

Sistem Akusisi Data untuk *Polling* berbasis Media Sosial dan Autentikasi *User* menggunakan reCaptcha

Tugas Akhir

**diajukan untuk memenuhi salah satu syarat
memperoleh gelar sarjana
dari Program Studi Sarjana Teknik Informatika
Fakultas Informatika
Universitas Telkom**

1301154390

Chando Anggara Natanael Batubara



**Program Studi Sarjana Teknik Informatika
Fakultas Informatika
Universitas Telkom
Bandung
2019**

LEMBAR PENGESAHAN

**Sistem Akuisisi Data untuk *Polling* berbasis Media Sosial
dan Autentikasi *User* menggunakan reCaptcha**

*Data Acquisition Sistem for Sosial Media-based Polling
and User Authentication using reCaptcha*

NIM : 1301154390

Chando Anggara Natanael Batubara

Tugas akhir ini telah diterima dan disahkan untuk memenuhi sebagian syarat memperoleh gelar pada Program Studi Sarjana Teknik Informatika

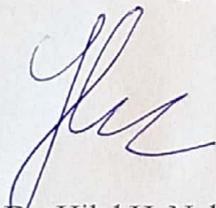
Fakultas Informatika

Universitas Telkom

Bandung, 15 Juli 2019

Menyetujui

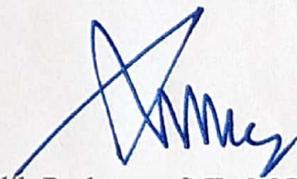
Pembimbing I,



Dr. Hilal H. Nuha

NIP : 13860093

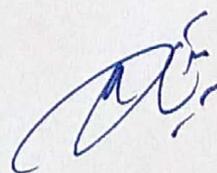
Pembimbing II,



Sidik Prabowo, S.T., M.T.

NIP : 15870072

Ketua Program Studi
Sarjana Teknik Informatika,



Niken Dwi Wahyu Cahyani, S.T, M.Kom., Ph.D.

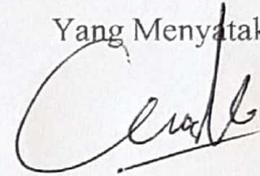
NIP : 00750052

LEMBAR PERNYATAAN

Dengan ini saya, Chando Anggara Natanael Batubara, menyatakan sesungguhnya bahwa Tugas Akhir saya dengan judul Sistem Akuisisi Data untuk *Polling* berbasis Media Sosial dan Autentikasi *User* menggunakan reCaptcha beserta dengan seluruh isinya adalah merupakan hasil karya sendiri, dan saya tidak melakukan penjiplakan yang tidak sesuai dengan etika keilmuan yang belaku dalam masyarakat keilmuan. Saya siap menanggung resiko/sanksi yang diberikan jika di kemudian hari ditemukan pelanggaran terhadap etika keilmuan dalam buku TA atau jika ada klaim dari pihak lain terhadap keaslian karya,

Bandung, 15 Juli 2019

Yang Menyatakan



Chando Anggara Natanael Batubara

Sistem Akuisi Data untuk *Polling* berbasis Media Sosial dan Autentikasi *User* menggunakan reCaptcha

Chando Anggara Natanael Batubara¹, Hilal Hudan Nuha², Sidik Prabowo³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹chandonatanal@students.telkomuniversity.ac.id, ²hilalnuha@telkomuniversity.ac.id,

³pakwowo@telkomuniversity.ac.id

Abstrak

Setiap lima tahun sekali kita melakukan pemilihan umum untuk menentukan calon pemimpin periode selanjutnya. Mendekati hari dimana pemilu tersebut dilaksanakan, banyak lembaga survei berlomba – lomba menyediakan informasi mengenai pemilu. Salah satunya dengan cara membuat sistem *polling*. Hasil *polling* yang dikeluarkan dari tiap lembaga survei pun berbeda – beda dan mengklaim bahwa hasil *polling* yang mereka keluarkan merupakan yang paling akurat. Oleh karena itu, sistem *polling* yang akurat dan cepat dengan memanfaatkan media sosial sangat dibutuhkan. Pada penelitian ini, sistem *polling* yang dibuat berbasis media sosial twitter, dengan memanfaatkan API untuk pengambilan data peserta *polling*. Dalam sistem yang dibangun, proses autentikasi terdiri dari dua tahap. Pada tahap pertama, sistem menggunakan autentikasi API media sosial untuk menyaring *bot*. Sedangkan, tahap kedua menggunakan reCaptcha yang merupakan sebuah program yang dapat membedakan manusia dengan program komputer atau *bot*. Sistem *polling* yang dibangun dengan menerapkan reCaptcha sebagai *autentikasi user* berhasil mencegah 100% akun *bot* untuk mengikuti *polling*. 70% akun *bot* ditangguhkan oleh media sosial (twitter) sementara sisanya diblokir oleh reCaptcha.

