



รายงานวิจัยฉบับสมบูรณ์

การประเมินคุณภาพการแปลบทคัดย่อวิทยานิพนธ์
จากภาษาไทยเป็นภาษาอังกฤษโดยมนุษย์และ ChatGPT
มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

Evaluating the Quality of Thai - to - English Thesis Abstract Translations by
Humans and ChatGPT at Rajamangala University of Technology Thanyaburi

ฉันทลักษณ์ สารชาติ
สำนักบัณฑิตศึกษา

ได้รับการสนับสนุนจากกองทุนส่งเสริมงานวิจัย
มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ประจำปี พ.ศ. 2568

ชื่องานวิจัย	การประเมินคุณภาพการแปลบทคัดย่อวิทยานิพนธ์จากภาษาไทย เป็นภาษาอังกฤษโดยมนุษย์และ ChatGPT มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ชื่อผู้วิจัย	นางสาวธันัญลักษณ์ สารชาติ
หน่วยงาน	สำนักบัณฑิตศึกษา มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ปีที่ทำการวิจัย	2568

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อ 1) ประเมินคุณภาพการแปลบทคัดย่อภาษาไทย เป็นภาษาอังกฤษ ตามเกณฑ์ในด้าน Accuracy (ความถูกต้องของความหมาย) Fluency (ความคล่องแคล่วทางภาษา) และ Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) ของเกณฑ์การประเมิน Multidimensional Quality Metrics (MQM) โดยพิจารณาความแตกต่างของคุณภาพระหว่างการแปลโดยมนุษย์ และการแปลโดย ChatGPT 2) เปรียบเทียบคุณภาพการแปล ระหว่างการแปลโดยมนุษย์และการแปลโดย ChatGPT จากคะแนนการประเมินของผู้เชี่ยวชาญ ด้านภาษาอังกฤษ และ 3) สังเคราะห์ความคิดเห็นของผู้เชี่ยวชาญเกี่ยวกับจุดเด่นและข้อจำกัด ของการแปลทั้งสองรูปแบบ

การวิจัยครั้งนี้ เป็นการวิจัยแบบผสมผสาน แหล่งข้อมูลประกอบด้วยบทคัดย่อวิทยานิพนธ์ ระดับปริญญาโท ของมหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี ปีการศึกษา 2567 จำนวน 10 เรื่อง ซึ่งคัดเลือกแบบเจาะจง จากต้นฉบับภาษาไทยเดียวกัน และจัดทำเป็นบทแปลภาษาอังกฤษ 2 รูปแบบ ได้แก่ การแปลโดยมนุษย์และการแปลโดย ChatGPT ผู้ให้ข้อมูลสำคัญ คือ ผู้เชี่ยวชาญด้าน ภาษาอังกฤษ จำนวน 18 คน เครื่องมือที่ใช้ในการวิจัย ได้แก่ แบบประเมินคุณภาพการแปลตามเกณฑ์ MQM และแบบสอบถามความคิดเห็นของผู้เชี่ยวชาญ การวิเคราะห์ข้อมูลเชิงปริมาณใช้สถิติเชิง พรรณนาและการทดสอบความแตกต่างด้วยสถิติ Wilcoxon Signed - Rank Test ส่วนข้อมูลเชิง คุณภาพใช้การวิเคราะห์เนื้อหา

ผลการวิจัย พบว่า 1) ผลการประเมินคุณภาพการแปลตามเกณฑ์ Multidimensional Quality Metrics (MQM) พบว่า การแปลโดย ChatGPT มีค่าเฉลี่ยคุณภาพการแปลโดยรวม อยู่ในระดับ ดีเยี่ยม ขณะที่การแปลโดยมนุษย์มีค่าเฉลี่ยอยู่ในระดับดี โดย ChatGPT มีคะแนน สูงกว่ามนุษย์ในทุกองค์ประกอบ ได้แก่ Accuracy (ความถูกต้องของความหมาย), Fluency (ความคล่องแคล่วทางภาษา) และ Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) 2) ผลการเปรียบเทียบคุณภาพการแปลระหว่างมนุษย์และการแปลโดย ChatGPT พบว่า คะแนน

คุณภาพการแปลของทั้งสองรูปแบบ แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ในทุกองค์ประกอบของเกณฑ์ MQM และคะแนนรวม โดยคะแนนของ ChatGPT สูงกว่าการแปลโดยมนุษย์อย่างสม่ำเสมอ และ 3) ผลการสังเคราะห์ความคิดเห็นของผู้เชี่ยวชาญ พบว่า การแปลโดย ChatGPT มีจุดเด่นด้านการสร้างประโยคภาษาอังกฤษที่เป็นระบบ มีความลื่นไหลของภาษา และมีความสม่ำเสมอในการใช้ศัพท์เฉพาะ อย่างไรก็ตาม ยังมีข้อจำกัดด้านความคลาดเคลื่อนของความหมายในบางบริบท จึงจำเป็นต้องมีการตรวจทานและปรับแก้โดยมนุษย์ เพื่อให้เหมาะสมกับมาตรฐานการเขียนบทความวิจัยวิชาการ สำหรับผลการวิจัยสามารถนำไปใช้เป็นแนวทางในการพัฒนากระบวนการแปลบทความวิทยานิพนธ์จากภาษาไทยเป็นภาษาอังกฤษของนักศึกษา ระดับบัณฑิตศึกษา รวมถึงใช้เป็นข้อมูลประกอบการกำหนดแนวปฏิบัติในการใช้เครื่องมือปัญญาประดิษฐ์เพื่อสนับสนุนการแปลเชิงวิชาการ ควบคู่กับการตรวจทานโดยผู้เชี่ยวชาญ เพื่อให้การเขียนบทความภาษาอังกฤษมีคุณภาพและสอดคล้องกับมาตรฐานทางวิชาการมากยิ่งขึ้น

คำสำคัญ: การแปลบทความวิจัย การแปลโดยมนุษย์ การแปลโดย ChatGPT เกณฑ์ MQM



Research Title	Evaluating the Quality of Thai - to - English Thesis Abstract Translations by Humans and ChatGPT at Rajamangala University of Technology Thanyaburi
Researcher's Name	Miss Thanyaphak Sarachart
Institute	Office of Graduate Studies, Rajamangala University of Technology Thanyaburi
Year of Research	2025

ABSTRACT

This study aimed to: 1) evaluate the quality of Thai-to-English translations of thesis abstracts based on the Multidimensional Quality Metrics (MQM), focusing on Accuracy, Fluency, and Terminology Consistency, by comparing human translation and ChatGPT translation; 2) compare the translation quality between human translation and ChatGPT translation based on evaluations by English language experts; and 3) synthesize expert opinions regarding the strengths and limitations of both translation approaches.

This study employed a mixed-methods research approach. The data sources consisted of ten master's thesis abstracts from Rajamangala University of Technology Thanyaburi in the academic year 2024, which were purposively selected from the same original Thai texts. These abstracts were translated into English in two versions, namely human translation and ChatGPT translation. The key informants were 18 experts in English language. The research instruments included a translation quality assessment form based on the Multidimensional Quality Metrics (MQM) framework and a questionnaire on experts' opinions. Quantitative data were analyzed using descriptive statistics and the Wilcoxon Signed-Rank Test, while qualitative data were analyzed using content analysis.

The results revealed that: 1) the evaluation of translation quality based on the Multidimensional Quality Metrics (MQM) criteria indicated that the translation produced by ChatGPT had an overall mean translation quality at the excellent level, whereas the human translation had a mean quality at the good level. ChatGPT obtained

higher scores than human translation in all components, namely Accuracy (correctness of meaning), Fluency (language fluency), and Terminology Consistency (consistency of terminology). 2) The comparison of translation quality between human translation and ChatGPT translation showed that the translation quality scores of the two methods were significantly different at the .05 level in all components of the MQM criteria and in the overall score, with ChatGPT consistently receiving higher scores than human translation, and 3) the synthesis of expert opinions revealed that ChatGPT translation demonstrated strengths in producing systematically structured English sentences, providing fluent language use, and maintaining consistent terminology. However, there were still some limitations related to meaning inaccuracies in certain contexts. Therefore, human review and revision are necessary to ensure that the translated abstracts conform to academic abstract writing standards. The findings of this research can be used as guidelines for developing the process of translating thesis abstracts from Thai into English for graduate students, as well as supporting the establishment of practical guidelines for using artificial intelligence tools to assist academic translation together with expert review, in order to improve the quality of English abstract writing and ensure compliance with academic standards.

Keywords: abstract translation, human translation, ChatGPT translation, MQM criteria



กิตติกรรมประกาศ

การวิจัยฉบับนี้สำเร็จลุล่วงไปได้ด้วยดี ด้วยความตระหนักถึงความสำคัญของการวิจัยในการพัฒนาการศึกษาระดับบัณฑิตศึกษา ผู้วิจัยขอขอบพระคุณมหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี ที่ให้การสนับสนุนทุนวิจัยประจำปีงบประมาณ พ.ศ. 2568 ซึ่งเป็นปัจจัยสำคัญที่เอื้อให้การดำเนินการวิจัยเป็นไปอย่างมีประสิทธิภาพ

ผู้วิจัยขอขอบพระคุณศาสตราจารย์ ดร.พิเชษฐ ลิ้มสุวรรณ ที่กรุณาให้คำแนะนำ แนวคิด และข้อเสนอแนะอันเป็นประโยชน์ต่อการดำเนินการวิจัยมาโดยตลอด ขอขอบพระคุณผู้ช่วยศาสตราจารย์ ดร.สโรชรัตน์ ธราพร และอาจารย์ประภาพร เล็กดำรงศักดิ์ ที่กรุณาให้คำแนะนำและแนวทางในการดำเนินการวิจัย รวมถึงนายรุ่งโรจน์ สุทธิสุข บุคลากรชำนาญการพิเศษ ที่กรุณาให้คำปรึกษาและข้อเสนอแนะที่เป็นประโยชน์ต่อการวิจัยฉบับนี้

นอกจากนี้ ผู้วิจัยขอขอบพระคุณรองศาสตราจารย์ ดร. พิมพันธ์ ภูมิภิตติพิชญ์ ผู้ช่วยศาสตราจารย์ ดร. สุรศักดิ์ โจถาวร และ ดร. เบญจวรรณ รุ่งเรืองสุภรัตน์ ที่กรุณาเป็นผู้ทรงคุณวุฒิในการให้คำแนะนำเกี่ยวกับการสร้างเครื่องมือวิจัย ตลอดจนสละเวลาอันมีค่าในการให้ข้อมูลและข้อเสนอแนะอันทรงคุณค่า ซึ่งมีส่วนสำคัญในการยกระดับคุณภาพของงานวิจัยฉบับนี้

สุดท้ายนี้ ผู้วิจัยหวังเป็นอย่างยิ่งว่างานวิจัยฉบับนี้จะเป็นประโยชน์ต่อการพัฒนาการจัดการศึกษาระดับบัณฑิตศึกษาของนักศึกษา อาจารย์ หน่วยงาน และมหาวิทยาลัย หากมีข้อบกพร่องหรือข้อผิดพลาดประการใด ผู้วิจัยขอน้อมรับไว้และจะนำไปใช้เป็นแนวทางในการปรับปรุงและพัฒนางานวิจัยให้ดียิ่งขึ้นต่อไป

นางสาวธัญลักษณ์ สารชาติ

ผู้วิจัย

มกราคม 2568

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ก
บทคัดย่อภาษาอังกฤษ.....	ค
กิตติกรรมประกาศ.....	จ
สารบัญ.....	ฉ
สารบัญตาราง.....	ช
สารบัญภาพ.....	ฌ
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของการวิจัย.....	2
1.3 คำถามงานวิจัย.....	3
1.4 สมมติฐานงานวิจัย.....	3
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	3
1.6 ขอบเขตการวิจัย.....	4
1.7 กรอบแนวคิดการวิจัย.....	5
1.8 ข้อยกเว้นการวิจัย.....	5
1.9 คำนิยามศัพท์เฉพาะ.....	6
บทที่ 2 แนวคิดทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	9
2.1 เกณฑ์การประเมินคุณภาพการแปล Multidimensional Quality Metrics (MQM)	9
2.2 การแปลบทคัดย่อจากภาษาไทยเป็นภาษาอังกฤษ.....	17
2.3 การแปลโดยมนุษย์.....	19
2.4 การแปลโดย ChatGPT.....	20
2.5 การประยุกต์ใช้เกณฑ์ MQM กับการวิเคราะห์คุณภาพการแปลภาษาไทย เป็นภาษาอังกฤษ.....	23
2.6 เทคนิคการควบคุมคุณภาพการแปล.....	25
2.7 งานวิจัยที่เกี่ยวข้องในประเทศไทย.....	26
2.8 งานวิจัยที่เกี่ยวข้องในต่างประเทศ.....	29
บทที่ 3 วิธีดำเนินการวิจัย.....	33
3.1 รูปแบบการวิจัย.....	33
3.2 แหล่งข้อมูลและผู้ให้ข้อมูล.....	34
3.3 ตัวแปรการวิจัย.....	35

สารบัญ (ต่อ)

	หน้า
3.4 เครื่องมือที่ใช้ในการวิจัย.....	35
3.5 การสร้างและตรวจสอบคุณภาพของเครื่องมือวิจัย	37
3.6 การเก็บรวบรวมข้อมูล.....	41
3.7 การวิเคราะห์ข้อมูล	42
บทที่ 4 ผลการวิจัย.....	44
4.1 ผลการประเมินคุณภาพการแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษตามเกณฑ์ MQM.....	44
4.2 ผลการเปรียบเทียบคุณภาพการแปลระหว่างการแปลโดยมนุษย์ และการแปลโดย ChatGPT	46
4.3 ผลการสังเคราะห์ความคิดเห็นของผู้เชี่ยวชาญเกี่ยวกับการแปลบทคัดย่อโดยมนุษย์ และการแปลโดย ChatGPT	48
4.4 สรุปผลการวิจัยในภาพรวม	56
บทที่ 5 สรุปผลการวิจัย อภิปรายผลและข้อเสนอแนะ	59
5.1 สรุปผลการวิจัย.....	59
5.2 อภิปรายผลการวิจัย	61
5.3 ข้อเสนอแนะในการวิจัย	63
5.4 สรุปประโยชน์ที่ได้จากการวิจัย	66
บรรณานุกรม.....	68
ภาคผนวก.....	73
ภาคผนวก ก เครื่องมือที่ใช้ในการวิจัย	74
ภาคผนวก ข ผลการวิเคราะห์ค่าดัชนีความสอดคล้อง IOC	78
ภาคผนวก ค ตารางข้อมูลดิบข้อเสนอแนะผู้เชี่ยวชาญตอบคำถามปลายเปิด 4 ข้อ.....	84
ภาคผนวก ง ตารางข้อมูลดิบความเห็นเพิ่มเติมของผู้เชี่ยวชาญในการประเมินบทคัดย่อรายฉบับ	108
ประวัติผู้วิจัย.....	131

