Research Article
Journal of Advanced Development in Engineering and Science
Vol. 13 ● No. 38 • September - December 2023

Development of Process for Generating Thai Audio Image Description

Wasin Pirom[1*]

Recived: 14 May 2023
Revised: 13 August 2023
Accepted: 11 September 2023

ABSTRACT

The motivation for this research began with the requirement to develop a tool for visually impaired and blind people to improve their quality of life by allowing them to access visual information and travel safely. The process for generating Thai audio image descriptions was developed. The research has developed an image detection system that classifies objects in detected images and generates Thai language descriptions without translation from English. The DEtection TRansformer (DETR) is accurately and quickly applied to detect the objects in the image; after that, the Thai sentences or phrases are composed using the Thai Text Generator with the Thai model of WangchanBERTa datasets. The important features of the image description in this research are the ability to indicate the number of objects of the same kind and to select appropriate noun classifiers. In this developed image detection system, objects within the detected image are automatically divided into different images with classification, and then the number of images in each category is counted. The suitable noun classifier is chosen using the Masked Token Prediction. This system can reclassify using zero-shot learning when more different images are added. This allows for more flexibility in use and saves a significant amount of time in creating an image database. The Thai image description consists of the details, including the type of object, number, and Thai noun classifier of objects in the images; the generated sentence to describe the image; and the predicted photo shoot location. After that, the Thai sentences or phases of captured image description are transformed to be the voice in the Thai language using the

[1] Department of Software Engineering and Information System, Faculty of Science and Technology, Pathumwan Institute of Technology, Bangkok, Thailand 10330

*Corresponding author e-mail: wasinpirom@pit.ac.th

VAJA Text-to-Speech Engine integrated image detection system to enable visually impaired and blind people to recognize the details of the image in front. The results showed a good performance of the developed process for generating Thai audio image descriptions. The input image file can be transformed into the image descriptions and the Thai audio descriptions.

Keywords: Image detection; Thai audio image description; Zero-shot learning

Introduction

The quality of life for visually impaired and blind (VIB) people, who face problems in daily life as they lack the ability to recognize visual information, has developed with many applications in smart phones to know objects [1]. For example, TapTapSee [2] and Aipoly Vision [3] are applications that allow VIB people to recognize objects by using the mobile phone camera to take a photo and hearing the description of the object read back to the user. Many real-time apps are in English or other languages; there are a limited number of applications that support Thai. Opportunely, the Digital Service for Disability has promoted a few applications supporting Thai, including TAB2Read for audio books [4], Smart Eye for transforming original document images into audio, such as receipts, notice boards, lottery tickets, and banknotes [5], and Navilens for scanning the QR code at a distance of up to 12 meters and detecting multiple tags at the same time [6]. However, there was no development of Thai audio image descriptions consisting of the type of object, number, and Thai noun classifier of objects generated from the image via the smartphone camera. This research has developed the Thai audio image description of the image detection system to support Thai VIB people in recognizing the details of the image on the book and the environment around them. The previous research developed the object position detection for VIB people to locate the position of the finding object using a vision-language pre-training model as CLIP (Contrastive Language-Image Pre-training) with a voice command in the Thai language [7]. Moreover, Nimmolrat et al. [8] developed a mobile pharmaceutical application with functions that are appropriate for visually impaired users and a voice command function.

Numerous approaches for generating image descriptions to specify the details and characteristics of images in the English language have been continuously developed, for example, D´ej`a Image-Captions: the naturally existing image descriptions that are repeated almost verbatim [9], a distributed representation based query expansion approach [10], Collective Generation of Natural Image Descriptions using ILP and HMM formulation [11], Web-scale N-grams [12], Corpus-Guided predictions [13], visual dependency representation (VDR) [14-

15], Kernel Canonical Correlation Analysis (KCCA) -based baseline systems [16], linear phrase-based model [17], and semantic hierarchies and zero-shot recognition [18].

