รายงานการวิจัย

การพัฒนาพจนานุกรมอิเล็คทรอนิกส์ ไทย-ญี่ปุ่น

รศ.กฤษดา วิศวธีรานนท์

รศ.ดร.บัณฑิต โรจน์อารยานนท์

Prof.Dr.Minegishi Makoto

ดร.สุรพันธ์ เมฆนาวิน

ดร.ดัชกรณ์ ตันเจริญ

อ.ดร.วรรณวิมล รุ่งธีระ

อ.ธิติสรณ์ แสงอุไร

อ.Shuichi Sano

อ.บัญชา บริคุต

อ.ณัฐพล จารัตน์

คุณสมใจ แซ่อึ้ง

คุณสลีลา บุญตระกูลสุข

สถาบันเทคโนโลยีไทย-ญี่ปุ่น

พฤษภาคม 2554

สารบัญ

สารบัญ	ก
สารบัญตาราง	ง
สารบัญรูป	จ
กิตติกรรมประกาศ	ฉ
บทคัดย่อภาษาไทย	ช
บทคัดย่อภาษาอังกฤษ	ซ
บทที่ 1 บทนำ	1
1.1 ความเป็นมาของโครงการวิจัย	1
1.2 วัตถุประสงค์ของโครงการวิจัย	2
1.3 ขอบเขตของโครงการวิจัย	2
1.4 ขั้นตอนและแผนการดำเนินการวิจัย	2
1.5 ประโยชน์ที่จะได้รับ	3
บทที่2 พจนานุกรมอิเล็คทรอนิกส์	4
2.1 แนวคิดในการวิจัย	4
2.2 e-Dictionary ที่มีการให้บริการ	5
2.3 การศึกษาวิจัยเรื่อง e-Dictionary ของ Tokyo University of Foreign Study	6
2.4 โครงสร้างฐานข้อมูลXMLของพจนานุกรมญี่ปุ่นไทย	7
2.5 โครงสร้างฐานข้อมูลXMLของพจนานุกรมไทยญี่ปุ่น	9
บทที่3 การสร้างพจนานุกรมอิเล็คทรอนิกส์ ไทย-ญี่ปุ่น	11
3.1 ฐานข้อมูล พจนานุกรม ไทย-ญี่ปุ่น	11
3.2 ระบบงานพจนานุกรมอิเล็คทรอนิกส์	14
3.3 อุปกรณ์และโปรแกรมที่ใช้ในการสร้างพจนานุกรมอิเล็คทรอนิกส์	14
3.4 ผังงานการทำงานของระบบ	15
3.5 แผนผังบริบท (Context Diagram)	18
3.6 แผนผังกระแสข้อมูล (Data Flow Diagram)	19
บทที่4 การใช้งาน TNI e-Dictionary	20
4.1 แนะนำโปรแกรม	20
4.2 การค้นหาศัพท์	21
4.5 การเพิ่มศัพท์	25

4.6 การแก้ไขศัพท์ 27
4.7 การ Approve คำศัพท์ 28

บทที่5 การเพิ่มคำศัพท์ในพจนานุกรม 29
5.1 การเพิ่มศัพท์ให้พจนานุกรมญี่ปุ่น-ไทย 29
5.2 การเพิ่มคำศัพท์ให้พจนานุกรมไทย-ญี่ปุ่น 34

บทที่6 การทดลองใช้งาน 36
6.1 ผลการทดสอบทดลองใช้ระบบ 36
6.2 สิ่งที่ต้องแก้ไขหรือทำเพิ่มเติม 37

บทที่7 สรุปการวิจัย 38

บรรณานุกรม 40

ภาคผนวก 42
ตารางเทียบเสียง Romaji 43
Development of Thai-Japanese e-Dictionary 46
เอกสารประกอบ 51

สารบัญตาราง

ตาราง

1	ตาราง 1.1 แผนดำเนินการวิจัย	3
2	ตาราง 3.1 สัญลักษณ์ที่ใช้ในแผนผังบริบท	18
3	ตาราง 5.1 Excel ศัพท์ญี่ปุ่น-ไทย (JP2TH)	29
4	ตาราง 5.2 Excel ศัพท์ไทย-ญี่ปุ่น (TH2JP)	30

สถาบันเทคโนโลยีไทย-ญี่ปุ่น THAI - NICHI INSTITUTE OF TECHNOLOGY TNI

สารบัญรูป

รูป		
1	รายละเอียดโครงสร้างฐานข้อมูล XML ของพจนานุกรมญี่ปุ่น-ไทย	8
2	รายละเอียดโครงสร้างฐานข้อมูล XML ของพจนานุกรมไทย-ญี่ปุ่น	10
3	โครงสร้างความสัมพันธ์ของฐานข้อมูลคำศัพท์ญี่ปุ่น-ไทย	12
4	โครงสร้างความสัมพันธ์ของฐานข้อมูลคำศัพท์ไทย-ญี่ปุ่น	13
5	ผังระบบงานการค้นหาข้อมูลสำหรับผู้ใช้งานทั่วไป	15
6	แผนผังระบบงานในการป้อนคำศัพท์ใหม่สำหรับสมาชิก	16
7	ผังระบบงานการตรวจสอบและอนุมัติคำศัพท์ใหม่	17
8	แผนผังบริบทของระบบ TNI e-Dictionary	18
9	แผนผังกระแสข้อมูลระบบ TNI e-Dictionary	19
10	ตัวอย่างส่วนหนึ่งของ Main Menu	20
11	รูปแสดงการค้นหาศัพท์ด้วย Romaj	21
12	รูปแสดงศัพท์จากการค้นหาด้วย Romaji	21
13	รูปแสดงการค้นหาศัพท์ด้วย Kana	22
14	รูปแสดงศัพท์จากการค้นหาด้วย Kana	22
15	รูปแสดงการค้นหาศัพท์ด้วย Kanji	23
16	รูปแสดงศัพท์จากการค้นหาด้วย Kanji	23
17	รูปแสดงการค้นหาศัพท์ด้วยคำไทย	24
18	รูปแสดงศัพท์จากการค้นหาด้วยคำไทย	24
19	รูปแสดงหน้าจอการเพิ่มศัพท์	25
20	รูปแสดงขั้นตอนการเพิ่มศัพท์	26
21	รูปแสดงศัพท์ที่รอการเพิ่ม	26
22	รูปแสดงรายละเอียดศัพท์ที่จะแก้ไข	27
23	รูปแสดงหน้าจอศัพท์ที่รอการแก้ไข	28
24	ศัพท์ที่ค้นหามากสุดเรียงลำดับ	32
25	ศัพท์ที่ค้นไม่พบในระบบฐานข้อมูลพจนานุกรม	32
26	กราฟสถิติจำนวนผู้ใช้งาน TNI e-Dict	33
27	ประเภทของผู้ใช้งาน	34

กิตติกรรมประกาศ

พจนานุกรมอิเล็คทรอนิกส์นี้ ได้จัดทำขึ้นเนื่องจากปัจจุบันมีนักศึกษาชาวไทยศึกษาภาษาญี่ปุ่นมากขึ้น และมีชาวญี่ปุ่นศึกษาภาษาไทยมากขึ้น พจนานุกรมจึงมีความจำเป็นในการศึกษาเป็นอย่างมาก พจนานุกรมอิเล็คทรอนิกส์นี้สำเร็จลงได้ด้วยความร่วมมือกันหลายฝ่าย ทั้งจากไทยและจากญี่ปุ่น รวมถึงการช่วยเหลือด้านข้อมูลคำศัพท์ และโปรแกรมในเบื้องต้นจาก Prof.Dr.Minegishi Makoto จากมหาวิทยาลัย Tokyo University of Foreign Studies ที่ได้มอบข้อมูลคำศัพท์เพื่อให้ทุกฝ่ายได้ร่วมกันพัฒนาต่อไปส่งผลให้การทำโปรแกรมสำเร็จได้ด้วยดี ผู้ร่วมวิจัยทุกคนได้ซาบซึ้งและขอขอบพระคุณอย่างสูงมา ณ ที่นี้

ขอขอบคุณ สมาคมส่งเสริมเทคโนโลยี(ไทย-ญี่ปุ่น) ที่ได้กรุณาให้ข้อมูลเกี่ยวกับพจนานุกรมต่างๆ เพื่อเป็นประโยชน์แก่การทำพจนานุกรมอิเล็คทรอนิกส์นี้

ขอขอบคุณอาจารย์บัญชา บริคุต ที่ช่วยพัฒนาฐานข้อมูลพจนานุกรมโปรแกรมค้นหาคำศัพท์ และพัฒนาเวปไซต์ที่ให้บริการพจนานุกรมอิเล็คทรอนิกส์ออนไลน์ ให้กับผู้ใช้งานทั่วไปนอกสถาบัน

สุดท้ายขอขอบคุณผู้ร่วมวิจัยทุกท่าน ที่ได้เสียสละเวลา ระดมความคิดเห็น และเสนอแนะข้อมูลที่เป็นประโยชน์จนทำให้พจนานุกรมอิเล็คทรอนิกส์นี้เสร็จสมบูรณ์

ผู้จัดทำรายงานวิจัย

บทคัดย่อไทย

จุดมุ่งหมายของการวิจัยนี้คือ สร้างพจนานุกรมอิเล็คทรอนิกส์ ญี่ปุ่นไทย สำหรับนักศึกษาสถาบันเทคโนโลยีไทยญี่ปุ่นที่ ต้องเรียนภาษาญี่ปุ่นและศัพท์เทคนิค ฐานข้อมูลของพจนานุกรมเริ่มแรกประกอบด้วยคำศัพท์จำนวนประมาณ 50,000 คำจากการแปลงข้อมูล XML ของมหาวิทยาลัยโตเกียวภาษาต่างประเทศ(TUFS)บวกด้วยคำศัพท์จากพจนานุกรมศัพท์ช่าง และ ศัพท์ท่องเที่ยวของสมาคมส่งเสริมเทคโนโลยี (ไทย-ญี่ปุ่น) โครงสร้างของฐานข้อมูลพจนานุกรมประกอบด้วยหลายฟิลด์ได้แก่ โรมาจิ คานะ คันจิ คำแปลภาษาไทยคำแปลภาษาอังกฤษ คำเหมือน หมวดคำ ที่มาของคำศัพท์ ชนิดคำ และตัวอย่างประโยค ได้ติดตั้งพจนานุกรมอิเล็คทรอนิกส์ในเวปไซต์ ของสถาบัน เพื่อใช้ในการทดสอบและบริการให้กับนักศึกษา สมาชิกสมาคมสสท.และบุคคลทั่วไป การค้นหาคำแปลของศัพท์ญี่ปุ่นสามารถทำได้จากการป้อนคำศัพท์ในรูปของโรมาจิ หรือ คานะ หรือ คันจิให้กับระบบกรณีค้นหาคำแปลของคำไทยก็ให้ป้อนภาษาไทยเข้าที่ระบบก็จะได้คำแปลเป็นภาษาญี่ปุ่นออกมา การเพิ่มคำศัพท์ทั้งไทยญี่ปุ่นให้กับพจนานุกรมก็ทำโดยนักวิจัย บุคลากรและนักศึกษาร่วมกันเสนอคำศัพท์ใหม่ ซึ่งระบบมีฟังก์ชั่นให้แก้ไขและอนุมัติคำศัพท์ใหม่ และการเพิ่มศัพท์เข้าในระบบได้ศัพท์ใหม่ด้านวิศวกรรมศาสตร์ เทคโนโลยีสารสนเทศ และบริหารธุรกิจ จำนวนกว่า 3,000 คำถูกเพิ่มเข้าไปในฐานข้อมูลพจนานุกรมโดยผ่านทางเวปเซอร์วิสนี้

Keywords--- พจนานุกรมไทย-ญี่ปุ่น พจนานุกรมญี่ปุ่น-ไทย พจนานุกรมอิเล็คทรอนิกส์ พจนานุกรมออนไลน์

บทคัดย่ออังกฤษ

Development of Thai-Japanese e-Dictionary

Abstract--The purpose of this project is to create a Thai-Japanese dictionary for students at Thai-Nichi Institute of Technology (TNI) who are studying Japanese language and technical Japanese words. The dictionary's database initially contains about 50,000 Japanese words which were created from Tokyo University of Foreign Studies (TUFS)'s XML database of Japanese-Thai dictionary, the Technology Promotion Association (Thai-Japan) (TPA)'s Japanese-Thai Technical words dictionary and Thai – Japanese Tourism dictionary. The structure of the database is divided to multiple fields including Romaji, Kana, Kanji, Thai, English, synonym, sample sentence, parts, field, etymology and style. Our e-dictionary uses the TNI's website for testing the system while at the same time provide service to the students and general Thai and Japanese users. Accessing of Japanese word can be done by enter Romaji, Kana or Kanji. Accessing of Thai words can be done by enter Thai characters. The development of new Thai and Japanese words is done by a group of researchers, staffs and students who involve in proposing, editing and approving of new words. More than 3,000 technical words in Engineering, Information Technology and Business Administration field are being added to the database via TNI web service.

Keywords---Thai-Japanese dictionary, Japanese-Thai dictionary, e-Dictionary, online dictionary

บทที่ 1
บทนำ

1.1 ความเป็นมาของโครงการวิจัย

สถาบันเทคโนโลยีไทย – ญี่ปุ่น มีการเรียนการสอนภาษาญี่ปุ่นสำหรับนักศึกษาทุกคณะและทุกชั้นปี เนื่องจากได้บรรจุวิชาภาษาญี่ปุ่นลงในหลักสูตรของทุกคณะ การสอนมีตั้งแต่ระดับพื้นฐาน จนถึงระดับสูงที่มีการประยุกต์ในสาขาวิชาต่าง ๆเช่น วิศวกรรมศาสตร์ เทคโนโลยีสารสนเทศ และ การบริหารธุรกิจ สถาบันได้มีการผลิตตำราเรียนภาษาญี่ปุ่น สำหรับนักศึกษาของสถาบัน เพื่อประกอบการเรียนสอน สถาบันมีความต้องการพจนานุกรมไทย – ญี่ปุ่น และญี่ปุ่น – ไทย ให้นักศึกษาใช้งานประกอบการเรียนภาษาญี่ปุ่นทั้งระดับพื้นฐาน และ ระดับประยุกต์ในสาขาวิชาต่าง ๆ สถาบันจึงต้องการพัฒนาพจนานุกรมศัพท์ด้านวิทยาศาสตร์ และ เทคโนโลยีสำหรับใช้ในการเรียนการสอนสาขาวิชาด้านวิศวกรรมศาสตร์ และเทคโนโลยีสารสนเทศ

โรงเรียนภาษาสมาคมส่งเสริมเทคโนโลยี(ไทย-ญี่ปุ่น)ตั้งอยู่ที่สุขุมวิท29 และมีสาขาหลายแห่งที่มีการสอนภาษาญี่ปุ่นสำหรับบุคคลทั่วไป มีการสอนตั้งแต่ระดับเริ่มต้น ระดับกลางจนถึงระดับขั้นสูง มีบุคคลทั่วไปและนักศึกษามาเรียนที่โรงเรียนจำนวนมาก

สมาคมได้จัดให้มีการพิมพ์พจนานุกรมญี่ปุ่น – ไทยหลายฉบับเพื่อให้นักศึกษาประกอบการเรียนรู้ภาษาและวัฒนธรรมญี่ปุ่น รวมทั้งจัดพิมพ์พจนานุกรมศัพท์เทคนิคญี่ปุ่นไทย[7] เพื่อการเรียนรู้ทางสายช่างเทคนิค พจนานุกรมนี้เริ่มพิมพ์ในปี 2522 มีคำศัพท์ประมาณ 3,000 คำ เหมาะสำหรับสมาชิกของสมาคมที่ต้องการค้นหาความหมายของศัพท์ช่างที่เกี่ยวของกับการทำงานในโรงงาน สมาคมส่งเสริมเทคโนโลยีไทย-ญี่ปุ่นก็มีความต้องการให้มีการพัฒนาพจนานุกรมไทย - ญี่ปุ่นด้านศัพท์ทางวิทยาศาสตร์และเทคโนโลยี เพิ่มเติมจากพจนานุกรมที่มีอยู่ นอกจากนั้นสมาคมยังพิมพ์หนังสือแปลจากภาษาญี่ปุ่นอีกจำนวนมาก จึงมีความจำเป็นในการสร้างพจนานุกรมญี่ปุ่นไทยร่วมกันใช้งานสำหรับผู้แปลเอกสารและตำราของ สสท. ด้วย

ภาควิชาภาษาไทย สถาบันวิจัยเอเซีย และ อาฟริกัน มหาวิทยาลัยโตเกียวการศึกษาต่างประเทศ (University of Tokyo Foreign Studies) เป็นแหล่งศึกษาวิจัยภาษาชั้นนำของประเทศญี่ปุ่น มีนักศึกษาเรียนภาษาไทยอยู่จำนวนหนึ่ง และเปิดสอนภาษาไทยให้กับประชาชนทั่วไปด้วย ในโครงการวิจัย Corpus-based Linguistics and language Education ได้ทดลองสร้างพจนานุกรมอิเล็คทรอนิกส์ ไทย – ญี่ปุ่นและญี่ปุ่น – ไทย[6] ขึ้นโดยมีคำศัพท์จำนวนประมาณ 30,000 คำ เพื่อใช้ในการศึกษาวิจัยภาษาไทย และบริการให้กับคนญี่ปุ่น ที่ต้องการเรียนภาษาไทย สถาบันวิจัยมีความต้องการ ที่จะพัฒนาและปรับปรุงพจนานุกรมอิเล็กทรอนิกส์ให้มีคำศัพท์เพิ่มมากขึ้นและสามารถให้บริการในรูปของเวปไซต์เพื่อให้สามารถใช้งานได้ในวงกว้างและยังต้องเพิ่มศัพท์ทางเทคนิคและวิทยาศาสตร์ที่ยังขาดอยู่ จึงได้ติดต่อขอทำการวิจัยร่วมกับสถาบันเทคโนโลยีไทย – ญี่ปุ่น ในการพัฒนาพจนานุกรมอิเล็คทรอนิกส์ไทย – ญี่ปุ่นนี้เพื่อให้สามารถบริการให้ทั้งคนไทยที่อยากเรียนภาษาญี่ปุ่นและคนญี่ปุ่นที่ต้องการเรียนภาษาไทย ทั้งที่อยู่ในประเทศไทยหรือประเทศญี่ปุ่นก็สามารถใช้งานได้เท่าเทียมกัน

มีเวปไซต์ที่ให้บริการพจนานุกรมญี่ปุ่นไทยในปัจจุบันอยู่บ้างเช่น longdo.com[1] และ jtdic.com[2] แต่มีจำนวนศัพท์ยังไม่เพียงพอ กับการใช้งานของนักศึกษา มีศัพท์เทคนิคและวิทยาศาสตร์น้อย ยังมีพจนานุกรมญี่ปุ่นไทยศัพท์สาขาวิศวกรรมศาสตร์ของมหาวิทยาลัย Nagaoka University of Technology[11]

ซึ่งมีคำศัพท์อยู่ 10,000 คำให้บริการสำหรับนักศึกษาที่เรียนที่มหาวิทยาลัยเป็นหลักโดยพิมพ์เป็นหนังสือรวมศัพท์วิศวกรรมศาสตร์ พจนานุกรมนี้ก็ยังไม่พร้อมที่จะให้บริการกับคนทั่วไปเนื่องจากขาดศัพท์ทั่วไปสำหรับการเรียนรู้ภาษาญี่ปุ่น

1.2 วัตถุประสงค์ของโครงการวิจัย

1. เพื่อสร้างพจนานุกรมอิเล็คทรอนิกส์ไทย – ญี่ปุ่น และญี่ปุ่น – ไทย สำหรับใช้งานในสถาบันเทคโนโลยีไทย – ญี่ปุ่นและสมาคมส่งเสริมเทคโนโลยี(ไทย-ญี่ปุ่น)
2. ให้บริการพจนานุกรมอิเล็คทรอนิกส์ในสถาบันและสมาคมส่งเสริมเทคโนโลยีไทย-ญี่ปุ่น และมหาวิทยาลัยในประเทศญี่ปุ่น

1.3 ขอบเขตของโครงการวิจัย

คำศัพท์ที่บรรจุในโปรแกรม e-Dictionary ประกอบด้วย

1. การสร้างฐานข้อมูลของคำศัพท์ ญี่ปุ่น-ไทย โดยใช้ข้อมูลจาก

1.1 ฐานข้อมูลศัพท์ไทย-ญี่ปุ่น และญี่ปุ่น-ไทย จากมหาวิทยาลัยTokyo University of Foreign Studies

1.2 รวบรวมศัพท์ทั้งหมดในตำราเรียนภาษาญี่ปุ่นของสถาบัน

1.3 หนังสือพจนานุกรมจาก ส.ส.ท. ที่ได้รับอนุญาต เช่น ปทานุกรมศัพท์ช่าง และพจนานุกรมท่องเที่ยวไทย-ญี่ปุ่น

1.4 ศัพท์เฉพาะทางด้านเทคนิค เช่น วิศวกรรมศาสตร์ เทคโนโลยีสารสนเทศ และบริหารธุรกิจ

2. เพิ่มศัพท์ทางด้านเทคนิคเข้าไปในฐานข้อมูล ประมาณ 3,000 คำ ประกอบด้วยศัพท์ทางด้าน

- วิศวกรรมศาสตร์
- เทคโนโลยีสารสนเทศ
- บริหารธุรกิจ

3. ให้บริการ e-Dictionary จากฐานข้อมูลคำศัพท์ โดยสร้างเป็นเวบไซต์ TNI e-Dictionary ให้บริการแก่นักศึกษาสถาบัน สมาชิก ส.ส.ท. และบุคคลทั่วไปทั้งในประเทศไทยและญี่ปุ่น

1.4 ขั้นตอนและแผนการดำเนินการวิจัย

วิธีการวิจัย

1. ศึกษาและกำหนดหารูปแบบของฐานข้อมูลพจนานุกรมอิเล็คทรอนิกส์ ไทย - ญี่ปุ่นที่เหมาะสมสำหรับการสร้างฐานคำศัพท์
2. สร้างฐานข้อมูลพจนานุกรมอิเลคทรอนิกส์ และแปลงข้อมูลเดิมของพจนานุกรมมหาวิทยาลัยโตเกียวเป็นฐานข้อมูลใหม่
3. การสร้างศัพท์ด้านวิทยาศาสตร์และเทคโนโลยีเข้าไปในฐานข้อมูลพจนานุกรม
4. การทดสอบการใช้งานพจนานุกรมกับนักศึกษาสถาบันเทคโนโลยีไทย – ญี่ปุ่น

แผนดำเนินการงานวิจัย

กิจกรรม	เมย - มิย 52	กค- กย 52	ตค- ธค 52	มค- มีค 53	เมย- มิย 53	กค- กย 53	ตค- ธค 53	มค- มีค 54	เมย- มิย 54
1.ประชุมวางแผนดำเนินงาน	x								
2.รับข้อมูลจาก TUFS	x	x							
3.กำหนดรูปแบบฐานข้อมูล		x	x						
4.Data convertion สร้างฐานข้อมูล			x	x					
5.สร้างเวปและเปิดให้บริการ				x	x	x	x	x	
6.สร้างศัพท์ทางเทคนิคเพิ่มเติม				x	x	x	x		
7.ทดสอบและประเมินผลการใช้งาน						x	x	x	x