สารบัญตาราง

ตารางที่	หน้า
2.1 การสังเคราะห์แนวคิดและการนำไปใช้ในการวิจัย	32
4.1 แสดงค่าเฉลี่ย (μ) ส่วนเบี่ยงเบนมาตรฐาน (σ) และระดับคุณภาพการแปลโดยภาพรวม และรายด้าน ของมนุษย์และ ChatGPT (N = 10 บทคัดย่อ)	45
4.2 ผลการเปรียบเทียบความแตกต่างของคุณภาพการแปลระหว่างมนุษย์และ ChatGPT	47



สารบัญภาพ

ภาพที่	หน้า
1.1 กรอบแนวคิดการวิจัย.....	5
4.1 แสดงแผนภาพสรุปผลการวิจัยในภาพรวม	57
5.1 แสดงแผนภาพสรุปประโยชน์จากการวิจัยเรื่องการประเมินคุณภาพการแปลบทคัดย่อ ภาษาไทย – ภาษาอังกฤษด้วย MQM.....	67



บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ปัจจุบันมหาวิทยาลัยในประเทศไทยส่วนใหญ่กำหนดให้นักศึกษาระดับบัณฑิตศึกษาต้องเผยแพร่งานวิจัยในฐานข้อมูลวิชาการระดับชาติหรือระดับนานาชาติเป็นเงื่อนไขในการสำเร็จการศึกษา โดยภาษาอังกฤษเป็นภาษาหลักที่ใช้ในการเผยแพร่ผลงานวิจัยถือเป็นปัจจัยกำหนดคุณภาพและการยอมรับในระดับสากล Stockemer และ Wigginton (2019) ชี้ว่าการเผยแพร่งานวิจัยเป็นภาษาอังกฤษช่วยให้ผลงานได้รับการเข้าถึงและอ้างอิงอย่างกว้างขวาง ขณะที่ Alspach (2017) อธิบายว่าบทความย่อเป็นส่วนแรกๆ ที่ผู้อ่านใช้ในการประเมินคุณค่าและความน่าสนใจของงานวิจัย ดังนั้นบทความย่อภาษาอังกฤษที่ถูกต้องตามหลักไวยากรณ์ ชัดเจนด้านความหมาย และสื่อสารสาระสำคัญของงานได้อย่างครบถ้วนจึงมีความสำคัญต่อคุณภาพงานวิจัยโดยรวม มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรีได้มีมาตรการกำกับคุณภาพบทความย่อภาษาอังกฤษอย่างเป็นทางการโดยสภาวิชาการได้แต่งตั้งคณะกรรมการกลั่นกรองเพื่อควบคุมคุณภาพการเขียนบทความย่อและบทความวิจัยสำหรับนักศึกษาระดับบัณฑิตศึกษา (สภาวิชาการมหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี, 2559) อย่างไรก็ตาม รายงานผลการตรวจสอบบทความย่อของสำนักบัณฑิตศึกษา ประจำปีการศึกษา 2565 ระบุว่า ผู้สำเร็จการศึกษาระดับปริญญาโทจำนวน 144 ราย มีถึง 111 ราย หรือคิดเป็นร้อยละ 77.08 ที่ต้องแก้ไขบทความย่อภาษาอังกฤษก่อนเผยแพร่ (มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี สำนักบัณฑิตศึกษา, 2565) ซึ่งสะท้อนว่าปัญหาการเขียนบทความย่อภาษาอังกฤษยังเป็นประเด็นที่ต้องให้ความสำคัญอย่างเร่งด่วน

ผู้วิจัยในฐานะนักวิชาการศึกษาที่ทำหน้าที่ควบคุมคุณภาพมาตรฐานการศึกษา และตรวจสอบบทความย่อภาษาอังกฤษของนักศึกษาระดับบัณฑิตศึกษา พบว่า ข้อผิดพลาดในการเขียนบทความย่อยังคงเกิดขึ้นซ้ำอย่างต่อเนื่อง โดยงานของธัญลักษณ์ สารชาติ (2566) ซึ่งศึกษาข้อผิดพลาดในบทความย่อภาษาอังกฤษของนักศึกษาปริญญาโท มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี พบว่า ข้อผิดพลาดระดับคำมีสัดส่วนสูงถึงร้อยละ 61.24 มากกว่าข้อผิดพลาดระดับประโยคที่ร้อยละ 38.76 และยังพบว่าสาขามนุษยศาสตร์และสังคมศาสตร์มีค่าเฉลี่ยข้อผิดพลาดสูงกว่าสาขาวิทยาศาสตร์และเทคโนโลยี ซึ่งผู้เชี่ยวชาญด้านภาษาเห็นพ้องว่าปัญหาสำคัญส่วนหนึ่งมาจากการพึ่งพาโปรแกรมแปลภาษาอัตโนมัติ (Machine Translation: MT) ซึ่งมักไม่สามารถรองรับความซับซ้อนเชิงโครงสร้างของภาษาไทยได้อย่างเหมาะสม โดยเฉพาะการไม่มีการเว้นวรรคระหว่างคำ ความหมายที่อาศัยบริบทสูง และความแตกต่างด้านลักษณะทางไวยากรณ์ ทำให้ผลลัพธ์จากการแปลมักคลาดเคลื่อนทั้งด้านรูปแบบและความหมาย ซึ่งงานของ Patanasorn (2017) ระบุว่าบทความย่อที่แปลด้วยโปรแกรม MT โดยทั่วไปมีคุณภาพเพียงระดับปานกลางและไม่เหมาะสมแก่การเผยแพร่หากไม่มีการตรวจทาน และแก้ไขความถูกต้องด้านความหมาย ไวยากรณ์ และศัพท์เฉพาะโดยมนุษย์ ขณะที่ Pewnim (2020) ย้ำว่าความคลาดเคลื่อนจากโครงสร้างคู่ภาษายังคงเป็นอุปสรรคสำคัญที่ทำให้ MT ไม่สามารถให้ผลลัพธ์ที่ถูกต้องและสอดคล้องกับงานวิชาการได้อย่างสมบูรณ์

ในช่วงไม่กี่ปีที่ผ่านมา เทคโนโลยีปัญญาประดิษฐ์เชิงภาษาได้พัฒนาไปอย่างก้าวกระโดด โดยเฉพาะแบบจำลองภาษาขนาดใหญ่ (Large Language Models) เช่น ChatGPT ซึ่งมีความสามารถด้านการประมวลผลภาษาเชิงลึก สามารถแปลความหมายตามบริบท สร้างประโยคที่มีความลื่นไหลทางภาษา และตรวจจับข้อผิดพลาดบางประเภทได้ ซึ่ง Human Science Co., Ltd. (2024) รายงานว่า ChatGPT มีศักยภาพในการแปลภาษาอังกฤษได้อย่างถูกต้องและเป็นธรรมชาติยิ่งกว่าโปรแกรมแปลภาษาแบบเดิม ส่งผลให้นักวิจัยและนักศึกษาจำนวนมากเริ่มนำ ChatGPT มาใช้เป็นเครื่องมือช่วยแปลบทความภาษาอังกฤษแทน MT แบบดั้งเดิม อย่างไรก็ตาม แม้ประสิทธิภาพของ ChatGPT จะได้รับการยอมรับในระดับหนึ่ง แต่ประเด็นสำคัญที่ยังขาดหลักฐานเชิงประจักษ์ คือ ChatGPT สามารถแปลบทความเชิงวิชาการได้ “เทียบเท่า” หรือ “ดีกว่า” การแปลโดยมนุษย์หรือไม่ โดยเฉพาะในด้านความถูกต้องของความหมาย (Accuracy) ความคล่องแคล่วของภาษา (Fluency) และความสอดคล้องของศัพท์เฉพาะ (Terminology Consistency) ซึ่งเป็นองค์ประกอบสำคัญของคุณภาพบทความในงานวิจัยระดับบัณฑิตศึกษา

จากความจำเป็นอย่างยิ่ง งานวิจัยนี้จึงมุ่งตรวจสอบและเปรียบเทียบคุณภาพการแปลบทความวิทยานิพนธ์จากภาษาไทยเป็นภาษาอังกฤษที่แปลโดยมนุษย์และที่แปลโดย ChatGPT โดยใช้เกณฑ์การประเมิน Multidimensional Quality Metrics (MQM) ของ Lommel et al. (2014) ซึ่งประกอบด้วยมิติด้าน Accuracy, Fluency และ Terminology Consistency เพื่อให้ได้ข้อมูลเชิงปริมาณที่สะท้อนคุณภาพของการแปลอย่างเป็นระบบ นอกจากนี้ยังใช้การวิเคราะห์เนื้อหา (Content Analysis) จากข้อเสนอแนะของผู้เชี่ยวชาญ เพื่อทำความเข้าใจลักษณะข้อผิดพลาด จุดเด่น และข้อจำกัดของการแปลทั้งสองรูปแบบอย่างรอบด้าน ผลการศึกษาคาดว่าจะช่วยให้หน่วยงานด้านภาษาและนักศึกษาเห็นแนวทางการใช้เครื่องมือ AI อย่างเหมาะสม รวมทั้งสนับสนุนการพัฒนามาตรฐานการเขียนบทความภาษาอังกฤษของนักศึกษาระดับบัณฑิตศึกษาในอนาคต

1.2 วัตถุประสงค์ของการวิจัย

1.2.1 เพื่อประเมินคุณภาพการแปลบทความภาษาไทยเป็นภาษาอังกฤษ ตามเกณฑ์ในด้าน Accuracy (ความถูกต้องของความหมาย) Fluency (ความคล่องแคล่วทางภาษา) และ Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) ของเกณฑ์การประเมิน Multidimensional Quality Metrics (MQM) โดยพิจารณาความแตกต่างของคุณภาพระหว่างการแปลโดยมนุษย์ และการแปลโดย ChatGPT

1.2.2 เพื่อเปรียบเทียบคุณภาพการแปลระหว่างการแปลโดยมนุษย์และการแปลโดย ChatGPT จากคะแนนการประเมินของผู้เชี่ยวชาญด้านภาษาอังกฤษ

1.2.3 เพื่อสังเคราะห์ความคิดเห็นของผู้เชี่ยวชาญเกี่ยวกับจุดเด่นและข้อจำกัดของการแปลทั้งสองรูปแบบ

1.3 คำถามงานวิจัย

จากวัตถุประสงค์ของการวิจัยทั้งสามประการ งานวิจัยนี้มีคำถามการวิจัยดังต่อไปนี้

1.3.1 คุณภาพการแปลบทคัดย่อวิทยานิพนธ์จากภาษาไทยเป็นภาษาอังกฤษตามกรอบการประเมิน Multidimensional Quality Metrics (MQM) ในด้าน Accuracy (ความถูกต้องของความหมาย) Fluency (ความคล่องแคล่วทางภาษา) Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) ของฉบับที่แปลโดยมนุษย์และที่แปลโดย ChatGPT เป็นอย่างไร

1.3.2 คุณภาพการแปลบทคัดย่อวิทยานิพนธ์จากภาษาไทยเป็นภาษาอังกฤษที่แปลโดยมนุษย์ และที่แปลโดย ChatGPT มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติหรือไม่ เมื่อพิจารณาตามเกณฑ์การประเมิน MQM ทั้งสามด้าน

1.3.3 ผู้เชี่ยวชาญด้านภาษาอังกฤษมีความคิดเห็นอย่างไรเกี่ยวกับจุดเด่น จุดจำกัด และข้อเสนอแนะ ของการแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษ ระหว่างการแปลโดยมนุษย์ และการแปลโดย ChatGPT

1.4 สมมติฐานงานวิจัย

จากการดำเนินการวิจัยที่มุ่งเปรียบเทียบคุณภาพการแปลระหว่างการแปลโดยมนุษย์ และ ChatGPT งานวิจัยนี้ตั้งสมมติฐานดังต่อไปนี้

- คุณภาพการแปลบทคัดย่อวิทยานิพนธ์จากภาษาไทยเป็นภาษาอังกฤษที่แปลโดยมนุษย์ และที่แปลโดย ChatGPT มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 เมื่อพิจารณาตามองค์ประกอบของเกณฑ์การประเมิน MQM ได้แก่ Accuracy (ความถูกต้องของความหมาย) Fluency (ความคล่องแคล่วทางภาษา) Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ)

1.5 ประโยชน์ที่คาดว่าจะได้รับ

งานวิจัยนี้คาดว่าจะก่อให้เกิดประโยชน์ ดังต่อไปนี้

1.5.1 ทำให้ทราบคุณภาพและข้อจำกัดของการแปลบทคัดย่อโดยมนุษย์ และ ChatGPT ตามเกณฑ์ MQM ซึ่งเป็นประโยชน์ต่อการพัฒนามาตรฐานการแปลบทคัดย่อในระดับบัณฑิตศึกษา

1.5.2 เป็นข้อมูลเชิงประจักษ์ที่ช่วยให้หน่วยงานด้านภาษาและนักวิจัยทราบว่า ChatGPT สามารถใช้เป็นเครื่องมือช่วยแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษได้ในระดับใด และควรมีการตรวจทานเพิ่มเติมในด้านใดบ้าง

1.5.3 ช่วยให้ผู้ที่สอนที่ดูแลการเขียนวิทยานิพนธ์สามารถแนะนำแนวทางการใช้เครื่องมือ AI และการตรวจปรับแก้ภาษาที่ถูกต้องให้นักศึกษาระดับบัณฑิตศึกษาได้อย่างเหมาะสม

1.5.4 เป็นองค์ความรู้พื้นฐานสำหรับการพัฒนาระบบตรวจคุณภาพการแปลหรือโมเดลการประเมินอัตโนมัติในอนาคต โดยอ้างอิงเกณฑ์ MQM ที่เป็นมาตรฐานสากล

1.6 ขอบเขตการวิจัย

1.6.1 ขอบเขตด้านเนื้อหา

การวิจัยฉบับนี้มุ่งประเมินและเปรียบเทียบคุณภาพการแปลบทคัดย่อวิทยานิพนธ์จากภาษาไทยเป็นภาษาอังกฤษ โดยใช้เกณฑ์การประเมิน Multidimensional Quality Metrics (MQM) ใน 3 มิติหลัก ดังนี้

1) ด้าน Accuracy (ความถูกต้องของความหมาย) ได้แก่ ความครบถ้วนของเนื้อหา ความถูกต้องของข้อมูลเชิงตัวเลข หน่วยวัด และการถ่ายทอดความหมายจากต้นฉบับ

