

Analiza ruchu w aplikacji webowej

Projektowanie i programowanie systemów internetowych II

mgr inż. Krzysztof Rewak

9 stycznia 2019

Wydział Nauk Technicznych i Ekonomicznych

Państwowa Wyższa Szkoła Zawodowa im. Witelona w Legnicy

Plan prezentacji

1. Klasyczna analityka
2. Przydatna analityka
3. Straszna analityka
4. Podsumowanie

Vinaora Visitors Counter

012462

| | | |
|---|------------|-------|
|  | Today | 11 |
|  | Yesterday | 33 |
|  | This week | 554 |
|  | Last week | 751 |
|  | This month | 1791 |
|  | Last month | 3674 |
|  | All | 12462 |

Online now: 2

Your IP: 127.0.0.1

FIREFOX 3.0.8, WINDOWS

Now: 2009-04-17 20:03

Visitors Counter 1.6

Czasy, w których takie liczniki gości można było znaleźć bardzo często, słusznie już minęły.

Ludzie wykształcili od tego czasu wiele przydatnych narzędzi.
Przynajmniej część z nich warto znać.

Jak to ugryźć?

Gdzie można podpiąć nasłuchiwanie ruchu w naszej aplikacji?

Praktycznie rzecz biorąc wszędzie: zarówno na webowym backendzie, webowym frontendzie czy aplikacji mobilnej.

Co ugryźć?

Czego można dowiedzieć się podczas nasłuchiwania ruchu w naszej aplikacji?

Wszystkiego. W granicach rozsądku. Zazwyczaj.

Klasyczna analityka

Badanie zapytań HTTP

Większość frameworków oferuje proste sposoby na wydobycie informacji o IP klienta.

Przykładowo:

```
public function returnIP(Request $request): JsonResponse {  
    return response()->json(["ip" => $request->ip()]);  
}
```


A znając IP można zbadać już naprawdę wiele ciekawych rzeczy.

← → ⓘ ip-api.com/json/186.27.58.141

JSON Nieprzetworzone dane Nagłówki

Zapisz Kopiuuj Zwiń wszystkie Rozwiń wszystkie

```
as: "AS28024 Nuevatel PCS de Bolivia S.A."
city: "Villa Imperial de Potosí"
country: "Bolivia"
countryCode: "BO"
isp: "Nuevatel PCS de Bolivia S.A."
lat: -19.5892
lon: -65.7509
org: "NuevatelV"
query: "186.27.58.141"
region: "P"
regionName: "Potosí"
status: "success"
timezone: "America/La_Paz"
zip: ""
```

Logowanie zdarzeń

Warto rozważyć logowanie ruchu na stronie wewnątrz własnej aplikacji.
Jest to proste do wykonania, ale warto pamiętać o kilku szczegółach.

Logowanie zdarzeń

| | | | |
|---------------------|----------------|--------|---|
| 2019-01-09 16:16:06 | 37.47.130.112 | ● INFO | Courses list requested and delivered. |
| 2019-01-09 16:13:31 | 37.47.130.112 | ● INFO | Authentication status requested and delivered. |
| 2019-01-09 13:30:49 | 156.17.194.43 | ● INFO | Grades requested for group { id: 16 } by student { no: █████ } and delivered. |
| 2019-01-09 13:30:49 | 156.17.194.43 | ● INFO | Groups requested for course { id: 7 } and delivered. |
| 2019-01-09 13:30:49 | 156.17.194.43 | ● INFO | Courses requested for semester { id: 3 } and delivered. |
| 2019-01-09 13:30:49 | 156.17.194.43 | ● INFO | Semesters requested and delivered. |
| 2019-01-09 13:30:46 | 156.17.194.43 | ● INFO | Courses requested for semester { id: 3 } and delivered. |
| 2019-01-09 13:30:46 | 156.17.194.43 | ● INFO | Groups requested for course { id: 7 } and delivered. |
| 2019-01-09 13:30:46 | 156.17.194.43 | ● INFO | Semesters requested and delivered. |
| 2019-01-09 13:30:42 | 156.17.194.43 | ● INFO | Groups requested for course { id: 7 } and delivered. |
| 2019-01-09 13:30:42 | 156.17.194.43 | ● INFO | Semesters requested and delivered. |
| 2019-01-09 13:30:42 | 156.17.194.43 | ● INFO | Courses requested for semester { id: 3 } and delivered. |
| 2019-01-09 13:30:42 | 156.17.194.43 | ● INFO | Groups requested for course { id: 7 } and delivered. |
| 2019-01-09 13:30:39 | 156.17.194.43 | ● INFO | Courses requested for semester { id: 3 } and delivered. |
| 2019-01-09 13:30:39 | 156.17.194.43 | ● INFO | Semesters requested and delivered. |
| 2019-01-09 13:30:39 | 156.17.194.43 | ● INFO | Courses requested for semester { id: 3 } and delivered. |
| 2019-01-09 13:30:36 | 156.17.194.43 | ● INFO | Semesters requested and delivered. |
| 2019-01-09 13:30:31 | 156.17.194.43 | ● INFO | Authentication status requested and delivered. |
| 2019-01-09 12:05:31 | 77.255.141.192 | ● INFO | Courses requested for semester { id: 3 } and delivered. |