ตาราง 1.1 แผนการดำเนินงานวิจัย

1.5 ประโยชน์ที่จะได้รับ

1. พจนานุกรมอิเล็คทรอนิกส์ไทย – ญี่ปุ่น และ ญี่ปุ่น – ไทย ประมาณ 50,000 คำให้บริการอยู่ในเวปของสถาบัน สามารถใช้งานในสถาบันเทคโนโลยีไทย – ญี่ปุ่นและสมาคมส่งเสริมเทคโนโลยีไทย(ไทย-ญี่ปุ่น)
2. สร้างกลไกและคณะบุคคลที่สามารถเพิ่มจำนวนศัพท์ให้กับพจนานุกรมอิเล็คทรอนิกส์ในอนาคต

บทที่ 2
พจนานุกรมอิเล็คทรอนิกส์
e-Dictionary

2.1 แนวคิดในการวิจัย

พจนานุกรมถือเป็นสิ่งสำคัญในการเรียนภาษา โดยเฉพาะการเรียนภาษาต่างประเทศ ซึ่งจุดเด่นของสถาบันเทคโนโลยีไทย-ญี่ปุ่นอยู่ที่การมีการเรียนการสอนด้านภาษาญี่ปุ่น ดังนั้นโครงการวิจัย e-Dictionary จึงเกิดขึ้นเพื่อรองรับการเรียนการสอนและเพื่อช่วยเหลือนักศึกษาให้มีเครื่องมือที่ดีในการค้นหาคำศัพท์ญี่ปุ่น

ในปัจจุบัน มีพจนานุกรมมากมายในท้องตลาด ทั้งพจนานุกรมรูปเล่มและพจนานุกรมอิเล็กทรอนิกส์ในเครื่อง Talking Dic แต่พจนานุกรมจำนวนมาก มีความจำเพาะในการค้นหาศัพท์ เช่น พจนานุกรมญี่ปุ่น-ไทยใช้ในการค้นหาคำแปลของศัพท์ญี่ปุ่นพจนานุกรมไทย – ญี่ปุ่นเพื่อค้นหาการแปลคำญี่ปุ่นของคำไทย พจนานุกรมคันจิ ใช้ในการค้นหาคำอ่านและความหมายของตัวอักษรญี่ปุ่นเป็นต้น ทำให้นักศึกษาจำเป็นต้องใช้พจนานุกรมหลายเล่ม แต่โครงการวิจัย e-Dictionary มุ่งตอบสนองความต้องการของนักศึกษาที่สามารถใช้พจนานุกรม TNI e-Dictionary เพียงอย่างเดียวก็สามารถครอบคลุมได้ครบความต้องการเนื่องจากมีคำศัพท์ทุกรูปแบบทั้งญี่ปุ่น-ไทย ไทย-ญี่ปุน มีคำศัพท์หลายด้านทั้งศัพท์ทั่วไปและศัพท์เทคนิควิทยาศาสตร์

TNI e-Dictionary หรือ พจนานุกรมอิเล็คทรอนิกส์ TNI เป็นพจนานุกรมสื่ออิเล็คทรอนิกส์ ไทย – ญี่ปุ่น ที่ผู้ใช้สามารถค้นหาข้อมูลคำศัพท์ทั้งภาษาไทยและภาษาญี่ปุ่นที่ต้องการ ผ่านเครือข่ายอินเตอร์เนตโดยไม่เสียค่าใช้จ่าย นักศึกษาของสถาบัน และนักเรียนโรงเรียนภาษาของสมาคมส่งเสริมเทคโนโลยี (ไทย-ญี่ปุ่น) สามารถใช้ประกอบการเรียนภาษาญี่ปุ่น โดยการค้นหาคำศัพท์ด้วย Romaji หรือ ฮิรางานะ เพื่อทราบคำแปลเป็นภาษาไทยและอังกฤษ

พจนานุกรมอิเล็คทรอนิกส์ TNI จะมีคำศัพท์ทั่วไปในทุกสาขา รวมคำศัพท์ประมาณ 50,000 คำ ทั้ง ไทย – ญี่ปุ่น และญี่ปุ่น – ไทย ยังมีคำศัพท์เฉพาะทางเทคนิคในสาขาต่างๆ เช่น วิศวกรรมศาสตร์ เทคโนโลยีสารสนเทศ การบริหารธุรกิจ เพิ่มเติมเข้าไปในฐานข้อมูลคำศัพท์ ซึ่งถือเป็นจุดเด่นของโครงการวิจัยนี้ ซึ่งพบได้ยากในพจนานุกรมออนไลน์ที่มีอยู่ในปัจจุบัน ซึ่งถ้าพจนานุกรมออนไลน์ทั่วไปจะไม่มีศัพท์เทคนิคหากเป็นพจนานุกรมออนไลน์ด้านศัพท์เฉพาะทางก็มีเฉพาะศัพท์เฉพาะทางเท่านั้น

พจนานุกรมอิเล็คทรอนิกส์ TNI จะเปิดให้สมาชิกของสมาคม ส.ส.ท. และบุคคลทั่วไปทั้งไทยและญี่ปุ่นได้ใช้งาน ยังใช้เป็นเวทีที่เปิดโอกาสให้สมาชิกและผู้ใช้งานช่วยกันพัฒนาคลังข้อมูลคำศัพท์ โดยการเสนอแนะคำศัพท์ใหม่ ที่ไม่มีในพจนนุกรม การแปลคำศัพท์ที่เหมาะสม หรือเสนอแนะคำแปลอื่นๆ เพิ่มเติมจากในคลังข้อมูลทำให้มีการพัฒนาฐานข้อมูลคำศัพท์เพิ่มเติม โดยผู้ใช้สามารถมีส่วนร่วมด้วย

2.2 e-Dictionary ที่มีการให้บริการ

ในปัจจุบันมีหลายเวปไซต์ที่ให้บริการพจนานุกรมญี่ปุ่น-ไทยและ ญี่ปุ่น-อังกฤษ เช่น

พจนานุกรมไทย-ญี่ปุ่น

1.Longdo Dictionary ที่เวปไซต์ http://www.longdo.com

เป็นพจนานุกรมที่สามารถค้นหาศัพท์ได้จาก ภาษาอังกฤษ ญี่ปุ่น ไทย เพื่อให้ได้คำแปลที่มาจากพจนานุกรมต่างๆ ที่มีข้อมูลเชื่อมต่อกันด้วย เช่น HOPE Dictionary ,CMU English Pronouncing Dictionary , Oxford Advanced Learners Dictionary (Pronunciation guide only) , Tanaka JP-EN Corpus (ตัวอย่างประโยค) และ Foreign Dictionary เป็นต้น คำศัพท์บางคำไม่สามารถพบคำแปลเนื่องจากยังไม่มีการเพิ่มเข้าไปในฐานข้อมูลและผู้ใช้สามารถเพิ่มคำนั้นลงไปได้ บางครั้งผู้ใช้จึงไม่สะดวกนักเนื่องจากไม่สามารถคาดการณ์ได้ว่าศัพท์ที่กำลังค้นหาอยู่นั้นจะพบความหมายหรือไม่ และจากการค้นหาศัพท์เทคนิคด้านเทคโนโลยีสารสนเทศ เช่น nishinsu ที่แปลว่า เลขฐาน 2 ไม่พบคำแปลแต่อย่างใด

2.Japanese Thai Dictionary ที่เวปไซต์ http://www.jtdic.com

เป็นพจนานุกรมที่สามารถค้นหาศัพท์ได้จาก ภาษาไทย ฮิรางานะ คะตาคานะ รวมถึงการค้นจาก โรมาจิ เพื่ออำนวยความสะดวกแก่ผู้ที่ไม่สามารถพิมพ์ภาษาญี่ปุ่นได้ เนื่องจากการค้นหาศัพท์ญี่ปุ่น – ไทย นั้นในฐานข้อมูลมีน้อยเกินไป จึงได้ใช้ข้อมูลฐานเปิดของมหาวิทยาลัยโมนาช (ออสเตรเลีย) ซึ่งอ้างว่ามีฐานคำศัพท์สูงถึง 120,000 คำ ทำให้การค้นศัพท์มีประสิทธิภาพยิ่งขึ้น ทั้งนี้ได้ลองค้นหาศัพท์เทคนิคด้านเทคโนโลยีสารสนเทศ เช่น nishinsu ที่แปลว่า เลขฐาน 2 ไม่พบคำแปลแต่อย่างใด

พจนานุกรมญี่ปุ่น-อังกฤษ

1.Denshi Jisho ที่เวปไซต์ http://jisho.org

เป็นพจนานุกรมที่สามารถค้นหาศัพท์ได้จาก ภาษาอังกฤษ และญี่ปุ่น โรมันจิ ฮิรากานะ คะตาคานะ โดยสามารถหาคำแปลเป็นอักษรคันจิ และสามารถตรวจสอบความหมายของคันจิทั้งแบบเดี่ยวๆ และที่ประกอบเป็นคำได้ และสามารถดาวน์โหลดโปรแกรมนี้ไปใช้งานใน iPhone ได้อีกด้วย ทั้งนี้ได้ลองค้นหาศัพท์เทคนิคด้านเทคโนโลยีสารสนเทศ เช่น nishinsu ที่แปลว่า เลขฐาน 2 ไม่พบคำแปลแต่อย่างใด

2.Hiragana Megane (How to Read Japanese) ที่เวปไซต์ http://www.hiragana.jp

เป็นการบริการการแปลภาษาญี่ปุ่นเป็นภาษาอังกฤษ โดยใส่ URL ของข้อมูลที่มีภาษาญี่ปุ่นที่ต้องการแปล แล้วระบบจะบริการแปลออกมาเป็นภาษาอังกฤษ เป็นการอำนวยความสะดวกในการแปลประโยคยาวที่มีความซับซ้อนได้ ซึ่งเน้นการแปลภาษาของประโยคมากกว่าการหาความหมายของศัพท์หรือพจนานุกรม

3.Kanji lookup ที่เวปไซต์ www.jedict.com

เป็นพจนานุกรมที่ใช้หาความหมายของศัพท์จาก kanji ผู้ใช้สามารถหาคำแปลของอักษร Kanji ของคำศัพท์ ซึ่งไม่สามารถหาคำแปลจากคำศัพท์แบบอื่นๆที่มิใช่ตัวอักษรคันจิได้ ทำให้การใช้งานไม่ค่อยสะดวกเท่าไรนัก

พจนานุกรมไทย-อังกฤษ

1.Lexitron พจนานุกรมอิเล็คทรอนิกส์ ไทย-อังกฤษ ที่เวปไซต์ lexitron.nectec.co.th

เป็นพจนานุกรมสำหรับค้นหาศัพท์ทางวิทยาศาสตร์ภาษาไทย และอังกฤษพัฒนาโดยศูนย์วิจัยอิเล็กทรอนิกส์แห่งชาติ เหมาะสำหรับผู้ใช้ที่ไม่ต้องการรู้ภาษาญี่ปุ่นเนื่องจากมีเพียงภาษาไทยและอังกฤษเท่านั้น สำหรับผู้เรียนภาษาญี่ปุ่นนั้นไม่สามารถใช้งานได้สะดวกนัก

2.3 การศึกษาวิจัยเรื่อง e-Dictionary ของ Tokyo University of Foreign Study[6]

สถาบันวิจัยเอเซียและอาฟริกันแห่งมหาวิทยาลัย Tokyo University of Foreign Study ได้ทำการวิจัยและพัฒนา พจนานุกรมอิเล็คทรอนิกส์ภายใต้โครงการวิจัย Corpus-based Linguistics and language Education โครงการวิจัยนี้ได้รับการสนับสนุนจากกระทรวงศึกษาธิการและวิทยาศาสตร์ของญี่ปุ่น โดยเป็นโครงการที่เป็น Global COE programs โครงการหนึ่งเริ่มจากปี 2007 เป็นเวลา 5 ปี

โครงการพัฒนา Thai Dictionary เป็นโครงการย่อยหนึ่งและเป็นส่วนหนึ่งของการศึกษา Asian language ของมหาวิทยาลัย การพัฒนาพจนานุกรมภาษาไทยนับเป็นโครงการแรกของการวิจัยเนื่องจากข้อมูลพจนานุกรมไทยในรูปแบบอิเล็คทรอนิกส์สามารถหาได้ง่ายกว่าภาษาอาเซียนอื่นมาก พจนานุกรมไทย-อังกฤษ พจนานุกรมอังกฤษ-ไทยก็มีอยู่มาก แม้พจนานุกรมไทย-ญี่ปุ่นยังมีน้อยก็ตาม แต่ก็มีความต้องการใช้งานกันมากพอสมควร เนื่องจากมีผู้เรียนภาษาญี่ปุ่นเป็นจำนวนมาก ในประเทศไทย มีโรงเรียนและสถาบันการศึกษา ที่มีการสอนภาษาญี่ปุ่นที่ดีที่สุดในกลุ่มประเทศอาเซียน และนอกจากนั้นยังมีคนญี่ปุ่นจำนวนมากในไทยที่ต้องการเรียนภาษาไทยด้วย

การสร้างพจนานุกรมอิเล็คทรอนิกส์ไทย-ญี่ปุ่นของ TUFS เริ่มจากการแปลงข้อมูลของพจนานุกรมไทย-ญี่ปุ่นและญี่ปุ่น-ไทยที่แต่งโดยอาจารย์ Tomita Tekejiro[5] ซึ่งเป็นอาจารย์ที่มหาวิทยาลัยโอซาก้าแต่งใน ปี 2003 จากรูปของกระดาษเป็นอิเล็คทรอนิกส์ไฟล์ ซึ่งเรียกว่าเป็นการ digitize ให้อยู่ในรูปแบบ XML database ฐานข้อมูลศัพท์ไทย-ญี่ปุ่นนี้จะมีจำนวนคำศัพท์อยู่ประมาณ 20000 คำซึ่งเพียงพอกับการใช้งานของผู้เริ่มศึกษาภาษาญี่ปุ่นหรือไทย ฐานข้อมูลนี้สามารถเรียกใช้งานได้ที่เวปไซต์มี URL address [3]ดังนี้

http://cblle.tufs.ac.jp/dic/th/jath และ

http://cblle.tufs.ac.jp/dic/th/thja

ในขั้นต้นฐานข้อมูลนี้สามารถใช้ในการเรียนรู้ภาษาญี่ปุ่น ของนักศึกษาไทยได้ การค้นหาคำสามารถทำได้โดยการพิมพ์ป้อนคำศัพท์ญี่ปุ่นในรูป hiragana หรือ Kanji หรือ Romanได้ ต่อมาTUFS จะเริ่มขั้นตอนที่จะพัฒนาพจนานุกรมไทย-ญี่ปุ่น สำหรับให้คนญี่ปุ่นใช้ในการเรียนภาษาไทยโครงการต่อจากนั้นจะพัฒนาพจนานุกรมที่เป็น Machine readable dictionary(MDR) สำหรับการแปลภาษาไทยโดยอัตโนมัติต่อไป

ในขณะเดียวกันโครงการพัฒนาพจนานุกรมไทยญี่ปุ่นก็จะขยายให้มีจำนวนคำศัพท์มากขึ้นโดยการ digitize พจนานุกรมฉบับใหญ่ของอาจารย์ Tomita (1997)ซึ่งมีคำศัพท์ถึง 50,000 คำและมีข้อมูลเกี่ยวกับการออกเสียง ที่มาของคำ (Etymology) ซึ่งจะมีการอธิบายถึงวัฒนธรรมไทยอย่างละเอียดอีกด้วย พจนานุกรมนี้จะเป็นประโยชน์อย่างมากต่อการเรียนรู้ภาษาไทยของคนญี่ปุ่น

2.4 โครงสร้างฐานข้อมูลXMLของพจนานุกรมญี่ปุ่นไทย

TUFS ได้จัดฐานข้อมูลพจนานุกรมญี่ปุ่นไทยในรูปXML แต่ละคำศัพท์จะมี 13 field มีหมายเลขคำศัพท์ (dic item id no.) ซึ่งสามารถชี้ไปที่คำศัพท์ลูก (subdic item) และคำศัพท์หลาน (sub2dic item)ได้ แต่ละคำศัพท์มี13 ฟิลด์ได้แก่

1. คำอ่านโรมาจิ (Romaji)
2. ฮิรางานะ/คะตาคานะ (Hiragana / Katakana)
3. คันจิ (Kanji)
4. คำอ่านออกเสียง (phonetics)
5. ตามตัวอักษร (literally)
6. คำเหมือน (synonym)
7. ที่มาของคำ (etymology)
8. หมวดคำ (field)
9. รูปแบบภาษา (style)
10. ชนิดคำ (parts of speech)
11. สุภาษิต (proverb)
12. ความหมายภาษาไทย (Thai meaning)
13. อ้างอิง (see)

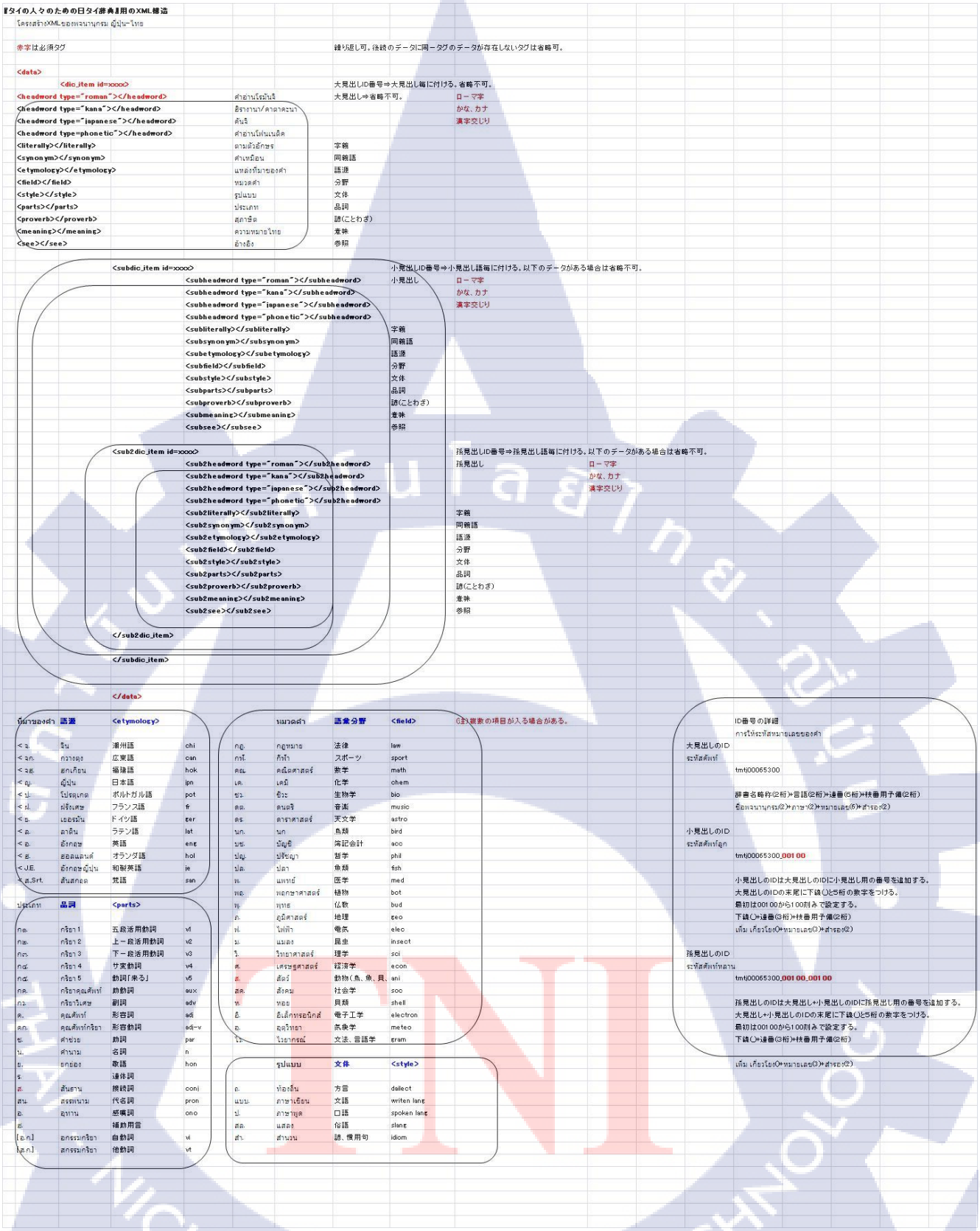
รูป 2.1 แสดงรายละเอียดของโครงสร้างฐานข้อมูลXMLของพจนานุกรมญี่ปุ่นไทยของ TUFS นี้ บางฟิลด์ยังไม่มีข้อมูลใส่ไว้ในฐานข้อมูลเช่น ฟิลด์ ตามตัวอักษร คำเหมือน สุภาษิต และอ้างอิง

ฟิลด์ที่มาของคำจะเป็นข้อมูล ของแหล่งที่มาของคำศัพท์นั้นเช่นมาจากคำภาษาอังกฤษ ลาติน ฝรั่งเศษ จีน เป็นต้น คำศัพท์ที่เป็นคะตาคานะ จะเป็นคำศัพท์ที่มาจากภาษาต่างประเทศทั้งหมด ดังนั้นข้อมูลในฟิลด์นี้จะบอกคำศัพท์เดิมที่มาของคำนั้น

ฟิลด์หมวดคำ เป็นการแบ่งคำศัพท์เป็นสาขาต่างๆ เช่น คณิตศาสตร์ เคมี ชีววิทยา ปรัชญา แพทย์ วิทยาศาสตร์ เป็นต้น มีการแบ่งหมวดคำไว้ละเอียด มีจำนวนถึง 21 หมวด

ฟิลด์รูปแบบภาษา เป็นข้อมูลที่บอกว่าคำศัพท์นั้นเป็น ภาษาท้องถิ่น ภาษาพูด ภาษาเขียน หรือแสลง เป็นต้น

ฟิลด์ชนิดของคำเป็นข้อมูลการแบ่งประเภทของคำ ได้แบ่งประเภทคำญี่ปุ่นออกเป็น 19 ประเภท ได้แก่ กริยา 7 ประเภท ก1 ก2 ก3 ก4 ก5 กค กว. คำคุณศัพท์ ค คก ช คำนาม คำยกย่อง สันธาน สรรพนาม และอุทานเป็นต้น



รูป 2.1 รายละเอียดโครงสร้างฐานข้อมูล XML ของพจนานุกรมญี่ปุ่นไทย

2.5 โครงสร้างฐานข้อมูลXMLของพจนานุกรมไทย-ญี่ปุ่น

ฐานข้อมูลพจนานุกรมไทย-ญี่ปุ่นเป็นฐานข้อมูลที่คนญี่ปุ่นใช้ในการเรียนภาษาไทย ค้นหาศัพท์ไทยโดยมีคำแปลญี่ปุ่นให้ความหมายของศัพท์นั้น แต่ละคำศัพท์จะมี 8 ฟิลด์ มีหมายเลขคำศัพท์(dic item id no.)ซึ่งสามารถชี้ไปที่คำศัพท์ลูก(subdic item) และคำศัพท์หลาน (sub2dic item)ได้

แต่ละคำศัพท์มี 8 ฟิลด์ได้แก่

1 คำศัพท์ไทย (Headword)
2 ตามตัวอักษร (literally)
3 คำเหมือน (synonym)
4 ที่มาของคำ (etymology)
5 หมวดคำ (field)
6 รูปแบบภาษา (style)
7 สุภาษิต (proverb)
8 ความหมายภาษาญี่ปุ่น (Japanese meaning)