2) ด้าน Fluency (ความคล่องแคล่วทางภาษา) ได้แก่ ความถูกต้องตามหลักไวยากรณ์ ความเป็นธรรมชาติของภาษา โครงสร้างประโยค และระดับภาษาที่เหมาะสมกับงานวิชาการ

3) ด้าน Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) ได้แก่ การใช้คำศัพท์ทางวิชาการอย่างถูกต้อง สม่่าเสมอ และตรงตามบริบทของสาขาวิชา

ทั้งนี้ การประเมินคุณภาพการแปลมุ่งเน้นเฉพาะบทคัดย่อวิทยานิพนธ์ และไม่ครอบคลุมเนื้อหาในส่วนอื่นของวิทยานิพนธ์ เช่น บทนำ วิธีวิจัย หรือผลการวิจัย

1.6.2 ขอบเขตด้านแหล่งข้อมูลและผู้ให้ข้อมูลสำคัญ

1) ขอบเขตด้านแหล่งข้อมูล

แหล่งข้อมูลในการวิจัยครั้งนี้ ได้แก่ บทคัดย่อวิทยานิพนธ์ระดับปริญญาโทของมหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี ปีการศึกษา 2567 ซึ่งเขียนเป็นภาษาไทย และมีฉบับแปลภาษาอังกฤษ ผู้วิจัยคัดเลือกบทคัดย่อแบบเจาะจง (Purposive Selection) จำนวน 10 เรื่อง เพื่อนำมาใช้เป็นหน่วยข้อมูลในการประเมินคุณภาพการแปล

2) ขอบเขตด้านผู้ให้ข้อมูลสำคัญ

ผู้ให้ข้อมูลในการวิจัยครั้งนี้ คือ อาจารย์ผู้เชี่ยวชาญด้านภาษาอังกฤษ จำนวน 18 คน ซึ่งมีประสบการณ์ในการสอนภาษาอังกฤษ การแปลภาษาอังกฤษ หรือการตรวจบทคัดย่อวิทยานิพนธ์ระดับบัณฑิตศึกษา ทำหน้าที่เป็นผู้ประเมินคุณภาพการแปลตามเกณฑ์ Multidimensional Quality Metrics (MQM)

1.6.3 ขอบเขตด้านหน่วยข้อมูลและหน่วยวิเคราะห์

1) หน่วยข้อมูลหลัก ได้แก่ บทคัดย่อวิทยานิพนธ์แต่ละเรื่อง ซึ่งประกอบด้วยต้นฉบับภาษาไทย และฉบับแปลภาษาอังกฤษ 2 รูปแบบ (มนุษย์และ ChatGPT)

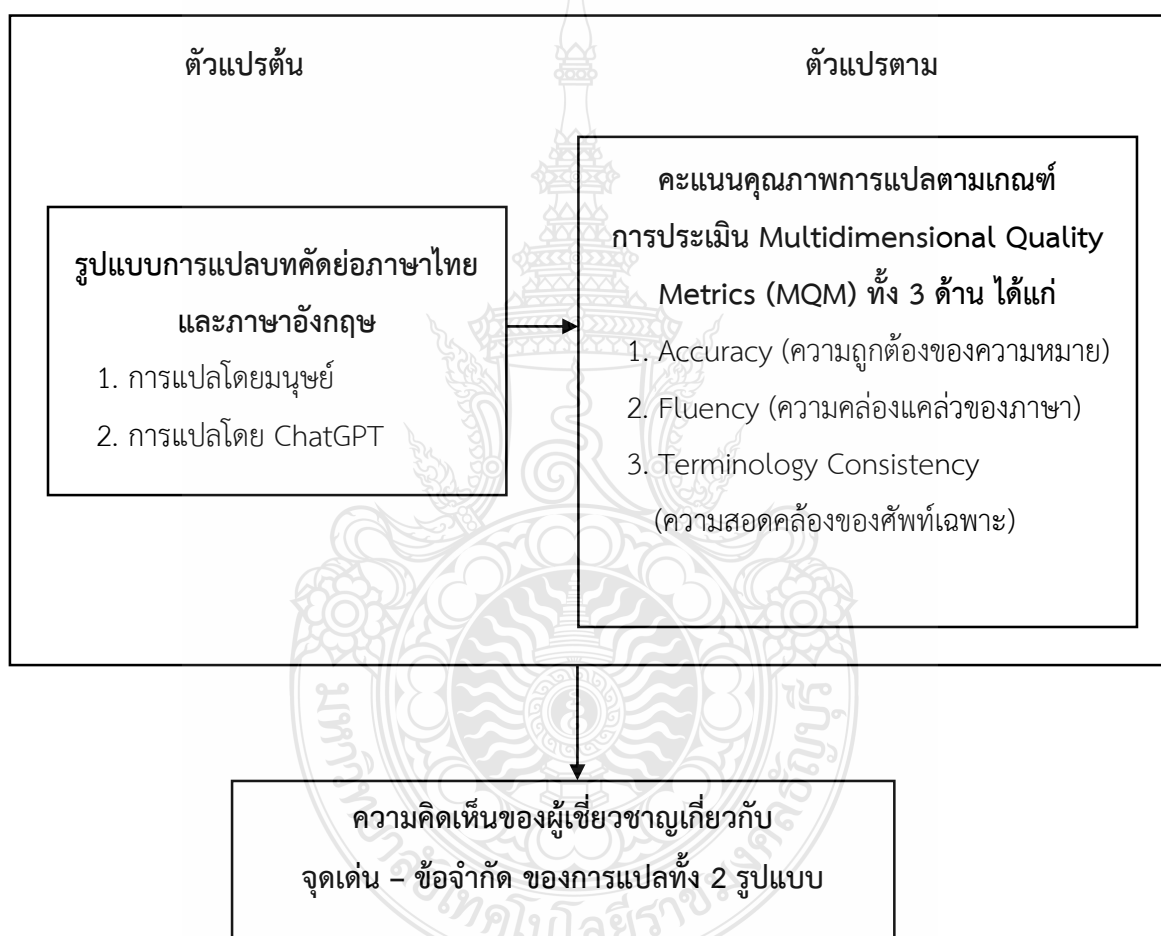
2) หน่วยวิเคราะห์เชิงปริมาณ ได้แก่ คะแนนการประเมินคุณภาพการแปลในแต่ละมิติของเกณฑ์ MQM ที่ผู้เชี่ยวชาญให้คะแนน เพื่อนำมาเปรียบเทียบความแตกต่างระหว่างการแปลโดยมนุษย์และการแปลโดย ChatGPT ด้วยสถิติ Wilcoxon Signed-Rank Test

3) หน่วยวิเคราะห์เชิงคุณภาพ ได้แก่ ข้อเสนอแนะและความคิดเห็นของผู้เชี่ยวชาญ ที่ได้จากคำถามปลายเปิดและช่องแสดงความคิดเห็น ซึ่งนำมาวิเคราะห์ด้วยวิธีการวิเคราะห์เนื้อหา (Content Analysis)

1.6.4 ขอบเขตด้านช่วงเวลาและสถานที่ศึกษา

การวิจัยดำเนินการเก็บรวบรวมข้อมูลในปีการศึกษา 2567 ณ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

1.7 กรอบแนวคิดการวิจัย



ภาพที่ 1.1 กรอบแนวคิดการวิจัย

1.8 ข้อจำกัดของการวิจัย

การวิจัยครั้งนี้มีข้อจำกัดที่ควรพิจารณาดังต่อไปนี้

1.8.1 จำนวนบทความที่ใช้ในการประเมินจำกัดเพียง 10 เรื่อง ซึ่งอาจไม่ครอบคลุมความหลากหลายของรูปแบบภาษาและเนื้อหาของบทความวิทยานิพนธ์ทั้งหมดในระดับบัณฑิตศึกษา

1.8.2 กลุ่มผู้ประเมินประกอบด้วยอาจารย์ผู้เชี่ยวชาญด้านภาษาอังกฤษจำนวน 18 คน ซึ่งแม้มีประสบการณ์ตรง แต่ยังคงมีความแตกต่างเชิงอัตนัยด้านมาตรฐานการให้คะแนนหรือมุมมองด้านคุณภาพของการแปล

1.8.3 การประเมินคุณภาพการแปลโดยใช้เกณฑ์ Multidimensional Quality Metrics (MQM) เน้นเฉพาะ 3 ด้าน ได้แก่ Accuracy (ความถูกต้องของความหมาย) Fluency (ความคล่องแคล่วทางภาษา) Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) ซึ่งอาจไม่ครอบคลุมองค์ประกอบด้านสไตล์ การสื่อความหมายเชิงบริบท และความเหมาะสมของสำนวนภาษาในภาพรวม

1.8.4 การแปลบทความภาษาไทยเป็นภาษาอังกฤษโดย ChatGPT อาจขึ้นอยู่กับเวอร์ชันของระบบและคำสั่ง (Prompt) ที่ใช้ในระยะเวลาวิจัย ทำให้ผลลัพธ์อาจแตกต่างจากเวอร์ชันในอนาคตและไม่สามารถสรุปทั่วไปได้ทั้งหมด

1.9 คำนิยามศัพท์เฉพาะ

เพื่อความเข้าใจที่ถูกต้องและสอดคล้องกันของผู้อ่าน งานวิจัยนี้กำหนดความหมายของคำนิยามศัพท์เฉพาะ ดังต่อไปนี้

การแปลโดยมนุษย์ หมายถึง การแปลบทความภาษาไทยเป็นภาษาอังกฤษโดยผู้แปลที่เป็นมนุษย์ เช่น อาจารย์ผู้เชี่ยวชาญด้านภาษา ผู้แปลมืออาชีพ หรือผู้มีความรู้ด้านการแปล โดยใช้ทักษะภาษาตามความสามารถของบุคคล ไม่ใช่ระบบปัญญาประดิษฐ์หรือเครื่องมือแปลอัตโนมัติ

การแปลโดย ChatGPT หมายถึง การแปลบทความภาษาไทยเป็นภาษาอังกฤษด้วยโปรแกรมปัญญาประดิษฐ์ “ChatGPT” โดยผู้วิจัยกำหนดข้อความป้อนเข้า (Prompt) ที่เหมือนกันในทุกบทความ เพื่อให้ได้ผลการแปลภายใต้เงื่อนไขเดียวกัน และไม่ผ่านการปรับแก้เพิ่มเติมหลังจากระบบแปลแล้ว

Multidimensional Quality Metrics (MQM) หมายถึง เกณฑ์การประเมินที่มีความครอบคลุมละเอียด และยืดหยุ่นสูง สามารถปรับใช้ได้หลายบริบท โดยในการศึกษาครั้งนี้ นักวิจัยได้นำหลักการของ MQM มาประยุกต์ใช้เป็นกรอบในการประเมินคุณภาพการแปลบทความภาษาไทยเป็นภาษาอังกฤษ โดยพิจารณาจากเกณฑ์หลัก 3 ประการ ได้แก่ Accuracy (ความถูกต้องของความหมาย) Fluency (ความคล่องแคล่วของภาษา) และ Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ)

Accuracy (ความถูกต้องของความหมาย) หมายถึง ระดับความถูกต้องแม่นยำของการแปลโดยพิจารณาว่าการแปลสามารถสื่อความหมายของต้นฉบับได้อย่างครบถ้วนและถูกต้อง โดยพิจารณาจากองค์ประกอบย่อยดังต่อไปนี้

Mistranslation (การแปลผิด) หมายถึง การถ่ายทอดความหมายของข้อความต้นฉบับคลาดเคลื่อนจากเจตนาหรือเนื้อหาที่แท้จริงของต้นฉบับ

Omission (การละเลยเนื้อหา) หมายถึง การละหรือไม่แปลส่วนของข้อความที่ปรากฏในต้นฉบับ ทำให้สาระสำคัญบางส่วนหายไป

Addition (การเพิ่มเนื้อหา) หมายถึง การเพิ่มคำหรือข้อความที่ไม่ได้ปรากฏในต้นฉบับ ซึ่งอาจทำให้ความหมายของข้อความเปลี่ยนไป

Untranslated Text (ข้อความที่ไม่ได้แปล) หมายถึง การคงข้อความบางส่วนของต้นฉบับไว้โดยไม่แปลเป็นภาษาเป้าหมาย

Fluency (ความคล่องแคล่วของภาษา) หมายถึง ระดับความเป็นธรรมชาติและความอ่านง่ายของการแปล โดยพิจารณาว่าการแปลมีความสละสลวยและเป็นไปตามหลักไวยากรณ์ของภาษาเป้าหมาย โดยพิจารณาจากองค์ประกอบย่อยดังต่อไปนี้

Punctuation (การใช้เครื่องหมายวรรคตอน) หมายถึง การใช้เครื่องหมายวรรคตอนในข้อความแปลให้ถูกต้องและเหมาะสมตามหลักภาษา

Spelling (การสะกดคำ) หมายถึง ความถูกต้องของการสะกดคำในภาษาเป้าหมาย

Grammar (ไวยากรณ์) หมายถึง ความถูกต้องของโครงสร้างประโยคและกฎไวยากรณ์ของภาษาเป้าหมาย

Register (ระดับภาษา) หมายถึง ความเหมาะสมของระดับภาษาและสำนวนที่ใช้ให้สอดคล้องกับบริบทของงานเขียนเชิงวิชาการ

Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) หมายถึง ระดับความถูกต้องและสอดคล้องในการใช้ศัพท์เฉพาะทางเทคนิค โดยพิจารณาว่ามีการใช้ศัพท์เฉพาะอย่างถูกต้องและสอดคล้องกันตลอดทั้งข้อความ โดยพิจารณาจากองค์ประกอบย่อยดังต่อไปนี้

Incorrect Term (การใช้ศัพท์ไม่ตรง) หมายถึง การเลือกใช้คำศัพท์ที่ไม่ถูกต้องหรือไม่เหมาะสมกับบริบทของเนื้อหา

Inconsistent Term (การใช้ศัพท์ไม่สอดคล้องกัน) หมายถึง การใช้คำศัพท์เฉพาะที่แตกต่างกันสำหรับความหมายเดียวกันภายในข้อความเดียวกัน ทำให้เกิดความไม่ต่อเนื่องของความหมาย

บทคัดย่อวิทยานิพนธ์ หมายถึง ส่วนสรุปเนื้อหาสำคัญของงานวิทยานิพนธ์ระดับบัณฑิตศึกษา ซึ่งประกอบด้วย วัตถุประสงค์ วิธีการดำเนินการวิจัย ผลการวิจัย และข้อค้นพบสำคัญ ใช้เป็นต้นฉบับสำหรับการแปลในการวิจัยครั้งนี้