In recent years, image descriptions in the Thai language were also introduced [19, 20]; however, the detected image model to generate image descriptions in the Thai language is mostly translated via English. Khuphiran et al. [19] developed a scene graph generator tool from a single image in 3 steps: image captioning, scene graph parser, and machine translation from English to Thai by a neural machine translator. Mookdarsanit et al. [20] proposed the model to generate the Thai image caption with Convolutional Neural Network (CNN) for the encoding stage and Recurrent Neural Network (RNN) for the decoding stage, and introduced the English captions directly translated into Thai texts by VISTEC thai2nmt in which this machine translation based on Transformer reduces the time for manual captioning. For Thai descriptions, there was no research emphasizing Thai noun classifiers and the exact quantity of objects in an image, whereas various researches [21-23] presented the generated image descriptions in English consisting of the number of objects in images. Nevertheless, the sentence structure of the Thai language, especially the noun classifiers, is different from the English language; the Thai language has noun classifiers used when the noun is being counted; that is, it appears with a numeral [24].

As mentioned above, the research concerned with the generation of the Thai image description was limited, especially the Thai audio image description. Consequently, the main objective of this research was to develop a process to generate the image descriptions in Thai without translation from English to Thai and combine the text-to-speech engine for creating Thai audio descriptions. In this research, the Thai datasets of WangchanBERTa were used [25], and Masked Token Prediction was used to predict a suitable noun classifier [26]. The integrated system between DEtection TRansformer or DETR [27] for detecting objects in the images and the Thai datasets of WangchanBERTa was developed to generate the Thai texts and further compose the Thai sentences by using Thai Text Generator, an open-source tool for composing Thai sentences [28], in order to decrease a training step and obtain the Thai language description of the image in a short time. In addition, zero-shot learning, which is a new conceptual learning technique without accessing any exemplars of the unseen categories during training and can construct recognition models with the assistance of transferring knowledge from previously seen categories and auxiliary information [29], was also utilized to predict the photoshoot location of the detected image.The generated image description was immediately demonstrated without converting from English to Thai, and after that, Thai language audio descriptions were generated to describe details of the captured image.

Process Design

This research developed a process for converting images into Thai-language descriptions without translating via the English language, and Thai descriptions generated as text were converted to Thai audio using a text-to-speech engine. The Python programming in Colab Pro was used, and the results were generated by using a GPU (Graphics Processing Unit) hardware accelerator. There were five steps in the developed process for generating Thai audio image description, as demonstrated in Figure 1.

In the first step, the image file with .JPG extensions (including .jpg and .jpeg) was loaded into the developed process. After that, the system started to preprocess the image to a suitable size (enlarge the image size eight times) for sending it to the process of image detection.

In the second step, objects in the image were detected using DEtection TRansformer (DETR) which immediately transfers to Thai text and demonstrates the number of objects by counting the repeated words of the same type of objects in the image detected by DERT as shown in Figure 2. It can predict up to 80 different objects (since the data trains are from MS-COCO with 80 Classes) and reveal the open-source code that can be used to train it to know more objects as needed. In this step, the names of objects in the image and the numbers of objects were obtained.