Logowanie zdarzeń

- serwery HTTP (Apache, nginx) najczęściej robią to już za nas w postaci access- i errorlogów;
- warto zastanowić się czy lepiej logować dane do pliku czy do bazy danych;
- jeszcze bardziej warto się zastanowić po jakim czasie dane należy usuwać;

Istnieje nawet standard RFC 5424, który opisuje poziomy logów:

- 100 DEBUG
- 200 INFO
- 250 NOTICE
- 300 WARNING
- 400 ERROR
- 500 CRITICAL
- 550 ALERT
- 600 EMERGENCY

Co można zrobić z takimi logami?

- zbadać zerwane linki;
- sprawdzić czy ktoś nie próbował się włamać do zastrzeżonych miejsc;
- sprawdzić popularność produktów;
- zbadać w którym miejscu procesu zakupowego klient najczęściej rezygnuje;
- i setki innych zastosowań...

Przydatna analityka

Królem webowej analityki od lat jest **Google Analytics**.

Stosowane na frontendzie, Google Analytics jest bardzo proste w użyciu. W bazowej wersji wymaga jedynie załączenia odpowiedniego skryptu, a następnie wywołania jednej metody `ga()`.

Google Analytics (a!ć!)

```
<script>
(function(i,s,o,g,r,a,m){
i['GoogleAnalyticsObject']=r;i[r]=i[r]||function(){
(i[r].q=i[r].q||[]).push(arguments)
},i[r].l=1*new Date();a=s.createElement(o),
m=s.getElementsByTagName(o)[0];a.async=1;
a.src=g;m.parentNode.insertBefore(a,m)
})(window,document,'script',
'https://www.google-analytics.com/analytics.js','ga');

ga('create', 'UA-4815162342-1', 'auto');
ga('send', 'pageview');
</script>
```

Oczywiście każdy sensowniejszy javascriptowy framework już od dawna ma własne wrappery, które niestraszą zminifikowanym kodem.

Przykładowa implementacja GA w Vue.js:

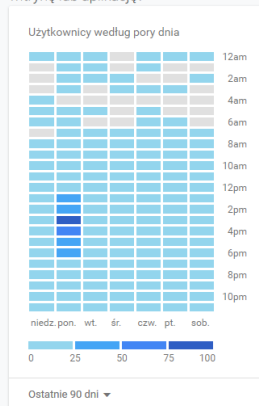
```
Vue.use(VueAnalytics, {  
  id: "UA-4815162342-1",  
  checkDuplicatedScript: true,  
  router  
})
```

Google Analytics

Strona główna Google Analytics



Kiedy użytkownicy odwiedzają Twoją witrynę lub aplikację?



Google Analytics

Użytkownicy

1 275



Nowi użytkownicy

1 280



Sesje

5 903



Sesje na użytkownika

4,63



Odsłony

28 588



Strony/sesja

4,84



Śr. czas trwania sesji

00:03:37

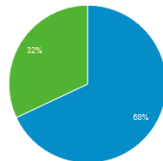


Współczynnik odrzuceń

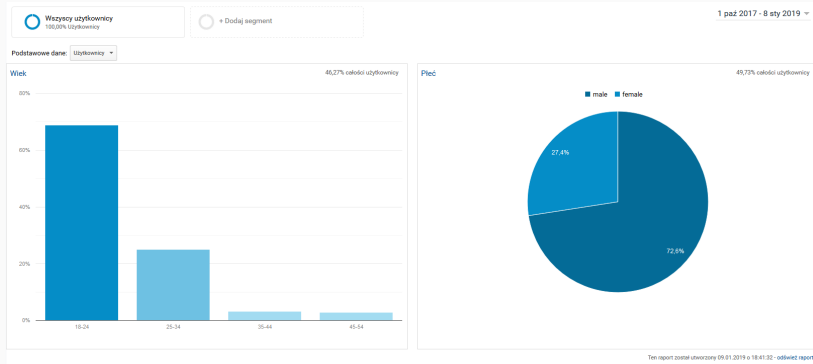
21,85%



■ New Visitor ■ Returning Visitor



Google Analytics



Google Analytics

Podobne kategorie (zasięg)