ผู้เชี่ยวชาญด้านภาษาอังกฤษ หมายถึง อาจารย์หรือผู้ทรงคุณวุฒิที่มีประสบการณ์
ด้านการสอนภาษาอังกฤษ การตรวจสอบบทคัดย่อ หรือการประเมินคุณภาพการแปล
ในระดับอุดมศึกษา ซึ่งทำหน้าที่ให้คะแนนและให้ความเห็นต่อคุณภาพการแปลในงานวิจัยนี้



บทที่ 2

แนวคิดทฤษฎีและงานวิจัยที่เกี่ยวข้อง

การวิจัยเรื่อง “ การประเมินคุณภาพการแปลบทคัดย่อวิทยานิพนธ์จากภาษาไทย เป็นภาษาอังกฤษโดยมนุษย์ และ ChatGPT ที่มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี ” มีเป้าหมาย เพื่อศึกษาและเปรียบเทียบคุณภาพของการแปลระหว่างมนุษย์และปัญญาประดิษฐ์ โดยใช้เกณฑ์ การประเมินคุณภาพการแปล Multidimensional Quality Metrics (MQM) (Lommel et al., 2014) ในงานนี้ ผู้วิจัยได้รวบรวมและวิเคราะห์เอกสาร ตำรา บทความวิชาการ และงานวิจัยที่เกี่ยวข้อง ซึ่งครอบคลุมประเด็นเกี่ยวกับเกณฑ์การประเมินคุณภาพการแปล Multidimensional Quality Metrics (MQM) การแปลภาษาไทยเป็นภาษาอังกฤษ การแปลบทคัดย่อ การแปลโดยมนุษย์ และการแปลโดย ChatGPT ตลอดจนงานวิจัยในประเทศและต่างประเทศที่เกี่ยวข้องกับประเด็น ดังกล่าว เพื่อนำมาเป็นกรอบแนวคิดและพื้นฐานในการออกแบบการวิจัย รวมถึงใช้เป็นแนวทาง ในการวิเคราะห์และอภิปรายผลการศึกษา โดยจัดลำดับหัวข้อในการนำเสนอไว้ดังนี้

- 2.1 เกณฑ์การประเมินคุณภาพการแปล Multidimensional Quality Metrics (MQM)
- 2.2 การแปลบทคัดย่อจากภาษาไทยเป็นภาษาอังกฤษ
- 2.3 การแปลโดยมนุษย์
- 2.4 การแปลโดยปัญญาประดิษฐ์ ChatGPT
- 2.5 การประยุกต์ใช้เกณฑ์ MQM กับการวิเคราะห์คุณภาพการแปลภาษาไทย เป็นภาษาอังกฤษ
- 2.6 เทคนิคการควบคุมคุณภาพการแปล
- 2.7 งานวิจัยที่เกี่ยวข้องในประเทศไทย
- 2.8 งานวิจัยที่เกี่ยวข้องในต่างประเทศ

2.1 เกณฑ์การประเมินคุณภาพการแปล Multidimensional Quality Metrics (MQM)

Multidimensional Quality Metrics (MQM) ถูกพัฒนาในปี 2013 และเผยแพร่ รายละเอียดแนวคิดในปี 2014 โดย Arle Lommel, Hans Uszkoreit และ Aljoscha Burchardt เพื่อแก้ปัญหาคงไม่สอดคล้องกันในการประเมินคุณภาพการแปล ซึ่งเป็นปัญหาสำคัญสำหรับผู้ว่าจ้าง และผู้ให้บริการในอดีต การประเมินมีความแตกต่างระหว่างการแปลโดยมนุษย์และเครื่องแปลภาษา มุมมองด้านคุณภาพหลากหลาย และความต้องการของผู้ว่าจ้างไม่ชัดเจน MQM จึงถูกสร้างจาก LISA QA Model และได้รับการพัฒนาต่อยอดเพื่อให้มีความยืดหยุ่นและสามารถทำงานร่วมกับมาตรฐาน อื่น ๆ โดยมุ่งเน้นการออกแบบโครงสร้างการประเมินตามหมวดหมู่ข้อผิดพลาด (Error Typology) ในรูปแบบลำดับชั้น (Hierarchical Structure) ที่ช่วยให้การวิเคราะห์ข้อผิดพลาดมีความละเอียด ชัดเจน และรอบด้าน ซึ่งประกอบไปด้วยคุณลักษณะสำคัญในเชิงแนวคิดพื้นฐานด้วยดังต่อไปนี้

MQM ให้ความสำคัญกับการประเมินคุณภาพการแปลโดยอาศัยโครงสร้างการจัดหมวดหมู่ของข้อผิดพลาด (Error typology) ซึ่งถูกจัดเรียงในลักษณะลำดับชั้น (Hierarchical Structure) ทำให้สามารถจำแนกและวิเคราะห์ข้อผิดพลาดในการแปลได้อย่างเป็นระบบ ผู้ประเมินสามารถเลือกใช้หมวดหมู่จากคลังข้อผิดพลาดหลัก (Master Error Catalog) ซึ่งประกอบด้วยข้อผิดพลาดมากกว่า 100 รายการ หรือเลือกใช้เฉพาะชุดหลัก (Core Set) ซึ่งประกอบด้วยมิติระดับสูงจำนวน 8 มิติ และมีประเภทข้อผิดพลาดย่อยรวม 19 ประเภท ทั้งนี้ หมวดหมู่มุ่งกล่าวครอบคลุมการวิเคราะห์คุณภาพการแปลในมิติต่าง ๆ อย่างละเอียดและรอบด้าน โดยสามารถแบ่งกลุ่มข้อผิดพลาดหลักออกเป็นด้าน Accuracy (ความถูกต้อง) ซึ่งรวมถึง Mistranslation (การแปลผิด) Omission (การละเลยเนื้อหา) Addition (การเพิ่มเนื้อหา) Untranslated (ข้อความที่ไม่ได้แปล) และ Over/Under – Translation (การแปลมากหรือน้อยเกินไป) Fluency (ด้านความคล่องแคล่ว) ซึ่งครอบคลุมข้อผิดพลาดด้านไวยากรณ์ การสะกดคำ เครื่องหมายวรรคตอน และระดับภาษา ตลอดจนมิติด้าน Terminology (ศัพท์เฉพาะ) ที่มุ่งเน้นการใช้คำศัพท์ให้ถูกต้องและตรงตามบริบทของสาขาวิชา

นอกจากนี้ MQM ยังครอบคลุมมิติที่สะท้อนความเหมาะสมด้านวัฒนธรรมและบริบทของการใช้ภาษา เช่น ธรรมเนียมท้องถิ่น (Locale – Convention) การใช้รูปแบบภาษา (Style) ความถูกต้องของข้อเท็จจริง (Verity) การจัดรูปแบบเอกสาร (Design) และความสามารถในการใช้งานในหลายภาษา (Internationalization) ซึ่งช่วยให้การประเมินคุณภาพการแปลมีความครอบคลุมทั้งในเชิงภาษา เนื้อหา และการใช้งานจริง

อีกหนึ่งจุดเด่นของ MQM คือความยืดหยุ่นในการนำไปประยุกต์ใช้กับเนื้อหาหลากหลายประเภท ไม่ว่าจะเป็นงานแปลเอกสารทางวิชาการ เอกสารกฎหมาย งานแปลเชิงพาณิชย์ หรือการแปลบทสนทนาในแอปพลิเคชัน โดยผู้ใช้สามารถเพิ่มหรือลดหมวดหมู่ข้อผิดพลาดให้เหมาะสมกับลักษณะของเอกสารหรือบริบทของการสื่อสารได้ตามความจำเป็น ตัวอย่างเช่น ในการแปลเอกสารทางกฎหมาย อาจต้องเพิ่มหมวดข้อผิดพลาดที่เกี่ยวข้องกับศัพท์เฉพาะทางกฎหมายหรือโครงสร้างเชิงตรรกะของข้อความ ซึ่ง MQM สามารถปรับใช้ให้สอดคล้องกับบริบทดังกล่าวได้อย่างเหมาะสม (Melby & Foster, 2014)

นอกจากนี้ MQM ยังรองรับทั้งการประเมินเชิงองค์รวม (Holistic Evaluation) และการประเมินแบบแยกส่วน (Analytic Evaluation) โดยผู้ประเมินสามารถเลือกประเมินคุณภาพการแปลในระดับข้อความโดยรวม (Macro - level) หรือประเมินในระดับประโยคและกลุ่มคำ (Micro - Level) ซึ่งเหมาะสมทั้งในบริบทของงานวิจัยที่ต้องการความแม่นยำเชิงประจักษ์ และการฝึกฝนนักแปลหรือผู้เรียนภาษาให้สามารถมองเห็นจุดแข็งและจุดอ่อนของงานแปลได้อย่างชัดเจน (Lommel et al., 2014) แนวคิดดังกล่าวยังส่งเสริมกระบวนการสะท้อนคิด (Reflective Practice) ในหมู่นักแปล ซึ่งมีบทบาทสำคัญต่อการพัฒนาทักษะการแปลในระยะยาว

ในระดับสากล MQM ได้รับการยอมรับอย่างแพร่หลายและถูกนำไปใช้เป็นเกณฑ์มาตรฐานในการประเมินคุณภาพการแปลในหลายบริบท โดยเฉพาะในการประเมินระบบการแปลด้วยเครื่อง (Machine Translation) เช่น การแข่งขัน Workshop on Machine Translation (WMT) ซึ่งนำ MQM มาใช้ในการจัดอันดับคุณภาพของระบบแปลภาษา (Freitag et al., 2021) นอกจากนี้ MQM ยังถูกนำไปพัฒนาระบบการให้คะแนนข้อผิดพลาดทั้งในรูปแบบเชิงเส้นและไม่เชิงเส้นเพื่อใช้วิเคราะห์ข้อมูลในระดับสถิติ อันช่วยเพิ่มความแม่นยำ ความน่าเชื่อถือของผลการประเมิน และสะท้อนให้เห็นถึงจุดอ่อนเชิงโครงสร้างของระบบแปลภาษาได้อย่างชัดเจน ซึ่งเป็นประโยชน์อย่างยิ่งต่อการพัฒนาระบบการแปลภาษาในอนาคต

ด้วยเหตุนี้ MQM จึงถือเป็นเกณฑ์การประเมินที่มีความครอบคลุม ละเอียด และยืดหยุ่นสูง สามารถปรับใช้ได้หลายบริบท โดยในการศึกษาครั้งนี้ นักวิจัยได้นำหลักการของ MQM มาประยุกต์ใช้เป็นกรอบในการประเมินคุณภาพการแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษ โดยพิจารณาจากเกณฑ์หลัก 3 ประการ ได้แก่ **Accuracy (ความถูกต้องของความหมาย)** **Fluency (ความคล่องแคล่วของภาษา)** และ **Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ)** ซึ่งแต่ละเกณฑ์ยังประกอบด้วยเกณฑ์ย่อยเพื่อให้สามารถประเมินข้อผิดพลาดได้อย่างละเอียด ผู้วิจัยได้ศึกษาความหมายของเกณฑ์แต่ละด้านจากแนวคิดของ Lommel et al. (2014) รวมทั้งศึกษาปัจจัยตัวอย่างการแปลภาษาไทยเป็นภาษาอังกฤษจากงานวิจัยของ Tongpoon - Patanasorn และ Griffith (2020) ซึ่งเป็นการวิเคราะห์การแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษด้วย Google Translate เพื่อใช้เป็นแนวทางในการอธิบายลักษณะของข้อผิดพลาดในแต่ละประเภท โดยมีรายละเอียดดังต่อไปนี้

2.1.1 Accuracy (ความถูกต้องของความหมาย) หมายถึง ระดับความถูกต้องแม่นยำของการแปลโดยพิจารณาว่าการแปลสามารถสื่อความหมายของต้นฉบับได้อย่างครบถ้วนและถูกต้อง (Lommel et al., 2014) เกณฑ์ย่อยที่นำมาพิจารณาในด้านความถูกต้อง (Accuracy) ได้แก่

1) Mistranslation (การแปลผิด) หมายถึง ระดับความถูกต้องแม่นยำของการแปล โดยพิจารณาว่าการแปลสามารถสื่อความหมายของต้นฉบับได้อย่างครบถ้วนและถูกต้องหรือไม่ ซึ่งหากการแปลคลาดเคลื่อนจากความหมายในต้นฉบับ จะถือเป็นข้อผิดพลาดด้าน “การแปลผิด” (Mistranslation) ตัวอย่างประโยคเช่น “The natural environment in tourist attractions is an important factor that attracts tourists **to go**.” เป็นการแปลที่คลาดเคลื่อนจากต้นฉบับภาษาไทยซึ่งมีความหมายว่า “สิ่งแวดล้อมตามธรรมชาติในแหล่งท่องเที่ยวเป็นปัจจัยสำคัญยิ่งที่ดึงดูดนักท่องเที่ยวให้ไป **ท่องเที่ยว**” โดยคำว่า **to go** ในภาษาอังกฤษไม่สามารถถ่ายทอดความหมายของ “**ไปเที่ยว**” ได้อย่างถูกต้องในบริบททางการท่องเที่ยว ทำให้ความหมายไม่สมบูรณ์และขาดความเป็นธรรมชาติ สามารถแก้ไขเป็น “The natural environment in tourist attractions is an important factor that attracts tourists **to travel**.” ซึ่งสื่อความหมาย

ได้ตรงและถูกต้องมากกว่า ตัวอย่างนี้สะท้อนให้เห็นว่าข้อผิดพลาดประเภท การแปลผิด (Mistranslation) เป็นข้อผิดพลาดที่เกิดขึ้นบ่อยในการใช้เครื่องแปลภาษาอัตโนมัติ โดยเฉพาะเมื่อต้องแปลข้อความที่มีความหมายเชิงบริบทหรืออุปมาอุปไมย ซึ่งเครื่องแปลมักจะแปลตามตัวอักษร (Literal Translation) มากกว่าการตีความตามบริบท ทำให้เกิดการสูญเสียความหมายสำคัญของต้นฉบับ (Tongpoon-Patanasorn & Griffith, 2020)

