

**LAPORAN PENELITIAN**

**Karakteristik Situs Web Populer Indonesia Dilihat dari Sudut Pandang  
Keamanan**



**Disusun Oleh:  
Pascal Alfadian Nugroho, S.Kom, M.Comp**

**Lembaga Penelitian dan Pengabdian kepada Masyarakat  
Universitas Katolik Parahyangan  
2020**

## Daftar Isi

Abstrak.....	2
Bab 1: Pendahuluan .....	2
Bab 2: Tinjauan Pustaka.....	3
Bab 3: Metode Penelitian .....	4
Bab 4: Jadwal Pelaksanaan .....	5
Bab 5: Hasil dan Pembahasan .....	5
Bab 6: Kesimpulan dan Saran.....	7
Daftar Pustaka .....	7

## Abstrak

Penelitian ini bertujuan untuk melihat karakteristik 1.500 situs web populer di Indonesia dari sisi keamanan. Target penelitian ini adalah dapatnya karakteristik situs web yang digunakan oleh masyarakat Indonesia dari sisi keamanan.

Penelitian ini diawali dengan mengumpulkan daftar situs web terpopuler. Kemudian, dilanjutkan dengan melakukan prapengujian untuk melihat teknologi-teknologi yang digunakan. Dari teknologi-teknologi tersebut, dilihat berapakah versi yang terakhir didukung oleh pembuatnya. Pada akhirnya, pada 1.500 situs web yang tersebut, dilihat seberapa banyak penggunaan teknologi yang masih / sudah tidak lagi didukung pembuatnya.

## Bab 1: Pendahuluan

Dalam kurun beberapa tahun terakhir, kehidupan manusia di dunia pada umumnya dan Indonesia pada khususnya, beralih dari aktivitas luring (luar jaringan) menjadi daring. Mulai dari situs-situs *e-commerce* seperti Bukalapak<sup>1</sup> dan Blibli<sup>2</sup>, aplikasi transportasi seperti Go-Jek<sup>3</sup> dan Grab<sup>4</sup>, sampai situs web dan aplikasi yang dibuat pemerintah seperti DJP Online<sup>5</sup> untuk pelaporan pajak DPMPTSP Kota Bandung<sup>6</sup> untuk mengajukan perizinan usaha.

Sama seperti aktivitas luring yang membutuhkan jaminan keamanan dari pihak berwenang, manusia pun membutuhkan jaminan yang sama pada saat beraktivitas daring. Dalam hal ini, penyedia jasa daring menjadi pihak yang paling bertanggung jawab atas rasa aman penggunanya.

Atas dasar hal tersebut, dirumuskan bahwa penelitian ini mencoba untuk menjawab pertanyaan berikut:

*Seberapa siapkah penyedia jasa daring yang digunakan oleh masyarakat Indonesia dalam menjamin keamanan bagi penggunanya?*

---

<sup>1</sup> <https://www.bukalapak.com/>

<sup>2</sup> <https://www.blibli.com/>

<sup>3</sup> <https://www.gojek.com/>

<sup>4</sup> <https://www.grab.com/id/>

<sup>5</sup> <https://djponline.pajak.go.id/>

<sup>6</sup> <https://dpmptsp.bandung.go.id/>

Pada penelitian ini, “penyedia jasa daring” dibatasi pada penyedia jasa situs web, tidak termasuk aplikasi gawai. Tidak seperti aplikasi gawai, pemeriksaan keamanan pada situs web jauh lebih mudah dilakukan dan diotomatisasi.

Urgensi penelitian ini dikaitkan dengan bidang unggulan penelitian Universitas Katolik Parahyangan, yaitu “Pemanfaatan teknologi informasi dan komunikasi”. Dengan maraknya kasus kebocoran data oleh penyedia jasa seperti bocornya data penumpang Malindo Air<sup>7</sup> dan data registrasi kartu prabayar<sup>8</sup>, dibutuhkan evaluasi terhadap seberapa siap penyedia jasa terkait masalah keamanan ini.

Temuan pada penelitian ini diharapkan untuk mendukung poin masyarakat inklusif (*inclusive society*) di mana setiap warga berhak mendapatkan jaminan keamanan dalam aktivitas daring mereka, sehingga meningkatkan efektivitas dan efisiensi masing-masing individu.