Kata kunci : *polling, media sosial, autentikasi, API, reCaptcha, bot.*

Abstract

Every five years, we held general elections to determine the future leaders of the next period. Approaching the election day, many survey institutions competed to provide information related to the elections. One of them is by creating a polling system. The results of the polls issued by each survey institute were different and claimed that the poll results they issued were the most accurate. Therefore, a faster polling system with higher accuracy by utilizing social media is required. In this study, a polling system was created based on Twitter social media, using the API for polling participants' data collection. In the developed system, the authentication process consists of two stages. In the first stage, the system utilizes social media authentication API to filter out the bots. Whereas, the second stage uses reCaptcha which is a program that can distinguish humans from computer programs or bots. The polling system that was built by applying reCaptcha as user authentication was able to prevent 100% bot accounts from participating the polling. 70% of bot accounts were suspended by social media (twitter) while the remaining were blocked by reCaptcha.

Keywords: *polling, social media, authentication, API, reCaptcha, bot.*

1. Pendahuluan

Latar Belakang

Pemilihan Umum yang selanjutnya disebut Pemilu adalah sarana kedaulatan rakyat untuk memilih anggota Dewan Perwakilan Rakyat (DPR), anggota Dewan Perwakilan Daerah (DPD), Presiden dan Wakil presiden, dan untuk memilih anggota Dewan Perwakilan Rakyat Daerah (DPRD), yang dilaksanakan secara langsung, umum, bebas, rahasia, jujur, dan adil dalam Negara Kesatuan Republik Indonesia (NKRI) berdasarkan Pancasila dan Undang-Undang Dasar Negara Republik Indonesia Tahun 1945, (UUD 1945) Pasal 1 ayat 1 [1]. Mendekati hari dimana pemilu tersebut dilaksanakan, banyak lembaga survei berlomba – lomba menyediakan informasi tentang pemilu. Dalam proses penyediaannya banyak metode yang dapat digunakan, salah satunya dengan cara *polling* atau jajak pendapat. Meskipun demikian, tidak sedikit dari mereka mengklaim bahwa informasi yang mereka miliki paling akurat dan hal ini cukup membuat publik kebingungan.

Media sosial sebagai sarana komunikasi *online* berpengaruh sangat besar dalam kehidupan masyarakat. Begitu juga Indonesia sebagai merupakan pengguna terbesar ke-tiga media sosial di dunia [2]. Oleh karena itu, penulis berkesimpulan untuk membuat sebuah sistem *polling* berbasis media sosial yaitu twitter, yang digunakan sebagai wadah untuk melakukan pengumpulan informasi tentang Pemilu. Dalam sistem yang dibangun, proses

otentikasi dari peserta *polling* menggunakan reCaptcha yang merupakan sebuah program yang dipakai untuk keamanan aplikasi agar terhindar dari serangan. Salah satu sumber serangan autentikasi adalah software jahat yang disebut *bot*. Sistem *polling* dibuat bertujuan untuk mengatasi permasalahan buruknya kualitas survei atau *polling* yang ada di media sosial saat ini. Hal ini dikarenakan *polling* yang dilakukan di media sosial masih ada *bot account* yang bisa ikut serta dalam jajak pendapat tersebut.

Topik dan Batasan Penelitian

Adapun perumusan masalah yang didapat berdasarkan latar belakang masalah yaitu, bagaimana cara membangun sebuah sistem *polling* berupa aplikasi yang berdiri diatas media sosial Twitter dengan autentikasi menggunakan reCaptcha. Penelitian ini memiliki batasan penelitian sebagai berikut :

1. Media sosial yang digunakan adalah Twitter.
2. Autentikasi *User* dengan menggunakan reCaptcha berbasis gambar.
3. Fokus autentikasi *User* adalah mencegah *bot* untuk ikut serta dalam *polling*.