2) Omission (การละเลยเนื้อหา) หมายถึง การละบางประโยคหรือข้อมูลสำคัญในต้นฉบับระหว่างกระบวนการแปล ซึ่งส่งผลให้สาระสำคัญของข้อความหายไป ทำให้ผู้อ่านไม่สามารถเข้าใจเนื้อหาฉบับแปลได้ครบถ้วน (Lommel et al., 2014) ตัวอย่างต้นฉบับภาษาไทย “...รีสอร์ทที่ใช้ทรัพยากรธรรมชาติของท้องถิ่นควรมีความรับผิดชอบต่อสิ่งแวดล้อมและชุมชนท้องถิ่น...” แต่ข้อความที่ Google Translate แปลได้ คือ “The resort business is based on a very local natural environment.” ซึ่งตัดประโยคส่วนที่กล่าวถึง “ความรับผิดชอบต่อสิ่งแวดล้อมและชุมชน” ออกไปทั้งหมด ทำให้สาระสำคัญด้านจริยธรรมของผู้ประกอบการหายไปจากเนื้อหา นักวิจัยจึงเสนอการปรับแก้เป็น “The resort business is primarily based on the local natural environment, and it should be responsible for preserving the environment and local community.” Tongpoon-Patanasorn และ Griffith (2020) กล่าวว่า “การละเลยเนื้อหา (Omission)” เป็นข้อผิดพลาดที่พบบ่อยในเครื่องแปลอัตโนมัติ เนื่องจากระบบมีแนวโน้มเลือกตัดข้อความที่ตีความยากหรือมีโครงสร้างซับซ้อนออกโดยอัตโนมัติ เพื่อให้ได้ประโยคที่ดูสมบูรณ์ทางรูปแบบ แต่ขาดความสมบูรณ์ทางเนื้อหา

3) Addition (การเพิ่มเนื้อหา) หมายถึง การเพิ่มข้อความ คำ หรือวลีที่ไม่ได้ปรากฏในต้นฉบับ ซึ่งทำให้ความหมายของต้นฉบับเปลี่ยนไปหรือนำไปสู่การตีความที่คลาดเคลื่อน (Lommel et al., 2014) ตัวอย่างเช่น ต้นฉบับภาษาไทย “เมื่อรีสอร์ทได้รับประโยชน์จากทรัพยากรธรรมชาติในท้องถิ่นควรมีความรับผิดชอบต่อสิ่งแวดล้อม” ถูกแปลโดย Google Translate เป็น “When resort business **Benefit** from local natural resources.” ซึ่งนอกจากจะใช้ตัวพิมพ์ใหญ่ของคำว่า Benefit ผิดแล้ว ระบบยัง “เพิ่ม” คำว่า Benefit เข้ามาในตำแหน่งที่ไม่ถูกต้อง ทั้งที่ในต้นฉบับไม่ได้มีคำกริยาเดี่ยวนี้อยู่ ทำให้ประโยคไม่สมบูรณ์เชิงความหมายและคลาดเคลื่อนจากเจตนาของผู้เขียน โดยปรับแก้ไขเป็น “When a resort business **benefits** from local natural resources, it should be responsible for environmental protection.” ซึ่งเป็นการปรับให้ถูกต้องตามหลักไวยากรณ์และคงความหมายของต้นฉบับไว้อย่างครบถ้วน กรณีนี้แสดงให้เห็นว่าการเพิ่มเนื้อหาโดยไม่ได้ตั้งใจของเครื่องแปลภาษาอัตโนมัติอาจทำให้ผู้อ่านเข้าใจว่าข้อความมีเจตนาหรือสารใหม่ที่ไม่ได้มีอยู่ในต้นฉบับ (Tongpoon - Patanasorn & Griffith, 2020)

4) Untranslated Text (ข้อความที่ไม่ได้แปล) หมายถึง การที่บางส่วนของข้อความต้นฉบับไม่ได้รับการแปลและยังคงปรากฏในภาษาต้นฉบับ (Lommel et al., 2014) ตัวอย่างเช่น ในการศึกษาของ Tongpoon - Patanasorn และ Griffith (2020) พบว่าการแปลบทคัดย่อภาษาไทยเกี่ยวกับการจัดการท่องเที่ยวด้วย Google Translate มีบางส่วนที่ระบบคงไว้เป็นภาษาไทยโดยไม่แปล เช่น ต้นฉบับภาษาไทย “การท่องเที่ยวโดยชุมชนเป็นกิจกรรมที่ส่งเสริมให้เกิดการพัฒนาอย่างยั่งยืนในพื้นที่” ผลการแปลของ Google Translate “Community-based tourism is an activity that promotes การพัฒนาอย่างยั่งยืน in the area.” จะเห็นได้ว่าระบบแปลอัตโนมัติไม่ได้แปลวลี “การพัฒนาอย่างยั่งยืน” เป็นภาษาอังกฤษ แต่คงไว้ในรูปภาษาไทย ซึ่งทำให้ผู้อ่านภาษาอังกฤษไม่สามารถเข้าใจเนื้อหาส่วนนี้ได้อย่างครบถ้วน ผู้วิจัยจึงเสนอการแก้ไขเป็น “Community-based tourism is an activity that promotes **sustainable development** in the area.” Tongpoon - Patanasorn และ Griffith (2020) กล่าวว่า ปัญหาการคงคำไทยไว้ในผลลัพธ์การแปลมักเกิดขึ้นกับคำศัพท์เชิงวิชาการหรือคำที่มีบริบทเฉพาะ เช่น “ภูมิปัญญาท้องถิ่น” หรือ “ร่วางมาตรฐาน” ซึ่งไม่มีคำเทียบในฐานข้อมูลของระบบแปลอัตโนมัติ ส่งผลให้ Google Translate ไม่สามารถตีความหรือถอดความหมายเป็นภาษาอังกฤษได้อย่างถูกต้อง จึงเลือกคงคำไทยไว้ในผลลัพธ์แทน ทั้งนี้ ข้อจำกัดดังกล่าวสะท้อนถึงสถานะของเทคโนโลยีการแปลอัตโนมัติในช่วงเวลานั้น (ปี 2020) ที่ฐานข้อมูลคำศัพท์ภาษาไทยยังไม่สมบูรณ์ โดยเฉพาะคำที่เกี่ยวข้องกับวัฒนธรรม คำเฉพาะทางวิชาการ และคำประสมยาว ซึ่งระบบยังไม่สามารถประมวลผลเชิงบริบทได้อย่างมีประสิทธิภาพ

ผู้วิจัยจึงสรุปได้ว่า Accuracy (ความถูกต้องของความหมาย) หมายถึง ระดับความถูกต้องแม่นยำของการแปลโดยพิจารณาว่าการแปลสามารถสื่อความหมายของต้นฉบับได้อย่างครบถ้วนและถูกต้อง โดยพิจารณาจากองค์ประกอบย่อย ได้แก่ Mistranslation (การแปลผิด) Omission (การละเลยเนื้อหา) Addition (การเพิ่มเนื้อหา) และ Untranslated Text (ข้อความที่ไม่ได้แปล)

2.1.2 Fluency (ความคล่องแคล่วของภาษา) หมายถึง ระดับความเป็นธรรมชาติและความอ่านง่ายของการแปล โดยพิจารณาว่าการแปลมีความสละสลวยและเป็นไปตามหลักไวยากรณ์ของภาษาเป้าหมาย (Lommel et al., 2014) เกณฑ์ย่อยที่นำมาพิจารณาในด้านความคล่องแคล่ว (Fluency) ได้แก่

1) Punctuation (การใช้เครื่องหมายวรรคตอน) หมายถึง การใช้เครื่องหมายวรรคตอนผิดหรือขาดหายไปทำให้ประโยคไม่เป็นธรรมชาติและอาจเปลี่ยนจังหวะการอ่านของผู้อ่าน (Lommel et al., 2014) ตัวอย่างเช่น ข้อความต้นฉบับภาษาไทย “การพัฒนาการท่องเที่ยวโดยชุมชนมีส่วนช่วยในการสร้างรายได้ให้กับคนในพื้นที่” แปลโดย Google Translate เป็น “Community - based tourism development. helps generate income for local people.” โดยระบบได้แทรก

จุด (.) กลางประโยค ซึ่งผิดหลักไวยากรณ์ของภาษาอังกฤษและทำให้ข้อความขาดความต่อเนื่อง ผู้วิจัยเสนอการแก้ไขให้ถูกต้องเป็น “Community-based tourism development helps generate income for local people.” (Tongpoon - Patanasorn & Griffith, 2020)

2) Spelling (การสะกดคำ) หมายถึง สะกดคำไม่ถูกต้อง (Lommel et al., 2014) ยกตัวอย่างเช่น คำว่า “environmental conservation” ถูกแปลผิดเป็น “environmental conversation” ซึ่งแม้จะสะกดใกล้เคียงกันแต่ความหมายต่างกันโดยสิ้นเชิง คำว่า conversation หมายถึง “การสนทนา” ไม่ใช่ “การอนุรักษ์” การแก้ไขเป็น “environmental conservation” (Tongpoon-Patanasorn & Griffith, 2020)

3) Grammar (ไวยากรณ์) หมายถึง โครงสร้างประโยคไม่ถูกต้อง (Lommel et al., 2014) ยกตัวอย่างเช่น “นักท่องเที่ยวมีความพึงพอใจต่อการบริการของรีสอร์ทในระดับมาก” ถูกแปลโดย Google Translate เป็น “Tourists satisfaction is high with the service of resort.” ซึ่งมีข้อผิดพลาดหลายจุด ได้แก่ การละ article (“the resort”) การใช้คำนามซ้อนโดยไม่มีเครื่องหมายแสดงความเป็นเจ้าของ (“Tourists satisfaction”) และลำดับคำที่ไม่เป็นธรรมชาติ การแก้ไขให้เหมาะสมคือ “Tourists are highly satisfied with the service of the resort.” (Tongpoon - Patanasorn & Griffith, 2020)

4) Register (ระดับภาษา) หมายถึง ใช้ภาษาที่ไม่เหมาะสมกับบริบทเชิงวิชาการ (Lommel et al., 2014) ยกตัวอย่างเช่น “การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อวิเคราะห์ปัจจัยที่ส่งผลต่อการตัดสินใจของนักท่องเที่ยว” แปลเป็น “This study wants to look at the factors that make tourists decide.” จะเห็นได้ว่า Google Translate ใช้สำนวน “wants to look at” ซึ่งเป็นภาษาพูด ไม่เหมาะสมกับบริบทเชิงวิชาการ และใช้คำว่า “make” แทน “influence” ซึ่งทำให้ความหมายไม่เป็นทางการและคลาดเคลื่อน การแก้ไขให้เหมาะสมคือ “This study aims to analyze the factors influencing tourists’ decision-making.” (Tongpoon - Patanasorn & Griffith, 2020)

ผู้วิจัยจึงสรุปได้ว่า Fluency (ความคล่องแคล่วของภาษา) หมายถึง ระดับความเป็นธรรมชาติและความสะดวกของการแปล โดยพิจารณาว่าการแปลมีความสละสลวยและเป็นไปตามหลักไวยากรณ์ของภาษาเป้าหมาย โดยพิจารณาจากองค์ประกอบย่อย ได้แก่ Punctuation (การใช้เครื่องหมายวรรคตอน) Spelling (การสะกดคำ) Grammar (ไวยากรณ์) และ Register (ระดับภาษา)

2.1.3 Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) หมายถึง ระดับความถูกต้องและสอดคล้องในการใช้ศัพท์เฉพาะทางเทคนิค โดยพิจารณาว่ามีการใช้ศัพท์เฉพาะอย่างถูกต้องและสอดคล้องกันตลอดทั้งข้อความ (Lommel et al., 2014) เกณฑ์ย่อยที่นำมาพิจารณาในด้าน Terminology Consistency ได้แก่

1) Incorrect Term (การใช้ศัพท์ไม่ตรง) หมายถึง การเลือกใช้คำศัพท์ที่ไม่เหมาะสมกับบริบททางวิชาการ ทำให้ความหมายของประโยคคลาดเคลื่อนหรือผิดเพี้ยนไปจากต้นฉบับ (Lommel et al., 2014) ตัวอย่างเช่น ข้อความต้นฉบับภาษาไทยว่า “การจัดการการท่องเที่ยวเชิงนิเวศมีบทบาทสำคัญในการพัฒนาอย่างยั่งยืนของชุมชน” แปลเป็น “Ecotourism management has an important role in the **development of the community’s durability.**” จะเห็นว่าระบบใช้คำว่า durability แทน sustainability ซึ่งไม่ถูกต้องในบริบทของงานวิชาการ เนื่องจาก durability หมายถึง “ความทนทานของวัตถุหรือสิ่งของ” ในขณะที่ sustainability หมายถึง “ความยั่งยืน” ในทางสังคมหรือสิ่งแวดล้อม การแก้ไขที่เหมาะสม คือ “Ecotourism management plays an important role in the **sustainable development** of the community.” ตัวอย่างนี้แสดงให้เห็นว่าเครื่องแปลภาษาอัตโนมัติมักเลือกคำแปลตามความใกล้เคียงทางรูปคำ (form-based translation) แทนการพิจารณาตามบริบท (context-based translation) ซึ่งเป็นจุดอ่อนด้าน terminological accuracy ที่พบได้บ่อย (Tongpoon-Patanasorn & Griffith, 2020)

2) Inconsistent Term (การใช้ศัพท์ไม่สอดคล้องกัน) หมายถึง ใช้คำศัพท์เฉพาะไม่เหมือนกันในเนื้อหาเดียวกัน ทำให้ความหมายไม่ต่อเนื่องหรือเกิดความสับสน (Lommel et al., 2014) ตัวอย่างเช่น Tongpoon - Patanasorn และ Griffith (2020) พบว่า ในการแปลบทคัดย่อภาษาไทยเกี่ยวกับ “การจัดการการท่องเที่ยว (Tourism Management)” ระบบ Google Translate ใช้คำไม่สม่ำเสมอระหว่าง “tourism management” และ “travel management” ซึ่งทำให้เกิดความไม่ต่อเนื่องในความหมาย ตัวอย่างเช่น ต้นฉบับภาษาไทย: “การจัดการการท่องเที่ยวที่ดีช่วยสร้างรายได้ให้แก่ชุมชน และส่งเสริมความยั่งยืนในระยะยาว” ผลการแปลของ Google Translate คือ “Good **travel management** helps generate income for the community and promotes long-term sustainability.” แต่ในประโยคถัดมา ระบบกลับแปลคำเดิมเป็น “The effectiveness of **tourism management** depends on cooperation among stakeholders.” การใช้คำสองแบบนี้ในเอกสารเดียวกัน (**travel management** กับ **tourism management**) ทำให้เกิดความไม่สอดคล้องเชิงศัพท์และลดความชัดเจนทางวิชาการ เนื่องจาก “travel management” มีความหมายเฉพาะด้านธุรกิจการเดินทาง (Corporate Travel) ในขณะที่ “tourism management” หมายถึง การบริหารจัดการกิจกรรมการท่องเที่ยวโดยรวม การแก้ไขควรเลือกใช้คำเดียวกันตลอดทั้งข้อความเพื่อกงความหมายให้สอดคล้อง โดยใช้คำว่า “Good **tourism management**

helps generate income for the community and promotes long-term sustainability.”

