

PENGECAMAN PERTUTURAN DALAM ROBOT PERKHIDMATAN

Kong Yee Teng

Assoc. Prof. Dr. Shahnorbanun Sahran

*Fakulti Teknologi & Sains Maklumat, Universiti Kebangsaan Malaysia***ABSTRAK**

Robot ibotn merupakan sejenis robot daripada Taiwan yang dapat menjaga kanak-kanak dan mengekori kanak-kanak sepanjang masa. Ia sebagai CCTV yang akan mengikuti kanak-kanak, selain menjamin keselamatan kanak-kanak di dalam rumah, tetapi juga merapatkan hubungan antara kanak-kanak dan juga ibu bapa yang berada di tempat kerja. Namun begitu, robot ini yang hanya dapat bertutur dalam bahasa Mandarin tidak sesuai untuk pemasaran ke Malaysia. Kebanyakan rakyat Malaysia tidak memahami Bahasa Mandarin dan hal ini akan menyebabkan pemasaran robot ini tidak lancar. Untuk meningkatkan peluang pemasaran robot tersebut dalam Malaysia, cara pengecaman pertuturan robot perlu dijadikan bahasa yang senang difahami oleh rakyat Malaysia. Pengecaman pertuturan dalam robot perkhidmatan merupakan satu kajian yang dihasilkan untuk meningkatkan prestasi robot ibotn yang bertutur bahasa Mandarin iaitu membolehkan robot ibotn bertutur dalam bahasa Inggeris. Kajian ini juga dihasilkan untuk mengenalpasti cara-cara yang akan digunakan untuk pengenalan suara oleh robot dan seterusnya robot dapat menjawab balik soalan yang akan diberikan oleh pengguna. Untuk menghasilkan sistem ini, PyAMIL dan Python merupakan bahasa pengaturcaraan yang amat penting kerana ia merupakan bahagian pengetahuan robot. Hasil bagi kajian ini dijangka dapat mengecam pertuturan Bahasa Inggeris daripada pengguna, dan seterusnya menjawab balik dengan jawapan yang bersesuaian.

PENGENALAN

Pengecaman pertuturan, atau dikenali sebagai *speech recognition* merupakan satu kebolehan dalam mesin atau program untuk mengenalpasti perkataan dan ayat dalam pertuturan dan menukarkannya kepada format yang boleh dibaca oleh mesin. Perisian pengecaman pertuturan perdana mempunyai kata-kata dan frasa yang terhad, dan ia hanya boleh dikenalpasti jika mereka bercakap dengan sangat jelas. (WhatIs.com, 2016). Menurut kepada *An overview of modern speech recognition* (Huang & Deng 2010) pula, tugas pengecaman pertuturan adalah untuk menukar ucapan ke dalam urutan kata-kata oleh program komputer. Pada hakikatnya, pengecaman pertuturan membolehkan manusia berkomunikasi secara semula jadi dan lebih berkesan.

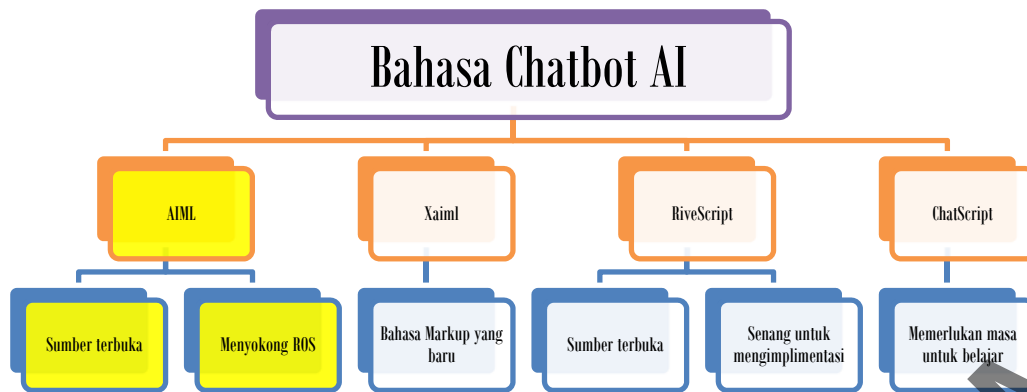
Berdasarkan Jurnal Teknologi Pengecaman Pertuturan Antarabangsa (*International Journal of Speech Technology*), terdapat tiga teknologi utama dalam teknologi pengecaman pertuturan, iaitu sintesis teks ke ucapan, pengecaman pertuturan automatik, dan ucapan yang tersimpan (dalam digital). Antara tiga teknologi ini, ucapan yang tersimpan (*stored digitized speech*) merupakan salah satu teknologi yang akan digunakan dalam kajian ini.