Tujuan

Tujuan yang ingin dicapai dalam Tugas Akhir (TA) ini adalah membangun sebuah sistem pengumpulan data *polling* berbasis media sosial dengan memanfaatkan Twitter API, serta reCaptcha berbasis gambar sebagai autentikasi peserta *polling*.

Organisasi Tulisan

Penelitian TA ini disusun dengan struktur sebagai berikut : Bagian pertama yaitu pendahuluan berisi mengenai latar belakang, topik dan batasan, dan tujuan serta organisasi penulisan. Bagian kedua membahas tentang studi terkait dan referensi-referensi yang berkaitan dengan TA penulis. Bagian ketiga berisi mengenai penjelasan sistem yang dibangun. Bagian keempat berisi analisis dan evaluasi penelitian. Bagian kelima mengenai kesimpulan dari penelitian yang dilakukan.

2. Studi Terkait

Bab ini membahas tentang konsep dasar dan istilah yang dipakai dalam TA ini. Sub-bab yang pertama berisi tentang konsep *polling* dan penelitian yang berhubungan dengannya. Sub-bab yang kedua membahas tentang definisi media sosial dan penggunaan *polling* berbasis media sosial. Application programming interface (API) media sosial yang dipakai dalam TA ini dibahas pada subbab tiga termasuk penelitian yang menggunakan API ini. Bab ini diakhiri dengan pembahasan tentang reCaptcha yang merupakan metode utama dalam TA ini.

Polling

Polling adalah suatu kerja pengumpulan pendapat umum dengan menggunakan teknik dan prosedur ilmiah. Lake et.al [3] mendefinisikan *polling* sebagai cara sistematis, ilmiah dan terpercaya mengumpulkan informasi dari sampel orang yang digunakan untuk menggeneralisasikan pada kelompok atau populasi yang lebih luas darimana sampel itu diambil. Dalam hal ini, desain penelitian *polling* merupakan suatu pengukuran pada satu waktu untuk mengetahui sikap, perilaku, kepercayaan, dan hubungan diantara semua parameter [3]. Dengan berkembangnya media sosial, banyak penelitian dan proyek ilmiah muncul untuk *polling* dengan menggunakan media sosial seperti pada [4] yang menggunakan twitter untuk jajak pendapat di amerika serikat. Contoh yang lain bisa dilihat pada [5] dimana peneliti menggunakan facebook untuk prediksi pemilihan anggota kongres.

Media Sosial

Media sosial atau yang dikenal juga dengan jejaring sosial merupakan sebuah media online, dengan para penggunanya bisa dengan mudah berpartisipasi, berbagi, dan menciptakan isi meliputi blog, jejaring sosial, wiki, forum dan dunia virtual. Blog, jejaring sosial dan wiki merupakan bentuk media sosial yang paling umum digunakan oleh masyarakat di seluruh dunia. Komunikasi di media sosial tidak dibatasi oleh jarak, waktu, dan ruang. Itu bisa terjadi di mana saja, kapan saja, tanpa harus bertatap muka. Bahkan media sosial dapat meniadakan status sosial yang seringkali sebagai penghambat dalam komunikasi [6].

Media sosial tidak hanya memberikan konektivitas kepada orang-orang tetapi juga menawarkan peluang dalam sektor pengetahuan dan bisnis. Data yang tersedia di media sosial dapat digunakan untuk menghasilkan beberapa informasi [7] dan informasi ini dapat berguna bagi organisasi mana pun untuk keperluan pengambilan keputusan dan memenuhi tujuan mereka.

API

API pada dasarnya adalah antarmuka dari program komputer yang memungkinkan perangkat lunak untuk "berbicara" dengan perangkat lunak lain [8]. Beberapa perusahaan media sosial, membuat bank data tentang data

user dan pola penggunaannya melalui APIs yang mereka miliki. Sehingga APIs juga merupakan sebuah sarana bagi para peneliti untuk mengumpulkan data.