53,02% całości użytkowników

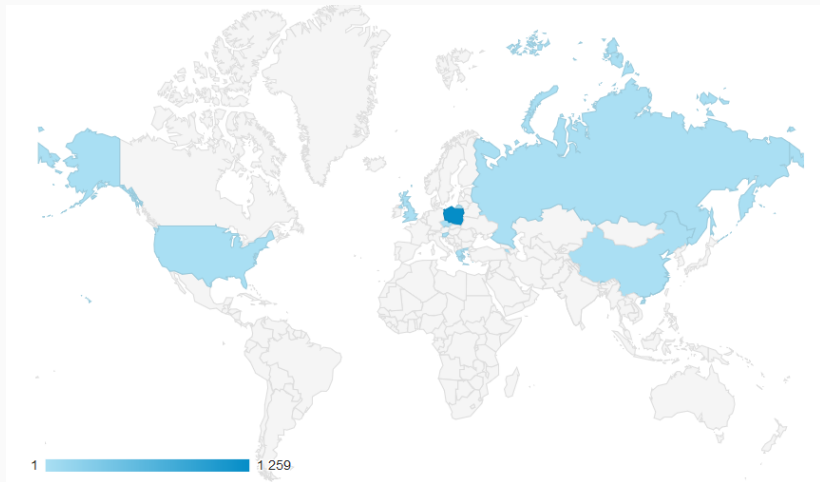
| | | |
|-------|-------------|--|
| 4,37% | <div></div> | Technology/Technophiles |
| 4,29% | <div></div> | Media & Entertainment/Gamers |
| 3,51% | <div></div> | Media & Entertainment/Movie Lovers |
| 3,39% | <div></div> | Shoppers/Value Shoppers |
| 3,17% | <div></div> | Lifestyles & Hobbies/Shutterbugs |
| 3,02% | <div></div> | Media & Entertainment/Music Lovers |
| 2,98% | <div></div> | Food & Dining/Fast Food Cravers |
| 2,89% | <div></div> | Sports & Fitness/Health & Fitness Buffs |
| 2,86% | <div></div> | Media & Entertainment/Gamers/Hardcore Gamers |
| 2,85% | <div></div> | Sports & Fitness/Sports Fans |

Inna kategoria

53,33% całości użytkowników

| | | |
|-------|-------------|--|
| 5,74% | <div></div> | Arts & Entertainment/Humor/Funny Pictures & Videos |
| 3,76% | <div></div> | Arts & Entertainment/Celebrities & Entertainment News |
| 3,35% | <div></div> | Science/Mathematics |
| 3,08% | <div></div> | Reference/General Reference/Dictionaries & Encyclopedias |
| 3,03% | <div></div> | Arts & Entertainment/TV & Video/Online Video |
| 3,01% | <div></div> | Games/Computer & Video Games/Browser Games |
| 2,92% | <div></div> | Online Communities/Photo & Video Sharing/Photo & Image Sharing |
| 2,62% | <div></div> | Internet & Telecom/Email & Messaging |
| 2,58% | <div></div> | Games/Roleplaying Games |
| 2,44% | <div></div> | Sports/Team Sports/Soccer |

Google Analytics



Google Analytics

| <input type="checkbox"/> | System operacyjny ? | Pozyskiwanie | | |
|--------------------------|---------------------|---|---|---|
| | | Użytkownicy ? ↓ | Nowi użytkownicy ? | Sesje ? |
| | | 1 275 % całości: 100,00% (1 275) | 1 282 % całości: 100,16% (1 280) | 5 903 % całości: 100,00% (5 903) |
| <input type="checkbox"/> | 1. Windows | 999 (77,68%) | 1 005 (78,39%) | 4 333 (73,40%) |
| <input type="checkbox"/> | 2. Android | 237 (18,43%) | 235 (18,33%) | 1 293 (21,90%) |
| <input type="checkbox"/> | 3. iOS | 22 (1,71%) | 22 (1,72%) | 102 (1,73%) |
| <input type="checkbox"/> | 4. Linux | 19 (1,48%) | 11 (0,86%) | 72 (1,22%) |
| <input type="checkbox"/> | 5. Macintosh | 8 (0,62%) | 8 (0,62%) | 102 (1,73%) |
| <input type="checkbox"/> | 6. (not set) | 1 (0,08%) | 1 (0,08%) | 1 (0,02%) |