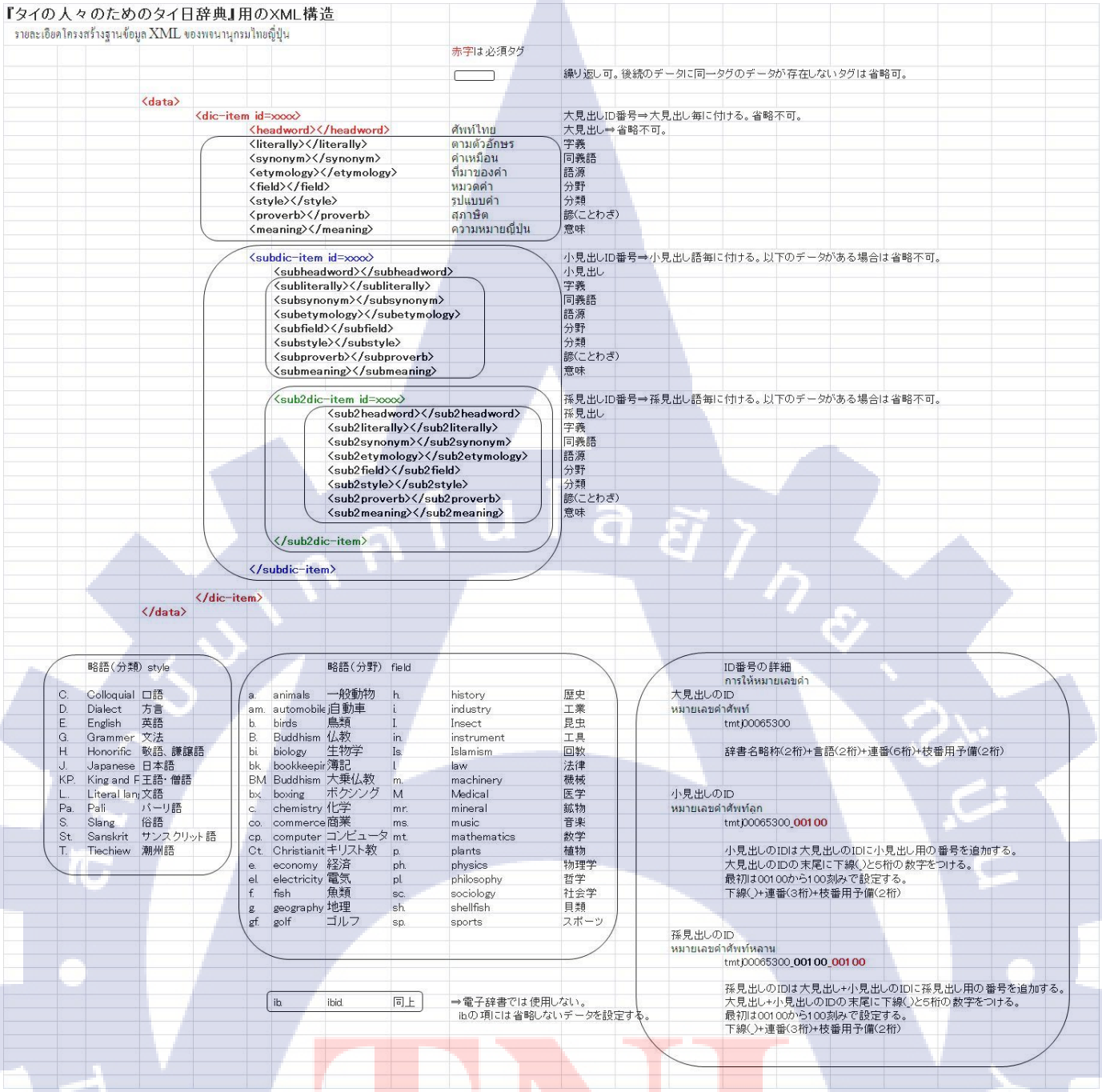
รูป 2.2 แสดงรายละเอียดของโครงสร้างฐานข้อมูลXMLของพจนานุกรมไทย-ญี่ปุ่นของ TUFS นี้ บางฟิลด์ยังไม่มีข้อมูลใส่ไว้ในฐานข้อมูลเช่น ฟิลด์ ตามตัวอักษร คำเหมือน และสุภาษิต

ฟิลด์ที่มาของคำจะเป็นข้อมูลของแหล่งที่มาของคำศัพท์นั้นเช่นมาจากคำภาษาอังกฤษ จีนแต้จิ๋ว บาลี สันสกฤต ญี่ปุ่นเป็นต้น

ฟิลด์หมวดคำ เป็นการแบ่งคำศัพท์เป็นสาขาต่างๆ เช่น สัตว์ รถยนต์ นก คณิตศาสตร์ เคมี ชีววิทยา ปรัชญา แพทย์ กีฬา เป็นต้น มีการแบ่งหมวดคำไว้ละเอียด มีจำนวนถึง 34 หมวด

ฟิลด์รูปแบบภาษา เป็นข้อมูลที่บอกว่าคำศัพท์นั้นเป็น ภาษาท้องถิ่น ภาษาพูด ภาษาเขียน ราชาศัพท์ ภาษาสุภาพ เป็นต้น

ไม่มีฟิลด์ชนิดของคำในฐานข้อมูลนี้



รูป 2.2 รายละเอียดโครงสร้างฐานข้อมูล XML ของพจนานุกรมไทยญี่ปุ่น

บทที่ 3
การสร้างพจนานุกรมอิเล็คทรอนิกส์ไทย-ญี่ปุ่น

3.1 โครงสร้างของ ฐานข้อมูลพจนานุกรม

3.1.1 โครงสร้างของฐานข้อมูลพจนานุกรม Japanese to Thai database ประกอบด้วย field ต่างๆ ดังนี้

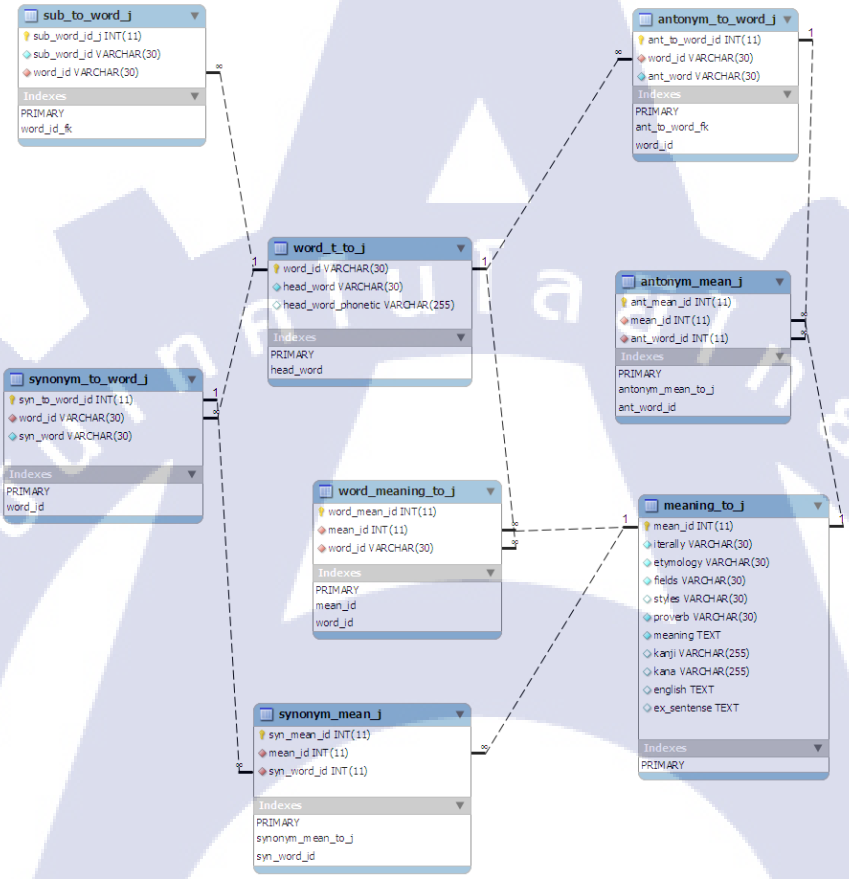
- mean_id เป็นหมายเลขรหัสประจำคำศัพท์เช่น tmtj00065300 หมายเลขคำศัพท์ t00065300_00100 หมายเลขคำศัพท์ลูก tmtj0006500_001_00100 หมายเลขคำศัพท์หลาน
- Head word-roman เป็นคำอ่านภาษาญี่ปุ่นในรูปของ Romaji
- iterally เป็นฟิลด์ตามตัวอักษร ยังไม่มีข้อมูลในฟิลด์นี้
- etymology เป็นที่มาของคำศัพท์ญี่ปุ่นที่เป็น Katagana จะมาจากภาษาต่างประเทศ เช่น อังกฤษ ฝรั่งเศส
- Field เป็นหมวดของคำ แบ่งเป็น

  Eng (วิศวกรรม) ,me (เครื่องกล) , ee (ไฟฟ้า อิเล็คทรอนิกส์) , pe (วิศวกรรมการผลิต อุตสาหการ), IT (เทคโนโลยีสารสนเทศ), aut (ยานยนต์), ba (บริหารธุรกิจ เศรฐศาสตร์), mat (คณิตศาสตร์), chm (เคมี) , sci (วิทยาศาสตร์) , bio (ชีววิทยา), Med (การแพทย์), ani (animal สัตว์), pln (plant ต้นไม้ ผลไม้), law (กฎหมาย การปกครอง), Rel (ศาสนา), Gen (ทั่วไป)
- Styles เป็นฟิลด์บอกรูปแบบของคำศัพท์ เช่นคำศัพท์เป็น ภาษาท้องถิ่น ภาษาพูด ภาษาเขียน คำแสลง สำนวน เป็นต้น
- Proverb เป็นฟิลด์ที่ระบุว่าคำศัพท์นี้มีสุภาษิตที่เกี่ยวข้อง
- Kana เป็นคำอ่านภาษาญี่ปุ่นในรูปของอักษร hiragana หรือ katagana เป็นคำอ่านหลักที่ใช้ออกเสียงที่ถูกต้องและเป็นหลักในการค้นหาคำ
- Kanji เป็นคำศัพท์ญี่ปุ่นซึ่งประกอบด้วย Kanji และ Hiragana หรือ Katagana เป็นคำศัพท์หลัก
- synonym คำเหมือนของคำศัพท์ญี่ปุ่นมีความหมายเหมือนกัน สามารถค้นหาในพจนานุกรมได้
- Parts คือ Parts of speech ชนิดของคำศัพท์ แบ่งเป็น

  n(คำนาม), v (คำกริยา), a (คำคุณศัพท์), ad (กริยาวิเศษ), conj (คำสันธาน)
- meaning –th เป็นคำแปลศัพท์เป็นภาษาไทย อาจมีหลายความหมาย

- meaning – en เป็นคำแปลศัพท์เป็นภาษาอังกฤษ
- Sample sentence เป็นตัวอย่างประโยคของคำศัพท์นั้น เป็นประโยคญี่ปุ่น พร้อมคำแปลไทย อาจมีหลายตัวอย่างในคำศัพท์หนึ่ง

รูป 3.1 แสดงตารางในฐานข้อมูลต่างๆและความสัมพันธ์



รูป 3.1 โครงสร้างความสัมพันธ์ของฐานข้อมูลคำศัพท์ญี่ปุ่น-ไทย

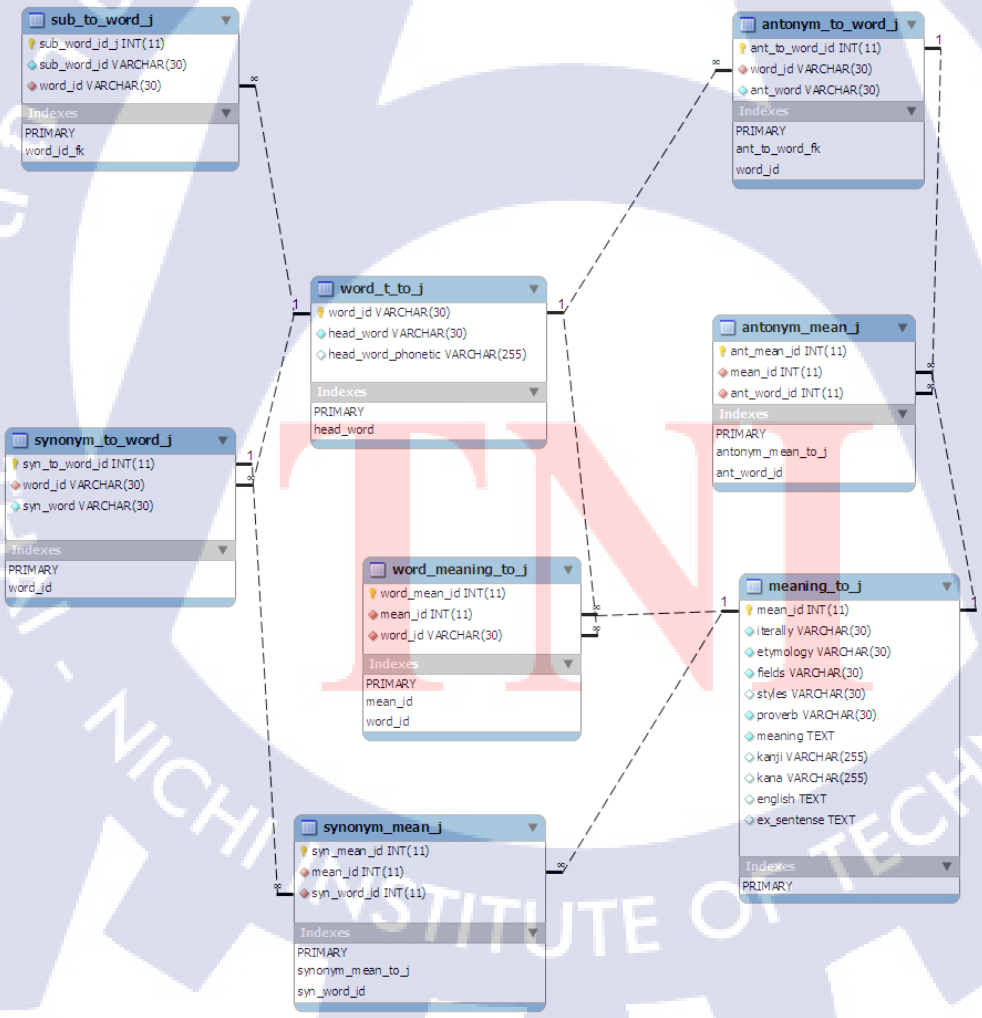
3.1.2 โครงสร้างของฐานข้อมูลพจนานุกรม Thai to Japanese database ประกอบด้วย field ต่างๆ ดังนี้

- mean_id เป็นหมายเลขรหัสประจำคำศัพท์ เช่น tmtj00065300 หมายเลขคำศัพท์ tmt00065300_00100 หมายเลขคำศัพท์ลูก tmtj0006500_001_00100 หมายเลขคำศัพท์หลาน
- Head word-thai เป็นคำศัพท์ภาษาไทย
- iterally เป็นฟิลด์ตามตัวอักษร ยังไม่มีข้อมูลในฟิลด์นี้
- etymology เป็นที่มาของคำศัพท์ญี่ปุ่นที่เป็น Katagana จะมาจากภาษาต่างประเทศ เช่น อังกฤษ ฝรั่งเศส
- Fields เป็นหมวดของคำ แบ่งเป็น Eng (วิศวกรรม) ,me (เครื่องกล) , ee (ไฟฟ้าอิเล็คทรอนิกส์) , pe (วิศวกรรมการผลิต อุตสาหการ), IT (เทคโนโลยีสารสนเทศ), aut (ยานยนต์), ba (บริหารธุรกิจ เศรฐศาสตร์), mat (คณิตศาสตร์), chm (เคมี), sci

(วิทยาศาสตร์), bio (ชีววิทยา), Med (การแพทย์), ani (animal สัตว์), pln (plant ต้นไม้ ผลไม้), law (กฎหมาย การปกครอง), Rel (ศาสนา), Gen (ทั่วไป)

- Styles เป็นฟิลด์บอกรูปแบบของคำศัพท์ เช่นคำศัพท์เป็น ภาษาท้องถิ่น ภาษาพูด ภาษาเขียน คำแสลง สำนวน เป็นต้น
- Proverb เป็นฟิลด์ที่ระบุว่าคำศัพท์นี้มีสุภาษิตที่เกี่ยวข้อง
- meaning –jp เป็นคำแปลศัพท์เป็นภาษาญี่ปุ่นซึ่งประกอบด้วย Kanji และ Hiragana หรือ Katagana อาจมีหลายความหมาย
- kana เป็นคำอ่านภาษาญี่ปุ่นในรูปของอักษร hiragana หรือ katagana เป็นคำอ่านหลักที่ใช้ออกเสียงที่ถูกต้องและเป็นหลักในการค้นหาคำ
- roman เป็นคำอ่านภาษาญี่ปุ่นในรูปของ Romaji
- meaning – en เป็นคำแปลศัพท์เป็นภาษาอังกฤษ
- Sample sentence เป็นตัวอย่างประโยคของคำศัพท์นั้น เป็นประโยคไทย พร้อมคำแปลญี่ปุ่น อาจมีหลาย ตัวอย่างในคำศัพท์หนึ่ง

รูป 3.2 แสดงตารางในฐานข้อมูลต่างๆและความสัมพันธ์



รูป 3.2 โครงสร้างความสัมพันธ์ของฐานข้อมูลคำศัพท์ ไทย-ญี่ปุ่น

3.2 ระบบงานพจนานุกรมอิเล็คทรอนิกส์

1) ระบบงานหลัก

TNI e-Dictionary เป็นระบบเวปเซอร์วิสบริการช่วยในการค้นหาคำศัพท์ภาษาไทยเป็นภาษาญี่ปุ่น และภาษาญี่ปุ่นเป็นภาษาไทยระดับเบื้องต้นพร้อมทั้งตัวอย่างการใช้งาน เหมาะสำหรับบุคคลทั่วไปและผู้ที่สนใจในภาษาญี่ปุ่น มีการแบ่งแยกระดับการใช้งานออกแบบไว้ 4 ระดับ คือ

- ผู้ดูและระบบ เป็นผู้จัดการปรับปรุงแก้ไขข้อมูลต่าง ๆ ของระบบ เช่น ข้อมูลสมาชิก เป็นต้น
- ผู้อนุมัติศัพท์เป็นผู้ใช้งานที่มีความชำนาญด้านภาษา และเป็นผู้ดูแลจัดการ ตรวจสอบข้อมูลคำศัพท์ที่สมาชิกได้เพิ่มเข้ามา เพื่อจัดการเพิ่มเข้าสู่ระบบฐานข้อมูลคำศัพท์ให้เป็นไปอย่างถูกต้อง
- สมาชิก เป็นผู้ใช้งานที่มีสิทธิ์ในการเพิ่มคำศัพท์และความหมาย และสามารถใช้งานระบบได้อย่างเต็มรูปแบบ
- ผู้ใช้งานทั่วไป เป็นผู้ใช้งานเวปทั่วไป ที่ไม่ได้เป็นสมาชิกกับทางเวบ ไม่สามารถเพิ่มและแก้ไขคำศัพท์ได้ แต่จะสามารถใช้งานได้เฉพาะส่วนของการค้นหาความหมายได้

2) ระบบงานอื่นๆ

นอกจากนี้ระบบงานยังประกอบระบบอื่นๆที่อำนวยความสะดวกในการใช้งานอีกได้แก่

- ระบบเก็บสถิติสมาชิก เป็นระบบการเก็บสถิติการเพิ่มศัพท์ใหม่ของสมาชิก
- ระบบแจ้งเตือนข้อผิดพลาดของคำศัพท์
- ระบบบันทึกสถิติการค้นหาศัพท์ของผู้ใช้งานทั่วไปและเก็บข้อมูลคำศัพท์ที่ค้นหาไม่พบไว้อัตโนมัติ

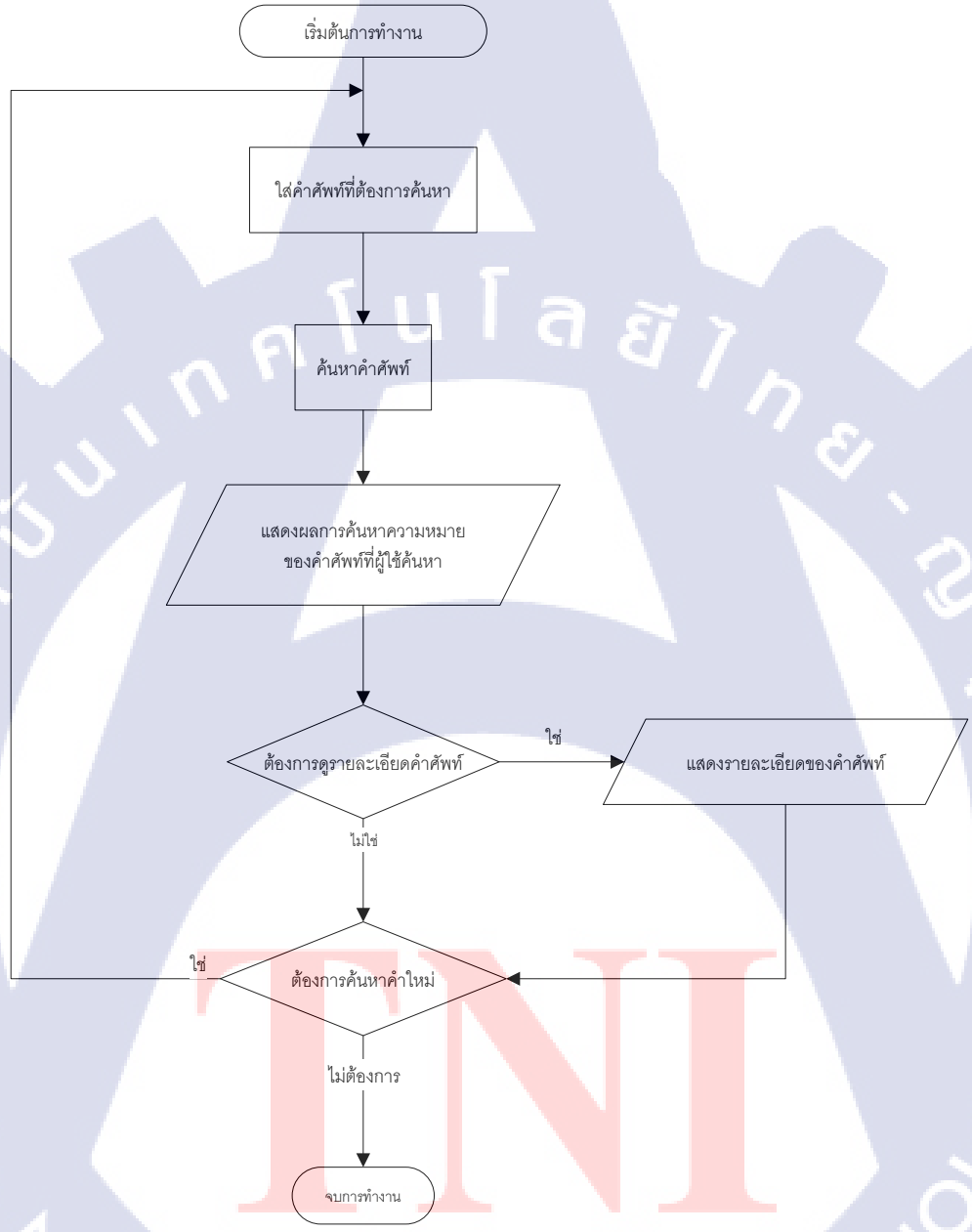
3.3 อุปกรณ์และโปรแกรมที่ใช้ในการสร้างพจนานุกรมอิเล็คทรอนิกส์

