i zmienną treścią, na cały ekran itp. Uproszczoną wersją, a właściwie składową motiontrackera jest kontroler clicktracker.

Poniżej znajduje się rysunek z modelem aplikacji, typ *swimlanes*.



W przypadku ObiTracks, clicktracking składa się z dwóch mniejszych aplikacji – skryptu zbierającego kliknięcia ze strony klienta oraz kontrolera, który zebrane dane przetwarza ze strony serwera, gdzie jest wdrożony cały system ObiTracks. Dzięki tak zebrany danym, procesy w tym kontrolerze, odpowiadające za generowanie mapy kliknięć (ClickMap) oraz mapy ciepłej (HeatMap), są w stanie wizualizować akcje użytkownika.

W ramach badań przeprowadziłem analizę porównawczą jakości uzyskanych wyników na tych samych zestawach zadań dla aplikacji webowej oraz Eyetrackingu, za pomocą urządzenia do badań Head Mounted Tracker, model Eyetrack 6000 oraz oprogramowania GTAnaly.

6. Case study: Czytodziala.pl.

W listopadzie 2007 roku nastąpiła zmiana wyglądu serwisu monitorującego działanie usług internetowych Czytodziala.pl. Wymagana była ewaluacja wprowadzonych zmian.

Zadania dla użytkowników.

Aby uzyskać dane użytkowników dla porównania skuteczności aplikacji webowej z clicktrackingiem i heatmapą uzyskiwaną z *eyetrackera*, grupa 10 osób została poproszona o wykonanie poniższych zadań.

Zadanie 1

1. Wejdź na stronę <http://czytodziala.pl>.
2. Spróbuj zapisać się na monitoring nasza-klasa.pl dla adresu Gadu-Gadu.
3. Powtórz Krok nr 2 dla niepoprawnych danych.

Zadanie 2

1. Wejdź na stronę <http://czytodziala.pl>.
2. Załóż nowe konto w serwisie.
3. Dodaj jeden kontakt mailowy oraz jeden test na dowolny protokół z dowolnymi ustawieniami.
4. Przypisz kontakt RSS do testu.

Zadanie 3

1. Wejdź na stronę <http://czytodziala.pl>.
2. Zaloguj się na swoje konto.
3. Odpowiedz na trzy pytania:
 - Czy wszystko, co widzę, jest dla mnie jasne? Wiesz gdzie jesteś?
 - Gdybym był w stanie, co bym poprawił na tej stronie?
 - Co przykuwa największą uwagę?

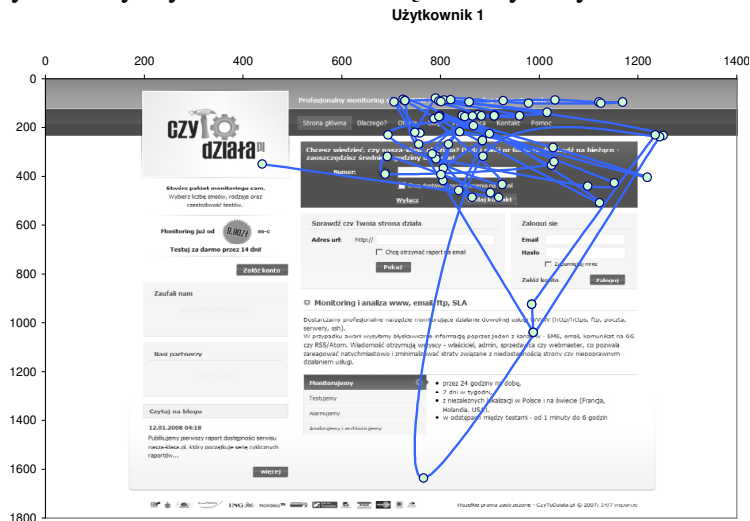
Uwagi i sugestie zostały przesłane za pomocą oprogramowania LiveChat Software, zainstalowanego w roli kanału kontaktowego dla Czytodziala.pl.

Na poniższym rysunku jest mapa kliknięć uzyskana podczas badania.



Wyniki uzyskano 15 stycznia 2008 roku, w laboratorium Politechniki Wrocławskiej. Dane ObiTracks.com zostały zebrane od 8 użytkowników. Najczęściej pracowali oni z przeglądarką Firefox, w jednym przypadku był to MS Explorer. Rozdzielczość nie przewyższała 1280px na szerokość, gdzie najczęściej była to 1024 px na szerokość.