Penelitian ini menggunakan media sosial Twitter. Twitter memiliki layanan untuk mengakses bank data mereka melalui *Twitter Developer*. *Twitter Developer* merupakan sebuah *platform* yang menyediakan alat untuk membangun sebuah produk agar dapat terintegrasi dengan twitter [9].

reCaptcha

reCaptcha merupakan pengembangan dari *Completely automated public Turing test to tell computers and humans apart* (Captcha) atau uji Turing terotomasi untuk membedakan Manusia dan Robot yang banyak digunakan di Internet saat ini [10]. Fungsi utamanya adalah melindungi situs web dari serangan *bot* web dengan membuat sebuah tes yang mana komputer itu tidak bisa mengatasinya. reCaptcha terdiri dari beberapa jenis yaitu :

- a. reCaptcha berbasis teks

reCaptcha berbasis teks sangat mudah diterapkan dan sangat efektif. Hanya saja membutuhkan bank soal yang besar untuk menampilkan teks berbeda.

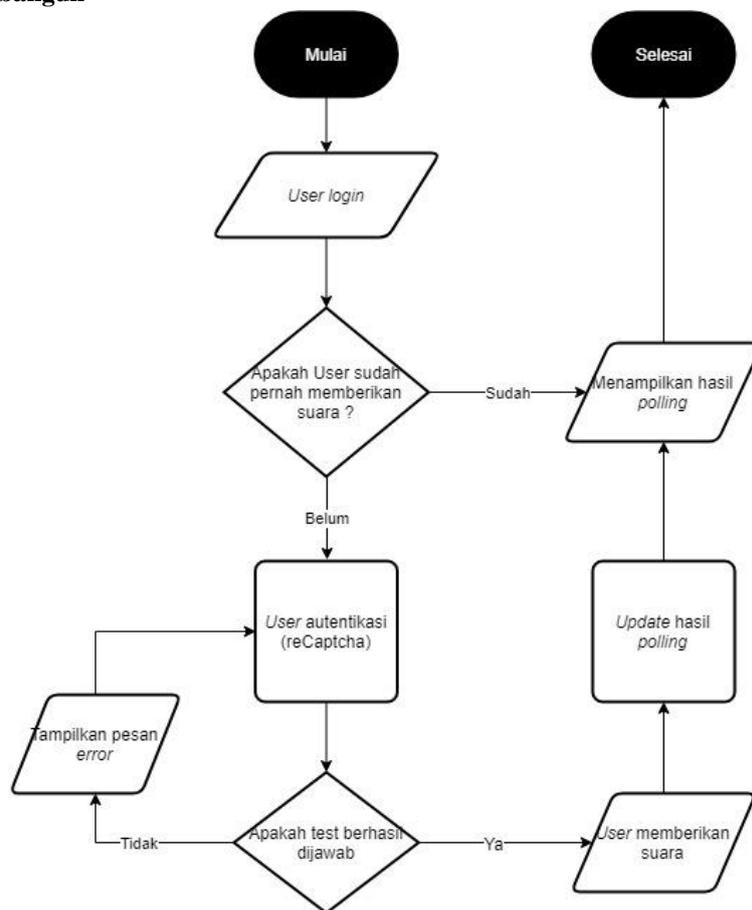
- b. reCaptcha berbasis gambar

reCaptcha berbasis gambar adalah sebuah tes di mana pengguna harus menebak gambar-gambar yang memiliki beberapa kesamaan.

- c. reCaptcha berbasis audio

reCaptcha berbasis audio dikembangkan untuk pengguna dengan gangguan visual. reCaptcha ini berisi klip audio yang dapat di-*download*. Cara kerjanya yaitu pertama pengguna mendengarkan sebuah klip audio dan setelah itu mengirimkan kata yang diucapkan.

3. Sistem yang Dibangun



Gambar 1. Flowchart sistem polling secara umum

Sistem *Polling* yang dibangun pada TA ini dibuat dalam bentuk aplikasi web dengan menerapkan reCaptcha sebagai keamanan sistemnya. Adapun proses yang terjadi pada sistem yang dibuat antara lain, pertama, sebelum peserta *polling* masuk akan diminta *login* menggunakan akun twitter, kedua, proses autentikasi *user* dengan menggunakan reCaptcha, dan ketiga, apabila berhasil melewati proses autentikasi *user* peserta dapat melakukan *polling*. Terakhir, setelah *polling* telah dilakukan, peserta dapat melihat hasil *polling*. Sistem yang dibangun bisa diilustrasikan pada Gambar 1.