ในการวิจัยได้ใช้อุปกรณ์ในการสร้างโปรแกรมทั้งฮาร์ดแวร์และซอฟท์แวร์ดังนี้

1. ฮาร์ดแวร์(Hardware)
   - เครื่องคอมพิวเตอร์ส่วนบุคคล ใช้ในการพัฒนาโปรแกรมและสร้างระบบ
   - เครื่องคอมพิวเตอร์แม่ข่าย ใช้ในการเก็บฐานข้อมูลพจนานุกรม
2. ซอฟต์แวร์(Software) ที่ใช้ในการพัฒนาระบบได้แก่
   - โปรแกรมระบบปฏิบัติการวินโดว์เอ็กส์พี
   - โปรแกรม adobe photoshop cs3
   - โปรแกรม adobe dreamweaver cs3
   - โปรแกรม Editplus or Text editor
   - โปรแกรมภาษา PHP5.0
   - ระบบฐานข้อมูล Mysql Server 5.0
   - โปรแกรม Apache Web Server

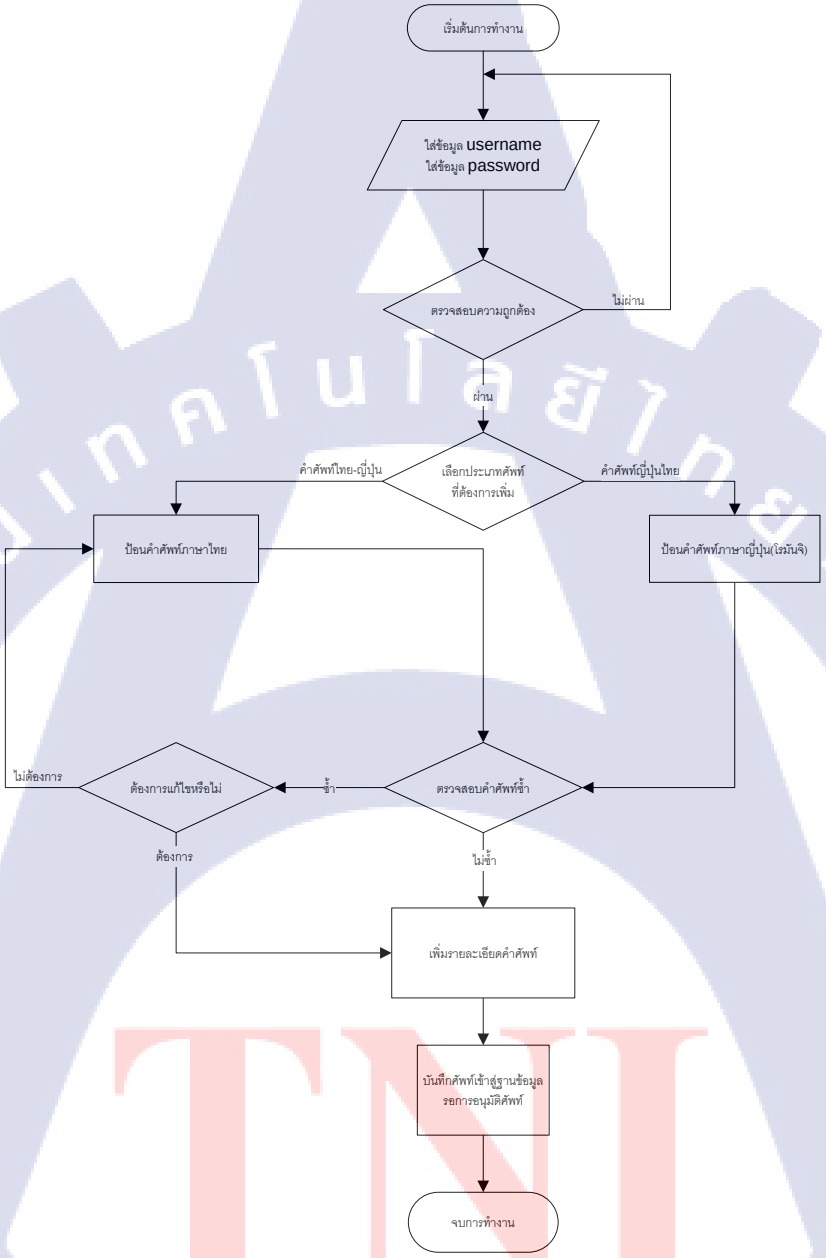
3.4 ผังงานการทำงานของระบบ

3.4.1 ผังระบบงานสำหรับการค้นหาคำศัพท์ของผู้ใช้งานทั่วไป จะบอกถึงการทำงานตั้งแต่การป้อนคำศัพท์ที่ต้องการค้นหา การค้นหาคำศัพท์จากฐานข้อมูล การแสดงผลข้อมูลคำศัพท์ให้ผู้ใช้ และการแสดงรายละเอียดเพิ่มเติมของคำศัพท์นั้น ดังรูปที่ 3.3



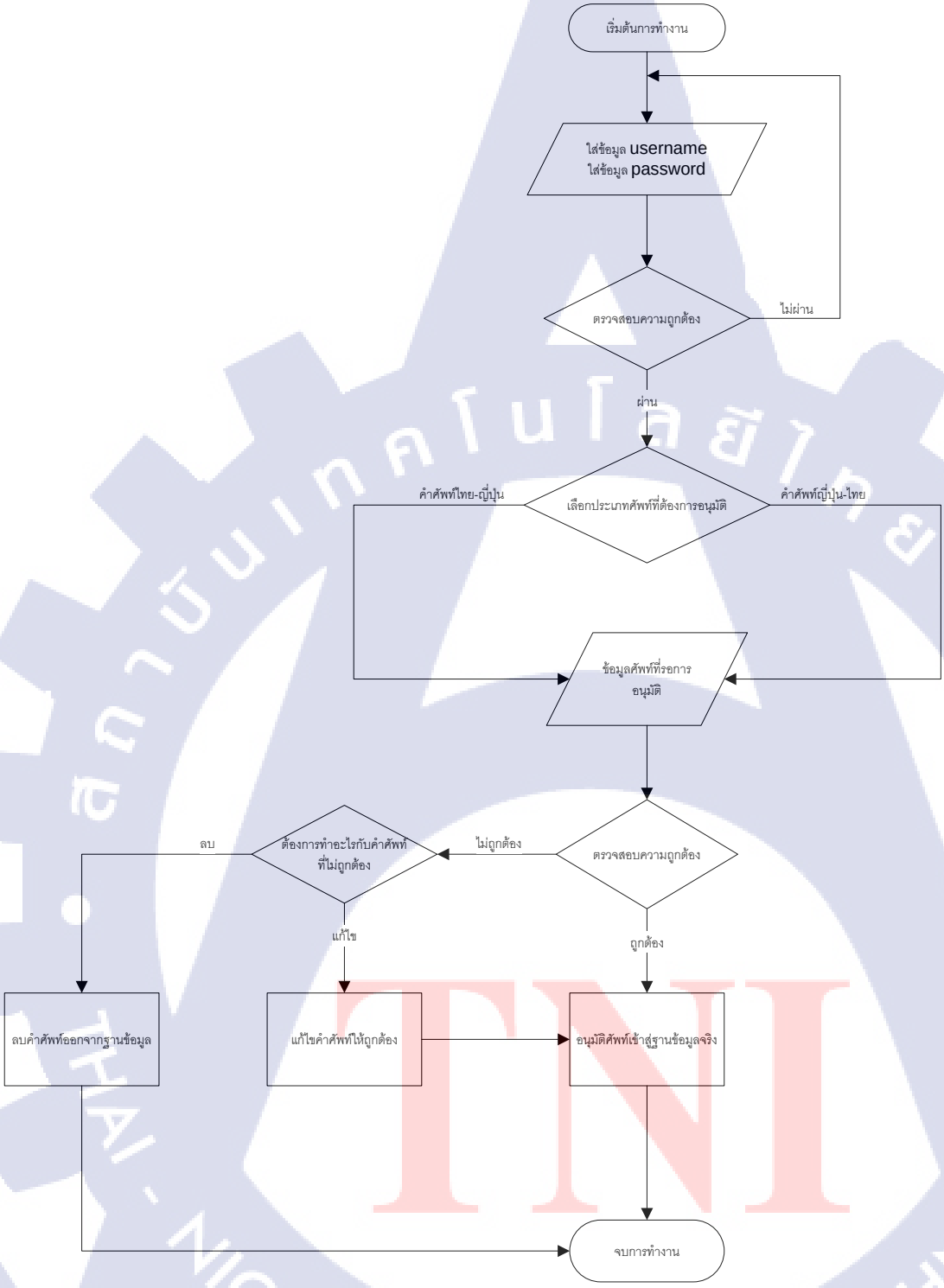
รูปที่ 3.3 ผังระบบงานการค้นหาข้อมูลสำหรับผู้ใช้งานทั่วไป

3.4.2 ผังระบบงานสำหรับสมาชิกในการป้อนคำศัพท์ใหม่ เพื่อนำคำศัพท์เข้าสู่ระบบฐานข้อมูล ตั้งแต่การป้อนและตรวจสอบ user name และ password การป้อนคำศัพท์ใหม่ การค้นหาคำศัพท์ว่าซ้ำกับในฐานข้อมูล การเพิ่มคำศัพท์ในฐานข้อมูลชั่วคราว ขั้นตอนต่างๆดังรูปที่ 3.4



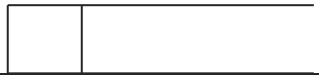
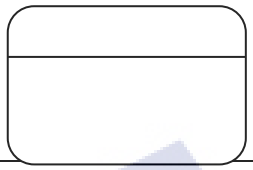
รูปที่ 3.4 แผนผังระบบงานในการป้อนคำศัพท์ใหม่สำหรับสมาชิก

3.4.3 ผังระบบงานสำหรับผู้ดูแลระบบ และผู้ได้รับหมอบหมายให้ทำการตรวจสอบและอนุมัติคำศัพท์ ดังรูปที่ 3.5



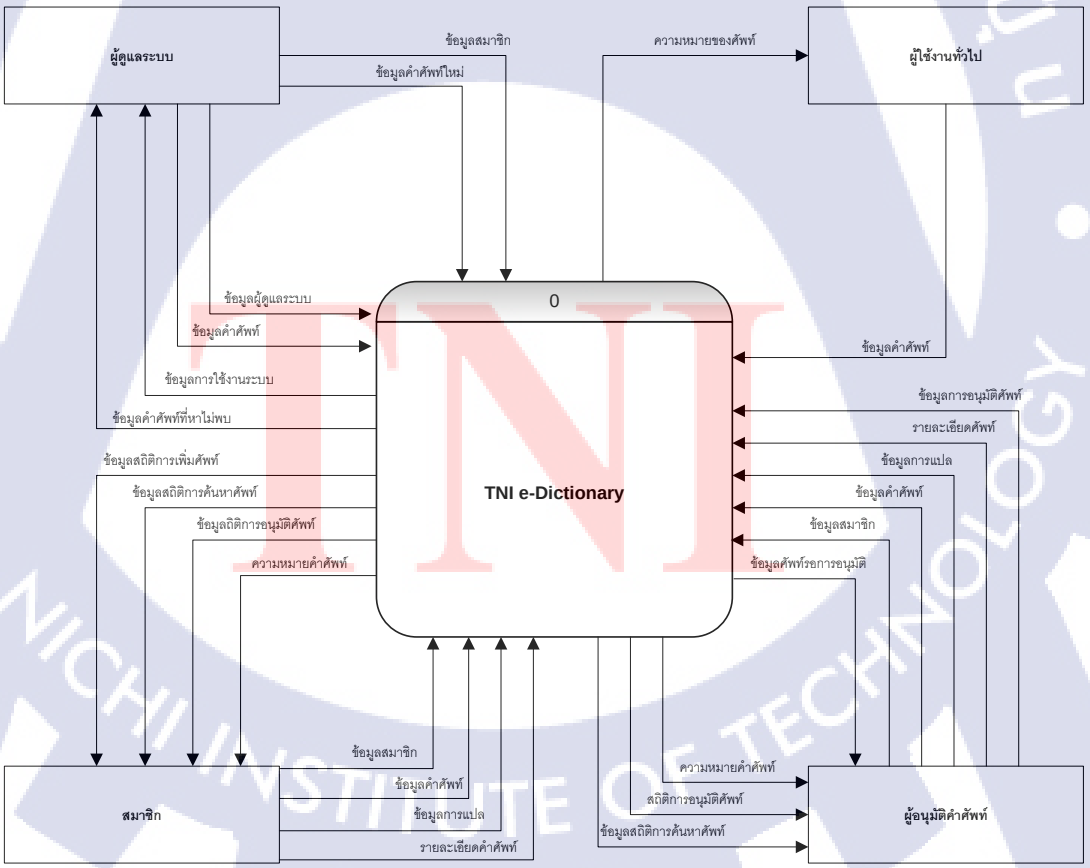
รูปที่ 3.5 ผังระบบงานการตรวจสอบและอนุมัติคำศัพท์ใหม่

3.5 แผนผังบริบท (Context Diagram)

สัญลักษณ์	ชื่อ	ความหมาย
	External Entity	ปัจจัยหรือสภาวะแวดล้อมภายนอกที่มีผลกระทบต่อระบบ
	Data Store	แหล่งข้อมูล
	Process	กระบวนการทำงาน
	Data Flow	เส้นทางการไหลของข้อมูล

ตาราง 3.1 สัญลักษณ์ที่ใช้ในแผนผังบริบท

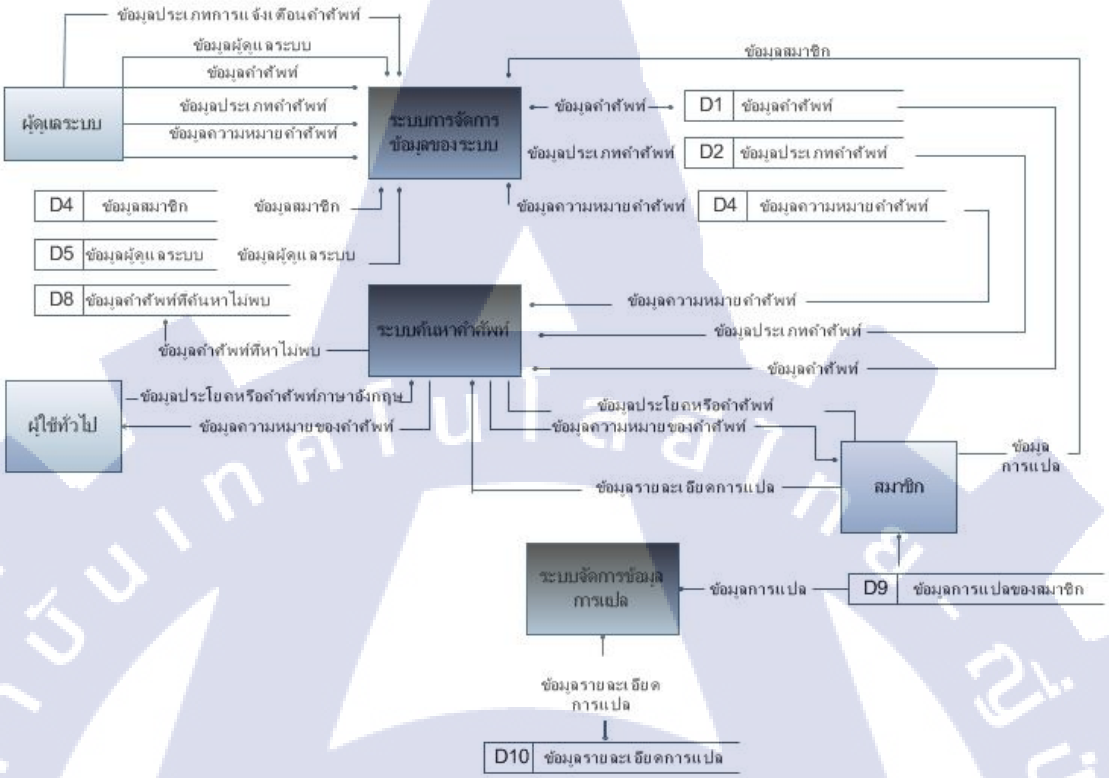
การวิเคราะห์ระบบเพื่อให้ทราบภาพรวมของระบบและความสัมพันธ์ของผู้ใช้ระบบทั่วไป สมาชิก ผู้อนุมัติศัพท์ และผู้ดูแลระบบที่ทำงานเกี่ยวข้อง กับระบบพจนานุกรมอิเล็คทรอนิกส์โดยใช้แผนผังบริบท ดังแสดงในรูป 3.6



รูป 3.6 แผนผังบริบทของระบบ TNI e-Dictionary

3.6 แผนผังกระแสข้อมูล (Data Flow Diagram)

แสดงข้อมูลที่วิ่งระหว่างผู้ใช้งานประเภทต่างๆ กับระบบงานต่างๆ ดังในรูป 3.7



รูปที่ 3.7 แผนผังกระแสข้อมูล ระบบ TNI e-Dictionary

บทที่ 4
การใช้งาน TNI e-Dictionary

แนะนำโปรแกรม

เป็นโปรแกรมการค้นหาศัพท์ที่สามารถค้นได้จาก คำไทย เสียงอ่าน (Romaji) อักษรคันจิ อักษร kana และสามารถค้นหาศัพท์เฉพาะทาง และศัพท์เทคนิค ด้านบริหารธุรกิจ ด้านวิศวกรรมศาสตร์ ด้านเทคโนโลยีสารสนเทศได้ เป็นต้น

4.1 ตัวอย่างส่วนหนึ่งของ Main Menu

การค้นหาศัพท์

การใช้โปรแกรมค้นหาคำศัพท์

1. การค้นหาคำศัพท์ โดยการใส่ศัพท์ที่ต้องการค้นหาลงในช่อง

1.1 การค้นหาศัพท์ด้วย Romaji

4.2 รูปแสดงการค้นหาศัพท์ด้วย Romaji

4.3 รูปแสดงศัพท์จากการค้นหาด้วย Romaji

1.2 การค้นหาศัพท์ด้วย Kana

4.4 รูปแสดงการค้นหาศัพท์ด้วย Kana

4.5 รูปแสดงศัพท์จากการค้นหาด้วย Kana

1.3 การค้นศัพท์ด้วย Kanji

4.6 รูปแสดงการค้นหาศัพท์ด้วย Kanji

4.7 รูปแสดงศัพท์จากการค้นหาด้วย Kanji

1.4 การค้นศัพท์ด้วยคำไทย

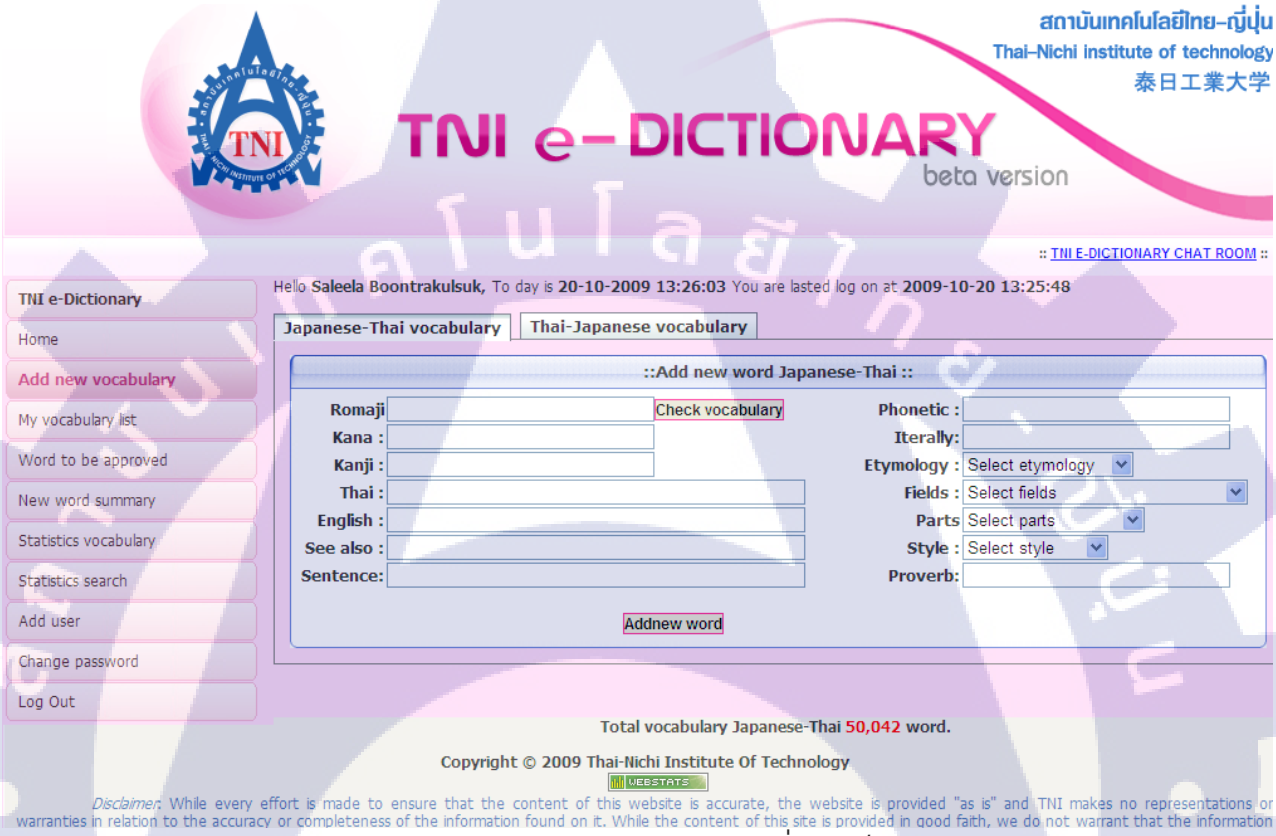
4.8 รูปแสดงการค้นหาศัพท์ด้วยคำไทย

4.9 รูปแสดงศัพท์จากการค้นหาด้วยคำไทย

ลักษณะพิเศษของโปรแกรม TNI e-Dictionary คือ ให้ผู้ใช้สามารถมีส่วนร่วมในการตรวจสอบ เพิ่มศัพท์และแก้ไขได้ โดยเมื่อผู้ใช้ค้นหาศัพท์แล้วพบว่าไม่มีในฐานข้อมูล หากมีความประสงค์จะเพิ่มเข้าไปในฐานข้อมูล หรือกรณีที่ที่ผู้ใช้พบว่าคำแปลไม่ถูกต้อง ก็สามารถเข้าไปในโปรแกรมเพื่อแก้ไขได้