Dari rumusan masalah yang dijelaskan di sebelumnya, penelitian ini bertujuan untuk menjadi langkah awal dalam memperbaiki kualitas keamanan penyedia layanan daring di Indonesia. Penelitian ini tidak bertujuan untuk memperbaiki celah keamanan yang ada, kecuali jika ditemukan celah yang sangat kritis.

Hipotesis peneliti adalah bahwa paling tidak 50% situs web yang menjadi sampel (akan dijelaskan pada bagian metode penelitian) berpotensi untuk ditingkatkan dari sisi keamanannya. Hal ini disebabkan oleh keinginan kuat berbagai organisasi untuk berlomba-lomba berpindah ke daring (*go online*), seringkali tanpa persiapan yang matang.

Rencana luaran dari penelitian ini adalah:

1. Basis data situs web yang populer digunakan masyarakat Indonesia, yang berisi informasi:
  - a. Teknologi yang digunakan oleh situs web tersebut beserta versinya
  - b. Ada atau tidaknya pengamanan standar seperti HTTPS
  - c. Waktu pengambilan sampel (*timestamp*)
2. Data agregat dari basis data tersebut dalam bentuk yang mudah dipahami seperti *chart*
3. Program komputer yang mampu mereproduksi pembuatan poin 1 di atas dengan mudah

## Bab 2: Tinjauan Pustaka

(Sudarno & Purnama, 2012) melakukan survei teknologi yang digunakan 40 situs web portal perguruan tinggi negeri dan swasta dan dari luar negeri. Walaupun jumlah sampel relatif sedikit dan survei tidak melihat dari sisi keamanan, penelitian ini memberikan sedikit gambaran terhadap teknologi yang digunakan, yaitu: Google Analytics, Java Script Framework, bahasa pemrograman PHP, *content management system* Joomla, Apache web server, dan sistem operasi CentOS.

(Duarte, et al., 2016) memeriksa 1.669 laman web dari berbagai domain, berfokus pada kategori bahasa pemrograman, *framework* web dan *JavaScript*, serta *content management system*. Penelitian ini juga menggunakan kakas evaluasi aksesibilitas terotomatisasi untuk mengukur tingkat aksesibilitas laman-laman tersebut.

(Alvarez, Correa, & Arango, 2016) memeriksa berbagai situs web di Kolombia terhadap serangan *XSS*, *CSRF*, dan *SQL Injection*. Tujuan penelitian tersebut adalah untuk mendapatkan informasi bagaimana

---

<sup>7</sup> <https://money.kompas.com/read/2019/09/20/062900626/malindo-air-sebut-jadi-korban-kebocoran-data-penumpang>

<sup>8</sup> <https://tekno.kompas.com/read/2018/04/09/12160057/siang-ini-dukcapil-jelaskan-isu-kebocoran-data-registrasi-prabayar>

perusahaan-perusahaan di Kolombia memberikan perhatian terhadap faktor keamanan dan bagaimana pengguna dapat terdampak.

(He, Chen, & Guo, 2017) membuat sistem pemindai celah keamanan berdasarkan fitur *fingerprint*, di mana kakas Wappalyzer digunakan untuk mengekstraksi fitur *fingerprint* dari aplikasi web. Dari situ, peneliti menggunakan algoritme *euclidian distance KMN* untuk mencocokkan data *fingerprint* tersebut ke basis data. Efisiensi sistem tersebut diklaim bertambah baik sebesar 3% dan presisi bertambah baik 6%, dari percobaan terhadap 1000 situs web. Performansi sistem bertambah baik sebesar 6%.

(Rakhmawati, Haris, Hermansyah, & Furqon, 2018) melakukan survei terhadap 548 situs web pemerintah daerah yang ada di Indonesia, untuk mengetahui teknologi yang digunakan pada setiap situs web tersebut. Teknologi yang dilihat pada survei ini adalah: server web, pemrograman web, CSS, sistem manajemen konten, *web framework*, dan web 3.0.

(HTTP Archive, 2019) merilis keadaan dari web di tahun 2019, berdasarkan survey dari arsip web yang mereka miliki. Keadaan tersebut dilihat dari sudut pandang konten laman, pengalaman pengguna, publikasi konten, dan distribusi konten. Analisa keamanan termasuk dalam kategori pengalaman pengguna, dan banyak melihat apakah situs web banyak menggunakan fitur-fitur keamanan terbaru.