User Login

Agar bisa mengakses aplikasi *polling* ini, *user* harus *login* menggunakan akun twitter. Data akun *user* nantinya yang akan digunakan sebagai data peserta dalam aplikasi *polling* yang dibangun. API twitter digunakan untuk memperoleh data tersebut dengan memanfaatkan twitter *developer*.

Meminta izin akses profil user

Untuk mendapatkan data *user* seperti nama, umur dan kota, penulis membutuhkan izin dari pihak *user* yang bersangkutan. Hal ini perlu dilakukan mengingat persyaratan utama yang dibutuhkan agar dapat mengakses token dari API twitter.

Autentikasi User

Pada proses ini, *user* akan diberikan sebuah tes (reCaptcha) untuk memastikan bahwa *user* bukanlah *bot* atau program komputer. Apabila *user* berhasil melewati tes tersebut, *user* dapat melanjutkan ke tahap berikutnya dan apabila gagal reCaptcha akan men-*generate* tes baru. Untuk integrasi ke halaman website *polling* yang dibuat, perlu dimasukkan kode pada Gambar 2.

```
<script src="https://www.google.com/recaptcha/api.js"></script>
@if ($errors->has('g-recaptcha-response'))
    <div class="alert alert-danger mt-3" style="width: 100%" role="alert">
        <strong>{{ $errors->first('g-recaptcha-response') }}</strong>
    </div>
@endif

<div class="g-recaptcha" data-sitekey="{{ env('RECAPTCHA_SITE_KEY') }}" style="display: inline-block;">
</div>
```

Gambar 2. Source code widget reCaptcha

Untuk validasi captchanya perlu di tambahkan code seperti pada Gambar 3.

```
public function validate($attribute, $value, $parameters, $validator){

    $client = new Client();

    $response = $client->post('https://www.google.com/recaptcha/api/siteverify',
        ['form_params'=>
            [
                'secret'=>env('RECAPTCHA_SECRET_KEY'),
                'response'=>$value
            ]
        ]
    );

    $body = json_decode((string)$response->getBody());
    return $body->success;
}
```

Gambar 3. Source code validasi reCaptcha

User memberikan suara

Pada proses ini, *user* diberi kesempatan untuk memberikan pilihannya dan *user* hanya diperbolehkan satu kali. Apabila sistem mendeteksi *user* sudah pernah memilih maka akan muncul pesan bahwa *user* telah melakukan *polling* dan aplikasi langsung berakhir.