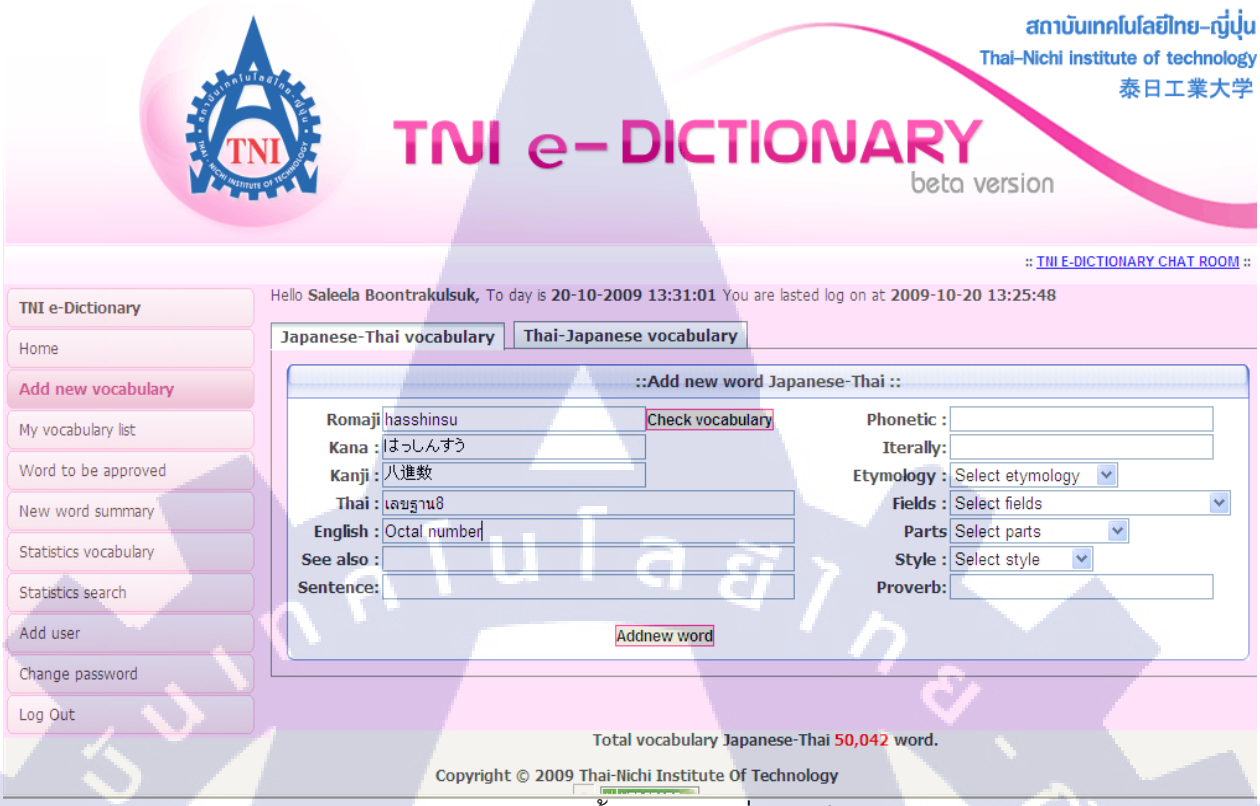
การเพิ่มศัพท์

1. เลือก MENU Add new vocabulary



4.10 รูปแสดงหน้าจอการเพิ่มศัพท์

2. กรอกรายละเอียดของคำที่ต้องการเพิ่ม แล้ว click ที่ Add new word



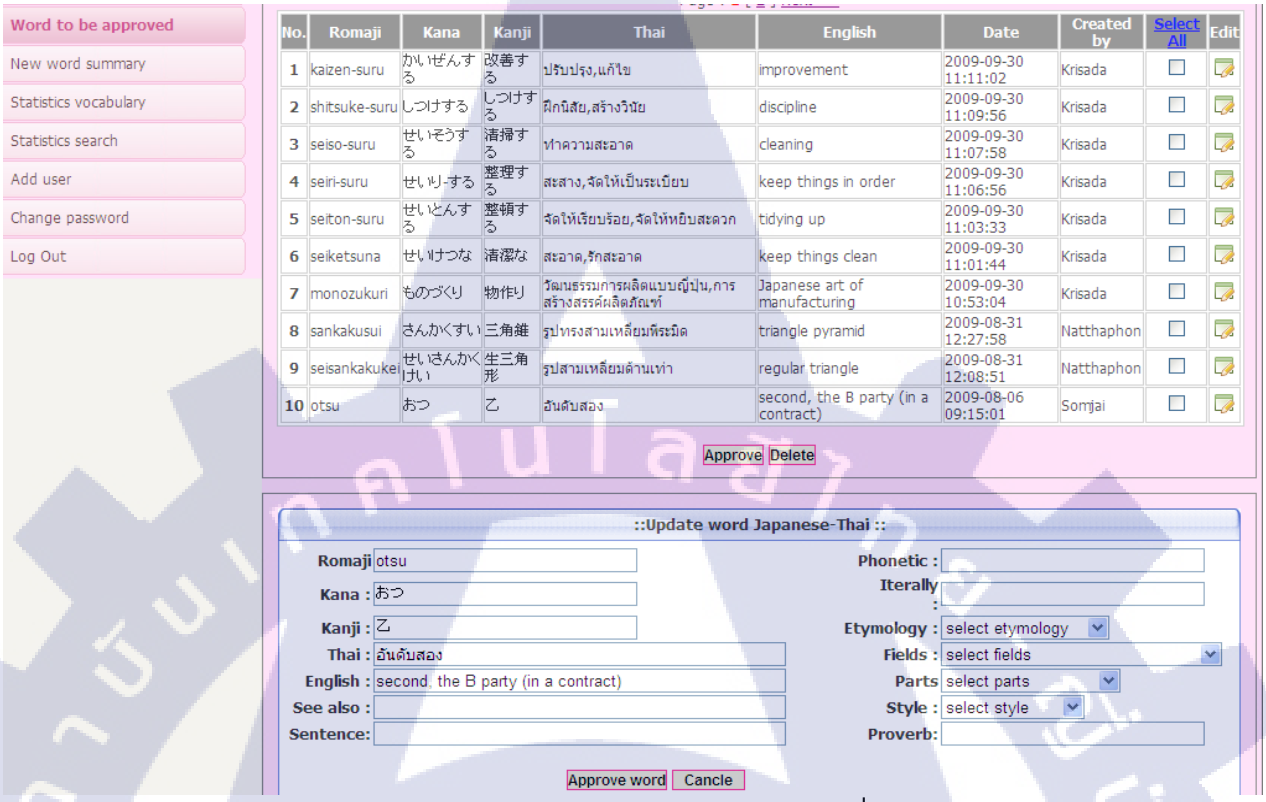
4.11 รูปแสดงขั้นตอนการเพิ่มศัพท์

3. คำศัพท์ที่เพิ่มมาจะอยู่ในฐานข้อมูลที่รอการตรวจสอบ

4.12 รูปแสดงศัพท์ที่รอการเพิ่ม

การแก้ไขศัพท์

1. เข้า Menu My vocabulary list เลือกคำศัพท์ที่ต้องการแก้ไข click Edit จะปรากฏรายละเอียดศัพท์ให้แก้ไข เมื่อแก้ไขแล้วจะได้รับการตรวจสอบก่อนที่จะแก้ไขในฐานข้อมูล

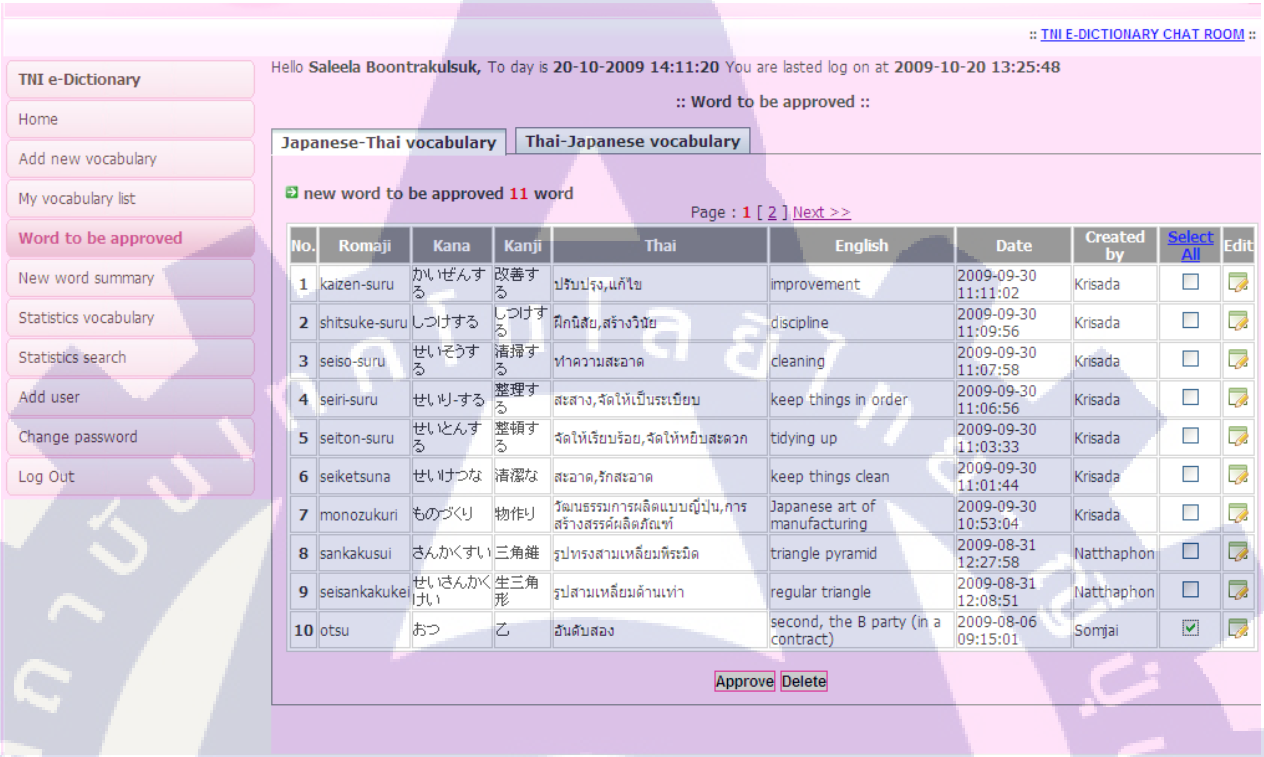


4.13 รูปแสดงรายละเอียดศัพท์ที่จะแก้ไข

การ Approve คำศัพท์

1. ดำเนินการโดยอาจารย์ที่ได้รับมอบหมายให้ตรวจสอบคำในหมวดนั้นๆ โดยเข้าไปที่ MENU Word to be approved Click เลือกคำศัพท์ที่จะ Approve ที่ช่องสี่เหลี่ยม และ click Approve คำศัพท์นั้นก็จะถูกเพิ่มหรือแก้ไขในฐานข้อมูล

ทั้งนี้ในอนาคตจะมีการเพิ่มศัพท์รวมทั้งสิ้นกว่า 50,000 คำ



4.14 รูปแสดงหน้าจอศัพท์ที่รอการแก้ไข

บทที่ 5
การเพิ่มศัพท์ในพจนานุกรม

5.1 การเพิ่มศัพท์ให้พจนานุกรมญี่ปุ่น-ไทย

พจนานุกรมญี่ปุ่น-ไทย เป็นพจนานุกรมที่มีนักศึกษาและบุคคลทั่วไปในไทยใช้กันมากที่สุด เพื่อเรียนภาษาญี่ปุ่นวิธีการเพิ่มศัพท์เข้าในฐานข้อมูลพจนานุกรมมี 2 วิธีคือ

1.การเพิ่มศัพท์โดยการป้อนผ่านตาราง Excel

เป็นการเพิ่มศัพท์โดยการสร้างเป็นตารางใน Excel ก่อนจากนั้นจึงแปลงข้อมูลเข้าในฐานข้อมูลพจนานุกรมญี่ปุ่น-ไทยอีกทีหนึ่ง การเพิ่มศัพท์โดยวิธีนี้มีข้อดีในกรณีที่ทำการเพิ่มศัพท์จำนวนมาก ศัพท์ที่เพิ่มจะมาจากพจนานุกรมต่างๆที่เกี่ยวข้องกับการเรียน การสอน เช่น ศัพท์โรงงาน ศัพท์วิศวกรรมศาสตร์ ศัพท์เทคโนโลยีสารสนเทศ ศัพท์บริหารธุรกิจ แต่สมาชิกไม่สามารถเพิ่มศัพท์โดยวิธีนี้ได้ จะต้องให้ผู้ควบคุมระบบเป็นผู้ทำการแปลงข้อมูลให้ทุกครั้ง

การโอนย้ายข้อมูลเข้าไปในฐานข้อมูล พจนานุกรมทีเดียวทั้งไฟล์ เงื่อนไขคือจะต้องป้อนข้อมูลคำศัพท์ให้มีฟิลด์ทุกฟิลด์ตรงกันในฐานข้อมูลจึงจะโอนย้ายข้อมูลได้

วิธีการเพิ่มศัพท์เริ่มจากการเตรียมตาราง Excel โดยแบ่งเป็นฟิลด์ต่างๆ

พจนานุกรมญี่ปุ่น-ไทย มีฟิลด์ Headword_Roman , Kana , Kanji , field , parts , etymology , mean_th , mean _en , sample sentence , ตามตัวอย่างในรูป 5.1

head_word_roman	head_word	head_word_japanese	head_word_phonetic	iterally
abaku	あばく	暴く	abaku	
abara	あばら	肋	abara	
abaraya	あばらや	あばら屋	abaraya	
abarenbo	あばれんぼう	暴れん坊	abarenbō	
abareru	あばれる	暴れる	abareru	

etymology	FIELDS	parts	styles	proverb	meaning	mean_en
		กด.			ทำให้เปิดออก	NULL
					ชี้โครง	NULL
					บ้านที่ชำรุดทร	NULL
			ป.		โฉเบ	NULL
		กต.			อาละวาด	NULL

รูป 5.1 ตาราง Excel ศัพท์ญี่ปุ่น-ไทย(JP2TH)

พจนานุกรมไทย-ญี่ปุ่น มีฟิลด์ Headword-th , phonetic , field , etymology , meaning_jp , Kana ,Romaji , meaning_en ตามตัวอย่างในรูป5.2

head_word_roman	head_word	head_word_japanese	head_word_phonetic	iterally
abaku	あばく	暴く	abaku	
abara	あばら	肋	abara	
abaraya	あばらや	あばら屋	abaraya	
abarenbo	あばれんぼう	暴れん坊	abarenbō	
abareru	あばれる	暴れる	abareru	

etymology	FIELDS	parts	styles	proverb	meaning	mean_en
		ก๓.			ทำให้เปิดออก	NULL
					ซี่โครง	NULL
					บ้านที่ช่ารุดทร	NULL
			ป.		โฉเบ๊	NULL
		ก๓.			อาละวาด	NULL

รูป 5.2 ตาราง Excel ศัพท์ไทย-ญี่ปุ่น(TH2JP)

แหล่งที่มีมาของคำศัพท์ แบ่งเป็นสาขาวิชาต่างๆ ได้แก่

1. ศัพท์เทคนิคโรงงานอุตสาหกรรม เป็นคำศัพท์ที่ใช้ในโรงงานอุตสาหกรรม ได้จากกลุ่มศึกษาคำศัพท์ญี่ปุ่นที่ใช้ในโรงงานอุตสาหกรรม มีคำศัพท์ประมาณ 3,000 คำ มีหมวดคำต่างๆ เช่น วิศวกรรมโรงงาน วิศวกรรมการผลิต วิศวกรรมไฟฟ้า การจัดการอุตสาหกรรม เคมี กฎหมาย และแพทย์ที่เหลือจะเป็นคำศัพท์ทั่วไปซึ่งจะซ้ำกับพจนานุกรม TUFS
2. ศัพท์วิศวกรรมเครื่องกล เป็นศัพท์ที่ได้จากพจนานุกรมรวมคำศัพท์วิชาการด้านวิศวกรรมเครื่องกลของสมาคมวิศวกรรมเครื่องกลแห่งประเทศญี่ปุ่น (Japanese scientific terms Mechanical Engineer. By The Japan Society of Mechanical Engineering) [11] คำศัพท์เดิมจะมีเฉพาะศัพท์ภาษาญี่ปุ่นและคำแปลเป็นภาษาอังกฤษ ได้ทำการเพิ่มเติมคำแปลภาษาไทยเข้าไปเพื่อให้สมบูรณ์ตามฟิลด์ต่างๆของฐานข้อมูล ศัพท์วิศวกรรมเครื่องกลจะมีประมาณ 4,000 คำ โดยเลือกคำที่เหมาะสมจากศัพท์ทั้งหมด 8,000 คำ
3. ศัพท์เทคโนโลยีสารสนเทศ เป็นศัพท์ที่ได้จากฐานข้อมูลคำศัพท์ในเวปไซต์ e-word.jp[13] ที่ได้รวบรวมคำศัพท์ IT ต่างๆ ไว้เป็นจำนวนมากจะมีประมาณ 3,000 คำ เป็นการอธิบาย

ความหมายของคำศัพท์ IT เพื่อให้เข้าใจและมีแปลเป็นภาษาอังกฤษ ในส่วนนี้ได้เพิ่มคำแปลภาษาไทยเพื่อนำใส่ในฐานข้อมูลได้ นอกจากนี้ยังคัดเลือกคำศัพท์จากพจนานุกรมคำศัพท์คอมพิวเตอร์อังกฤษญี่ปุ่นและญี่ปุ่นอังกฤษ[12] อีกด้วย คำศัพท์ IT ส่วนใหญ่จะเป็น Katakana เนื่องจากรากศัพท์มาจากคำอังกฤษ จึงเพิ่มส่วนที่เป็น Etymology ที่มาของคำในตาราง Excel ด้วย ได้เลือกคำศัพท์ที่เหมาะสมในตาราง Excel ประมาณ 1,200 คำ ให้เพียงพอกับการใช้งานในการเรียนคณะเทคโนโลยีสารสนเทศ

4. ศัพท์บริหารธุรกิจ เป็นศัพท์ที่ได้จากพจนานุกรมคำศัพท์ทางเศรษฐศาสตร์ บัญชี การเงิน บริหารธุรกิจต่างๆ ได้แก่ พจนานุกรมศัพท์บริหารธุรกิจไทย-ญี่ปุ่น[14] โดยคำศัพท์เดิมจะเป็นพจนานุกรมไทย-ญี่ปุ่นและ มีส่วนพจนานุกรมญี่ปุ่น-อังกฤษด้วย จึงต้องเพิ่มเติมคำแปลภาษาไทยหรือภาษาอังกฤษที่ขาดหายไป พจนานุกรมในส่วนนี้จะรวบรวมได้ประมาณ 2,000 คำ จึงครอบคลุมคำศัพท์ที่ใช้ในการเรียนบริหารธุรกิจของสถาบันได้

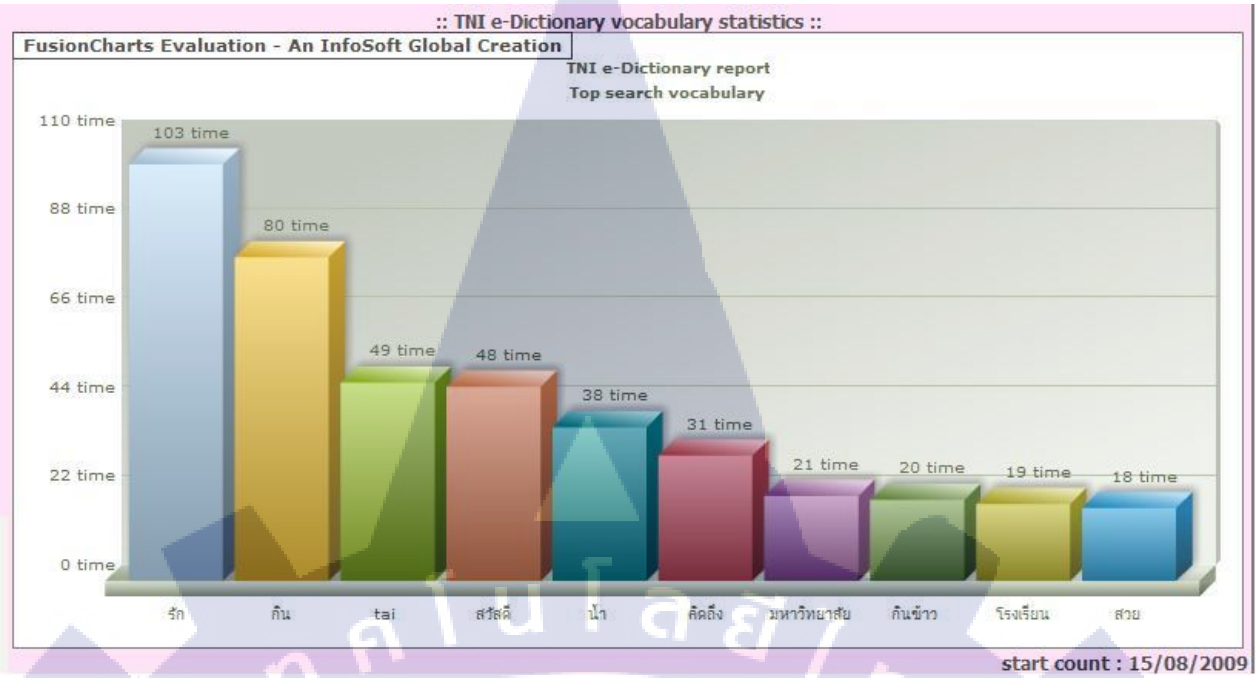
1. การเพิ่มศัพท์จากการเสนอศัพท์ใหม่เข้าทางระบบ

ผู้วิจัย อาจารย์ นักศึกษาที่ใช้งานระบบที่เป็นสมาชิก จะสามารถป้อนศัพท์ใหม่เข้าในระบบได้ การป้อนศัพท์ใหม่มีหลายจุดมุ่งหมายได้แก่

1. ศัพท์ใหม่ที่ยังไม่มีในฐานข้อมูล ซึ่งตรวจสอบจากการค้นหาศัพท์จากในระบบ
2. การแก้ไขศัพท์ที่มีในฐานข้อมูล โดยการปรับปรุงคำแปลให้เหมาะสม หรือเสนอคำแปลเพิ่มเติมจากที่ที่มีในฐานข้อมูล หรือเพิ่มเติมข้อมูลในส่วนของตัวอย่างประโยคที่ใช้คำศัพท์นั้น เป็นประโยชน์อย่างมากในการเรียนภาษา
3. การป้อนคำศัพท์ใหม่โดยการศึกษาจากคำที่หาไม่พบในระบบ (Unfound words) จากการทดลองใช้งานระบบในหมู่นักศึกษา TNI เป็นระยะเวลา 6 เดือน ได้ทำการเก็บสถิติเกี่ยวกับการใช้งานพจนานุกรมอิเล็คทรอนิกส์ มีการเก็บข้อมูลคำศัพท์ที่ค้นหาบ่อย คำศัพท์ที่ค้นหาไม่พบในระบบ จำพวกการใช้งานเป็นต้น รูป5.3แสดงคำศัพท์ที่ค้นหามากที่สุดเรียงลำดับกันมา
รูป 5.4 แสดงตัวอย่างรายงานของคำศัพท์ที่ค้นไม่พบในฐานข้อมูล คณะวิจัยได้เลือกคำศัพท์จากในรายการนี้และเพิ่มเข้าไปในพจนานุกรมโดยวิธีการป้อนผ่านระบบ