### Bab 3: Metode Penelitian

Penelitian dilakukan dengan terlebih dahulu mengumpulkan 1.500<sup>9</sup> alamat situs web terpopuler di Indonesia yang diambil dari situs web Alexa Top Sites untuk Indonesia<sup>10</sup>. Alternatif dari Alexa adalah Similar Web<sup>11</sup>, yang memberikan data serupa. Pada kedua situs web tersebut hanya dibatasi 50 situs web terpopuler, sedangkan sisanya bisa didapatkan dengan berbayar. Keputusan apakah dalam penelitian ini menggunakan Alexa atau Similar Web akan ditentukan kemudian, berdasarkan mana yang biayanya terjangkau dengan dana penelitian yang didapat.

Untuk setiap situs web yang didapat, akan dilakukan pra-pemindaian menggunakan *tool open source* Wappalyzer<sup>12</sup> untuk mengetahui teknologi yang digunakan dalam membangun ke-1.500 situs web tersebut. Teknologi-teknologi yang digunakan tersebut dikelompokkan, dan peneliti akan mencari tahu versi terakhir yang masih didukung oleh pembuatnya<sup>13</sup>. Versi ini nantinya akan digunakan untuk mengetahui karakteristik setiap situs web: apakah situs web tersebut menggunakan teknologi yang masih didukung atau sudah *end of life*. Penggunaan teknologi yang sudah *end of life* beresiko memiliki celah keamanan.

Pada akhirnya, pemindaian ke 1.500 situs web memeriksa seberapa besar penggunaan teknologi *end of life* pada situs web tersebut, seperti dijelaskan pada paragraf sebelumnya<sup>14</sup>.

---

<sup>9</sup> Pada awal penelitian, direncanakan 500 situs web, tetapi pada akhirnya diputuskan untuk dinaikkan menjadi 1.500 situs web.

<sup>10</sup> <https://www.alexa.com/topsites/countries/ID>

<sup>11</sup> <https://www.similarweb.com/top-websites/indonesia>

<sup>12</sup> <https://www.wappalyzer.com/download>

<sup>13</sup> Pada awal penelitian, direncanakan hanya 50 teknologi terpopuler saja, tetapi pada akhirnya diputuskan untuk melihat semua teknologi yang mengandung informasi versi yang digunakan.

<sup>14</sup> Pada awal penelitian, direncanakan juga untuk melihat pemanfaatan, tetapi pada akhirnya diputuskan untuk tidak mengukurnya.

Setelah dilakukan pemindaian, peneliti akan melakukan analisa terhadap hasilnya. Hasil analisa diharapkan mampu menjawab pertanyaan-pertanyaan berikut:

- Bagaimana karakteristik situs web populer di Indonesia, terutama dilihat dari kategorinya?
- Teknologi apa yang banyak digunakan situs web di Indonesia?
- Seberapa banyak situs web di Indonesia yang masih menggunakan teknologi yang *end of life*?

## Bab 4: Jadwal Pelaksanaan

Berikut adalah jadwal pelaksanaan kegiatan penelitian di tahun 2020. Sel tabel yang diarsir menandakan bahwa ada aktivitas pelaksanaan sebuah tahap pada bulan tersebut, sedangkan bilangan yang tertera di dalamnya menunjukkan jumlah *man-days* yang dibutuhkan dalam bulan tersebut.