Rajah 1: Kategori Pengecaman Pertuturan

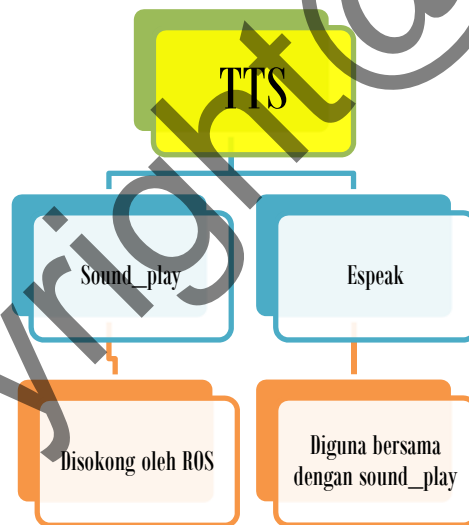
Pengecaman Pertuturan dapat dikategorikan kepada dua kategori, iaitu atas talian dan di luar talian. Seperti yang ditunjukkan dalam Rajah 1, dalam bahagian atas talian terdapat dua kategori lagi iaitu perkataan pendek dan juga data besar. Kategori perkataan pendek biasanya digunakan semasa pengecaman pertuturan untuk input yang pendek dan kategori data besar pula adalah untuk input yang banyak. Bahagian di luar talian pula terdapat dua kategori lagi, iaitu proprietari dan sumber terbuka. Kategori proprietary adalah bermaksud enjin pengecaman pertuturan ini adalah hak milik peribadi kepada pembangun enjin dan tidak boleh digunakan sesuka hati. Manakala kategori sumber terbuka pula, enjin di bawah kategori ini boleh digunakan dan diatur cara semula.

Pemilihan enjin pengecaman pertuturan adalah berdasarkan 3 kriteria, iaitu enjin haruslah dapat dijalankan dengan baik dalam ROS (*Robot Operating System*) dan Linux, dan juga dapat membuat pengecaman pertuturan di luar talian. Pocketsphinx telah dipilih sebagai enjin utama kerana ketiga-tiga kriteria dapat dicapai sekiranya menggunakan pocketsphinx. Pocketsphinx adalah satu enjin cabang CMU Sphinx.



Rajah 2: Jenis-jenis bahasa Chatbot

Dalam empat jenis bahasa Chatbot ini, RiveScript merupakan bahasa paling bagus untuk membina chatbot. Pengimplimentasi Rivescript adalah amat senang dalam usaha membangunkan pangkalan data untuk komunikasi. Akan tetapi, atas sebab RiveScript tidak menyokong ROS, oleh itu AIML dipilih sebagai bahasa chatbot yang utama dalam kajian ini.



Rajah 3: Jenis-jenis TTS

Dalam bahagian TTS, terdapat dua jenis enjin TTS untuk ROS, iaitu Sound_play dan Espeak. Kedua-dua enjin saling bergantung sekiranya dijalankan dalam ROS. Espeak merupakan satu jenis enjin TTS yang dapat digunakan dalam atau di luar ROS. Sound_play pula merupakan satu enjin TTS yang hanya dapat digunakan dalam ROS sahaja. Hal ini kerana Sound_play merupakan satu pakej ROS dan teks kepada suara boleh dilaksanakan hanya apabila diguna bersama dengan Espeak sahaja.