จำนวนศัพท์ที่ป้อนผ่านระบบมีจำนวนประมาณ 1,000 คำ ในทุกหมวดคำ เหตุที่ได้จำนวนศัพท์น้อยกว่าแบบการป้อนด้วยตาราง Excel เนื่องจากขาดผู้อนุมัติคำศัพท์และแก้ไขให้ถูกต้องก่อนจะอนุมัติให้บันทึกในพจนานุกรม ดังนั้นในการดำเนินงานระบบหลังจากติดตั้งใช้งานจริงจะต้องกำหนดหน้าที่ของทีมงานและงานในส่วนที่จะรับผิดชอบอย่างชัดเจนกว่านี้ เพื่อให้สามารถดำเนินงาน บำรุงรักษาระบบ และสามารถเพิ่มคำศัพท์ให้ระบบตลอดไปได้

รูป 5.5 -5.6 แสดงกราฟสถิติของการใช้งานนี้ จำนวนผู้ใช้งาน และประเภทผู้ใช้งานที่เป็นนักศึกษา บุคคลทั่วไป ในประเทศไทยหรือต่างประเทศเป็นต้น



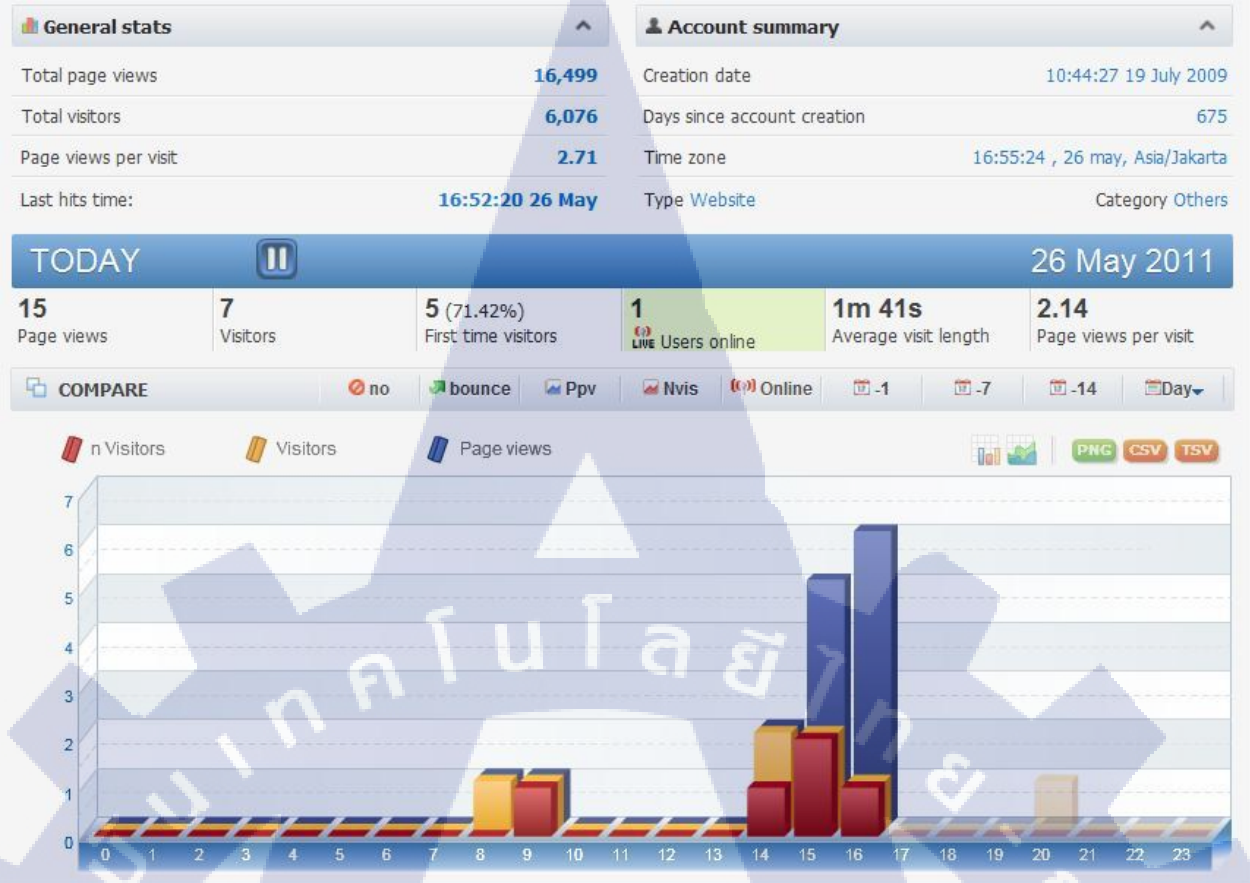
รูป 5.3 คำศัพท์ที่ค้นหามากที่สุดเรียงลำดับ

Lastest search missing **1834** vocabulary

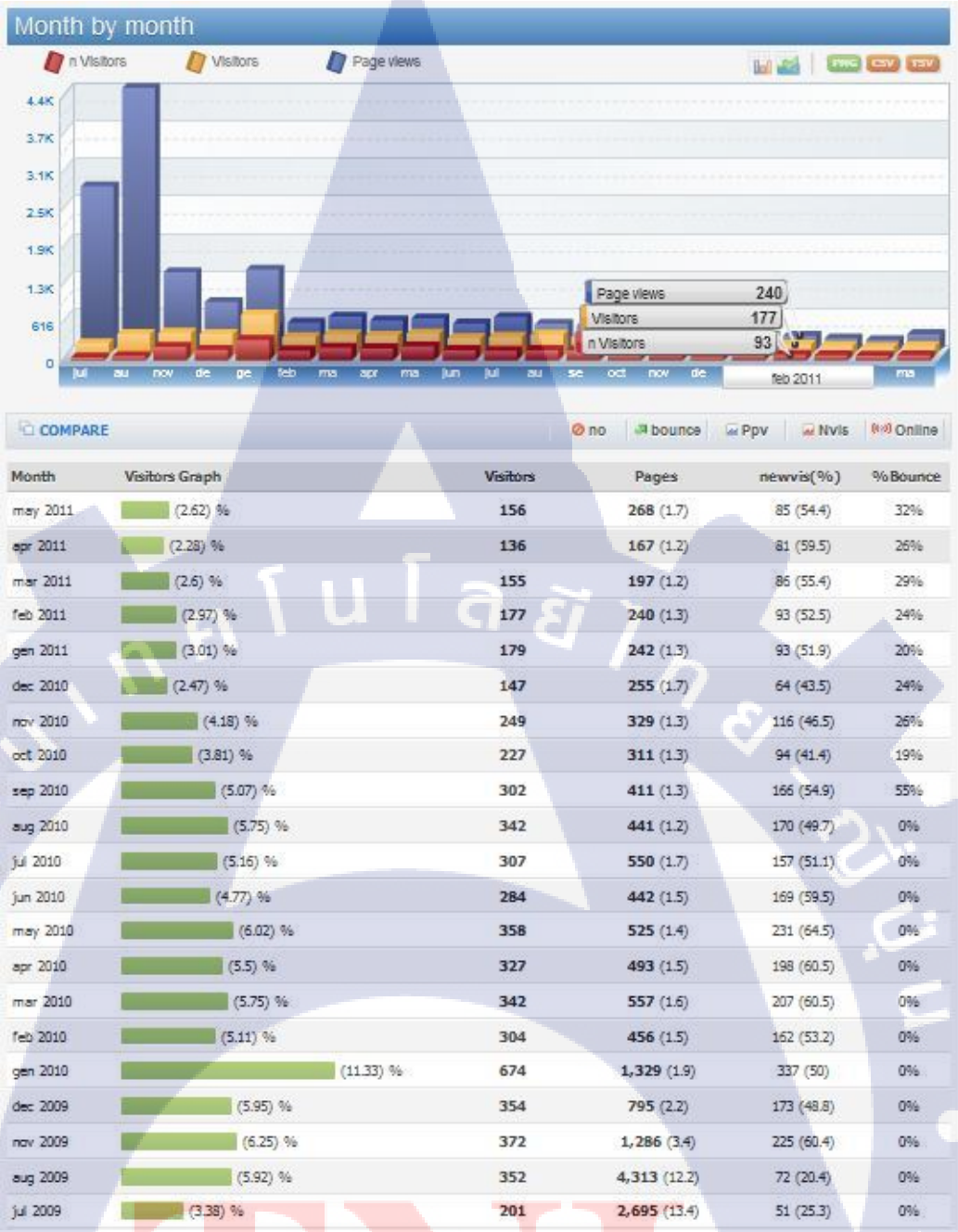
Page : **1** [2][3][4][5][6][7][8][9][10][11][12][13][14][15][16][17][18][19] Next >>

No.	Vocabulary	No.	Vocabulary	No.	Vocabulary	No.	Vocabulary
1834	aricato	1833	いちにちですこしれんしゅうするのはじょうずになって、わすれない	1832	nodoki	1831	カマじゃないって
1830	カマじゃないって！	1829	สภาวิชาชีพบัญชี	1828	号口	1827	ööönagai
1826	ใบลาป่วย	1825	แก้วใส	1824	แก้วน้ำ	1823	ついたし
1822	ついなじ	1821	ぎまをみろ	1820	きょうそお	1819	kyousoo
1818	kyoosoo	1817	ผู้ชนะ	1816	zennen	1815	ぜんねん
1814	売り切る	1813	売り飛ばす	1812	あぐねる	1811	ohayoo
1810	ทำอย่างไร	1809	กินยา	1808	NAMAA	1807	ชั่งน้ำหนัก
1806	เครื่องชั่งน้ำหนัก	1805	เครื่องชั่ง	1804	แถวตรง	1803	ยังมี
1802	กินข้าวหรือยัง	1801	enpitsu kezuri	1800	tabesugi	1799	imananishiteruno
1798	imanani	1797	ทำไร	1796	โปรดรอน	1795	shimashita
1794	tomodachi to hanashimashita	1793	ohanban	1792	ขนมหวาน	1791	แปลว่าอะไร
1790	sandwich	1789	sandoicchi	1788	lnhlkk	1787	ฉัน รัก เธอ
1786	panel	1785	front end	1784	member	1783	outrigger
1782	dd	1781	ちやらない	1780	ちやらないい	1779	ブワドン
1778	jyou-nin	1777	jyou	1776	jyu	1775	sensee
1774	jyuu	1773	jyuu-nin	1772	shoogi	1771	shoogu
1770	スペース	1769	スパン	1768	スパークアウト	1767	すのこ
1766	すみだし	1765	隅きり	1764	墨糸	1763	隅石
1762	スキュウ	1761	スキャン	1760	スクロール	1759	スクリーンセーバ
1758	スクレーバ	1757	スクランブル	1756	スコッチヨーク	1755	スケルチ
1754	スケールメリット	1753	スカラ	1752	すじかい	1751	水溶液
1750	水和	1749	水頭	1748	吹送	1747	水質
1746	垂心	1745	水線	1744	水生植物	1743	水砕
1742	推論	1741	水密	1740	吸込	1739	水管
1738	垂下	1737	水蒸気	1736	水封	1735	水撃

รูป 5.4 ศัพท์ที่ค้นไม่พบในระบบฐานข้อมูลพจนานุกรม



รูป 5.5 กราฟสถิติจำนวนผู้ใช้งาน TNI e-Dict



Month	Visitors Graph	Visitors	Pages	newvis(%)	%Bounce
may 2011	(2.62) %	156	268 (1.7)	85 (54.4)	32%
apr 2011	(2.28) %	136	167 (1.2)	81 (59.5)	26%
mar 2011	(2.6) %	155	197 (1.2)	86 (55.4)	29%
feb 2011	(2.97) %	177	240 (1.3)	93 (52.5)	24%
gen 2011	(3.01) %	179	242 (1.3)	93 (51.9)	20%
dec 2010	(2.47) %	147	255 (1.7)	64 (43.5)	24%
nov 2010	(4.18) %	249	329 (1.3)	116 (46.5)	26%
oct 2010	(3.81) %	227	311 (1.3)	94 (41.4)	19%
sep 2010	(5.07) %	302	411 (1.3)	166 (54.9)	55%
aug 2010	(5.75) %	342	441 (1.2)	170 (49.7)	0%
jul 2010	(5.16) %	307	550 (1.7)	157 (51.1)	0%
jun 2010	(4.77) %	284	442 (1.5)	169 (59.5)	0%
may 2010	(6.02) %	358	525 (1.4)	231 (64.5)	0%
apr 2010	(5.5) %	327	493 (1.5)	198 (60.5)	0%
mar 2010	(5.75) %	342	557 (1.6)	207 (60.5)	0%
feb 2010	(5.11) %	304	456 (1.5)	162 (53.2)	0%
gen 2010	(11.33) %	674	1,329 (1.9)	337 (50)	0%
dec 2009	(5.95) %	354	795 (2.2)	173 (48.8)	0%
nov 2009	(6.25) %	372	1,286 (3.4)	225 (60.4)	0%
aug 2009	(5.92) %	352	4,313 (12.2)	72 (20.4)	0%
jul 2009	(3.38) %	201	2,695 (13.4)	51 (25.3)	0%

รูป 5.6 ประเภทของผู้ใช้งาน

5.2 **การเพิ่มคำศัพท์ให้พจนานุกรมไทย-ญี่ปุ่น**

พจนานุกรมไทย-ญี่ปุ่น เป็นพจนานุกรมที่คนญี่ปุ่นจะใช้มากกว่าเพื่อการเรียนภาษาไทย ในฐานข้อมูลจะมีฟิลด์พิเศษเพิ่มขึ้นมาอีกหนึ่งฟิลด์คือ Phonetic หรือการออกเสียงคำศัพท์ภาษาไทย เพื่อช่วยให้คนญี่ปุ่นสามารถอ่านออกเสียงคำไทยให้ถูกต้องมากที่สุด เป็นประโยชน์ในการเรียนรู้ภาษาไทย

การเพิ่มศัพท์มี 2 วิธีเช่นเดียวกันคือ การเพิ่มจากตาราง Excel และการเพิ่มศัพท์ผ่านทางระบบ การเพิ่มโดยตาราง Excel มี 2 ไฟล์คือ

1. ศัพท์จากพจนานุกรมท่องเที่ยวของสมาคม สสท. รวบรวมโดย อ.ปราณี จงสุจริตธรรม[8] มีจำนวนศัพท์ประมาณ 6,800 คำ เป็นคำศัพท์ไทย มีคำแปลเป็นญี่ปุ่นพร้อมคำอ่าน Romaji เป็นคำศัพท์ทั่วไปที่ใช้ใน

ชีวิตประจำวันและเหมาะสำหรับการท่องเที่ยวสำหรับมัคคุเทศก์แนะนำให้กับนักท่องเที่ยวชาวญี่ปุ่น ส่วนนี้ได้เพิ่มเติมคำอ่านออกเสียง Hiragana เข้าไปในฐานข้อมูลด้วย

2. ศัพท์พจนานุกรม คำศัพท์เทคนิค วิศวกรรมศาสตร์ เทคโนโลยีสารสนเทศ และบริหารธุรกิจที่ได้จากเลือกคำศัพท์จากฐานข้อมูล JP2TH (ญี่ปุ่น-ไทย) เฉพาะคำที่ไม่ปรากฏในฐานข้อมูลเดิมและมีคำแปลเป็นหนึ่งต่อหนึ่งเท่านั้น ส่วนนี้มีจำนวนประมาณ 3,000 คำ

สำหรับการเพิ่มศัพท์ผ่านทางระบบยังมีน้อยมากจากการทดลองใช้งานระบบเป็นเวลา 6 เดือน อาจต้องเชิญชวนอาจารย์ชาวญี่ปุ่นที่รู้ภาษาไทยมาช่วยในการเพิ่มคำศัพท์

สถาบันเทคโนโลยีไทย-ญี่ปุ่น THAI - NICHI INSTITUTE OF TECHNOLOGY TNI

บทที่ 6
การทดลองใช้งาน

ได้สร้างพจนานุกรมอิเล็คทรอนิกส์ญี่ปุ่น-ไทยและไทย-ญี่ปุ่น โดยใช้บริการผ่านทางเวปไซต์ TNI e-dict.net โดยติดตั้งทดสอบระบบเพื่อทำการแก้ไขและทดลองใช้งาน เพื่อประเมินผลระบบ เป็นระยะเวลา 6 เดือน เริ่มจาก พฤษภาคม 2553 ถึง มกราคม 2554

คำศัพท์ที่ให้บริการได้แก่คำศัพท์ที่ได้จากพจนานุกรมไทย-ญี่ปุ่น TUFS ซึ่งมีจำนวนคำศัพท์รวมทั้งญี่ปุ่น-ไทย และไทย-ญี่ปุ่นประมาณ 50,000 คำ

การค้นหาคำศัพท์สามารถทำได้โดยการป้อนคำศัพท์ในรูป Romaji , Hiragana , Kanji กรณีค้นหาคำไทยก็ป้อนอักษรไทย ระบบก็จะสืบค้นคำศัพท์ในฐานข้อมูลและ แสดงผลบนจอภาพในรูปตารางซึ่งประกอบด้วย คำอ่าน Romaji ,Kana , Kanji , Thai และอังกฤษ คำศัพท์ที่แสดงจะรวมคำศัพท์ที่ใกล้เคียงกันแสดงผลออกมาด้วยเพื่อให้เห็นคำศัพท์ทั้งหมดเพื่อให้ผู้ค้นสามารถเห็นคำศัพท์ที่เกี่ยวข้องกัน ในกรณีที่ไม่พบคำศัพท์นั้นจะรายงานว่าไม่พบคำศัพท์และจะบันทึกศัพท์นั้นไว้เพื่อใช้เป็นข้อมูลในการเพิ่มคำศัพท์ต่อไป

6.1 ผลการทดสอบทดลองใช้ระบบ

1. จำนวนผู้เข้าใช้ระบบตลอดระยะเวลา 6 เดือน มีจำนวน 6,076 คำ
   เฉลี่ยวันละ 34 คน คำศัพท์ที่ถูกค้นหาทั้งสิ้น 16,499 คำ
2. สถิติที่สามารถเก็บจากการใช้ระบบ
   2.1 ศัพท์ที่นิยมค้นหาจากระบบมากที่สุด 10 อันดับต้นได้แก่ รัก กิน tai สวัสดี น้ำ คิดถึง มหาวิทยาลัย กินข้าว โรงเรียน สวย ตามลำดับ
   2.2 ศัพท์ที่ค้นไม่พบในระบบ มีรายงานจำนวนทั้งสิ้น 1,834 คำ
   มีคำศัพท์ที่เกิดจาการพิมพ์ผิด 1,230 คำ
   คำศัพท์ที่ค้นไม่พบ 1,834 คำ
   คำศัพท์ที่สามารถเพิ่มเข้าในระบบ 545 คำ
   2.3 ผู้ใช้งานฐานข้อมูล
   Page views 4,400
   Visitors 616
   New visitors 500
   2.4 คำศัพท์ภาษาที่มีคนค้นมากที่สุดคือ ศัพท์ไทยหาคำแปลญี่ปุ่น แสดงว่าต้องการหาคำญี่ปุ่นจากการป้อนภาษาไทยมากกว่า มักเป็นผู้เรียนเริ่มต้นยังไม่สามารถป้อนคำญี่ปุ่นได้
3. สิ่งที่ค้นพบจากการทดลองระบบ
   3.1 การค้นหาคำศัพท์ไม่พบเนื่องจากาการป้อน Romaji ที่มีคำอ่านผิดเนื่องจากภาษาญี่ปุ่นมีคำสั้น คำยาว ซึ่งผู้เรียนเริ่มต้นจะแยกแยะได้ลำบาก เมื่อป้อนคำศัพท์เป็น Romaji ก็ไม่สามารถแยกคำได้ เช่นคำว่า Gakugoo ซึ่งแปลว่าโรงเรียน ผู้ใช้มักพิมพ์ว่า Gakugo จึงค้นคำไม่พบ

3.2 การค้นหาคำศัพท์ไม่พบเนื่องจากการป้อน Romaji ไม่ได้เป็นไปตามแบบ Shinkanshiki และไม่ได้เตรียม Keyboard ที่จะพิมพ์คำยาวเป็น ā ī ū ē ō จึงไม่สามารถค้นหาคำได้ถูกต้อง เช่น พิมพ์ Gakugou แทน Gakugoo หรือ Gakugo เป็นต้น การแก้ไขต้องเพิ่ม Algorithm ในการสืบค้นหาคำใกล้เคียงให้มากกว่าเดิม คือพิมพ์ Gakugo ก็ให้หาคำ Gakugoo, Gakugou , Gakugo ออกมาทั้งหมดหรือถ้าพิมพ์คำอื่นก็ให้หาคำที่เหลือ ออกมาเช่นกัน เพื่อให้สามารถหาคำได้ทั้งหมด

6.2 สิ่งที่ต้องแก้ไขหรือทำเพิ่มเติม

1. การเพิ่มศัพท์ให้พจนานุกรมญี่ปุ่นไทยให้เพียงพอและครอบคลุมสำหรับการใช้งานของนักศึกษา โดยเฉพาะศัพท์ทางด้านวิทยาศาสตร์และเทคโนโลยี ศัพท์คณิตศาสตร์ ศัพท์เคมี ศัพท์ฟิสิกส์ ศัพท์ชีววิทยา รวมทั้งศัพท์ด้านการแพทย์ซึ่งมีใช้ในชีวิตประจำวัน

2. เพิ่มตัวอย่างประโยคสำหรับคำศัพท์ทั่วไปที่มีการใช้งานมาก เพื่อให้นักศึกษาเข้าใจวิธีการใช้คำหรือเข้าความหมายของคำในกรณีต่างๆ

3. เพิ่มข้อมูล phonetic สำหรับคำอ่านออกเสียงให้กับพจนานุกรมไทย-ญี่ปุ่นเพื่อให้คนญี่ปุ่นใช้ประโยชน์ในการเรียนภาษาไทยได้

4. ปรับปรุงระบบโปรแกรมให้สามารถใช้งานได้ง่ายขึ้น ในการค้นหาข้อมูล การแสดงผลคำศัพท์การเพิ่มคำศัพท์ หรือการแก้ไขคำศัพท์เดิม การพิจารณาอนุมัติคำศัพท์ เป็นต้น

5. พัฒนาโปรแกรมที่สามารถเพิ่มคำศัพท์จากตาราง Excel ได้อย่างอัตโนมัติโดยใช้คำสั่งจากผู้ดูแลระบบ ซึ่งจะทำให้การเพิ่มศัพท์จำนวนมากสามารถทำได้ง่ายยิ่งขึ้น

6. ได้ทำแบบสอบถามสำหรับผู้ใช้งานระบบเพื่อเป็นการประเมินผลการใช้งานสำหรับการปรับปรุงฐานข้อมูลและเวปไซต์ในเวปไซต์ให้ดียิ่งขึ้น