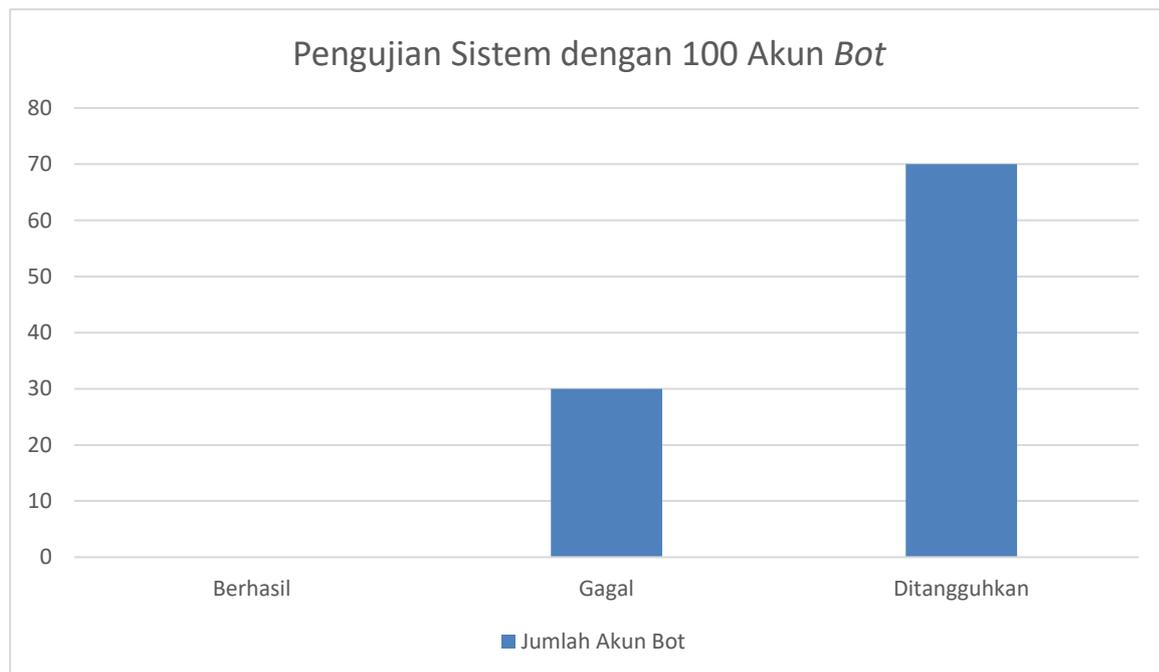
Update dan tampilkan hasil *polling*

Setelah data disimpan maka hasil *polling* di-update berdasarkan kalkulasi sistem, kemudian ditampilkan dalam bentuk presentasi berdasarkan banyaknya pilihan.

4. Evaluasi

4.1 Hasil Pengujian

Skenario pengujian sistem pada penelitian ini menggunakan 100 akun twitter yang akan dikontrol oleh 100 *bot* untuk mengikuti sebuah *polling* yang ada pada sistem *polling* yang dibangun. *Bot* yang dibuat memiliki kemampuan untuk memecahkan tes gambar dari reCaptcha, dengan cara mengklik secara *random* beberapa gambar atau potongan gambar yang di tampilkan. Adapun banyak percobaan untuk memecahkan tes gambar dari reCaptcha yaitu sebanyak dua kali percobaan tiap *bot* akun. Parameter yang akan diuji yaitu seberapa efektif sistem keamanan reCaptcha mengatasi akun *bot* yang mencoba mengikuti *polling* yang tersedia. Hasil yang diperoleh dari skenario pengujian yang dilakukan dapat dilihat pada Gambar 4.



Gambar 4. Hasil pengujian

Dari grafik diatas diperoleh data bahwa tidak ada akun *bot* yang berhasil mengikuti *polling*, 30% gagal mengikuti *polling* dan sebanyak 70% akun *bot* ditangguhkan.

Kemudian penulis melakukan skenario pengujian berikutnya dengan, menggunakan 30 akun twitter yang akan di kendalikan oleh *bot* untuk mengikuti sebuah *polling* yang ada pada sistem *polling* yang dibangun. Pada skenario pengujian ini, *bot* memiliki kemampuan yang sama seperti *bot* yang ada pada skenario sebelumnya, namun perbedaannya terletak pada percobaan untuk menyelesaikan tes pada reCaptcha di perbanyak hingga 100 kali percobaan. Adapun hasil dari skenario pengujian ini yaitu 30 akun *bot* yang di jalankan, tidak ada yang berhasil melewati reCaptcha.

4.2 Analisis Hasil Pengujian

Berdasarkan hasil pengujian yang diperoleh tidak ada akun *bot* yang berhasil mengikuti *polling*. Grafik perolehan jumlah akun *bot* yang ditangguhkan sangat dominan, dengan presentasi mencapai 70%. Hal ini dapat terjadi karena pihak media sosial Twitter, mendeteksi adanya aktifitas yang mencurigakan. Sesuai dengan aturan yang berlaku pada twitter, tentang kebijakan pengguna maka akun yang dicurigai keamanannya akan dilakukan tindakan penangguhan akun [11]. Tidak hanya faktor aktivitas yang mencurigakan saja yang membuat akun *bot* ditangguhkan, tetapi faktor dari pembuatan akun *bot* juga mempengaruhi tindakan tersebut. Secara umum, setiap akun media sosial yang baru dibuat akan sangat ketat dipantau kegiatannya oleh pengembang terkait. Hal ini dilakukan untuk mencegah maraknya penyalahgunaan media sosial.

Kemudian terdapat 30% akun *bot* yang gagal mengikuti *polling* yang ada. Dari 30% akun *bot* yang gagal tersebut, terdapat kondisi gagal yang berbeda, pertama 20% akun *bot* gagal mengikuti *polling* dikarenakan tes atau soal yang diberikan oleh reCaptcha tidak bisa diselesaikan dan 10% akun *bot* terdeteksi oleh reCaptcha lalu diblok.

Setelah itu diperoleh juga hasil yang sama pada skenario pengujian selanjutnya, di mana dari 30 akun *bot* yang dijalankan tidak satu pun ada yang berhasil melewati reCaptcha.