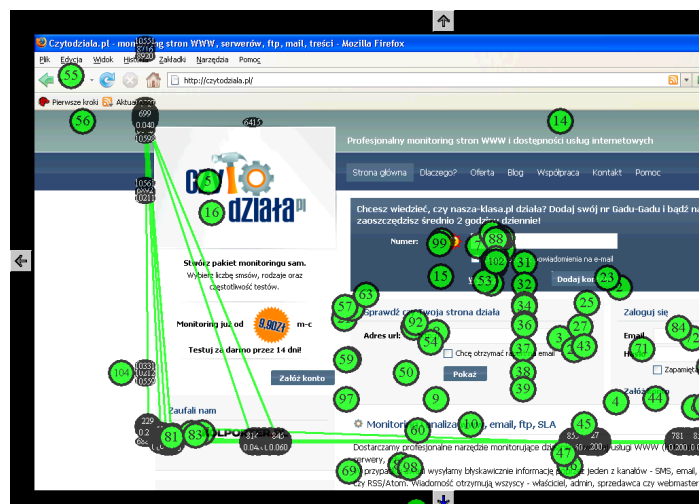
Poniżej przykładowy wynik ścieżki kliknięć z danych systemu obitracks.com.



Mapa cieplna dla wygenerowanego ruchu przedstawia się następująco:



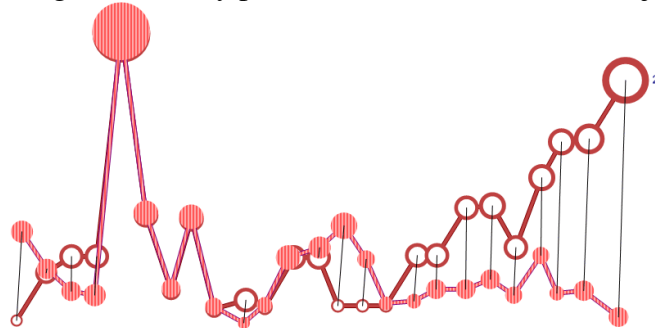
Zadania powtórzono w laboratorium ergonomii. Uzyskano wstępne wyniki, które wymagały przetworzenia, niżej zobrazowane obszary skupienia w kolejności zapisania (duża czułość).



7. Metoda porównania danych.

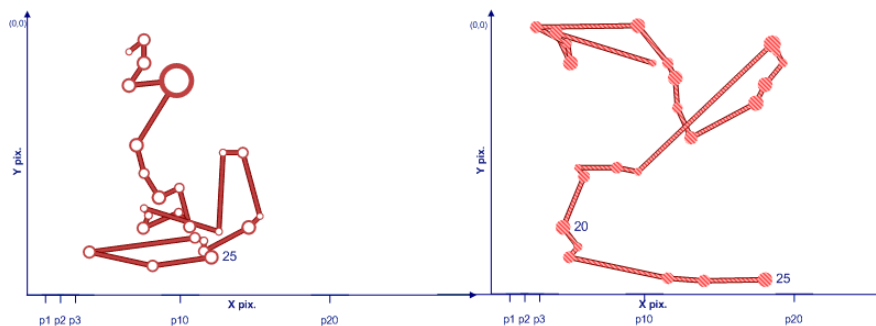
Uzyskane ścieżki ruchów użytkowników podczas badania zdalnego oraz testu z wykorzystaniem Eyetrackera, są porównywane za pomocą następujących opracowanych metod:

1. Suma różnic odległości między punktami w stosunku do łącznej długości ścieżki.



Zgodność ok. 60%

2. Różnica w położeniu kursora i czasie przestoju między poprzedzającym i następującym kliknięciem



Wykorzystując uzyskane dane, należy wyznaczyć następujące parametry:

- NOOP TIME DURATION (ms) – długość przestoju myszki w danej pozycji,
- NOOP POINT – punkt przestoju myszki, uwarunkowany stałą NOOP_LIM, która określa potrzebny czas do zmiany stanu z ACTIVE na NOOP,
- MAX_NORMALIZED_DIFF_LENGTH – stała obliczana poprzez pierwiastkowanie iloczynu szerokości i wysokości ekranu przeglądarki użytkownika. Definiuje, czy różnica długości dwóch ścieżek mieści się w zakresie podobieństwa, co skutkuje rozpoczęciem dokładnej punktowej analizy.