บทที่ 7
สรุปการวิจัย

รายงานการวิจัยนี้ได้กล่าวถึงการพัฒนาพจนานุกรมไทย-ญี่ปุ่น TNI e-Dictionary สำหรับใช้งานในสถาบันเทคโลยีไทย – ญี่ปุ่น บุคคลทั่วไปทั้งคนไทยและคนญี่ปุ่น โดยให้บริการในรูปของเวปไซต์ให้แก่บุคคลทั่วไปสามารถใช้งานในการเรียนภาษาได้โดยไม่เสียค่าใช้จ่าย พจนานุกรมที่พัฒนาขึ้นมามีทั้งพจนานุกรมญี่ปุ่นไทยและพจนานุกรมไทย-ญี่ปุ่น มีคำศัพท์รวมกันประมาณ 50,000 คำ เพียงพอสำหรับการเรียนการสอนภาษาญี่ปุ่นสำหรับนักศึกษาทุกคณะและทุกชั้นปี ประกอบด้วยคำศัพท์ทั่วไปสำหรับการเรียนภาษาเบื้องต้นถึงชั้นกลาง จนถึงระดับสูงที่มีการประยุกต์ในสาขาวิชาต่างๆ เช่น วิศวกรรมศาสตร์ เทคโนโลยีสารสนเทศ และ การบริหารธุรกิจ

TNI e-Dictionary หรือ พจนานุกรมอิเล็คทรอนิกส์ TNI เป็นพจนานุกรมสื่ออิเล็คทรอนิกส์ ไทย – ญี่ปุ่น ที่ผู้ใช้สามารถค้นหาข้อมูลคำศัพท์ทั้งภาษาไทยและภาษาญี่ปุ่นที่ต้องการ ผ่านเครือข่ายอินเตอร์เนตโดยไม่เสียค่าใช้จ่าย นักศึกษาของสถาบัน และนักเรียนโรงเรียนภาษาของสมาคมส่งเสริมเทคโนโลยี(ไทย-ญี่ปุ่น) สามารถใช้ประกอบการเรียนภาษาญี่ปุ่น โดยการค้นหาคำศัพท์ด้วย Romaji หรือ อิรางานะ เพื่อทราบคำแปลเป็นภาษาไทยและอังกฤษ

พจนานุกรมอิเล็คทรอนิกส์ TNI จะมีคำศัพท์ทั่วไปในทุกสาขา รวมคำศัพท์ประมาณ 50,000 คำ ทั้งไทย – ญี่ปุ่น และญี่ปุ่น – ไทย คำศัพท์ที่บรรจุในฐานข้อมูล ประกอบด้วยฐานข้อมูลต่างๆคือ

1 ฐานข้อมูลศัพท์ไทย-ญี่ปุ่น และญี่ปุ่น-ไทย จากมหาวิทยาลัยTokyo University of Foreign Studies

2 รวบรวมศัพท์ทั้งหมดในตำราเรียนภาษาญี่ปุ่นของสถาบัน

3 หนังสือพจนานุกรมจาก ส.ส.ท. ที่ได้รับอนุญาต เช่น ปทานุกรมศัพท์ช่าง และพจนานุกรมท่องเที่ยวไทย-ญี่ปุ่น

4 ยังมีคำศัพท์เฉพาะทางเทคนิคในสาขาต่างๆ เช่น วิศวกรรมศาสตร์ เทคโนโลยีสารสนเทศ การบริหารธุรกิจ เพิ่มเติมเข้าไปในฐานข้อมูลคำศัพท์ ซึ่งถือเป็นจุดเด่นของโครงการวิจัยนี้ ซึ่งพบได้ยากในพจนานุกรมออนไลน์ที่มีอยู่ในปัจจุบัน

พจนานุกรมอิเล็คทรอนิกส์ TNI จะเปิดให้สมาชิกของสมาคม ส.ส.ท. และบุคคลทั่วไปทั้งไทยและญี่ปุ่นได้ใช้งาน ให้บริการ e-Dictionary จากฐานข้อมูลคำศัพท์ โดยสร้างเป็นเวปไซต์ TNI e-Dictionary ให้บริการแก่นักศึกษาสถาบัน สมาชิก ส.ส.ท. และบุคคลทั่วไปทั้งในประเทศไทยและญี่ปุ่น นอกจากนั้นยังใช้เป็นเวทีที่เปิดโอกาสให้สมาชิกและผู้ใช้งานช่วยกันพัฒนาคลังข้อมูลคำศัพท์ โดยการเสนอแนะคำศัพท์ใหม่ที่ไม่มีในพจนานุกรม การแปลคำศัพท์ที่เหมาะสม หรือเสนอแนะคำแปลอื่นๆ เพิ่มเติมจากในคลังข้อมูลทำให้มีการพัฒนาฐานข้อมูลคำศัพท์เพิ่มเติม โดยผู้ใช้สามารถมีส่วนร่วมด้วย

ได้สร้างพจนานุกรมอิเล็คทรอนิกส์ญี่ปุ่น-ไทยและไทย-ญี่ปุ่นเสร็จและทดสอบระบบโดยเปิดให้บริการผ่านทางเวปไซต์ tni.ac.th โดยติดตั้งทดสอบระบบเพื่อทำการแก้ไขและทดลองใช้งาน เพื่อประเมินผลระบบ เป็นระยะเวลา 6 เดือน ได้ทำแบบสอบถามสำหรับผู้ใช้งานระบบเพื่อเป็นการประเมินผลการใช้งานสำหรับการปรับปรุงฐานข้อมูลและบริการในเวปไซต์ให้ดียิ่งขึ้น

สิ่งที่ควรจะปรับปรุงเพิ่มเติมในขั้นถัดไปได้แก่ การเพิ่มศัพท์ให้พจนานุกรมญี่ปุ่นไทยให้เพียงพอและ

ครอบคลุมสำหรับการใช้งานของนักศึกษา โดยเฉพาะศัพท์ทางด้านวิทยาศาสตร์และเทคโนโลยี ศัพท์คณิตศาสตร์ เคมี ฟิสิกส์ ชีววิทยา รวมทั้งศัพท์ด้านการแพทย์ซึ่งมีใช้ในชีวิตประจำวัน เพิ่มตัวอย่างประโยคสำหรับคำศัพท์ทั่วไปที่มีการใช้งานมาก เพื่อให้นักศึกษาเข้าวิธีการใช้คำ หรือเข้าความหมายของคำในกรณีต่างๆ ปรับปรุงระบบโปรแกรมให้สามารถใช้งานได้ง่ายขึ้นในการค้นหาข้อมูลการแสดงผลคำศัพท์ การเพิ่มคำศัพท์ หรือการแก้ไขคำศัพท์ เป็นต้น

การวิจัยในขั้นถัดไปได้แก่การพัฒนาพจนานุกรมไทย-ญี่ปุ่นให้สมบูรณ์ขึ้นโดยการเพิ่มข้อมูล phoneticสำหรับคำอ่านออกเสียงให้กับพจนานุกรมไทย-ญี่ปุ่นเพื่อให้คนญี่ปุ่นใช้ประโยชน์ในการเรียนภาษาไทยได้

บรรณานุกรม

[1] Londo, **http://www.longdo.com**

[2] JTDIC, **http://www.jtdic.com//jtdic.aspx/**

[3] TUFS,**http:/www.cblle.tufs.ac.jp./dic/th/jath/**

[4] NUT-MTD, **http://gii.nagaokaut.ac.jp/mtd/**

[5] Tomita Takejiro,2003, **"Japanese-Thai and Thai-Japanese Dictionary for Thai People"**, Revised Edition, 367pp+270pp, Bangkok : Association for Thai Education and Cultural Promotion.

[6] Minegishi Makoto and Akagi Osamu ,2009, **"Development of and Electronic Dictionary for Analyzing Linguistic Data**". Proceedings of Chula-Japan Linguistics Conference 2008, pp.11-16, Tokyo University of Foreign Studies, Tokyo

[7] Klanhan Voraputiporn, Phulporn Saenbanpra ,1990, **"Japanese-Thai Technical words Dictionary"**. Technology Promotion Association (Thai-Japan) press.

[8] Pranee Chongsucharitkul ,2006, **"Thai_Japanese Tourism words dictionary"**.Technology Promotion Association (Thai-Japan) press.

[9] TNI e-Dictionary: **http//dict.thainichi.net**

[10] K. Visavateeranon, et al., **" Development of Thai- Japanese e-Dictionary"** Proceeding of ICBIR TNI, 2008.

[11] The Japan Society of Mechanical Engineers **" Japanese Scientific terms, Mechanical Engineering"** Ministry of Eduaction, Culture, Sports, Science and Technology. Japan 1985.

[12] Oba Itsuo, **" Dictionary of Computer terms, English/Japanese and Japanese/English"** Paru Publishing 1991

[13] IT words dictionary, **http://www.e-words.jp/**

[14] JAL Academy,2010, **“พจนานุกรมศัพท์บริหารธุรกิจไทย-ญี่ปุ่น”** .สำนักพิมพ์สมาคมส่งเสริมเทคโนโลยี (ไทย-ญี่ปุ่น)

[15] คณะกรรมการจัดทำพจนานุกรม มหาวิทยาลัยเทคโนโลยีนางาโอกะ,2009, “**พจนานุกรมช่างภาษาญี่ปุ่น-ไทย สาขาวิศวกรรมศาสตร์”** มหาวิทยาลัยนางาโอกะ

ภาคผนวก

เสียงสั้น				เสียงยาว		
hiragana	katakana	romaji		hiragana	katakana	romaji
あ	ア	a		ああ	アー/アア	aa
い	イ	i		いい	イー/イイ	ii
う	ウ	u		うう	ウー/ウウ	uu
え	エ	e		えい/ええ	エー/エイ/エエ	ee
お	オ	o		おう/おお	オー/オウ/オオ	oo
か	カ	ka		かあ	カー/カア	kaa
き	キ	ki		きい	キー/キイ	kii
く	ク	ku		くう	クー/クウ	kuu
け	ケ	ke		けい/けえ	ケー/ケイ/ケエ	kee
こ	コ	ko		こう/こお	コー/コウ/コオ	koo
さ	サ	sa		さあ	サー/サア	saa
し	シ	shi		しい	シー/シイ	shii
す	ス	su		すう	スー/スウ	suu
せ	セ	se		せい/せえ	セー/セイ/セエ	see
そ	ソ	so		そう/そお	ソー/ソウ/ソオ	soo
た	タ	ta		たあ	ター/タア	taa
ち	チ	chi		ちい	チー/チイ	chii
つ	ツ	tsu		つう	ツー/ツウ	tsuu
て	テ	te		てい/てえ	テー/テイ/テエ	tee
と	ト	to		とう/とお	トー/トウ/トオ	too
な	ナ	na		なあ	ナー/ナア	naa
に	ニ	ni		にい	ニー/ニイ	nii
ぬ	ヌ	nu		ぬう	ヌー/ヌウ	nuu
ね	ネ	ne		ねい/ねえ	ネー/ネイ/ネエ	nee
の	ノ	no		のう/のお	ノー/ノウ/ノオ	noo
は	ハ	ha		はあ	ハー/ハア	haa
ひ	ヒ	hi		ひい	ヒー/ヒイ	hii
ふ	フ	fu		ふう	フー/フウ	fuu
へ	ヘ	he		へい/へえ	ヘー/ヘイ/ヘエ	hee
ほ	ホ	ho		ほう/ほお	ホー/ホウ/ホオ	hoo
ま	マ	ma		まあ	マー/マア	maa
み	ミ	mi		みい	ミー/ミイ	mii
む	ム	mu		むう	ムー/ムウ	muu
め	メ	me		めい/めえ	メー/メイ/メエ	mee
も	モ	mo		もう/もお	モー/モウ/モオ	moo
や	ヤ	ya		やあ	ヤー/ヤア	yaa
ゆ	ユ	yu		ゆう	ユー/ユウ	yuu
よ	ヨ	yo		よう/よお	ヨー/ヨウ/ヨオ	yoo
ら	ラ	ra		らあ	ラー/ラア	raa
り	リ	ri		りい	リー/リイ	rii
る	ル	ru		るう	ルー/ルウ	ruu
れ	レ	re		れい/れえ	レー/レイ/レエ	ree
ろ	ロ	ro		ろう/ろお	ロー/ロウ/ロオ	roo
わ	ワ	wa		わあ	ワー/ワア	waa
を	ヲ	o				
ん	ン	n				
が	ガ	ga		があ	ガー/ガア	gaa
ぎ	ギ	gi		ぎい	ギー/ギイ	gii
ぐ	グ	gu		ぐう	グー/グウ	guu
げ	ゲ	ge		げい/げえ	ゲー/ゲイ/ゲエ	gee
ご	ゴ	go		ごう/ごお	ゴー/ゴウ/ゴオ	goo
ざ	ザ	za		ざあ	ザー/ザア	zaa
じ	ジ	ji		じい	ジー/ジイ	jii
ず	ズ	zu		ずう	ズー/ズウ	zuu
ぜ	ゼ	ze		ぜい/ぜえ	ゼー/ゼイ/ゼエ	zee
ぞ	ゾ	zo		ぞう/ぞお	ゾー/ゾウ/ゾオ	zoo
だ	ダ	da		だあ	ダー/ダア	daa
ぢ	ヂ	ji		ぢい	ヂー/ヂイ	jii
づ	ヅ	zu		づう	ヅー/ヅウ	zuu
で	デ	de		でい/でえ	デー/デイ/デエ	dee
ど	ド	do		どう/どお	ドー/ドウ/ドオ	doo
ば	バ	ba		ばあ	バー/バア	baa
び	ビ	bi		びい	ビー/ビイ	bii
ぶ	ブ	bu		ぶう	ブー/ブウ	buu
べ	ベ	be		べい/べえ	ベー/ベイ/ベエ	bee
ぼ	ボ	bo		ぼう/ぼお	ボー/ボウ/ボオ	boo
ぱ	パ	pa		ぱあ	パー/パア	paa
ぴ	ピ	pi		ぴい	ピー/ピイ	pii
ぷ	プ	pu		ぷう	プー/プウ	puu
ぺ	ペ	pe		ぺい/ぺえ	ペー/ペイ/ペエ	pee
ぽ	ポ	po		ぽう/ぽお	ポー/ポウ/ポオ	poo
きゃ	キャ	kya		きゃあ	キャー/キャア	kyaa
きゅ	キュ	kyu		きゅう	キュー/キュウ	kyuu
きょ	キョ	kyo		きょう/きょお	キョー/キョウ/キョオ	kyoo
しゃ	シャ	sha		しゃあ	シャー/シャア	shaa
しゅ	シュ	shu		しゅう	シュー/シュウ	shuu
しょ	ショ	sho		しょう/しょお	ショー/ショウ/ショオ	shoo
にゃ	ニャ	nya		にゃあ	ニャー/ニャア	nyaa
にゅ	ニュ	nyu		にゅう	ニュー/ニュウ	nyuu
にょ	ニョ	nyo		にょう/にょお	ニョー/ニョウ/ニョオ	nyoo
ひゃ	ヒャ	hya		ひゃあ	ヒャー/ヒャア	hyaa
ひゅ	ヒュ	hyu		ひゅう	ヒュー/ヒュウ	hyuu
ひょ	ヒョ	hyo		ひょう/ひょお	ヒョ=ヒョウ/ヒョオ	hyoo
みゃ	ミャ	mya		みゃあ	ミャー/ミャア	myaa
みゅ	ミュ	myu		みゅう	ミュー/ミュウ	myuu
みょ	ミョ	myo		みょう/みょお	ミョー/ミョウ/ミョオ	myoo
りゃ	リャ	rya		りゃあ	リャー/リャア	ryaa
りゅ	リュ	ryu		りゅう	リュー/リュウ	ryuu
りょ	リョ	ryo		りょう/りょお	リョー/リョウ/リョオ	ryoo
ぎゃ	ギャ	gya		ぎゃあ	ギャー/ギャア	gyaa
ぎゅ	ギュ	gyu		ぎゅう	ギュー/ギュウ	gyuu
ぎょ	ギョ	gyo		ぎょう/ぎょお	ギョー/ギョウ/ギョオ	gyoo
じゃ	ジャ	ja		じゃあ	ジャー/ジャア	jaa
じゅ	ジュ	ju		じゅう	ジュー/ジュウ	juu
じょ	ジョ	jo		じょう/じょお	ジョー/ジョウ/ジョオ	joo
ぢゃ	ヂャ	ja		ぢゃあ	ヂャー/ヂャア	jaa
ぢゅ	ヂュ	ju		ぢゅう	ヂュー/ヂュウ	juu
ぢょ	ヂョ	jo		ぢょう/ぢょお	ヂョー/ヂョウ/ヂョオ	joo
びゃ	ビャ	bya		びゃあ	ビャー/ビャア	byaa
びゅ	ビュ	byu		びゅう	ビュー/ビュウ	byuu
びょ	ビョ	byo		びょう/びょお	ビョー/ビョウ/ビョオ	byoo
ぴゃ	ピャ	pya		ぴゃあ	ピャー/ピャア	pyaa
ぴゅ	ピュ	pyu		ぴゅう	ピュー/ピュウ	pyuu
ぴょ	ピョ	pyo		ぴょう/ぴょお	ピョー/ピョウ/ピョオ	pyoo

เสียงกัก				เสียงกัก+เสียงยาว			
hiragana	*katakana*	*romaji*		*hiragana*	*romaji*	*katakana*	*romaji*
						ディ	di
						ドゥ	du
						ファ	fa
						フェ	fe
						フィ	fi
っか	ッカ	kka		っかあ	kkaa	フォ	fo
っき	ッキ	kki		っきい	kkii	スィ	si
っく	ック	kku		っくう	kkuu	ティ	ti
っけ	ッケ	kke		っけい/っけえ	kkee	トゥ	tu
っこ	ッコ	kko		っこう/っこお	kkoo	ヴァ	va
っさ	ッサ	ssa		っさあ	ssaa	ヴェ	ve
っし	ッシ	sshi		っしい	sshii	ヴィ	vi
っす	ッス	ssu		っすう	ssuu	ヴォ	vo
っせ	ッセ	sse		っせい/っせえ	ssee	ヴ	vu
っそ	ッソ	sso		っそう/っそお	ssoo	ウェ	we
った	ッタ	tta		ったあ	ttaa	ウィ	wi
っち	ッチ	cchi		っちい	cchii	ウォ	wo
っつ	ッツ	ttsu		っつう	ttsuu	ズィ	zi
って	ッテ	tte		ってい/ってえ	ttee	ディー	dii
っと	ット	tto		っとう/っとお	ttoo	ドゥー	duu
っな	ッナ	tna		っなあ	tnaa	ファー	faa
っに	ッニ	tni		っにい	tnii	フェー	fee
っぬ	ッヌ	tnu		っぬう	tnuu	フィー	fii
っね	ッネ	tne		っねい/っねえ	tnee	フォー	foo
っの	ッノ	tno		っのう/っのお	tnoo	スィー	sii
っは	ッハ	hha		っはあ	hhaa	ティー	tii
っひ	ッヒ	hhi		っひい	hhii	トゥー	tuu
っふ	ッフ	ffu		っふう	ffuu	ヴァー	vaa
っへ	ッヘ	hhe		っへい/っへえ	hhee	ヴェー	vee
っほ	ッホ	hho		っほう/っほお	hhoo	ヴィー	vii
っま	ッマ	tma		っまあ	tmaa	ヴォー	voo
っみ	ッミ	tmi		っみい	tmii	ヴー	vuu
っむ	ッム	tmu		っむう	tmuu	ウェー	wee
っめ	ッメ	tme		っめい/っめえ	tmee	ウィー	wii
っも	ッモ	tmo		っもう/っもお	tmoo	ウォー	woo
っや	ッヤ	yya		っやあ	yyaa	ズィー	zii
っゆ	ッユ	yyu		っゆう	yyuu		
っよ	ッヨ	yyo		っよう/っよお	yyoo		
っら	ッラ	rra		っらあ	rraa		
っり	ッリ	rri		っりい	rrii	ッディ	ddi
っる	ッル	rru		っるう	rruu	ッドゥ	ddu
っれ	ッレ	rre		っれい/っれえ	rree	ッファ	ffa
っろ	ッロ	rro		っろう/っろお	rroo	ッフェ	ffe
っわ	ッワ	wwa		っわあ	wwaa	ッフィ	ffi
						ッフォ	ffo
						ッスィ	ssi
っが	ッガ	gga		っがあ	ggaa	ッティ	tti
っぎ	ッギ	ggi		っぎい	ggii	ットゥ	ttu
っぐ	ッグ	ggu		っぐう	gguu	ッヴァ	vva
っげ	ッゲ	gge		っげい/っげえ	ggee	ッヴェ	vve
っご	ッゴ	ggo		っごう/っごお	ggoo	ッヴィ	vvi
っざ	ッザ	zza		っざあ	zzaa	ッヴォ	vvo
っじ	ッジ	jji		っじい	jjii	ッヴ	vvu
っず	ッズ	zzu		っずう	zzuu	ッウェ	wwe
っぜ	ッゼ	zze		っぜい/っぜえ	zzee	ッウィ	wwi
っぞ	ッゾ	zzo		っぞう/っぞお	zzoo	ッズィ	zzi
っだ	ッダ	dda		っだあ	ddaa	ッウォ	wwo
っぢ	ッヂ	jji		っぢい	jjii	ッディー	ddii
っづ	ッヅ	zzu		っづう	zzuu	ッドゥー	dduu
っで	ッデ	dde		っでい/っでえ	ddee	ッファー	ffaa
っど	ッド	ddo		っどう/っどお	ddoo	ッフェー	ffee
っば	ッバ	bba		っばあ	bbaa	ッフィー	ffii
っび	ッビ	bbi		っびい	bbii	ッフォー	ffoo
っぶ	ッブ	bbu		っぶう	bbuu	ッスィー	ssii
っべ	ッベ	bbe		っべい/っべえ	bbee	ッティー	ttii
っぼ	ッボ	bbo		っぼう/っぼお	bboo	ットゥー	ttuu
っぱ	ッパ	ppa		っぱあ	ppaa	ッヴァー	vvaa
っぴ	ッピ	ppi		っぴい	ppii	ッヴェー	vvee
っぷ	ップ	ppu		っぷう	ppuu	ッヴィー	vvii
っぺ	ッペ	ppe		っぺい/っぺえ	ppee	ッヴォー	vvoo
っぽ	ッポ	ppo		っぽう/っぽお	ppoo	ッヴー	vvuu
っきゃ	ッキャ	kkya		っきゃあ	kkyaa	ッウェー	wwee
っきゅ	ッキュ	kkyu		っきゅう	kkyuu	ッウィー	wwii
っきょ	ッキョ	kkyo		っきょう/っきょお	kkyoo	ッウォー	wwoo
っしゃ	ッシャ	ssha		っしゃあ	sshaa	ッズィー	zzii
っしゅ	ッシュ	sshu		っしゅう	sshuu		
っしょ	ッショ	ssho		っしょう/っしょお	sshoo		
っにゃ	ッニャ	tnya		っにゃあ	tnyaa		
っにゅ	ッニュ	tnyu		っにゅう	tnyuu		
っにょ	ッニョ	tnyo		っにょう/っにょお	tnyoo		
っひゃ	ッヒャ	hhya		っひゃあ	hhyaa		
っひゅ	ッヒュ	hhyu		っひゅう	hhyuu		
っひょ	ッヒョ	hhyo		っひょう/っひょお	hhyoo		
っみゃ	ッミャ	tmya		っみゃあ	tmyaa		
っみゅ	ッミュ	tmyu		っみゅう	tmyuu		
っみょ	ッミョ	tmyo		っみょう/っみょお	tmyoo		
っりゃ	ッリャ	rrya		っりゃあ	rryaa		
っりゅ	ッリュ	rryu		っりゅう	rryuu		
っりょ	ッリョ	rryo		っりょう/っりょお	rryoo		
っぎゃ	ッギャ	ggya		っぎゃあ	ggyaa		
っぎゅ	ッギュ	ggyu		っぎゅう	ggyuu		
っぎょ	ッギョ	ggyo		っぎょう/っぎょお	ggyoo		
っじゃ	ッジャ	jja		っじゃあ	jjaa		
っじゅ	ッジュ	jju		っじゅう	jjuu		
っじょ	ッジョ	jjo		っじょう/っじょお	jjoo		
っぢゃ	ッヂャ	jja		っぢゃあ	jjaa		
っぢゅ	ッヂュ	jju		っぢゅう	jjuu		
っぢょ	ッヂョ	jjo		っぢょう/っぢょお	jjoo		
っびゃ	ッビャ	bbya		っびゃあ	bbyaa		
っびゅ	ッビュ	bbyu		っびゅう	bbyuu		
っびょ	ッビョ	bbyo		っびょう/っびょお	bbyoo		
っぴゃ	ッピャ	ppya		っぴゃあ	ppyaa		
っぴゅ	ッピュ	ppyu		っぴゅう	ppyuu		
っぴょ	ッピョ	ppyo		っぴょう/っぴょお	ppyoo		