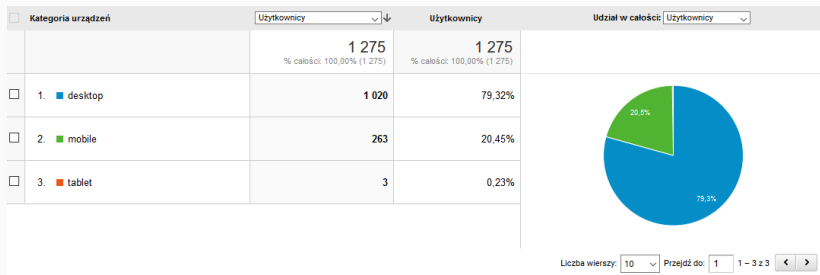
Google Analytics

| <input type="checkbox"/> | Przeglądarka ? | Pozyskiwanie | | |
|--------------------------|--------------------|---|---|---|
| | | Użytkownicy ? ↓ | Nowi użytkownicy ? | Sesje ? |
| | | 1 275 % całości: 100,00% (1 275) | 1 282 % całości: 100,16% (1 280) | 5 903 % całości: 100,00% (5 903) |
| <input type="checkbox"/> | 1. Chrome | 781 (61,21%) | 784 (61,15%) | 3 901 (66,09%) |
| <input type="checkbox"/> | 2. Firefox | 343 (26,88%) | 347 (27,07%) | 1 330 (22,53%) |
| <input type="checkbox"/> | 3. Opera | 43 (3,37%) | 42 (3,28%) | 430 (7,28%) |
| <input type="checkbox"/> | 4. Android Webview | 38 (2,98%) | 38 (2,96%) | 41 (0,69%) |
| <input type="checkbox"/> | 5. Edge | 30 (2,35%) | 30 (2,34%) | 71 (1,20%) |
| <input type="checkbox"/> | 6. Safari | 19 (1,49%) | 19 (1,48%) | 99 (1,68%) |





















Google Analytics

| <input type="checkbox"/> | Wersja dodatku Flash [?] | Pozyskiwanie | | |
|--------------------------|-----------------------------------|---|---|---|
| | | Użytkownicy [?] ↓ | Nowi użytkownicy [?] | Sesje [?] |
| | | 1 275 % całości: 100,00% (1 275) | 1 282 % całości: 100,16% (1 280) | 5 903 % całości: 100,00% (5 903) |
| <input type="checkbox"/> | 1. (not set) | 1 068 (80,06%) | 1 047 (81,67%) | 5 127 (86,85%) |
| <input type="checkbox"/> | 2. 27.0 r0 | 82 (6,15%) | 77 (6,01%) | 225 (3,81%) |
| <input type="checkbox"/> | 3. 29.0 r0 | 79 (5,92%) | 72 (5,62%) | 199 (3,37%) |
| <input type="checkbox"/> | 4. 28.0 r0 | 40 (3,00%) | 34 (2,65%) | 103 (1,74%) |
| <input type="checkbox"/> | 5. 31.0 r0 | 35 (2,62%) | 30 (2,34%) | 135 (2,29%) |
| <input type="checkbox"/> | 6. 32.0 r0 | 9 (0,67%) | 3 (0,23%) | 19 (0,32%) |

Google Analytics



Google Analytics

| Strona | | Odsłony | % Odsłony |
|-----------------|---|---------|--|
| 1. / |  | 5 696 |  19,92% |
| 2. /kursy |  | 5 410 |  18,92% |
| 3. /oceny |  | 3 490 |  12,21% |
| 4. /kursy/1 |  | 2 718 |  9,51% |
| 5. /aktualnosci |  | 1 348 |  4,72% |
| 6. /kursy/3 |  | 770 |  2,69% |
| 7. /kontakt |  | 684 |  2,39% |
| 8. /faq |  | 583 |  2,04% |
| 9. /zarzadzaj |  | 551 |  1,93% |
| 10. /kursy/4 |  | 502 |  1,76% |

wyświetl pełny raport

Google Analytics



Straszna analityka

Istnieje cała masa innych rozwiązań.

Popularnym może być New Relic (raczej wykorzystywany po stronie backendu), Kissmetrics (do śledzenia konkretnych użytkowników) czy CrazyEgg (do generowania heatmap).

Należy pamiętać, że granica prywatności to rzecz bardzo względna i może się okazać, że aplikacja śledzi naprawdę każdy ruch użytkownika.

Jakim problemem (od strony technicznej) jest sprawdzenie w jakie miejsca na stronie klika się najczęściej?

Jakim problemem (od strony technicznej) jest zrobienie zrzutu ekranu użytkownika?

Jakim problemem (od strony moralnej) jest nagranie całej sesji użytkownika?

Czas na eksperyment!

Podsumowanie

Bibliografia i ciekawe źródła



<https://tools.ietf.org/html/rfc5424>

Pytania?

Kod prezentacji dostępny jest w repozytorium git pod adresem
<https://bitbucket.org/krewak/pwsz-ppsi2>



Wszystkie informacje dot. kursu dostępne są pod adresem
<http://pwsz.rewak.pl/kursy/6>