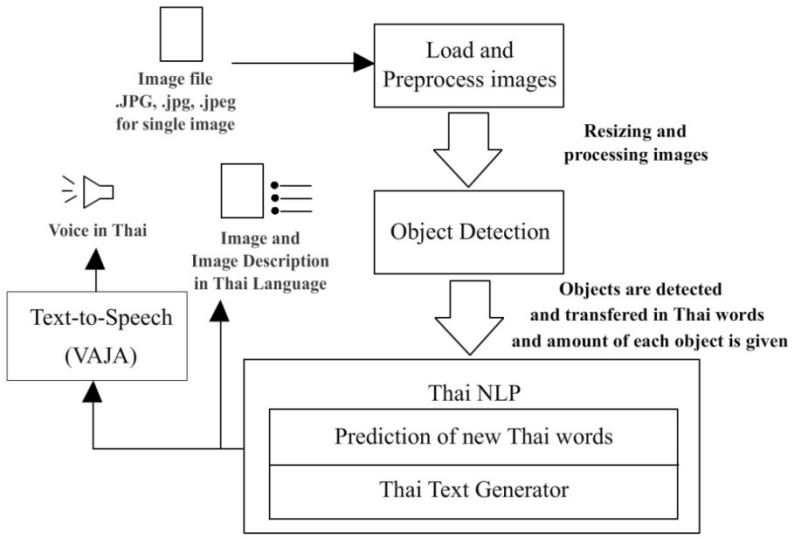

Figure 1 Developed Process for generating Thai audio image description

In the next step, the Thai Natural Language Processing (Thai NLP) was used to predict new related Thai words with similar meanings using WangchanBERTa datasets. The WangchanBERTa pre-trained language model on Thai datasets obtained from sources such as news, Wikipedia, social media messages, and information obtained from crawling websites on the Internet with a total data size of 78.5 GB is developed by Thailand Artificial Intelligence Research Institute [25]. The model Wangchanberta-base-att-spm-uncased trained on the 78.5 GB dataset outperforms strong baselines (NBSVM, CRF, and ULMFit) and multi-lingual models (XLMR and mBERT) on both sequence classification and token classification tasks in human-annotated, mono-lingual contexts [26]. In each group of similar words, only one word was chosen from the highest score, and the noun classifier of the chosen word was predicted using the Masked Token Prediction that a suitable noun classifier was chosen from the maximum probability score in the Thai text as illustrated in Figure 2. The Thai text of an object's name, a related word and a chosen noun classifier will be presented in 'the detail of the image'. The developed process can be used to segment words into sentences (word tokenization). The Masked Language Model was used to predict the missing words in the masked position to complete sentences.

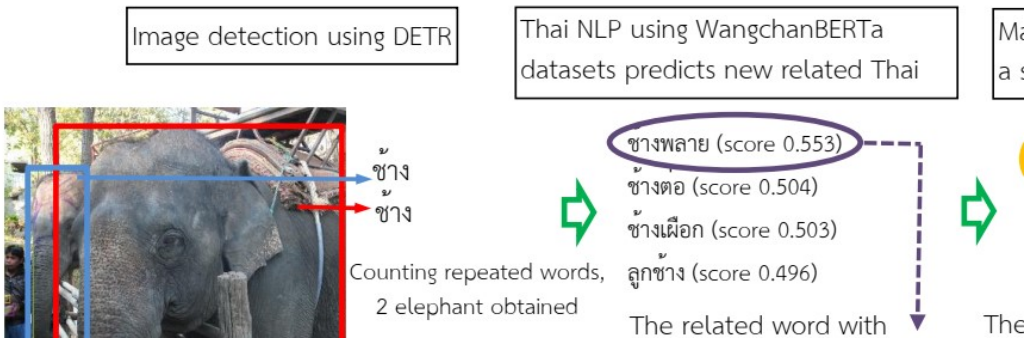

Figure 2 Image detection using DERT and the prediction of the related word and the suitable noun classifier

Then, the Thai Text Generator system was utilized to compose phrases or sentences [30]. Using the Markov chains [31-32], all the words are transferred to the system to create sentences of 2-10 words, as shown in 'the creating text caption'. Figure 3 shows an example of the word of 'ช้าง'elephant for phase or sentence generation using the Thai text generator. In addition, 'the prediction of the photoshoot location' of the detected image was presented

using zero-shot learning for text classification [33-34], which can reclassify words and sentences into new types without requiring any additional training. Figure 4 shows the highest score of the word of 'สวนสัตว์' zoo when using zero-shot learning to predict the photoshoot location.

Words from DETR		Group of phrase generation	English meaning
ช้าง (elephant)	Thai Text Generator	ช้างป่างาม ในประเทศไทย	Beautiful wild elephant In Thailand

Figure 3 Group of phrase generation using Thai text generator

Words from DETR		location	English meaning
ช้าง (elephant) คน (human) ชาม (bowl)	Zero-shot for text classification	**สวนสัตว์ (score: 0.470)** ท้องถนน (score: 0.294) สี่แยกไฟแดง (score: 0.149) ห้างสรรพสินค้า (score: 0.086)	Zoo Road Traffic light intersection Shopping mall

Figure 4 The prediction of the photoshoot location using zero-shot learning

After the generation process of Thai image description was completed, the image description was composed of the details of the image, including the type of object, amount of object, Thai noun classifier of an object in the images, the creating text caption, and the prediction of the photoshoot location. The Thai language image description was shown above the picture.