Robot perkhidmatan pula merupakan satu mesin yang dihasilkan untuk membantu manusia dalam banyak perkara. Seperti mana yang kita tahu, kebolehan manusia mempunyai batasan. Oleh itu, manusia menginginkan sesuatu untuk menyempurnakan ketidaksempurnaan manusia dalam pelbagai aspek. Di dalam era *4th Industrial Revolution*, penggunaan robot adalah sangat popular dan ia menjadi pilihan untuk membantu dalam perkara yang tidak ingin atau tidak mampu dilaksanakan seperti memindahkan barang berat, menyapu, mencuci pinggan mangkuk dan sebagainya.

Pengecaman pertuturan dalam robot perkhidmatan adalah salah satu topik hangat dalam dunia terkini. Kajian terhadap robot perkhidmatan masih diteruskan dalam aspek kemampuan untuk berkomunikasi dengan manusia.

Copyright@FITSM

iBotn adalah sejenis robot perkhidmatan yang dibina khusus untuk kanak-kanak berumur 1 hingga 3 tahun. Fungsi robot ini adalah untuk menjaga keselamatan kanak-kanak yang di tempat-tempat yang kemungkinan terdapat bahaya seperti rumah, kawasan lapang dan lain-lain lagi. Robot ini dapat membantu ibu bapa untuk menemani kanak-kanak dan juga sebagai alat komunikasi antara ibu bapa dan juga kanak-kanak. Selain itu, robot tersebut juga sebagai CCTV yang bergerak dan video kamera yang merekodkan saat-saat terbaik kanak-kanak supaya ibu bapa dapat melihat pembesaran anak mereka. Akan tetapi, robot ini mempunyai kekurangan dalam fungsinya iaitu, robot tersebut cuma dapat bertutur dalam bahasa Mandarin. Hal ini menjadi satu masalah sebab kebanyakan rakyat Malaysia tidak fasih dalam bahasa Mandarin.

Oleh itu, tindakan perlu dilakukan untuk meningkatkan penggunaan robot ini di kalangan rakyat Malaysia.

PENYATAAN MASALAH

Kajian ini difokuskan pada satu robot berjenama iBotn. Robot ini mempunyai kekurangan dalam cara berkomunikasi. Ia hanya berkomunikasi dalam bahasa tempatan Taiwan iaitu bahasa Mandarin sahaja. Berikut merupakan penjelasan tentang masalah yang dinyatakan di atas:

- i. Masalah implementasi dalam pengecaman dan sintesis pertuturan bahasa antarabangsa dan tempatan

Jenis robot tersebut hanya dapat berkomunikasi dalam Bahasa Mandarin sahaja. Hal ini menyebabkan kekurangan penggunaan robot ini di Malaysia atas factor pemahaman Bahasa.

OBJEKTIF KAJIAN

Usulan projek ini adalah bertujuan membangunkan sistem pengecaman pertuturan yang dapat memperbaiki kekurangan dalam ibotn. Pengecaman pertuturan dalam Bahasa Inggeris harus dihasilkan supaya robot ini dapat digunakan oleh keluarga yang tidak fasih dalam bahasa Mandarin.

Bagi mencapai matlamat tersebut, beberapa objektif khusus telah dipilih. Antaranya ialah:

- i.) Mengenalpasti kebolehan robot perkhidmatan.
- ii.) Membangunkan satu algoritma yang mempunyai pengecaman pertuturan dalam robot dengan menggunakan bahasa Inggeris.
- iii.) Membangunkan satu algoritma yang boleh menjawab balik soalan pengguna dengan pertuturan.

METOD KAJIAN

Metodologi yang akan digunakan dalam projek ini adalah amalan *Agile*. Amalan *Agile* merupakan salah satu metodologi yang paling popular. Keistimewaan amalan *Agile* ini ialah ia mempunyai fleksibiliti yang tinggi dan ia merupakan satu metodologi yang akan melibatkan pengguna untuk memberi pandangan dari masa ke semasa. Pengguna dapat memberi pandangan terhadap perisian yang telah dibangun pada masa awal dan dengan itulah keperluan pengguna dapat dijaga. Proses pembangunan aplikasi ini dapat dipercepatkan dan meminimumkan kos pembangunan dengan menggunakan metodologi ini.