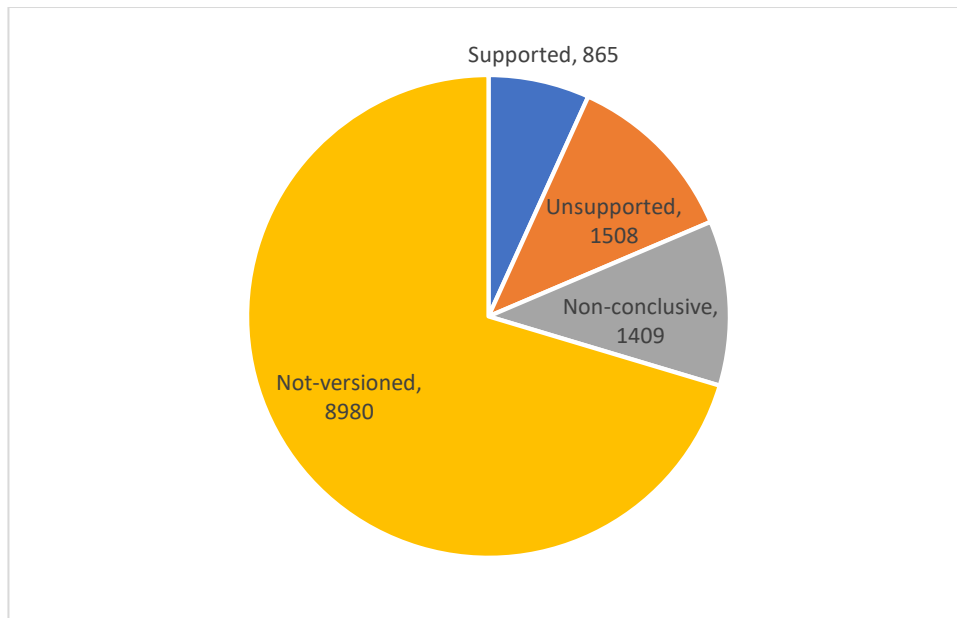
Tahap	Jan	Feb	Mar	Apr	Mei	Jun	Jul	Agt	Sep	Okt	Nov	Des
<b>Pengumpulan situs populer</b>	3											
<b>Pra-pemindaian</b>		7	7									
<b>Pemindaian</b>				7	7							
<b>Analisis Data</b>						3	3	3				
<b>Publikasi dan Pelaporan</b>	1	1	1	1	1	1	1	1	8	8	8	8

Penelitian ini relatif tidak terpengaruh oleh situasi pandemi COVID-19, karena seluruhnya dilakukan secara daring.

## Bab 5: Hasil dan Pembahasan

Di akhir penelitian ini, peneliti sudah mendapatkan basis data situs web populer di Indonesia, berdasarkan daftar 1.500 situs terpopuler dari Alexa. Walaupun demikian, peneliti belum mempublikasikan basis data ini. Dari hasil umpan balik pada seminar LPPM, peneliti mendapatkan ide untuk terlebih dahulu menginformasikan pengelola situs web terkait celah keamanan yang ada, sebelum mempublikasikannya.

Diagram di bawah ini memberikan gambaran tentang seberapa banyak teknologi yang digunakan oleh situs web, yang masih didukung oleh pembuatnya. Dari 12.762 teknologi yang digunakan oleh ke-1.500 situs web terpopuler, 8.980 di antaranya tidak mengandung informasi versi yang digunakan, sehingga tidak dapat disimpulkan. Sisanya, dari 3.782 teknologi didapatkan 865 di antaranya masih didukung oleh pembuatnya. 1.508 berada dalam keadaan tidak aman, karena tidak didukung oleh pembuatnya. Pada 1.409 teknologi sisanya, peneliti tidak dapat memutuskan apakah masih didukung atau tidak, walaupun versinya dapat dideteksi.