การปรับดังกล่าวช่วยให้เนื้อหาที่มีความต่อเนื่องและสอดคล้องกับบริบทของสาขาวิชาการท่องเที่ยว

Inconsistent Term (การใช้ศัพท์ไม่สอดคล้องกัน) เป็นองค์ประกอบสำคัญของการประเมินคุณภาพการแปลตามกรอบ MQM เพราะช่วยให้ข้อความมีความชัดเจนและสอดคล้องกันทางความหมาย (Lommel et al., 2014) จากผลการศึกษาของ Tongpoon - Patanasorn และ Griffith (2020) พบว่าเครื่องแปลภาษาอัตโนมัติอย่าง Google Translate ยังมีข้อจำกัดในการประมวลคำศัพท์ทางเทคนิคและบริบททางสาขาวิชา ซึ่งอาจทำให้เกิดการใช้คำที่ผิดความหมายหรือไม่สม่ำเสมอในเนื้อหาเดียวกัน ดังนั้น การตรวจแก้โดยผู้เชี่ยวชาญเฉพาะด้าน (Subject - Matter Expert) จึงมีความจำเป็นเพื่อปรับปรุงคุณภาพคำศัพท์ให้ถูกต้องและคงความหมายที่สอดคล้องกันตลอดทั้งบทความ

ผู้วิจัยจึงสรุปได้ว่า Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) หมายถึง ระดับความถูกต้องและสอดคล้องในการใช้ศัพท์เฉพาะทางเทคนิค โดยพิจารณาว่ามีการใช้ศัพท์เฉพาะอย่างถูกต้องและสอดคล้องกันตลอดทั้งข้อความ โดยพิจารณาจากองค์ประกอบย่อย ได้แก่ Incorrect Term (การใช้ศัพท์ไม่ตรง) และ Inconsistent Term (การใช้ศัพท์ไม่สอดคล้องกัน)

เกณฑ์การประเมินคุณภาพการแปล MQM ได้รับการนำไปใช้อย่างแพร่หลายในงานวิจัยร่วมสมัย ทั้งในการประเมินการแปลโดยมนุษย์และการแปลด้วยเครื่องแปลภาษา Mariana, Cox และ Melby (2025) ได้ศึกษาการประยุกต์ใช้เกณฑ์ MQM โดยมุ่งเน้นประเด็นสำคัญคือ ความเป็นไปได้ในการใช้งาน (Usability) และ ความสอดคล้องของผลการประเมิน (Rating Consistency) เมื่อผู้ประเมินไม่ใช่ผู้เชี่ยวชาญด้านการแปลโดยตรง โดยให้ผู้ประเมินมือใหม่ (Novice Raters) ใช้เกณฑ์ MQM ในการให้คะแนนงานแปลของผู้เรียน แล้วนำผลการประเมินไปเปรียบเทียบกับคะแนนที่ได้จากผู้ประเมินที่มีประสบการณ์หรือผู้เชี่ยวชาญ ผลการศึกษพบว่า คะแนนที่ได้จากผู้ประเมินมือใหม่มีแนวโน้มสอดคล้องกับคะแนนของผู้เชี่ยวชาญในระดับที่ยอมรับได้ และ MQM ช่วยให้ผู้ประเมินเข้าใจลักษณะข้อผิดพลาดของงานแปลได้ชัดเจนขึ้น เนื่องจากมีโครงสร้างการจำแนกข้อผิดพลาดที่เป็นระบบ สรุปได้ว่า MQM ไม่ได้จำกัดการใช้งานเฉพาะในกลุ่มผู้เชี่ยวชาญเท่านั้น แต่สามารถนำไปใช้ในบริบทการเรียนการสอน การฝึกอบรมนักแปล และการประเมินคุณภาพงานแปลในระดับการศึกษาได้อย่างมีประสิทธิภาพ Park และ Padó (2024) ได้นำเกณฑ์การประเมิน MQM ประเมินแปลภาษาอังกฤษเป็นภาษาเกาหลี เนื่องด้วยข้อจำกัดของวิธีการประเมินการแปลด้วยเครื่องแปลภาษาแบบเดิมที่ใช้คะแนนเดียวไม่สามารถสะท้อนมิติคุณภาพที่หลากหลายได้อย่างเพียงพอ ผลการทดลองพบว่าการประเมิน MQM ช่วยเพิ่มการตีความผลลัพธ์ได้ดีขึ้น ในขณะที่ยังรักษาประสิทธิภาพการประเมินคุณภาพโดยรวมได้ นอกจากนี้กรอบ MQM สามารถขยายไปยังคู่ภาษาอื่นได้ และสามารถพัฒนาระบบแปลอัตโนมัติให้ตอบโจทย์บริบทการใช้งานที่หลากหลายในอนาคต Freitag et al. (2021) ได้ศึกษาวิจัยเรื่องผู้เชี่ยวชาญ ข้อผิดพลาด และบริบท การศึกษาขนาดใหญ่เกี่ยวกับการประเมินของมนุษย์สำหรับการแปลโดยเครื่อง โดยใช้กรอบงานการวัดคุณภาพ

แบบหลายมิติ (MQM) เพื่อประเมินระบบการแปลด้วยเครื่องแปลภาษา ผ่านการวิเคราะห์ข้อผิดพลาดที่ชัดเจนและการให้คะแนนโดยนักแปลมืออาชีพพร้อมบริบทของเอกสารทั้งหมด MQM ช่วยระบุและจัดหมวดหมู่ข้อผิดพลาด เช่น การแปลผิดและปัญหาความถูกต้อง และเสนอรูปแบบการให้คะแนนมาตรฐานสำหรับการประเมินการแปลด้วยเครื่องที่ครอบคลุมกว้าง การศึกษาเผยให้เห็นว่าการให้คะแนน MQM ให้การจัดอันดับระบบการแปลด้วยเครื่องที่แตกต่างไปจากการประเมินครั้งก่อน ๆ โดยเน้นย้ำจุดแข็งและจุดอ่อนของระบบได้อย่างมีประสิทธิภาพมากกว่า Wang (2024) เน้นการประเมินคุณภาพงานแปลโดยใช้ ChatGPT - 4o ผ่านเกณฑ์การประเมินคุณภาพการแปล Multidimensional Quality Metrics (MQM) โดยทดสอบกับงานแปลต้นฉบับที่ประกอบด้วยงานเขียนประเภทนิยายและงานเขียนไม่ใช่นิยาย ซึ่งนักเรียนสาขา MTI ทำเอง แล้วเปรียบเทียบผลระหว่างเวอร์ชันของมนุษย์กับเวอร์ชันที่ปรับด้วย ChatGPT - 4o โดยใช้การให้คะแนนตาม MQM

แม้ว่าแนวคิด MQM จะได้รับการพัฒนาขึ้นตั้งแต่ปี ค.ศ. 2014 แต่หลักฐานจากงานวิจัยร่วมสมัยสะท้อนให้เห็นว่า MQM ยังคงมีความทันสมัยและมีความเหมาะสมอย่างยิ่งในการประเมินคุณภาพการแปลในบริบทปัจจุบัน งานวิจัยในช่วงไม่กี่ปีที่ผ่านมาได้แสดงให้เห็นว่า MQM สามารถนำไปประยุกต์ใช้ได้อย่างมีประสิทธิภาพทั้งในการประเมินการแปลโดยมนุษย์และการแปลด้วยระบบปัญญาประดิษฐ์ รวมถึงโมเดลภาษาเชิงกำเนิดรุ่นใหม่ เช่น ChatGPT และระบบแปลอัตโนมัติขั้นสูง โดย MQM ช่วยให้การประเมินคุณภาพการแปลสามารถสะท้อนมิติของคุณภาพได้อย่างรอบด้านมากกว่าการใช้คะแนนเชิงองค์รวมเพียงค่าเดียว นอกจากนี้ ลักษณะโครงสร้างแบบเปิดและยืดหยุ่นของ MQM ยังเอื้อให้สามารถปรับเลือกเกณฑ์ให้เหมาะสมกับบริบท ภาษา และประเภทของข้อความได้อย่างต่อเนื่องตามพัฒนาการของเทคโนโลยีการแปล ด้วยเหตุนี้ นักวิจัยจึงเห็นว่า MQM มิได้เป็นเกณฑ์แนวคิดที่ล้าสมัย แต่เป็นเกณฑ์การประเมินที่ได้รับการยอมรับและต่อยอดอย่างต่อเนื่องในงานวิจัยร่วมสมัย จึงมีความเหมาะสมอย่างยิ่งที่จะนำมาใช้เป็นกรอบในการประเมินคุณภาพการแปล บทความวิทยานิพนธ์ภาษาไทยเป็นภาษาอังกฤษในการศึกษาครั้งนี้ เนื่องจาก MQM มีโครงสร้างที่ชัดเจน ครอบคลุมประเด็นสำคัญด้านคุณภาพ และสามารถรองรับการเปรียบเทียบการแปลระหว่างมนุษย์และระบบปัญญาประดิษฐ์ได้อย่างเป็นระบบและน่าเชื่อถือ

2.2 การแปลบทคัดย่อจากภาษาไทยเป็นภาษาอังกฤษ

การแปลบทคัดย่อจากภาษาไทยเป็นภาษาอังกฤษเป็นกระบวนการที่ซับซ้อน เนื่องจากภาษาไทยและภาษาอังกฤษมีความแตกต่างอย่างมากในด้านโครงสร้างไวยากรณ์ การเรียงลำดับคำ และการใช้บริบททางวัฒนธรรม Jeharsae และ Boonsuk (2024) กล่าวว่า ภาษาไทยเป็นภาษาที่มีโครงสร้างที่ยืดหยุ่น โดยสามารถละประธานหรือกริยาในประโยคได้โดยไม่ทำให้ความหมายเสียหาย ขณะที่ภาษาอังกฤษต้องการโครงสร้างประโยคที่ครบถ้วน ได้แก่ ประธาน กริยา และกรรมอย่างชัดเจน ความแตกต่างนี้นำไปสู่ข้อผิดพลาดเชิงโครงสร้าง เช่น การละประธาน การเรียงลำดับคำผิด หรือการใช้

โครงสร้างที่ไม่เหมาะสมในภาษาอังกฤษ นอกจากนี้ ยังพบข้อผิดพลาดเชิงความหมาย (Semantic Errors) ซึ่งมักเกิดจากการตีความความหมายของต้นฉบับผิดหรือเลือกใช้คำไม่เหมาะสมในภาษาเป้าหมาย โดยเฉพาะเมื่อผู้เรียนขาดความเข้าใจด้านศัพท์เชิงบริบทและระดับภาษาที่ใช้ในงานเขียนวิชาการ Wongranu (2017) กล่าวว่า ความแตกต่างทางภาษาระหว่างไทยกับอังกฤษ เป็นปัจจัยสำคัญที่ก่อให้เกิดข้อผิดพลาดในการแปล โดยเฉพาะในเรื่องของลำดับคำและการใช้ไวยากรณ์ที่เหมาะสมในบริบททางวิชาการ ดังนั้น การแปลที่มีคุณภาพจึงจำเป็นต้องอาศัยความรู้ด้านโครงสร้างของทั้งสองภาษาและความสามารถในการวิเคราะห์บริบทของเนื้อหาอย่างรอบด้าน Kampookaew (2020) ซึ่งวิเคราะห์ข้อผิดพลาดทางไวยากรณ์ในการเขียนภาษาอังกฤษของนักศึกษาไทย พบว่า ปัญหาหลักได้แก่ การใช้คำนามและกริยาในรูปแบบที่ไม่ถูกต้อง การละการใช้คำนำหน้านาม 'the' และความไม่สอดคล้องระหว่างประธานและกริยา สาเหตุที่เกิดจากการขาดความเข้าใจเชิงโครงสร้างของภาษาอังกฤษ ซึ่งส่งผลให้คุณภาพของข้อความลดลง แม้สาระสำคัญจะยังสามารถถ่ายทอดได้ หากนำมาเปรียบเทียบกับบริบทของการแปลบทคัดย่อ จะเห็นได้ว่าข้อผิดพลาดเหล่านี้อาจลดทั้งความน่าเชื่อถือและความเป็นทางการของบทแปล ดังนั้น การพัฒนาเครื่องมือประเมินคุณภาพการแปลจึงควรคำนึงถึงการแก้ไขข้อผิดพลาดด้านโครงสร้างประโยคและไวยากรณ์ภาษาอังกฤษ Chuanpipatpong (2025) กล่าวว่า ภาษาไทยมีลักษณะโครงสร้างที่ยืดหยุ่นสูง โดยสามารถละประธานหรือกริยาได้ในบางบริบท และยังสามารถสื่อความหมายได้อย่างชัดเจนแม้ไม่มีโครงสร้างประโยคที่ครบถ้วน ขณะที่ภาษาอังกฤษมีโครงสร้างที่ชัดเจนและรัดกุมมากกว่า โดยจำเป็นต้องมีองค์ประกอบของประธาน กริยา และกรรมอย่างครบถ้วน พร้อมทั้งเคารพกฎทางไวยากรณ์อย่างเคร่งครัด เช่น ความสอดคล้องระหว่างประธานและกริยา กาล และการใช้คำนำหน้านาม ซึ่งมีรูปแบบที่หลากหลายและซับซ้อนกว่าภาษาไทย

การแปลจากภาษาไทยเป็นภาษาอังกฤษไม่ว่าจะในรูปแบบของบทคัดย่อ งานเขียนวิชาการหรือเนื้อหาอื่น ๆ ล้วนต้องอาศัยความเข้าใจเชิงลึกในโครงสร้างของทั้งสองภาษา ความแตกต่างทางไวยากรณ์ ลำดับคำ ระบบกาลเวลา และระดับภาษาล้วนเป็นปัจจัยสำคัญที่ส่งผลต่อคุณภาพของการแปล ผู้แปลจำเป็นต้องมีทักษะในการถ่ายทอดสาระสำคัญของเนื้อหาได้อย่างครบถ้วน โดยต้องสามารถเลือกใช้คำศัพท์ โครงสร้างประโยค และระดับภาษาให้เหมาะสมกับบริบทของงานเขียนแต่ละประเภท ความสามารถในการปรับบริบท (Contextualization) และการใช้ภาษาอังกฤษอย่างถูกต้องตามหลักไวยากรณ์จึงเป็นสิ่งจำเป็นอย่างยิ่งในการแปลจากภาษาไทยไปสู่ภาษาอังกฤษให้มีความแม่นยำ ชัดเจน และสื่อสารได้อย่างมีประสิทธิภาพ