(Mary 2013)(Dmitry 2015)

Antara fasa-fasa dalam kitaran lelaran projek *Agile* adalah seperti berikut:

- i.) Perancangan (*Planning*)

Fasa pertama iaitu fasa perancangan adalah untuk mengenalpasti objektif pelaksanaan projek supaya dapat menyelesaikan masalah yang dihadapi, mengkaji tentang masalah yang akan dihadapi semasa projek dilaksanakan, dan juga mencari cara penyelesaian masalah yang sesuai untuk menyelesaikan masalah yang telah dihadapi.

ii.) Analisis (*Analysis*)

Fasa analisis adalah untuk menganalisis keperluan spesifikasi untuk kajian ini. Perkara ini melibatkan banyak mesyuarat untuk membincang keperluan pengguna secara terperinci dan lebih mendalam. Selain itu, keperluan spesifikasi sistem juga perlu mengenalpasti supaya penghasilan kajian dapat berjalan dengan lancar.

iii.) Fasa Reka Bentuk (*Design*)

Fasa reka bentuk adalah untuk mereka bentuk antara muka aplikasi yang telah ditentukan. Dalam kajian ini, fasa reka bentuk tidak digunakan kerana sistem ini tidak mempunyai antara muka.

iv.) Fasa Pengekodan dan Pengujian (*Coding and Testing*)

Pembangunan berasaskan keperluan pengguna telah bermula dalam fasa ini. Pengujian akan dilakukan apabila sebahagian kod telah selesai untuk mengesan kecacatan dalam sistem. Apabila semua kod telah dihasilkan, pengujian ke atas pengguna seperti ibu bapa akan diadakan.

v.) Fasa Penggunaan

Hasil kajian akan dihantar kepada pelanggan dan akan mula digunakan. Komen daripada pelanggan akan dikumpulkan dan masalah akan diatasi.

HASIL KAJIAN

Kajian ini perlu membincangkan hasil kajian mengikut matlamat kajian iaitu membangunkan satu aplikasi yang mempunyai pengecaman pertuturan dalam robot dengan menggunakan bahasa Inggeris dan dapat menjawab balik soalan pengguna dengan pertuturan. Untuk mencapai matlamat kajian, proses pembangunan telah dimulakan pada semester 1 dengan mengenalpasti cara untuk membangun satu aplikasi yang dapat dijalankan dalam robot. Hal ini tercapai dengan perjumpaan bersama dengan Dr Abdul Hadi Abdul Rahman, iaitu pensyarah senior FTSM UKM yang pakar dalam aspek robot.

Sepanjang penghasilan hasil kajian ini, fungsi teks kepada teks dan fungsi teks kepada suara telah berjaya. Walaupun fungsi suara kepada teks juga telah berjaya, akan tetapi hasil untuk kajian ini masih tidak begitu tepat. Hal ini mungkin disebabkan perisian yang tidak cukup kualiti iaitu mic yang digunakan tidak dapat menerima suara dengan baik. Selain itu, hal ini juga mungkin disebabkan jumlah pangkalan data yang sebagai latihan untuk sistem ini tidak mencukupi dan juga kekurangan dalam kod dalam sistem ini. Walau bagaimanapun, kajian ini masih dapat diuji dengan cara memasukkan teks dan sistem akan membalas dengan suara dan juga teks.

CARA PENGUJIAN

Dalam fasa pengujian, terdapat tiga perkara yang perlu dibuat, hal ini adalah untuk memastikan kelancaran sistem dan ketepatan di setiap perkara. Perkara tersebut adalah seperti berikut:

- i. Teks kepada Teks
- ii. Teks kepada suara
- iii. Pengecaman pertuturan kepada suara

Untuk memulakan pengujian, komen berikut perlu dilakukan:

```
roscore
```

Roscore dipanggil adalah bertujuan untuk memulakan ROS dalam Linux.