เสียงกัก				เสียงกัก+เสียงยาว			เสียงกัก
hiragana	*katakana*	*romaji*		*hiragana*	*romaji*	*katakana*	*hiragana*
						ディ	
						ドゥ	
						ファ	
						フェ	
						フィ	
っか	ッカ	kka		っかあ	kkaa	フォ	っか
っき	ッキ	kki		っきい	kkii	スィ	っき
っく	ック	kku		っくう	kkuu	ティ	っく
っけ	ッケ	kke		っけい/っけえ	kkee	トゥ	っけ
っこ	ッコ	kko		っこう/っこお	kkoo	ヴァ	っこ
っさ	ッサ	ssa		っさあ	ssaa	ヴェ	っさ
っし	ッシ	sshi		っしい	sshii	ヴィ	っし
っす	ッス	ssu		っすう	ssuu	ヴォ	っす
っせ	ッセ	sse		っせい/っせえ	ssee	ヴ	っせ
っそ	ッソ	sso		っそう/っそお	ssoo	ウェ	っそ
った	ッタ	tta		ったあ	ttaa	ウィ	った
っち	ッチ	cchi		っちい	cchii	ウォ	っち
っつ	ッツ	ttsu		っつう	ttsuu	ズィ	っつ
って	ッテ	tte		ってい/ってえ	ttee	ディー	って
っと	ット	tto		っとう/っとお	ttoo	ドゥー	っと
っな	ッナ	tna		っなあ	tnaa	ファー	っな
っに	ッニ	tni		っにい	tnii	フェー	っに
っぬ	ッヌ	tnu		っぬう	tnuu	フィー	っぬ
っね	ッネ	tne		っねい/っねえ	tnee	フォー	っね
っの	ッノ	tno		っのう/っのお	tnoo	スィー	っの
っは	ッハ	hha		っはあ	hhaa	ティー	っは
っひ	ッヒ	hhi		っひい	hhii	トゥー	っひ
っふ	ッフ	ffu		っふう	ffuu	ヴァー	っふ
っへ	ッヘ	hhe		っへい/っへえ	hhee	ヴェー	っへ
っほ	ッホ	hho		っほう/っほお	hhoo	ヴィー	っほ
っま	ッマ	tma		っまあ	tmaa	ヴォー	っま
っみ	ッミ	tmi		っみい	tmii	ヴー	っみ
っむ	ッム	tmu		っむう	tmuu	ウェー	っむ
っめ	ッメ	tme		っめい/っめえ	tmee	ウィー	っめ
っも	ッモ	tmo		っもう/っもお	tmoo	ウォー	っも
っや	ッヤ	yya		っやあ	yyaa	ズィー	っや
っゆ	ッユ	yyu		っゆう	yyuu		っゆ
っよ	ッヨ	yyo		っよう/っよお	yyoo		っよ
っら	ッラ	rra		っらあ	rraa		っら
っり	ッリ	rri		っりい	rrii	ッディ	っり
っる	ッル	rru		っるう	rruu	ッドゥ	っる
っれ	ッレ	rre		っれい/っれえ	rree	ッファ	っれ
っろ	ッロ	rro		っろう/っろお	rroo	ッフェ	っろ
っわ	ッワ	wwa		っわあ	wwaa	ッフィ	っわ
						ッフォ	
						ッスィ	
っが	ッガ	gga		っがあ	ggaa	ッティ	っが
っぎ	ッギ	ggi		っぎい	ggii	ットゥ	っぎ
っぐ	ッグ	ggu		っぐう	gguu	ッヴァ	っぐ
っげ	ッゲ	gge		っげい/っげえ	ggee	ッヴェ	っげ
っご	ッゴ	ggo		っごう/っごお	ggoo	ッヴィ	っご
っざ	ッザ	zza		っざあ	zzaa	ッヴォ	っざ
っじ	ッジ	jji		っじい	jjii	ッヴ	っじ
っず	ッズ	zzu		っずう	zzuu	ッウェ	っず
っぜ	ッゼ	zze		っぜい/っぜえ	zzee	ッウィ	っぜ
っぞ	ッゾ	zzo		っぞう/っぞお	zzoo	ッズィ	っぞ
っだ	ッダ	dda		っだあ	ddaa	ッウォ	っだ
っぢ	ッヂ	jji		っぢい	jjii	ッディー	っぢ
っづ	ッヅ	zzu		っづう	zzuu	ッドゥー	っづ
っで	ッデ	dde		っでい/っでえ	ddee	ッファー	っで
っど	ッド	ddo		っどう/っどお	ddoo	ッフェー	っど
っば	ッバ	bba		っばあ	bbaa	ッフィー	っば
っび	ッビ	bbi		っびい	bbii	ッフォー	っび
っぶ	ッブ	bbu		っぶう	bbuu	ッスィー	っぶ
っべ	ッベ	bbe		っべい/っべえ	bbee	ッティー	っべ
っぼ	ッボ	bbo		っぼう/っぼお	bboo	ットゥー	っぼ
っぱ	ッパ	ppa		っぱあ	ppaa	ッヴァー	っぱ
っぴ	ッピ	ppi		っぴい	ppii	ッヴェー	っぴ
っぷ	ップ	ppu		っぷう	ppuu	ッヴィー	っぷ
っぺ	ッペ	ppe		っぺい/っぺえ	ppee	ッヴォー	っぺ
っぽ	ッポ	ppo		っぽう/っぽお	ppoo	ッヴー	っぽ
っきゃ	ッキャ	kkya		っきゃあ	kkyaa	ッウェー	っきゃ
っきゅ	ッキュ	kkyu		っきゅう	kkyuu	ッウィー	っきゅ
っきょ	ッキョ	kkyo		っきょう/っきょお	kkyoo	ッウォー	っきょ
っしゃ	ッシャ	ssha		っしゃあ	sshaa	ッズィー	っしゃ
っしゅ	ッシュ	sshu		っしゅう	sshuu		っしゅ
っしょ	ッショ	ssho		っしょう/っしょお	sshoo		っしょ
っにゃ	ッニャ	tnya		っにゃあ	tnyaa		っにゃ
っにゅ	ッニュ	tnyu		っにゅう	tnyuu		っにゅ
っにょ	ッニョ	tnyo		っにょう/っにょお	tnyoo		っにょ
っひゃ	ッヒャ	hhya		っひゃあ	hhyaa		っひゃ
っひゅ	ッヒュ	hhyu		っひゅう	hhyuu		っひゅ
っひょ	ッヒョ	hhyo		っひょう/っひょお	hhyoo		っひょ
っみゃ	ッミャ	tmya		っみゃあ	tmyaa		っみゃ
っみゅ	ッミュ	tmyu		っみゅう	tmyuu		っみゅ
っみょ	ッミョ	tmyo		っみょう/っみょお	tmyoo		っみょ
っりゃ	ッリャ	rrya		っりゃあ	rryaa		っりゃ
っりゅ	ッリュ	rryu		っりゅう	rryuu		っりゅ
っりょ	ッリョ	rryo		っりょう/っりょお	rryoo		っりょ
っぎゃ	ッギャ	ggya		っぎゃあ	ggyaa		っぎゃ
っぎゅ	ッギュ	ggyu		っぎゅう	ggyuu		っぎゅ
っぎょ	ッギョ	ggyo		っぎょう/っぎょお	ggyoo		っぎょ
っじゃ	ッジャ	jja		っじゃあ	jjaa		っじゃ
っじゅ	ッジュ	jju		っじゅう	jjuu		っじゅ
っじょ	ッジョ	jjo		っじょう/っじょお	jjoo		っじょ
っぢゃ	ッヂャ	jja		っぢゃあ	jjaa		っぢゃ
っぢゅ	ッヂュ	jju		っぢゅう	jjuu		っぢゅ
っぢょ	ッヂョ	jjo		っぢょう/っぢょお	jjoo		っぢょ
っびゃ	ッビャ	bbya		っびゃあ	bbyaa		っびゃ
っびゅ	ッビュ	bbyu		っびゅう	bbyuu		っびゅ
っびょ	ッビョ	bbyo		っびょう/っびょお	bbyoo		っびょ
っぴゃ	ッピャ	ppya		っぴゃあ	ppyaa		っぴゃ
っぴゅ	ッピュ	ppyu		っぴゅう	ppyuu		っぴゅ
っぴょ	ッピョ	ppyo		っぴょう/っぴょお	ppyoo		っぴょ

Development of Thai-Japanese e-Dictionary

Krisada Visavateeranon, Surapan Meknavin, Datchakorn Tancharoen,
Wanwimon Rungtheera, Sano Shuichi, Thitisorn Sang-urai,
Somjai Sae-ung, Bandhit Rojarayanont, Natthaphon Jarat,
Saleela Boontrakulsuk, Bancha Borikud

Thai-Nichi Institute of Technology,
1771/1 Pattanakarn Rd., Suanluang, Bangkok 10250 Thailand
Krisada.v@tni.ac.th
and Makoto Minegishi
Research Institute for Languages and Cultures of Asia and Africa,
Tokyo University of Foreign Studies
3-11-1 Asahi-cho, Fuchu-shi, Tokyo, 183-8534, Japan

Abstract--**The purpose of this project is to create a Thai-Japanese dictionary for students at Thai-Nichi Institute of Technology (TNI) who are studying Japanese language and technical Japanese words. The dictionary's database initially contains about 50,000 Japanese words which were created from Tokyo University of Foreign Studies (TUFS)'s XML database of Japanese-Thai dictionary, the Technology Promotion Association (Thai-Japan) (TPA)'s Japanese-Thai Technical words dictionary and Thai – Japanese Tourism dictionary. The structure of the database is divided to multiple fields including Romaji, Kana, Kanji, Thai, English, synonym, sample sentence, parts, field, etymology and style. Our e-dictionary uses the TNI's website for testing the system while at the same time provide service to the students and general Thai and Japanese users. Accessing of Japanese word can be done by enter Romaji, Kana or Kanji. Accessing of Thai words can be done by enter Thai characters. The development of new Thai and Japanese words is done by a group of researchers, staffs and students which involve in proposing, editing and approving of new words. More than 3,000 technical words in Engineering, Information Technology and Business Administration field are being added to the database via TNI web service.**

Keywords---**Thai-Japanese dictionary, Japanese-Thai dictionary, e-Dictionary, online dictionary**

I. INTRODUCTION

The development of a Thai - Japanese electronic dictionary is a joint research between Thai - Nichi Institute of Technology (TNI) and Tokyo University of Foreign Studies (TUFS). TUFS's research program on "Corpus-based Linguistics and Language Education" which is one of the Global Center of Excellent (COE) programs conducted under the auspices of the Japanese Ministry of Education Culture, Sports, Science and Technology (MEXT). The objective of this research project is to create a Thai-Japanese and Japanese-Thai electronic online dictionary for educational purpose and give service to TNI students and TPA's member who are studying Japanese and Japanese people who are learning Thai language. In this paper, we first describe the objectives and strategy of the project. Second, on the digitization of Thai-Japanese dictionary, the detail of e-dictionary database. We will then discuss about the further development of e-dictionary.

II. THAI-JAPANESE ONLINE DICTIONARY

The advent of the internet and the increasing popularity of the web have altered many aspect of language usage. At present, an online dictionary plays an essential role as a tool of understanding and studying foreign language. Many Thai are now learning Japanese and at the same time some Japanese are studying Thai language. The online Thai-Japanese dictionary is helpful for learning of both languages. Some online Thai-Japanese dictionaries are in services, for example Longdo [1], JTDIC [2], TUFS's Thai-Japanese and Japanese-Thai dictionary [3], and Nagaoka University of Technology's Multilingual Terminology Dictionary (NUT-MTD) [4] which is a web-based dictionary for students and engineers.

TUFS has completed creating the XML database of Tomita (2003) [5]. The number of Thai and Japanese headwords is approximately 20,000 each, which should suffice for foreign language learners of both the languages. The Japanese-Thai and Thai-Japanese dictionary is now accessible to the public via the following URLs [6].

http://cblle.tufs.ac.jp/dic/th/jath/
http://cblle.tufs.ac.jp/dic/th/thja/

As the original dictionaries are designed to help Thai people to learn Japanese, which is possible for Thai people to search for Japanese words using Japanese Hiragana, Kanji, or Romaji. On the other hand, Japanese learners can only search for Thai words in the Thai-Japanese part of the dictionary using Thai characters.

Since the original dictionary is designed to help Thai people to learn Japanese, thus it is not sufficient for Japanese who wish to learn Thai. In order to apply the data to both Thai and Japanese language learning, TUFS plan to add information that is useful for Japanese people, such as pronunciation and grammatical category of Thai words. By adding information on etymology or loan words from Pali- Sanskrit,

Khmer , Chinese, English, etc., is important for not only learning the language , but also for aiding further understanding of the cultural background of the society where the language is spoken.

III. DICTIONARY DATABASE

Creation of TNI dictionary database started from converting TUFS's dictionary which was available in form of a XML data to a database form. As the result from the conversion, two dictionary's databases were created. They are Thai – Japanese dictionary database for Japanese user and Japanese -Thai database for Thai user. The field of database consists of Romaji, Kana, and Kanji to describe Japanese words, Thai and phonetic for Thai words, English meaning of the words to understand Katakana words. Other fields which are essential for the language learner are also included in the database such as synonym , etymology, field, part of speech, style, and sample sentence. The database structure of both database is shown in Fig. 1

More data was added to the two databases by digitizing two TPA dictionaries. They are Japanese-Thai dictionary for technician [7] and Thai-Japanese dictionary for traveling [8]. The electronics data are then added to the dictionary database. The number of words in each dictionary is approximately 25,000 words.

IV. WEB SERVICE OF E-DICTIONARY

TNI e-Dictionary system is developed to be web service of the Thai-Japanese dictionary database for TNI student, TPA member and general user including Japanese user in Japan. Thai - Japanese dictionary is now accessible to the public via the following URLs [9].

http//dict.thainichi.net

The user of dictionary can access the database via the internet. The system help Thai student to search for a Japanese word in the dictionary to get the meaning in Thai and English and Japanese user can search to find the meaning of Thai words. The user of the system will be divided into 4 categories: general user, member, approver and system administrator.

General User: user who search for the meaning of a word.

Member: user who registered to system and can propose new words or edit the existing words in the dictionary database.

Approver: linguistics specialist who is assigned by the system to approve new words or edited words proposed by member.

System Administrator: person who does administration work, assign new member and approver, update the dictionary database, get all statistic data of the system usage and improve the system performance.

Structure of Japanese-Thai dictionary

dic_item ID	Romaji	Kana	Kanji	Thai	English	Synonym	Etymology	Field	Parts	See	Sample sentence
รหัสคำศัพท์	โรมาจิ	ฮิรางานะ	คันจิ	ไทย	อังกฤษ	คำเหมือน	ที่มาของคำ	สาขา	ชนิด	อ้างอิง	ตัวอย่างประโยค
tmjt00049400	anzen	あんぜん	安全	ปลอดภัย	safety			ind	n		安全弟一 ปลอดภัยไว้ก่อน
tmjt00049400_00300	anzen-chitai	あんぜんちたい	安全地帯	เขตปลอดภัย	safety zone			ind	n		

Structure of Thai-Japanese dictionary

dic_item ID	Thai	Phonetic	Romaji	Kanij(Hiragana)	English	Synonym	Etymology	Field	Parts	Style
รหัสคำศัพท์	ไทย	การออกเสียง	โรมาจิ	คันจิ	อังกฤษ	คำเหมือน	ที่มาของคำ	สาขา	ชนิด	
tmtj00108100	ข้าว		indo-ine	インドイネ	rice			pln	n	
tmtj00108100_00300	ข้าวของ		nimotsu	荷物 (にもつ)	thing	สัมภาระ		comm.	n	

Fig 1. The structure of Thai-Japanese and Japanese-Thai dictionaries.

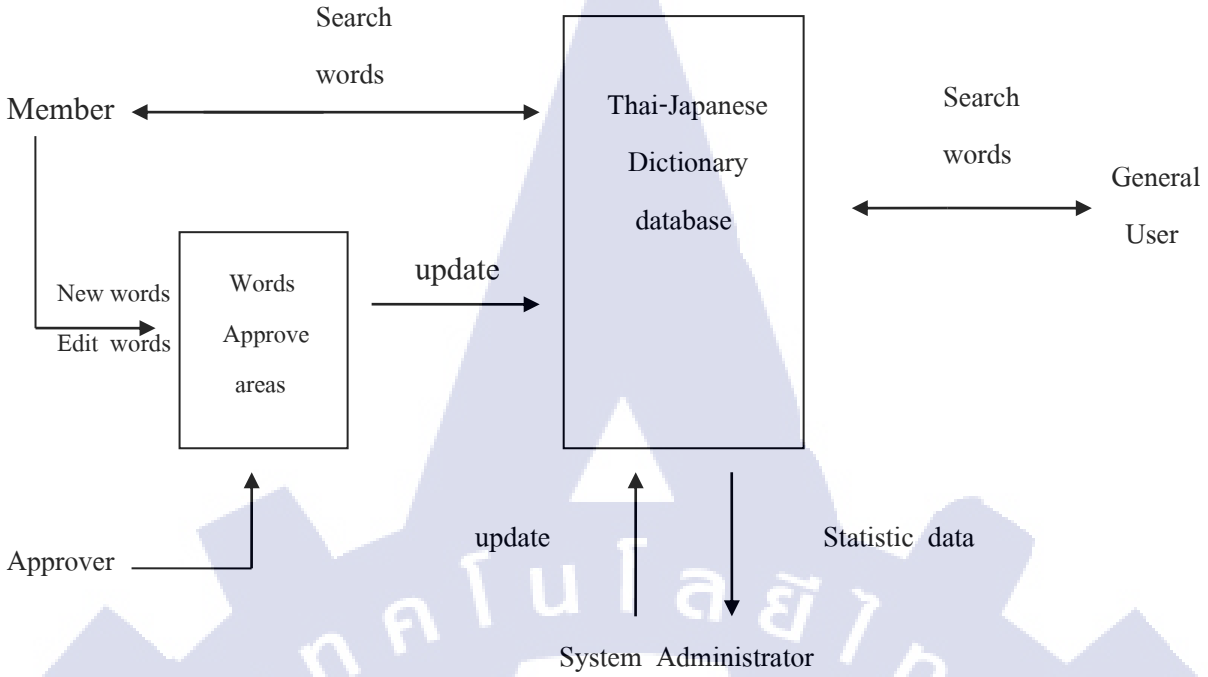


Fig. 2 Users of the TNI's e-Dictionary system

To search for a Japanese word, the word entry can be either Romaji, Hiragana, Katakana or Kanji. Single word and compound word can be searched by a single word entry, all the word will be displayed as the result. Compound words can be searched by using single word entry and select prefix or suffix search. Found words will be displayed on the screen in Romaji, Kana, Kanji, Thai meaning of the words and English meaning. More information of each word can be displayed by simply clicking at the selected word on the screen, the information are field, part of speech, antonym, and sample sentence.

To help the Japanese learner, when entry search in Romaji, all combination of short and long pronunciation will be displayed, for example 'gakuko' searching will result in 'gakuko' and gakukoo'. Some of the word has same pronunciation but different Romaji writing such as 'ou', 'oo' and 'ee' and 'ei', the e-Dictionary will search and display all case for the user

New words can be added to dictionary database by member proposing of new words or editing of existing words in the database. All new words will be recorded in the words approve area waiting to get approve from the approver who are linguistic specialist. The approver may edit the data to get more accuracy of the meaning. When a word is approved, it will be written into the dictionary database for further access by general user. Fig.4 and Fig.5 show the sample screen of adding new word and approve of word by approver.

Fig.3 Example of Japanese word search result

Fig. 4 Example of adding new words by a member

Currently the beta version of TNI e-Dictionary system is on service to general user since mid 2009. The number of search is roughly 1,000 search per day. Statistic data on the usage of the system will be simultaneously reported when requested by the member. The reported data are including the latest search, popular words, unfound words and most contribution member. By investigating these data, new words will be added to the system and the new service for the user and member will be improved and developed.

V. FUTHER DEVELOPMENT.

In this paper, we have described the efforts at the initial stage to develop a Thai-Japanese e-Dictionary and put in service to test the performance of the system and develop the method to input more words to the database. The purpose of the project is to create a Thai - Japanese and Japanese – Thai electronic online dictionary for educational purpose and give service to TNI students and TPA's member who are studying Japanese and

Fig. 5 Example of approving new words or edited words

Japanese people who are learning Thai language. Further development is to add more words especially technical words in the field of Engineering, Information Technology and Business Administration which are the major of students at TNI. The ultimate objective of the project is to promote language education and linguistic analysis in both Thailand and Japan. Since the demands of linguists, language educators, and language learners are going to continue increasing, more cooperation between Thai and Japanese Universities in making the Thai-Japanese and Japanese-Thai dictionary is necessary.

REFERENCES

[1] Londo, http://www.longdo.com
[2] JTDIC, http://www.jtdic.com//jtdic.aspx/
[3] TUFS,http:/www.cblle.tufs.ac.jp./dic/th/jath/
[4] NUT-MTD, http://gii.nagaokaut.ac.jp/mtd/
[5] Tomita Takejiro(2003): Japanese-Thai and Thai-Japanese Dictionary for Thai People, Revised Edition, 367pp+270pp, Bangkok: Association for Thai Education and Cultural Promotion.
[6] Minegishi Makoto and Akagi Osamu (2009): Development of and Electronic Dictionary for Analyzing Linguistic Data.. Proceedings of Chula-Japan Linguistics Conference 2008, pp.11-16, Tokyo University of Foreign Studies, Tokyo
[7] Klanhan Voraputiporn, Phulporn Saenbanpra (1990): Japanese-Thai Technical words dictionary. Technology Promotion Association (Thai-Japan) press.
[8] Pranee Chongsucharitkul (2006): Thai_Japanese Tourism words dictionary. Technology Promotion Association (Thai-Japan) press.
[9] TNI e-Dictionary: http//dict.thainichi.net

สถาบันเทคโนโลยีไทย - ญี่ปุ่น
THAI - NICHI INSTITUTE OF TECHNOLOGY
TNI