2.3 การแปลโดยมนุษย์

Lau et al. (2024) ได้กล่าวถึงการแปลโดยมนุษย์ว่าเป็นกระบวนการที่ยังคงมีความสำคัญอย่างยิ่งในการถ่ายทอดความหมายของข้อความจากภาษาหนึ่งไปยังอีกภาษาหนึ่ง โดยเฉพาะอย่างยิ่งในกรณีที่ข้อความนั้นมีความซับซ้อนทางวัฒนธรรมหรือมีความหมายที่ไม่ชัดเจน นักแปลมนุษย์สามารถตีความบริบท ปรับโทนภาษา และเลือกใช้ถ้อยคำที่เหมาะสมกับกลุ่มเป้าหมายได้อย่างยืดหยุ่น และเป็นธรรมชาติ ซึ่งเป็นข้อได้เปรียบที่ระบบปัญญาประดิษฐ์ยังไม่สามารถทดแทนได้อย่างสมบูรณ์ นอกจากนี้ ผู้วิจัยยังชี้ให้เห็นว่า การแปลโดยมนุษย์มักถูกใช้เป็นการประเมินคุณภาพของการแปลโดยเครื่องหรือ AI อีกด้วย ทั้งนี้ แสดงให้เห็นว่าแม้เทคโนโลยีการแปลอัตโนมัติจะมีการพัฒนาอย่างต่อเนื่อง แต่บทบาทของมนุษย์ในฐานะผู้แปลที่เข้าใจวัฒนธรรม บริบท และความละเอียดอ่อนของภาษา ยังคงเป็นองค์ประกอบที่ขาดไม่ได้ในกระบวนการแปล Herbig et al. (2019) กล่าวว่า การแปลโดยมนุษย์เป็นกระบวนการที่มีจุดเด่นด้านความสามารถในการเข้าใจบริบทของข้อความ ความหมายแฝง และเจตนาของผู้เขียน ซึ่งเป็นสิ่งที่ระบบแปลอัตโนมัติยังไม่สามารถทำได้อย่างมีประสิทธิภาพ นักแปลมนุษย์มีความสามารถในการตีความวลีที่คลุมเครือ ปรับโทนภาษา รูปแบบและระดับความเป็นทางการของสำนวนให้เหมาะสมกับกลุ่มผู้อ่าน รวมถึงเลือกใช้ถ้อยคำที่สอดคล้องกับบริบททางวัฒนธรรมของผู้รับสาร ทั้งนี้ ผู้วิจัยยังชี้ว่า แม้เทคโนโลยีการแปลโดยปัญญาประดิษฐ์จะสามารถผลิตร่างคำแปลได้อย่างรวดเร็ว แต่คุณภาพของคำแปลยังคงต้องได้รับการตรวจสอบและแก้ไขโดยมนุษย์ ซึ่งมีบทบาทในฐานะผู้ควบคุมคุณภาพสุดท้าย เพื่อให้คำแปลมีความถูกต้อง สั้นไหลและเหมาะสมตามหลักการสื่อสารในภาษาปลายทาง Tongpoon - Patanasorn และ Griffith (2020) กล่าวว่า การแปลโดยมนุษย์กับการใช้เครื่องแปลภาษา แม้เครื่องแปลภาษาจะมีข้อได้เปรียบในด้านความรวดเร็วและต้นทุนต่ำ แต่การแปลโดยมนุษย์ยังคงรักษาความแม่นยำของความหมาย ความชัดเจนของประโยค และความสั้นไหลของภาษาได้ดีกว่า โดยเฉพาะในบริบทของงานวิชาการ การแปลโดยมนุษย์สามารถปรับบทแปลให้เหมาะสมทั้งในระดับคำ วลี และประโยค โดยคำนึงถึงวัตถุประสงค์และกลุ่มเป้าหมายของข้อความ ส่งผลให้เนื้อหามีความเป็นธรรมชาติมากกว่า ขณะที่เครื่องแปลภาษามีข้อจำกัดในการแปลคำวัฒนธรรมและการแบ่งประโยคที่ถูกต้อง (กัญญาภักธสาสงเคราะห์, 2562) แม้เทคโนโลยีจะมีการพัฒนาขึ้นอย่างต่อเนื่อง แต่ความแม่นยำและความเหมาะสมทางบริบทของการแปลโดยมนุษย์ยังคงได้เปรียบอย่างมีนัยสำคัญ โดยเฉพาะเมื่อเนื้อหามีความซับซ้อนและมีลักษณะเฉพาะทางวัฒนธรรม (นรินทร์ทิพย์ ทองศรี, 2567)

ดังนั้นการแปลโดยมนุษย์ยังคงมีบทบาทสำคัญในการผลิตงานแปลที่มีคุณภาพ โดยเฉพาะในงานวิชาการที่ต้องการความถูกต้อง ชัดเจน และมีลักษณะเฉพาะของบริบท การใช้มนุษย์ในการแปลจึงเป็นแนวทางที่ยังคงมีความจำเป็นควบคู่กับการใช้เทคโนโลยีแปลภาษาในยุคปัจจุบัน

ผู้วิจัยจึงสรุปได้ว่า การแปลโดยมนุษย์ หมายถึง การแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษโดยผู้แปลที่เป็นมนุษย์ เช่น อาจารย์ผู้เชี่ยวชาญด้านภาษา ผู้แปลมืออาชีพ หรือผู้มีความรู้ด้านการแปลโดยใช้ทักษะภาษาตามความสามารถของบุคคล ไม่ใช่ระบบปัญญาประดิษฐ์หรือเครื่องมือแปลอัตโนมัติ

2.4 การแปลโดย ChatGPT

2.4.1 ศักยภาพและข้อจำกัดของ ChatGPT ในการแปลภาษา

การแปลโดยปัญญาประดิษฐ์ (Artificial Intelligence Translation) โดยเฉพาะ ChatGPT ซึ่งเป็นโมเดลภาษาแบบ (Large Language Model) ที่พัฒนาโดยบริษัท OpenAI ได้รับความนิยมอย่างแพร่หลายในด้านการแปลภาษาในช่วงหลัง เนื่องจากความสามารถในการประมวลผลภาษาธรรมชาติ การสังเคราะห์ข้อความ และความเร็วในการประมวลผล อย่างไรก็ตาม แม้จะมีความสามารถทางภาษาในระดับสูง แต่ก็ยังคงมีข้อจำกัดด้านความแม่นยำ ความเข้าใจบริบท และการตีความเชิงลึก

จากการศึกษาเปรียบเทียบระหว่างการแปลโดยมนุษย์และ ChatGPT ในบริบทของการแปลภาษาอังกฤษเป็นภาษาไทย โดยใช้กลุ่มตัวอย่างเป็นนักศึกษาระดับปริญญาตรี พบว่า แม้ ChatGPT จะสามารถสร้างข้อความที่มีโครงสร้างประโยคที่ถูกต้องในระดับหนึ่ง และสามารถสื่อสารได้อย่างชัดเจนพอสมควร แต่ยังคงปรากฏข้อผิดพลาดในเชิงโครงสร้างและไวยากรณ์ในสัดส่วนที่สูงกว่าการแปลของมนุษย์ โดยเฉพาะในระดับคำและวลี ซึ่งมักเกิดจากการเลือกใช้คำที่ไม่สอดคล้องกับบริบทวัฒนธรรม หรือไม่สามารถถ่ายทอดความหมายโดยนัยได้อย่างถูกต้อง นอกจากนี้ยังพบว่า ChatGPT มักแปลข้อความในลักษณะตรงตัวมากกว่าการตีความแบบองค์รวม ซึ่งทำให้ความหมายที่ถ่ายทอดออกมาขาดความลึกซึ้ง และไม่สามารถสะท้อนเจตนาารมณ์ของผู้เขียนต้นฉบับได้อย่างแท้จริง ตรงกันข้ามกับนักแปลมนุษย์ที่สามารถประเมินบริบทของเนื้อหา วัตถุประสงค์ของข้อความ และความเหมาะสมกับกลุ่มเป้าหมายได้อย่างแม่นยำ ผลลัพธ์จากการศึกษาแสดงให้เห็นถึงข้อจำกัดของระบบปัญญาประดิษฐ์ในการจัดการกับเนื้อหาที่มีความซับซ้อนสูง โดยเฉพาะเมื่อมีบริบททางวัฒนธรรมเข้ามาเกี่ยวข้อง โดยสรุป แม้ AI อย่าง ChatGPT จะมีศักยภาพในด้านการแปลภาษาในระดับพื้นฐานและสามารถใช้งานได้ในบางบริบท แต่ยังไม่สามารถทดแทนบทบาทของมนุษย์ได้ในด้านความแม่นยำทางภาษาและการตีความเชิงลึก โดยเฉพาะอย่างยิ่งในงานแปลวิชาการที่ต้องการความละเอียดรอบคอบและความเข้าใจทางวัฒนธรรม (นรินทร์ทิพย์ ทองศรี, 2567) ในอีกบริบทหนึ่งของการแปลเฉพาะทาง เช่น ข้อความทางกฎหมาย Altakhaineh et al. (2025) ได้ทำการศึกษาการแปลเอกสารกฎหมายจากภาษาอังกฤษเป็นภาษาอาหรับโดยใช้ ChatGPT รุ่น GPT - 4 ซึ่งพบว่าแม้ระบบปัญญาประดิษฐ์จะสามารถถ่ายทอดข้อความที่มีความถูกต้องทางไวยากรณ์ได้ในระดับหนึ่ง แต่ยังคงมีข้อจำกัดที่ชัดเจนในด้านการเลือกใช้คำศัพท์เฉพาะทางและการตีความความหมายในบริบทของภาษาเป้าหมาย โดยเฉพาะอย่างยิ่งเมื่อข้อความมีความซับซ้อนในเชิงกฎหมาย หรือมีความละเอียดอ่อนที่ต้องการความเข้าใจทางระบบกฎหมายของทั้งสองภาษาการแปลของ ChatGPT ในกรณีนี้แม้จะมีความคล่องแคล่วทางภาษาที่สูง แต่กลับขาดความแม่นยำในการถ่ายทอดเนื้อหาทางนิติศาสตร์ที่จำเป็นต้องอาศัยคำศัพท์เฉพาะ การใช้ภาษาที่เป็นทางการและโครงสร้างที่มีลำดับเหตุผลชัดเจน นักวิจัยยังชี้ให้เห็นว่าข้อผิดพลาดที่พบบ่อย ได้แก่ การใช้คำศัพท์กฎหมายผิดบริบท ความกำกวมใน

โครงสร้างประโยค และการละเลยบริบทสำคัญของข้อความต้นฉบับ ซึ่งทำให้ผู้อ่านอาจตีความเนื้อหาผิดพลาดได้ โดยเฉพาะอย่างยิ่งในเอกสารทางกฎหมายที่ต้องการความเที่ยงตรงสูง การแปลของ ChatGPT บางครั้งยังแสดงให้เห็นถึงการสลับตำแหน่งโครงสร้างของข้อความ ทำให้ความหมายเปลี่ยนแปลงไปจากต้นฉบับ ซึ่งสะท้อนถึงขีดจำกัดของระบบ AI ในการจัดการกับภาษาที่ต้องใช้การตีความแบบลำดับตรรกะอย่างแม่นยำ ผลการศึกษาชี้ว่า ChatGPT แม้จะสามารถใช้งานได้ดีในบริบททั่วไป แต่ยังไม่สามารถทดแทนมนุษย์ในงานแปลเฉพาะทางที่ต้องอาศัยทั้งความเข้าใจด้านภาษาและความรู้ในสาขาเฉพาะนั้น ๆ เช่น กฎหมายหรือวิชาชีพเฉพาะทางอื่น ๆ ดังนั้น การแปลเอกสารในบริบทที่มีความซับซ้อนทางความหมายสูงจึงยังควรอาศัยนักแปลมนุษย์เป็นหลัก เพื่อคงไว้ซึ่งความถูกต้องและความน่าเชื่อถือของข้อความในภาษาเป้าหมาย (Altakhaineh et al., 2025)

การแปลโดยปัญญาประดิษฐ์ เช่น ChatGPT เป็นเครื่องมือที่มีศักยภาพในการแปลภาษา โดยเฉพาะในบริบททั่วไปที่ไม่ซับซ้อนมากนัก แต่เมื่อพิจารณาในแง่ของความแม่นยำ ความเข้าใจทางวัฒนธรรม และการตีความความหมายเชิงลึกแล้วอาจจะยังไม่สามารถทดแทนมนุษย์ได้อย่างสมบูรณ์ โดยเฉพาะอย่างยิ่งในงานแปลวิชาการ หรือเอกสารเฉพาะทางที่ต้องการความละเอียดและความถูกต้องสูง จึงต้องมีการศึกษาแนวทางการใช้งานให้มีประสิทธิภาพมากที่สุดต่อไป

ผู้วิจัยจึงสรุปได้ว่า การแปลโดย ChatGPT หมายถึง การแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษ ด้วยโปรแกรมปัญญาประดิษฐ์ “ChatGPT” โดยผู้วิจัยกำหนดข้อความป้อนเข้า (Prompt) ที่เหมือนกันในทุกบทคัดย่อ เพื่อให้ได้ผลการแปลภายใต้เงื่อนไขเดียวกัน และไม่ผ่านการปรับแก้เพิ่มเติมหลังจากระบบแปลแล้ว

2.4.2 การออกแบบ Prompt เพื่อการแปลภาษาใน ChatGPT

การออกแบบคำสั่งหรือคำถามที่ใช้ในการสื่อสารกับระบบปัญญาประดิษฐ์ (Prompt) ถือเป็นองค์ประกอบสำคัญที่ช่วยให้ ChatGPT สามารถตอบสนองได้ตรงตามความต้องการของผู้ใช้มากยิ่งขึ้น การเรียนรู้เทคนิคในการเขียน Prompt อย่างมีประสิทธิภาพจึงมีความจำเป็นสำหรับผู้ที่ต้องการใช้งาน AI อย่างเต็มศักยภาพ โดย Prompt ทำหน้าที่เป็นแนวทางเชิงโครงสร้างที่ช่วยให้แบบจำลองทางภาษาสามารถเข้าใจภารกิจการแปลได้อย่างชัดเจน ทั้งในด้านทิศทางของภาษาและบริบทของเนื้อหา การปรับปรุง Prompt ให้มีคำแนะนำและตัวอย่างที่เฉพาะเจาะจง จะช่วยเพิ่มความถูกต้อง ความสั้นไหล และความสอดคล้องของผลลัพธ์จากการแปลที่ได้จาก ChatGPT ได้อย่างมีประสิทธิภาพ (Gao, Wang, & Hou, 2023) จากงานวิจัยของ He (2024) ได้มีการทดลองเชิงเปรียบเทียบ ประสิทธิภาพของ Prompt ที่ออกแบบตามแนวคิดด้านการแปลของมนุษย์ ได้แก่ (1) Basic Prompt, (2) Translation Brief Prompt, (3) Author Persona Prompt และ (4) Translator Persona Prompt ผลการวิจัยพบว่า Prompt แบบ Translator Persona ซึ่งออกแบบให้ ChatGPT แสดงบทบาทเสมือนนักแปลมืออาชีพ มีผลลัพธ์ที่ดีที่สุด ทั้งในแง่ของคะแนนประเมินอัตโนมัติ (BLEU, COMET - 2.2) และจากผู้ประเมินมนุษย์ ตัวอย่าง Translator Persona Prompt