Seterusnya, untuk pengujian pertama iaitu teks kepada teks, pengujian ini akan menguji tentang kelancaran AIML:

```
roslaunch ros_aiml start_chat.launch
```

```

/home/user/catkin_ws/src/ros_aiml/launch/start_speech_chat.launch http://localhost:11311
loading reduction3.safe.aiml... done (0.39 seconds)
loading pickup.aiml... done (0.00 seconds)
loading sex.aiml... done (0.01 seconds)
PARSE ERROR: Unexpected <category> tag (line 18, column 0)
PARSE ERROR: Unexpected </learn> tag (line 22, column 0)
PARSE ERROR: Unexpected </template> tag (line 35, column 0)
PARSE ERROR: Unexpected </category> tag (line 36, column 0)
loading primitive-math.aiml... done (0.00 seconds)
loading bot_profile.aiml... done (0.00 seconds)
loading psychology.aiml... done (0.05 seconds)
loading geography.aiml... done (0.08 seconds)
loading gossip.aiml... done (0.00 seconds)
loading personality.aiml... done (0.01 seconds)
PARSE ERROR: Unexpected text inside <random> element (line 4311, column 262)
PARSE ERROR: Unexpected text inside <random> element (line 4848, column 172)
PARSE ERROR: Unexpected text inside <random> element (line 8844, column 351)
loading default.aiml... done (0.52 seconds)
loading food.aiml... done (0.00 seconds)
loading reduction.names.aiml... done (0.35 seconds)
loading salutations.aiml... done (0.01 seconds)
loading computers.aiml... done (0.01 seconds)
loading mp2.aiml... done (0.58 seconds)
loading music.aiml... done (0.00 seconds)
loading mp5.aiml... done (0.23 seconds)
loading inquiry.aiml... done (0.01 seconds)
loading knowledge.aiml... done (0.14 seconds)
loading drugs.aiml... done (0.00 seconds)
loading reductionb.safe.aiml... done (0.76 seconds)
loading not_understand.aiml... done (0.00 seconds)
loading movies.aiml... done (0.00 seconds)
loading reductions.update.aiml... done (0.00 seconds)
loading looper10.aiml... done (0.00 seconds)
PARSE ERROR: <random> elements can only contain <li> subelements (line 443, column 495)
loading client.aiml... done (0.19 seconds)
loading wallace.aiml... done (0.06 seconds)
loading mp4.aiml... done (0.32 seconds)
loading continuation.aiml... done (0.00 seconds)
loading intersection.aiml... done (0.00 seconds)
loading humor.aiml... done (0.00 seconds)
loading politics.aiml... done (0.00 seconds)
loading mp3.aiml... done (0.62 seconds)
loading mp0.aiml... done (0.36 seconds)
Kernel bootstrap completed in 7.65 seconds
Saving brain to standard.brn... done (0.50 seconds)
[INFO] [1526226154.734236]: Starting ROS AIML Server