Działanie algorytmu:

- NOOP_LIM = 100; Limit czasu w ms, po którym kursor przechodzi w stan przestoju,
- MAX_NORMALIZED_DIFF_LENGTH – pierwiastek z iloczynu szerokości i wysokości okna przeglądarki użytkownika,
- porównanie następuje dopiero gdy wartość bezwzględna między długością ścieżek jest mniejsza niż MAX_NORMALIZED_DIFF_LENGTH,
- wyliczenie odległości między punktami dwóch ścieżek wykresu w stosunku do odległości od punktu (0,1) - początek ekranu,
- wyliczenie różnic w czasie przestoju między dwoma ścieżkami,
- uwzględnienie odchylenia standardowego,
- obliczenie iloczynu z poszczególnych współczynników podobieństwa – podobieństwo geometryczne.

Na podstawie przykładowych danych (ramka), można zobaczyć jak przebiega proces porównania ścieżek.

Wg pierwszej metody uzyskuje się większe podobieństwo ścieżek ruchu. Wadą tej metody jest zakładanie, że użytkownicy w sposób podobny wykonują zadania. Dlatego, może ona być stosowana w grupach użytkowników o podobnych profilach.

Druga metoda wprowadza z kolei zbyt restrykcyjne porównanie ścieżek, biorąc pod uwagę położenie punktu, kolejność oraz czas przestoju myszki w nim.

Bezwzględne podobieństwo geometryczne można obliczyć mnożąc przez siebie uzyskane wyniki z powyższych dwóch metod.

Przykładowe wyniki obliczeń:

```
document loaded, functions done
Path1 and Path2 are qualified for further analysis.
190,92,873,922,378,799,628,759,337,120,663,107,811,558,127,189,634,665,
404,720,731,69,844,765,746,566
standard deviation for path1: 279.568 ms
190,92,873,922,378,799,628,759,337,120,663,107,811,558,127,189,634,665,
404,720,731,69,844,765,746,566
676,657,168,903,902,699,637,937,951,917,396,438,77,453,379,611,174,866,
536,258,971,504,202,376,444,400
Noop point path similarity is: 77.3%
Position-Time point path similarity is: 58%
=====
190,92,873,922,378,799,628,759,337,120,663,107,811,558,127,189,634,665,
404,720,731,69,844,765,746,566
standard deviation for path1: 279.568 ms
190,92,873,922,378,799,628,759,337,120,663,107,811,558,127,189,634,665,
404,720,731,69,844,765,746,566
676,657,168,903,902,699,637,937,951,917,396,438,77,453,379,611,174,866,
536,258,971,504,202,376,444,400
Average similarity of two track pathes: 44.83%
```

Literatura.

1. HACKOS, J. T., REDISH J. C., *User and Task Analysis for Interface Design*, New York, John Wiley & Sons, 1998, 6-11, 52-71, 130-142, 189-192, 302-344.
2. HILBERT D. M., REDMILES D. F., *Why Let Perfectly Good Usability Data Go to Waste?* Human-Computer Interaction Consortium Meeting USA, 1998, 1-2.
3. NIELSEN, J., *Usability Engineering*. Boston, MA: Academic Press Professional, 1994.
4. NIELSEN, J., *Designing Web usability - the practice of simplicity*.
5. NIELSEN J., <http://www.useit.com/alertbox>, ISSN 1548-5552, 2003-2007.
6. NIELSEN, J., LORANGER H., *Prioritizing Web Usability*, New Riders Press, April 2006.
7. NORMAN K. L., *Levels of Automation and User Participation in Usability Testing*, 2004.
8. SOFTWARE QUALITY JOURNAL, Vol. 14, No. 2. (June 2006), 159-178.
9. ZELDMAN J., *Designing with Web Standards 2nd Edition*, Peachpit Press, 2006, 22-24, 53-57, 338-341.
10. Standard ISO 9126.
11. <http://www.w3c.org>
12. www.upassoc.org/upa_publications/