In the final step, this process added the VAJA Text-to-Speech Engine [35] to transform Thai text descriptions into the voice in Thai so that VIB people can access Thai image descriptions. Using the developed process as proposed in Figure 1, when the input image file was loaded, the output of this process was the image description and the Thai audio description.

Results and Discussion

The developed process to generate the Thai audio image description was performed within about a minute operating on the Google Colab. This research used three images to test the developed process, including two close-up photos as shown in Figures 5 and 6 and one

distant photo as shown in Figure 7. An example of results for an image in Figure 5 was discussed. In the step of image detection, five Thai words were generated as ‘ช้าง’elephant, ‘ช้าง’elephant, ‘ชาม’bowl, ‘ชาม’bowl, and ‘คน’human. The repeated words were counted to create the number of things in the image. The results showed that the detection step can identify the type and number of objects exactly, including 2 elephants, 2 bowls, and 1 human, as demonstrated in Figure 5.

The next step was the generation of similar words for each group, and the suitable noun classifier was created as shown in Table 1. In the third step, the groups of similar words generated in the second step of image detection were predicted, and their scores were shown. Each group chose one similar word with the highest score to present in the Thai language description. For example, the word of ‘ช้าง’ (elephant) was generated in the second step as mentioned above; a group of words for elephant consists of ‘ช้างพลาย’ (male elephant), ‘ช้างต่อ’ (decoy elephant), ‘ช้างเผือก’ (white elephant), and ‘ลูกช้าง’ (calf), in which the highest score of the male elephant was presented and chosen.

Figure 5 Thai language description withThai audio of the image shooting at the zoo as meaning:

‘**The detail of the image:** The picture consists of 2 elephants (or male elephants), 1 human (or group of humans) and 2 bowls (or dishes)

The creating of text caption: Human has relationship with beautiful wild elephant in Thailand.

The prediction of the photoshoot location: Zoo’

Moreover, in the Thai language, there was a noun classifier of different things, for example, a phrase of ‘2 elephants’ in English is often required to be expressed as ‘ช้าง 2 ตัว’ (Ch̄̂ang 2 Tạw), that ‘Tạw’ is the noun classifier of the wild elephant, with a higher score than ‘ช้าง 2 เชือก’ (Ch̄̂ang 2 Cheụ̄xk), that ‘Cheụ̄xk’ is the noun classifier of the domestic elephant. Therefore, the detail of the image in the Thai language presented ‘ช้าง (หรือช้างพลาย) 2 ตัว’ representing ‘2 elephants (or male elephants)’ in the English language.

As shown in Figure 5, all noun classifiers for objects were correctly chosen. The Thai sentence to describe the image was composed, and the photoshoot location of the zoo was

predicted using zero-shot learning. After Thai image descriptions were generated, the Thai audio description was then automatically generated and appeared under the picture.

Table 1 Words and noun classifiers of objects in image descriptions of Figure5

Groups of similar words and their scores	Classifiers and Scores	Chosen classifiers
Group of elephant ช้างพลาย score: 0.553 ช้างต่อ score: 0.504 ช้างเผือก score: 0.503 ลูกช้าง score: 0.496	- ช้าง 2 ตัว (Tạw) score: 0.751 / ช้าง 2 เชือก (Cheụ̄xk) score: 0.043	ตัว (Tạw)
Group of human กลุ่มคน score: 0.556 ผู้หญิง score: 0.510 คนไทย score: 0.477 ผู้ชาย score: 0.477	- คน 1 คน (khn) score: 0.131 / คน 1 ค่ะ (Khà) score: 0.051	คน (khn)
Group of bowl จาน score: 0.685 กะละมัง score: 0.650 ถ้วย score: 0.619 หม้อ score: 0.613	- ชาม 2 ใบ (Bı) score: 0.061 / ชาม 2 คน (khn) score: 0.035	ใบ (Bı)

In Figure 6, Although the image shows only part of a train, this system can detect the image and create the word correctly. Only one train in the image was detected and generated one word of ‘รถไฟ’train and a similar word of ‘ขบวนรถไฟ’railroad train. A correct noun classification was chosen. In addition, the train station was predicted as the photoshoot location.