Akun *bot* tidak bisa menyelesaikan tes atau soal yang dibuat karena kemampuan untuk memecahkan soal tersebut sangat kecil. Akun *bot* hanya bisa menjawab soal dengan mengklik secara *random* potongan gambar yang terdeteksi dan tidak bisa mengklik gambar sesuai soal yang ada. Yang membuat tes reCaptcha semakin sulit untuk dipecahkan akun *bot* yaitu adanya variasi potongan gambar dan variasi gambar yang disajikan. Sedangkan 10% akun *bot* terakhir diblok karena aktifitas yang dilakukan terindikasi oleh reCaptcha sebagai *bot*.

5. Kesimpulan

Polling merupakan salah satu cara untuk mengumpulkan pendapat umum tentang suatu objek. Pada saat ini, *Polling* paling banyak digunakan pada media online, seperti media sosial. Selain pembuatannya mudah *polling* dengan media online mudah untuk disebar ke khalayak umum. Ditambah lagi terdapat fakta bahwa Indonesia merupakan pengguna media sosial terbesar ketiga di dunia. Pada penelitian ini, penulis membangun Sistem akuisisi data dengan media sosial dan menggunakan reCaptcha sebagai autentikasi *user*. Kemudian penulis melakukan pengujian terhadap sistem yang dibangun dengan membuat 100 akun *bot* ikut serta ke dalam sebuah *polling* yang ada di dalam sistem. Dari hasil pengujian, data yang diperoleh menunjukkan tidak ada akun *bot* yang berhasil mengikuti *polling*, 30% gagal mengikuti *polling* dan sebanyak 70% akun *bot* ditanggihkan. Kesimpulan yang dapat diambil yaitu penelitian yang dilakukan menghasilkan sistem *polling* yang bekerja dengan baik, data umum pengunjung *polling* berhasil diperoleh dan sistem keamanan berhasil mencegah akun *bot* mengikuti *polling*. Namun tidak menutup kemungkinan bahwa sistem keamanan reCaptcha dapat ditembus oleh *bot* dengan kemampuan yang berbeda. Jadi untuk penelitian selanjutnya ada baiknya untuk mencoba menggunakan sistem keamanan yang berbeda dengan kualitas keamanan yang lebih baik tapi juga mudah untuk dihadapi pengguna.

Daftar Pustaka

- [1] "Jurnal Konstitusi," *P.U. Bengkulu*, vol. II, p. 97, 2009.
- [2] "datareportal.com," [Online]. Available: <https://datareportal.com/reports/digital-2019-global-digital-overview>. [Accessed 15 June 2019].
- [3] I. M. Hanika, "FENOMENA PHUBBING DI ERA MILENIA (Ketergantungan Seseorang pada Smartphone terhadap Lingkungannya)," *Jurnal Ilmu Komunikasi*, vol. IV, pp. 42-51, 2015.
- [4] N. Beauchamp, "Predicting and interpolating state-level polls using Twitter textual data," *American Journal of Political Science*, no. 61(2), pp. 490-503, 2017.
- [5] M. MacWilliams, "Forecasting congressional elections using facebook data," *Political Science & Politics*, no. 48(4), pp. 579-583, 2015.
- [6] E. D. S. Watie, "Komunikasi dan media sosial (communications and social media)," *Jurnal The Messenger*, vol. 3, pp. 69-74, 2016.
- [7] H. Gohel, "Automation of Social Media Analysis by Web Intelligence," *International Journal of Science Research and Technology*, vol. I, pp. 17-21, 2015.
- [8] Lomborg, Stine, Bechmann and Anja, "Using APIs for data collection on social media," *The Information Society*, vol. 30, pp. 256-265, 2014.
- [9] "Twitter Developer," Twitter, Inc., 2019. [Online]. Available: <https://developer.twitter.com/en/products/products-overview>. [Accessed 15 may 2019].
- [10] L. Von Ahn, B. Maurer, C. McMillen, D. Abraham and M. Blum, "recaptcha: Human-based character recognition via web security measures," *Science*, vol. 321, no. 5895, pp. 1465-1468, 2008.
- [11] "Twitter Rules and policies," Twitter, [Online]. Available: <https://help.twitter.com/id/managing-your-account/suspended-twitter-accounts>. [Accessed 3 July 2019].