```

Rajah 5.6: Proses membaca AIML fail

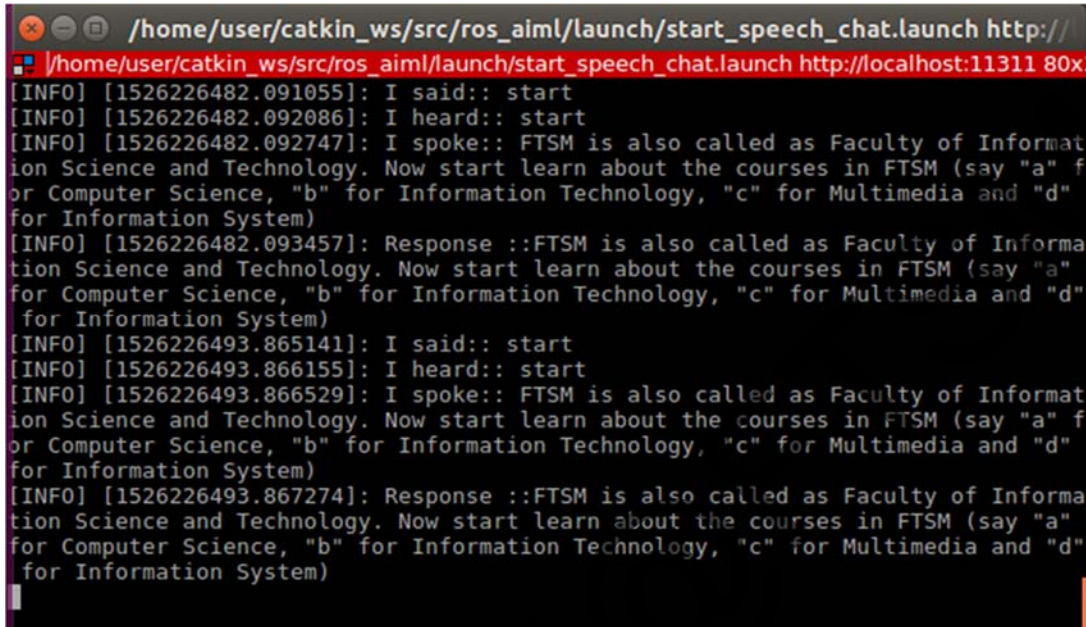
Untuk pengujian kedua iaitu teks kepada suara, pengujian ini akan menguji tentang kelancaran proses AIML kepada *sound_play*:

```
roslaunch ros_aiml start_tts_chat.launch
```

Dan akhir sekali, pengecaman pertuturan kepada suara, pengujian ini akan menguji tentang kelancaran proses pocketsphinx, AIML dan juga *sound_play*. Pocketsphinx perlu dipanggil terlebih dahulu supaya dapat menerima suara daripada pengguna.

```
roslaunch pocketsphinx pocketsphinx.launch
```

```
roslaunch ros_aiml start_speech_chat.launch
```



```

/home/user/catkin_ws/src/ros_aiml/launch/start_speech_chat.launch http://
/home/user/catkin_ws/src/ros_aiml/launch/start_speech_chat.launch http://localhost:11311 80x
[INFO] [1526226482.091055]: I said:: start
[INFO] [1526226482.092086]: I heard:: start
[INFO] [1526226482.092747]: I spoke:: FTSM is also called as Faculty of Informa
tion Science and Technology. Now start learn about the courses in FTSM (say "a" f
or Computer Science, "b" for Information Technology, "c" for Multimedia and "d"
for Information System)
[INFO] [1526226482.093457]: Response ::FTSM is also called as Faculty of Informa
tion Science and Technology. Now start learn about the courses in FTSM (say "a"
for Computer Science, "b" for Information Technology, "c" for Multimedia and "d"
for Information System)
[INFO] [1526226493.865141]: I said:: start
[INFO] [1526226493.866155]: I heard:: start
[INFO] [1526226493.866529]: I spoke:: FTSM is also called as Faculty of Informat
ion Science and Technology. Now start learn about the courses in FTSM (say "a" f
or Computer Science, "b" for Information Technology, "c" for Multimedia and "d"
for Information System)
[INFO] [1526226493.867274]: Response ::FTSM is also called as Faculty of Informa
tion Science and Technology. Now start learn about the courses in FTSM (say "a"
for Computer Science, "b" for Information Technology, "c" for Multimedia and "d"
for Information System)

```

Rajah 5.7: Proses pengecaman pertuturan kepada suara

Dalam proses pengecaman pertuturan, terdapat sebahagian daripada data perkataan yang digunakan untuk mengesan sama ada perkataan yang dicakap oleh pengguna dapat difahami oleh sistem. Berikut merupakan keputusan bagi peratusan untuk pengesanan dan pengecaman.