รายละเอียดของภาพ:
ในภาพมี รถไฟ(หรือ ขบวนรถไฟ) 1 ขบวน
การสร้างข้อความอธิบายภาพ:
รถไฟเชื่อมต่อระหว่างเมือง
การทำนายสถานที่ถ่ายภาพ: สถานีรถไฟ

Figure 6 Thai language description with Thai audio for the image shooting atthe train stationas meaning

'**Detail of image:** The picture consists of 1 train (or railroad train)
Creating of text caption: Train links intercity
Prediction of photoshoot location: Train station'

In the case of a remote photo, this system can detect the image to generate the Thai language description as illustrated in Figure 7. Moreover, small objects such as people can be detected and generated the words in detail of image. However, the different color of a trailer in the left hand side of image was detected as two trucks. However, other objects of airplane, human, and bus were precisely counted. The results showed that all noun classifiers of objects were correctly selected, and the location of the photo shooting of the airport was appropriately predicted.

Figure 7 Thai language description with Thai audio for the image shooting at the airport as meaning

'**Detail of image:** The picture consists of 1 airplane (or helicopter), 2 humans (or group of human), 1 bus (or coach), and 3 trucks (or ten-wheeled trucks)
Creating of text caption: Airplane contains fresh air, bus, transport, trucks
Prediction of photoshoot location: Airport'

The results showed that the developed process can generatethe Thai image descriptions and Thai audio descriptions. In this study, the developed process was the first step in progressing an application for the Thai VIB people, who are limited in using existing English language applications. Nowadays, there are mobile applications to recognize objects and colors for supporting VIB people, such as Aipoly Vision [3], which supports many languages but does not support Thai [3]. TapTapSee [2] is a mobile camera application designed specifically for VIB users, powered by the Cloud Sight Image Recognition API; the processing is available in Thai using Voice Over function for Thai audio. The results of processing Figures 5, 6, and 7 using TapTapSee were

illustrated in Figure 8 (a), (b), and (c), respectively when it can analyze and identify objects within seconds. Comparing the results of image processing using the process developed in this study with that of the TapTapSee application, it was found that the processing time in the developed process was slower than in the applicationdue to many steps for selecting of noun classifier, counting, composing and arranging sentences, and predicting the location. Highlights of this development are able to describe the number and noun classifier of objects correctly, and to predict the location properly,even images taken from a distance. However, the limitations of this developed process cannot describe the object colors.

(a) TapTapSee result of Figure 5 with the meaning ‘elephant on brown wooden fence during daytime’

(b) TapTapSee result of Figure 6 with the meaning ‘white and blue train in a train station’

(c) TapTapSee result of Figure 7 with the meaning ’white and black van in the parking lot‘

Figure 8 Thai audio image description using the TabTabSee mobile application

Conclusion

This research needs to develop the process for generating Thai audio image description. In this study, the developed process can generate the image details, which have the highlights of being able to accurately identify the number and Thai noun classifier of the objects. Moreover, the image detection step to generate the Thai language description without

translating it from the English language was designed. The image can be detected using DETR, and Thai sentences are generated using the Thai Text Generator and WangchanBERTa dataset. The Thai language description consists of the details of the image, including the type of object, number, and Thai noun classifier of objects in the images, the creating text caption, and the prediction of the photoshoot location. The developed process showed satisfactory performance and can specify the correct type and amount of objects and the noun classifiers (in Thai) of objects. Furthermore, the zero-shot text classification system can be used effectively for predicting the location of images. In the main idea for developing the process to generate the Thai audio description for supporting the VIB people, the Thai descriptions were transformed into the Thai audio by using VAJA Text-to-Speech. However, there are many steps in processing, thus causing the processing time to slow down, so faster processing will be improved in the future.

