

**AN IMPLEMENTATION OF WEB SERVICES FOR THAI TEXT  
LEXICONIZATION AND VISUALIZATION**



**SOMKIT KITTICHAJAROEN**

**A THEMATIC PAPER SUBMITTED IN PARTIAL  
FULFILLMENT OF THE REQUIREMENTS FOR  
THE DEGREE OF MASTER OF SCIENCE  
(INFORMATION TECHNOLOGY MANAGEMENT)  
FACULTY OF GRADUATE STUDIES  
MAHIDOL UNIVERSITY**

**2018**

**COPYRIGHT OF MAHIDOL UNIVERSITY**

AN IMPLEMENTATION OF WEB SERVICES FOR THAI TEXT LEXICONIZATION  
AND VISUALIZATION

SOMKIT KITTICHAJAROEN 5938523 EGIT/M

M.Sc. (INFORMATION TECHNOLOGY MANAGEMENT)

THEMATIC PAPER ADVISORY COMMITTEE: SOTARAT THAMMABOOSADEE,  
Ph.D., TAWEESEK SAMANCHUEN, Ph.D.

ABSTRACT

Analysis of text data in different languages uses different techniques for word segmentation or sentence division. Thai language is one of the languages that use different techniques from others because of the complexity in the language structure. The most popular current technique is word comparison by using dictionary and compares the longest word available. This research presents the web service for Thai word segmentation to work in the form of Client-Server to serve the users by providing services in the form of REST API via the protocol HTTP. Besides web services, this research also develops a system to test web service and word segmentation, and to show analysis results in the form of visualization by displaying various type of visualization results through web application. Alternative channels to data analysts are added as tools for analyzing additional information. The results of research show that the time spent for processing the Thai word of the Web service depends on the amount of data sent to the system, which is a satisfactory result. Web application used for words segmentation and visualization in Thai language can display 5 types of visuals from the evaluation of general users. Based on a 30 persons evaluation, the rating was rated 3.7 out of 5. The results are satisfactory.

KEY WORDS: THAI WORD SEGMENTATION / WEB SERVICE  
TEXT VISUALIZATION

59 pages

การดำเนินการเว็บเซอร์วิสสำหรับการตัดคำและทัศนภาพข้อความภาษาไทย

AN IMPLEMENTATION OF WEB SERVICES FOR THAI TEXT LEXICONIZATION AND VISUALIZATION

สมกิจ กิตติชัยเจริญ 5938523 EGIT/M

วท.ม. (การจัดการเทคโนโลยีสารสนเทศ)

คณะกรรมการที่ปรึกษาสารนิพนธ์: โยทศรัตต ธรรมบุษดี, Ph.D., ทวีศักดิ์ สมานชื่น, Ph.D.

บทคัดย่อ

การวิเคราะห์ข้อมูลข้อความในภาษาต่าง ๆ มีการใช้เทคนิคในการตัดคำหรือการแบ่งประโยคของแต่ละภาษาแตกต่างกัน ภาษาไทยเป็นภาษาหนึ่งที่ใช้เทคนิคในการตัดคำที่แตกต่างจากภาษาอื่น เนื่องจากโครงสร้างทางภาษามีความซับซ้อน โดยมีเทคนิคที่นิยมใช้กันคือ การเทียบคำที่ยาวที่สุดกับคลังข้อความ งานวิจัยนี้ทำการนำเสนอระบบ Web Service การตัดคำภาษาไทยให้สามารถทำงานในรูปแบบของ Client - Server เพื่อให้บริการแก่ผู้ใช้งาน โดยให้บริการในรูปแบบของ REST API ผ่านทาง Protocol HTTP และยังพัฒนาระบบที่ใช้ทดสอบ Web Service การตัดคำภาษาไทยและการวิเคราะห์ผลออกมาในรูปแบบของ Visualization ที่แตกต่างกัน โดยแสดงผลผ่าน Web Application เพิ่มช่องทางให้กับนักวิเคราะห์ข้อมูลใช้เป็นเครื่องมือในการวิเคราะห์ข้อมูลเพิ่มเติมได้ ผลจากงานวิจัยคือ เวลาที่ใช้สำหรับการประมวลผลการตัดคำภาษาไทยของ Web service ขึ้นอยู่กับจำนวนข้อมูลที่ส่งเข้าระบบ ซึ่งผลเป็นที่น่าพอใจและ Web application ที่ใช้ในการตัดคำและทัศนภาพข้อความภาษาไทย ซึ่งสามารถแสดงทัศนภาพได้ 5 แบบ จากการประเมินโดยผู้ใช้งานทั่วไปจำนวน 30 คน มีผลการประเมินอยู่ในระดับ 3.7 จาก 5 ผลการประเมินอยู่ในระดับที่น่าพอใจ

59 หน้า