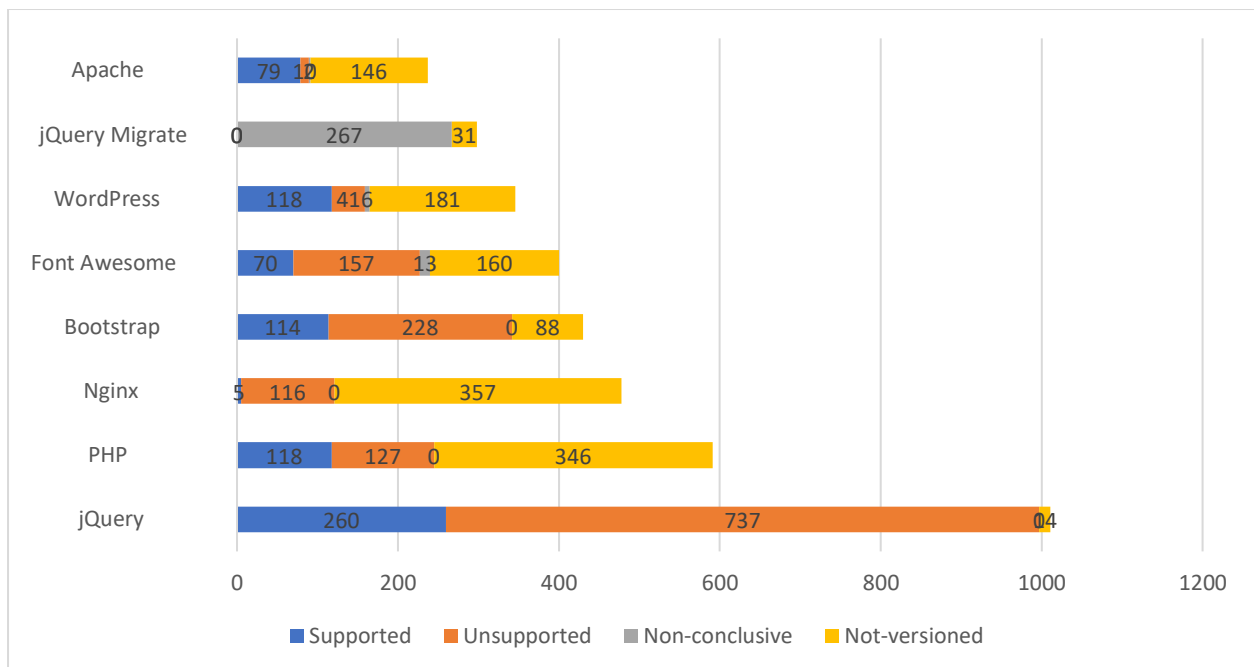


Berikut adalah informasi yang lebih detail untuk 8 teknologi terbanyak yang digunakan. jQuery menjadi teknologi yang paling banyak digunakan, dan 737 penggunaannya menggunakan versi yang sudah tidak didukung. Versi PHP yang tidak didukung lagi juga masih banyak digunakan, melebihi versi yang didukung. WordPress dan Apache merupakan 2 contoh teknologi yang berhasil dalam mendorong penggunaannya untuk selalu menggunakan versi yang terbaru atau masih didukung.

Secara khusus, untuk situs web UNPAR, dideteksi menggunakan teknologi-teknologi berikut:

- WordPress (versi tidak diketahui)
- Apache (versi tidak diketahui)
- PHP versi 5.6.40 (di bawah versi yang masih didukung 7.2)
- jQuery versi 1.12.4 (di bawah versi yang masih didukung 3)
- jQuery Migrate 1.4.1 (versi minimal tidak diketahui)

Seluruh program bantuan yang dibuat untuk mengumpulkan data, disimpan dalam repositori GitHub milik peneliti <https://github.com/pascalalfadian/IndonesiaWebsiteSurvey> namun aksesnya tidak dibuka secara publik.



## Bab 6: Kesimpulan dan Saran

Hasil penelitian ini menekankan perlunya pengelola situs web untuk menjaga situs web mereka up-to-date terhadap versi terbaru yang masih didukung oleh pembuat teknologinya. WordPress dan Apache terbilang sukses dalam mendorong penggunaannya untuk menggunakan versi terbaru yang masih didukung. Penelitian lebih lanjut dibutuhkan untuk mengidentifikasi faktor yang mendorong pengguna untuk melakukan upgrade ke versi terbaru, serta mengapa beberapa pengelola tidak terlalu ingin melakukan upgrade padahal aplikasinya terbilang penting, seperti contohnya Nginx.

## Daftar Pustaka

- Alvarez, D. E., Correa, D. B., & Arango, F. I. (2016). An analysis of XSS, CSRF and SQL injection in colombian software and web site development. *2016 8th Euro American Conference on Telematics and Information Systems (EATIS)*. Cartagena: IEEE.
- Duarte, C., Matos, I., Vicente, J., Salvado, A., Duarte, C. M., & Carriço, L. (2016). Development technologies impact in web accessibility. *Proceedings of the 13th Web for All Conference*. Montreal: ACM.
- He, H., Chen, L., & Guo, W. (2017). Research on Web Application Vulnerability Scanning System based on Fingerprint Feature. *Proceedings of the 2017 International Conference on Mechanical, Electronic, Control and Automation Engineering (MECAE 2017)*. Qingdao: Atlantis Press.
- HTTP Archive. (2019, December 1). Retrieved from The 2019 Web Almanac: <https://almanac.httparchive.org/en/2019/>
- Rakhmawati, N. A., Haris, S., Hermansyah, D., & Furqon, M. ' . (2018). A survey of Web Technologies used in Indonesia Local Governments. *Jurnal Sisfo*, 213-222.
- Sudarno, & Purnama, B. E. (2012). Analisis Penggunaan Tools Web Perguruan Tinggi. *Seminar Riset Unggulan Nasional Informatika dan Komputer*. Surakarta: Universitas Surakarta.