To be able to develop this process to make it easier to use for the blind in addition to using it on a computer, a mobile application will be developed in the future. Satisfaction surveys and feedback from VIB users will also be conducted.

Acknowledgements

An author gratefully acknowledges the Pathumwan Institute of Technology Research Fund for financial support in the research project of “A Design and Development of 360-degree Object Position Detection System to Assist Visually Impaired People.”

References

[1] Gitari, M. (2023). *The 7 Best Apps to Help People with Visual Impairments Recognize Objects*. Available from https://www.pathstoliteracy.org/resource/7-best-apps-help-people- visual -impairments-recognize-objects. Accessed date: 16 April 2023.

[2] TapTapSee. (2023). *TapTapSee: Assistive Technology for the Blind and Visually Impaired*. Available from https://taptapseeapp.com. Accessed date: 16 April 2023.

[3] Aipoly. (2017). *Aipoly Vision for Android*. Available from https://download.cnet.com/ Aipoly-Vision/3000-20432_4-77580945.html. Accessed date: 16 April 2023.

[4] D4D. (2021). *Digital Service for Disability*. Available from https://d4d.onde.go.th/. Accessed date: 16 April 2023.

[5] TAB2Read. (2023). *TAB2Read*. Available from https://d4d.onde.go.th/app-portal/47. Accessed date: 16 April 2023.

[6] Navilens. (2023). *Navilens*. Available fromhttps://www.navilens.com/en. Accessed date

16 April 2023.

[7] Pirom,W. (2022). Object Detection and Position using CLIP with Thai Voice Command for Thai Visually Impaired. In 37th International Technical Conference on Circuits/Systems, Computers and Communications (p.391-394). 5-8 July 2022, Phuket, Thailand.

[8] Nimmolrat, A., et al. (2021). Pharmaceutical mobile application for visually-impaired people in Thailand: development and implementation. *BMC Medical Informatics and Decision Making*, 21, 217(2021).

[9] Chen, J., et al. (2015). Déjà image-captions: A corpus of expressive descriptions in repetition. In *The 2015 Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (p.504-514). 31 May - 5 June 2015, Denver, Colorado, USA.

[10] Yagcioglu, S., et al. (2015). A Distributed Representation Based Query Expansion Approach for Image Captioning. In *the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing* (p.106-111). 26-31 July, 2015, Beijing, China.

[11] Kuznetsova, P., et al. (2012). Collective generation of natural image descriptions. In *The 50th Annual Meeting of the Association for Computational Linguistics* (p.359-368). 8-14 July, 2012, Jeju, Korea.

[12] Li, S., et al. (2011). Composing simple image descriptions using web-scale n-grams. In *The 15th Conference on Computational Natural Language Learning* (p.220-228). 23-24 June, 2011, Portland, Oregon, USA.

[13] Yang, Y., et al. (2011). Corpus-guided sentence generation of natural images. In *The 2011 Conference on Empirical Methods in Natural Language Processing* (p.444-454). 27-31 July, 2011, Edinburgh, Scotland.

[14] Elliott, D. & de Vries, A. P. (2015). Describing images using inferred visual dependency representations. In *The 53rd Annual Meeting of the Association for Computational Linguistics and The 7th International Joint Conference on Natural Language Processing* (p.42-52). 26-31 July, 2015, Beijing, China.

[15] Elliott, D. & Keller, F. (2013). Image Description using Visual Dependency Representations. In *2013 Conference on Empirical Methods in Natural Language Processing* (p.1292-1302). 18-21 October, 2013, Seattle, Washington, USA.

[16] Hodosh, M., et al. (2013). Framing Image Description as a Ranking Task: Data, Models and Evaluation Metrics. *Journal of Artificial Intelligence Research*, 47, 853-899.