No	Perkataan	Pengesanan (%)	Pengecaman(%)
1	Hi	50%	20%
2	Hello	50%	20%
3	Start	70%	50%
4	No	90%	70%
5	Yes	60%	20%
6	One	70%	50%
7	Two	70%	30%
8	Three	60%	30%
9	Four	70%	40%
10	A	50%	30%
11	B	50%	30%
12	C	80%	60%
13	D	60%	40%
14	Morning	50%	30%
15	Good	70%	50%
16	Bye	60%	30%

Jadual 1: Keputusan untuk pengecaman pertuturan

Pengiraan bagi pengesanan (%)

= (jumlah perkataan wujud dalam bahagian pengesanan/jumlah pengujian)*100

Pengiraan bagi pengecaman (%)

= (jumlah perkataan wujud dalam bahagian pengecaman/jumlah pengujian)*100

Dalam hasil kajian ini, terdapat dua aspek yang mengenalpasti ketepatan pengecaman pertuturan, iaitu pengesanan dan pengecaman. Pengesanan adalah untuk mengenalpasti ketepatan perkataan yang telah dikesan oleh mesin. Dalam bahagian pengesanan, segala kemungkinan bagi perkataan yang disebut oleh pengguna yang seiras akan dipamerkan. Oleh itu, bahagian pengesanan biasanya mendapat ketepatan yang agak tinggi. Manakala bahagian pengecaman pula, keputusannya merupakan salah satu keputusan daripada

bahagian pengesanan. Oleh kerana gelombang bunyi yang dihasilkan oleh manusia sentiasa berubah, keputusan bagi pengecaman sentiasa berubah. Keputusan bagi bahagian pengecaman merupakan input kepada teks kepada suara. Oleh itu, hal ini menyebabkan keputusan jawapan yang tidak tepat kerana input yang tidak tepat.

KESIMPULAN

Bab ini menunjukkan kebolegunaan sistem ini dalam bahagian teks kepada teks dan juga bahagian teks kepada suara, akan tetapi wujudnya kekurangan dalam pengecaman pertuturan kepada suara kerana kurang ketepatan perkataan yang telah dikecam itu.

RUJUKAN

- Dmitry Gurdendo. 2015. Scrum Methodology Phases which Help in Agile SDLC Process: 5 Key Steps. *June 30*,. <https://xbssoftware.com/blog/software-development-life-cycle-sdlc-scrum-step-step/> [30 September 2017].
- Huang, X. & Deng, L. 2010. An overview of modern speech recognition. *Handbook of Natural Language ...*, 339–367. doi:doi:10.1201/9781420085938-c15
- Joseph, L. 2017. *ROS Robotics Projects*. (K. Pandey, Ed.). Packt Publishing Ltd.
- Mary Lotz. 2013. Waterfall vs. Agile: Which Methodology is Right for Your Project? *July 5*,. <https://www.seguetech.com/waterfall-vs-agile-methodology/> [30 September 2017].
- Melorose, J., Perroy, R. & Careas, S. 2015. *Learning Robotics Using Python. Statewide Agricultural Land Use Baseline 2015*, hlm.Vol. 1. doi:10.1017/CBO9781107415324.004

Muchammad. 2013. Pengertian Linux, Sejarah Linux, Kelebihan dan Kekurangan dan Macam-macam Distro Linux - Distro. 20 February ,. <http://barkahtroox.blogspot.my/2013/02/pengertian-linux-sejarah-linux.html> [20 October 2017].

Open Source Robotic Foundation. 2014. ROS/Introduction - ROS Wiki. *Open Source Robotic Foundation*,. <http://wiki.ros.org/ROS/Introduction> [11 December 2017].

Perkenal, A., Perlindungan, S., Menyeluruh, T. & Dengan, T. M. 2017. Siaran Akhbar Anggaran Penduduk Semasa, 2016-2017, pp. 2–3. doi:<http://www.mida.gov.my/env3/uploads/PerformanceReport/2013/KenyataanMedia.pdf>

Severinson-Eklundh, K., Green, a & Huttenrauch, H. 2003. Social and Colaborative Aspects of Interaction with a Service Robot. *Robotics and Autonomous Systems*, 42, 223–234.