[17] Lebret, R., et al. (2015). Phrase-based image captioning. In *International Conference on Machine Learning* (p.2085-2094). 6-11 July 2015, Lille, France.

[18] Guadarrama, S., et al. (2013). Youtube2text: Recognizing and describing arbitrary activities using semantic hierarchies and zero-shot recognition. In *IEEE International Conference on Computer Vision* (p.2712-2719). 1-8 December, 2013, Sydney, Australia.

[19] Khuphiran, P., et al. (2019). Thai Scene Graph Generation from Images and Applications. In *23rd International Computer Science and Engineering Conference* (p.361-365). 30 October – 1 November 2019, Phuket, Thailand.

[20] Mookdarsanit, P. & Mookdarsanit, L. (2020). Thai-IC: Thai Image Captioning based on CNN-RNN Architecture. *International Journal of Applied Computer Technology and Information Systems*, 10, 40-45.

[21] He, S.,et al. (2020). Image Captioning through Image Transformer. In *16th Asian Conference on Computer Vision* (p.153-169). 4-8 December, 2022, Macao, China.

[22] Kulkarni, G., et al. (2013). Baby Talk: Understanding and Generating Simple Image Descriptions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35, 2891-2903.

[23] Elliott, D. & Keller, F. (2014). Comparing Automatic Evaluation Measures for Image Description. In *The 52nd Annual Meeting of the Association for Computational Linguistics* (p.452-457). 22-27 June 2014, Baltimore, Maryland.

[24] Wikipedia. (2023). *Classifier (linguistics).* Available from https://en.wikipedia.org/wiki/Classifier_(linguistics). Accessed date: 16 April 2023.

[25] VISTEC-depa Thailand AI Research Institute. (2021). *WangchanBERTa: Pre-trained Thai Language Model*. Available from https://airesearch.in.th/releases/wangchanberta-pre-trained-thai-language–model. Accessed date: 16 April 2023.

[26] Lowphansirikul, L., et al. (2021). *WangchanBERTa: Pre-training transformer-based Thai Language Models*. Available from https://arxiv.org/abs/2101.09635. Accessed date: 16 April 2023.

[27] Carion, N., et al. (2020). *End-to-end object detection with transformers*. Available from https://arxiv.org/abs/2005.12872.Accessed date: 16 April 2023.

[28] Phatthiyaphaibun, W. (2020). *TTG: Thai Text Generator*. Available from https://colab.research.google.com/drive/1X6D8J0sWNi8UgJi7Hk5YL4FqepZ7laxS?usp=sharing. Accessed date: 16 April 2023.

[29] Rezaei, M. & Shahidi, M. (2020). Zero-Shot Learning and its Applications from Autonomous Vehicles to COVID-19 Diagnosis: A Review. *Intelligence-Based Medicine*, 3-4, 100005.

[30] Github. (2021). *Thai Text Generator*. Available from https://github.com/ PyThaiNLP/ Thai-Text-Generator. Accessed date: 9 September 2022.

[31] Szymaṅski, G. & Ciota, Z. (2002). *Hidden Markov Models Suitable for Text Generation*.

Available from http://citeseerx.ist.psu.edu/viewdoc/summary? doi=10.1.1.335.93. Accessed date: 16 April 2023.

[32] Department of Statistics. (2014). *COURSE NOTES STATS 325 Stochastic Processes*. Auckland: University of Auckland.

[33] Pushp, P. K. & Srivastava, M. M.(2017). *Train once, test anywhere: Zero-shot learning for text classification*. Available from https://arxiv.org/abs/1712.05972. Accessed date: 16 April 2023.

[34] Puri, R. & Catanzaro, B. (2019). *Zero-shot text classification with generative language models*. Available from https://arxiv.org/abs/1912.10165. Accessed date: 16 April 2023.

[35] NECTEC. (2016). *VAJA Text-to-speech Engine*. Available from https://www.nectec.or.th/innovation/innovation-mobile-application/vaja.html date: 30 April 2023. (in Thai)