ที่มีประสิทธิภาพ ยกตัวอย่างเช่น “คุณเป็นนักแปลวิชาการมืออาชีพ โปรดแปลบทความย่อภาษาไทยต่อไปนี้เป็นภาษาอังกฤษที่คล่องแคล่วและถูกต้อง ซึ่งเหมาะสำหรับการนำเสนอต่อวารสารนานาชาติ โปรดให้ความสำคัญเป็นพิเศษกับความสอดคล้องของคำศัพท์ ความถูกต้องทางไวยากรณ์ และความชัดเจนของรูปแบบการเขียนเชิงวิชาการ” โครงสร้างของ Prompt ดังกล่าวมีองค์ประกอบสำคัญ 3 ประการ ได้แก่ 1) การระบุบทบาท (Role Declaration) เช่น “คุณเป็นนักแปลวิชาการมืออาชีพ” 2) การกำหนดบริบทการใช้งาน (Task Context) เช่น “เหมาะสำหรับการส่งไปตีพิมพ์ในวารสารนานาชาติ” 3) การเน้นคุณภาพเฉพาะทาง (Evaluation Focus): เช่น “ให้ความสำคัญเป็นพิเศษกับความสอดคล้องของคำศัพท์เฉพาะทาง” ผลการศึกษาชี้ให้เห็นว่า การกำหนดบทบาทอย่างชัดเจน (Persona) และบริบทที่สัมพันธ์กับลักษณะของงานแปล สามารถส่งผลต่อการเลือกคำ โครงสร้างประโยค และน้ำเสียงของภาษาเป้าหมายที่มีความเหมาะสมกับประเภทของเนื้อหานั้น ๆ ได้ดีกว่าการใช้คำสั่งทั่วไป ซึ่งมักขาดความเฉพาะเจาะจงในลักษณะของงานแปลทางวิชาการ Yamada (2023) ได้ศึกษาวิธีการออกแบบ Prompt (Prompt Engineering) เพื่อเพิ่มประสิทธิภาพของการแปลภาษาโดยใช้ ChatGPT โดยมุ่งเน้นที่ความสามารถในการ “ปรับแต่งตามบริบท” (Customizability) ของโมเดลเมื่อได้รับคำสั่งแปลข้อความในรูปแบบต่าง ๆ งานวิจัยนี้เน้นการใช้ Prompt ที่เจาะจงและมีบริบท เช่น การระบุ วัตถุประสงค์ของการแปล และ กลุ่มเป้าหมายไว้ในคำสั่งแปล ซึ่งพบว่าส่งผลต่อคุณภาพของข้อความแปลอย่างมีนัยสำคัญ ตัวอย่างเช่น ในการแปลเอกสารทางการแพทย์หรือเอกสารประชาสัมพันธ์ หากผู้ใช้ระบุวัตถุประสงค์ว่า “เพื่อดึงดูดใจผู้อ่าน” และกลุ่มเป้าหมายว่า “ลูกค้าชาวต่างชาติ” ใน Prompt ChatGPT จะสามารถเลือกใช้คำและโทนภาษาให้เหมาะสมมากกว่าการใช้ Prompt ทั่วไป เช่น “แปลประโยคต่อไปนี้เป็นภาษาอังกฤษ” ผลการทดลองพบว่า Prompt ที่มีรายละเอียดเกี่ยวกับบริบท เช่น “แปลข้อความต่อไปนี้เป็นภาษาอังกฤษสำหรับแคมเปญโฆษณาระดับนานาชาติที่มุ่งเป้าไปที่กลุ่มวัยรุ่น ตรวจสอบให้แน่ใจว่าน้ำเสียงมีความน่าสนใจและเป็นมิตร” ให้ผลลัพธ์การแปลที่ดีกว่าในแง่ของ ความเหมาะสมตามวัตถุประสงค์ และ ความเป็นธรรมชาติของภาษาเมื่อเทียบกับ Prompt พื้นฐานอย่างเช่น “แปลประโยคต่อไปนี้เป็นภาษาอังกฤษ”

การออกแบบ Prompt อย่างมีโครงสร้างและเฉพาะเจาะจง โดยเน้นบทบาทของผู้แปล บริบทของเนื้อหา และคุณภาพที่ต้องการ ส่งผลโดยตรงต่อประสิทธิภาพการแปลของ ChatGPT อย่างมีนัยสำคัญ งานวิจัยของ He (2024) และ Yamada (2023) ต่างยืนยันว่า Prompt ที่มีการระบุ Persona และ Task Context อย่างชัดเจน เช่น Prompt แบบ Translator Persona สามารถช่วยให้ ChatGPT แสดงผลลัพธ์การแปลที่มีความแม่นยำ คล่องแคล่ว และเหมาะสมกับบริบทได้มากกว่าการใช้คำสั่งพื้นฐานทั่วไป ดังนั้น การเรียนรู้และประยุกต์ใช้หลักการออกแบบ Prompt อย่างเหมาะสมจึงเป็นกลยุทธ์สำคัญสำหรับการใช้งาน ChatGPT ในงานแปลเชิงวิชาการและงานเฉพาะทางอย่างมีประสิทธิภาพ

2.5 การประยุกต์ใช้เกณฑ์ MQM กับการวิเคราะห์คุณภาพการแปลภาษาไทยเป็นภาษาอังกฤษ

แม้ว่าเกณฑ์การประเมินคุณภาพการแปล Multidimensional Quality Metrics (MQM) ที่พัฒนาโดย (Lommel et al., 2014) จะยังไม่ถูกนำมาใช้อย่างแพร่หลายในงานวิจัยด้านการแปลภาษาไทยเป็นภาษาอังกฤษในประเทศไทยโดยตรง แต่มีงานวิจัยจำนวนหนึ่งที่ได้ศึกษาข้อผิดพลาดในการแปลโดยใช้เกณฑ์อื่น หรือเกณฑ์ที่สามารถเทียบเคียงกับ MQM ได้ในทางปฏิบัติ ตัวอย่างเช่น การศึกษาของ พุทธชาติ ลิ้มศิริเรืองไร (2565) พบว่าข้อผิดพลาดด้านไวยากรณ์ โดยเฉพาะการใช้กาลเวลามีสัดส่วนสูงถึง 68.91% ของข้อผิดพลาดทั้งหมด รองลงมาคือการใช้คำศัพท์ที่ไม่เหมาะสม (24.39%) และการใช้ภาษาที่ไม่เป็นทางการ (6.7%) ซึ่งสะท้อนลักษณะปัญหาที่ตรงกับหมวด **Fluency และ Accuracy** ตามเกณฑ์ MQM ทั้งนี้ พุทธชาติ ลิ้มศิริเรืองไร (2565) ยังชี้ว่าข้อผิดพลาดเหล่านี้สะท้อนถึงปัญหาทั่วไปในการแปลบทความเชิงวิชาการของผู้เรียน เช่น การไม่เข้าใจโครงสร้างไวยากรณ์ของภาษาเป้าหมาย และการขาดความตระหนักถึงความสำคัญของการเลือกใช้ภาษาให้เหมาะสมกับบริบททางวิชาการ

ในทำนองเดียวกัน Jeharsae และ Boonsuk (2024) ได้ศึกษาข้อผิดพลาดในการแปลของนักศึกษาจากภาคใต้ของประเทศไทย และพบว่าข้อผิดพลาดส่วนใหญ่อยู่ในหมวดเชิงโครงสร้าง (Syntactic Errors) คิดเป็น 46.88% และข้อผิดพลาดเชิงความหมาย (Semantic Errors) คิดเป็น 40.76% ซึ่งสามารถเชื่อมโยงกับเกณฑ์ย่อยของ MQM ได้แก่ **Grammar, Mistranslation และ Untranslated Text ได้อย่างชัดเจน** จากตัวอย่างเหล่านี้แสดงให้เห็นว่า แม้งานวิจัยเหล่านี้จะไม่ได้ใช้ MQM โดยตรง แต่ประเด็นข้อผิดพลาดที่ศึกษาในแต่ละกรณีมีลักษณะ สอดคล้องกันกับองค์ประกอบของ MQM อย่างมีนัยสำคัญ จึงสามารถนำมาใช้เป็นพื้นฐานในการประยุกต์ใช้กรอบ MQM กับการประเมินคุณภาพการแปลภาษาไทยเป็นภาษาอังกฤษได้อย่างเหมาะสม

ในด้านความถูกต้อง (Accuracy) งานวิจัยในประเทศไทยให้ความสำคัญเป็นอย่างมาก เนื่องจากการถ่ายทอดความหมายจากภาษาไทยไปสู่ภาษาอังกฤษต้องอาศัยความเข้าใจลึกซึ้งทั้งด้านโครงสร้างและระบบความหมายที่แตกต่างกันระหว่างสองภาษา ผลการศึกษาของ พุทธชาติ ลิ้มศิริเรืองไร (2565) พบว่าข้อผิดพลาดประเภท Mistranslation หรือการแปลผิดความหมาย เป็นปัญหาหลัก คิดเป็นสัดส่วนสูงที่สุด (ร้อยละ 68.91) เมื่อเทียบกับการเลือกใช้คำศัพท์ที่ไม่เหมาะสม (ร้อยละ 24.39) และการใช้ระดับภาษาที่ไม่เป็นทางการ (ร้อยละ 6.7) ซึ่งชี้ให้เห็นถึงความเข้าใจผิดทั้งทางด้านไวยากรณ์และความหมายของภาษาเป้าหมาย โดยนักวิจัยเน้นว่าปัญหาเหล่านี้เป็นผลมาจากการขาดความรู้เชิงระบบของภาษาปลายทางและความไม่ชัดเจนของการเลือกใช้ภาษาตามบริบท Duklim (2022) ชี้ให้เห็นข้อผิดพลาดที่สอดคล้องกับเกณฑ์ของ MQM ได้อย่างชัดเจน โดยในด้าน Accuracy มีข้อผิดพลาดเชิงความหมาย เช่น การแปลผิด (Mistranslation) และการละเว้นเนื้อหา (Omission) ซึ่งส่งผลให้สาระสำคัญของต้นฉบับถูกบิดเบือนหรือลดทอนลง ขณะที่ด้าน Fluency พบข้อผิดพลาดเชิงโครงสร้าง เช่น การใช้ไวยากรณ์ไม่ถูกต้อง (Grammar) การใช้รูปแบบประโยคไม่เหมาะสมกับบริบททางวิชาการ ตลอดจนการใช้วรรคตอนและคำเชื่อมที่ไม่เหมาะสม

(Punctuation and Cohesion) ทำให้บทแปลขาดความลื่นไหลและความเป็นธรรมชาติ Duklim (2022) ได้วิเคราะห์สาเหตุของข้อผิดพลาดเหล่านี้ไว้อย่างละเอียด โดยระบุว่าปัจจัยสำคัญ 3 ประการ ได้แก่ (1) ความไม่เข้าใจความหมายของต้นฉบับอย่างถ่องแท้ (2) การไม่สามารถจับโครงสร้างของภาษาไทยและถ่ายทอดเป็นภาษาอังกฤษได้อย่างถูกต้อง และ (3) ความไม่คุ้นเคยกับรูปแบบภาษาอังกฤษที่ใช้ในบริบทวิชาการ โดยเฉพาะภาษาเขียนที่มีโครงสร้างซับซ้อนและระดับภาษาที่เป็นทางการ ทั้งหมดนี้แสดงให้เห็นถึงความจำเป็นในการพัฒนาทักษะการแปลแบบบูรณาการไม่เพียงแต่ในเชิงไวยากรณ์หรือคำศัพท์ แต่ต้องเน้นที่ความเข้าใจในระดับข้อความ (Discourse - Level Understanding) และการฝึกฝนการถ่ายทอดภาษาไทยให้เป็นภาษาอังกฤษที่มีความสมบูรณ์ ทั้งในด้านความหมายและรูปแบบตามมาตรฐานวิชาการ ซึ่งสอดคล้องกับจุดมุ่งหมายของ MQM ที่เน้นการประเมินอย่างรอบด้านทั้งด้านเนื้อหา ความสมบูรณ์ และความสละสลวยของภาษาที่ใช้ สำหรับด้านความคล่องแคล่ว (Fluency) ซึ่งประกอบด้วยความสละสลวยของภาษา การจัดวางโครงสร้างประโยค และความถูกต้องตามหลักไวยากรณ์นั้น มักถูกมองข้ามในการแปลบทคัดย่อเชิงวิชาการ โดยเฉพาะในกรณีของผู้แปลมีพื้นฐานภาษาอังกฤษไม่แข็งแรง แม้ว่าผู้แปลจะสามารถถ่ายทอดเนื้อหาหลักได้อย่างถูกต้อง แต่หากภาษาเป้าหมายขาดความลื่นไหลหรืออ่านไม่เป็นธรรมชาติ ก็จะส่งผลต่อความน่าเชื่อถือของบทคัดย่อโดยรวม งานวิจัยของ Tongpoon - Patanasorn และ Griffith (2020) พบว่า Google Translate ให้ผลการแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษในระดับปานกลาง โดยข้อผิดพลาดบ่อย ได้แก่ การใช้ตัวพิมพ์ใหญ่ไม่เหมาะสม การใช้วรรคตอนผิด และการตัดหรือแบ่งข้อความอย่างไม่เหมาะสม ซึ่งล้วนจัดอยู่ในหมวด Fluency ตามกรอบ MQM โดยเฉพาะด้านย่อย Punctuation และ Spelling ความผิดพลาดเหล่านี้สะท้อนให้เห็นข้อจำกัดของระบบแปลอัตโนมัติในแง่ของความแม่นยำทางโครงสร้างภาษา (Grammar) และความสละสลวยทางภาษา (Naturalness) ซึ่งเป็นองค์ประกอบสำคัญในการสร้างความเข้าใจที่ชัดเจนในงานเขียนทางวิชาการ

ด้านความสอดคล้องของศัพท์เฉพาะ (Terminology Consistency) ซึ่งหมายถึงความถูกต้องและความสอดคล้องในการใช้คำศัพท์เฉพาะทางในงานแปลหรือการเขียนเชิงวิชาการ หากใช้ศัพท์ไม่ถูกต้องตามบริบท หรือใช้ไม่สอดคล้องกันในบริบทเดียวกัน ย่อมทำให้เกิดความสับสนและลดความชัดเจนของสาร งานวิจัยของ Boonraksa และ Naisena (2022) ชี้ให้เห็นว่า นักศึกษาไทยที่เรียนภาษาอังกฤษในฐานะภาษาต่างประเทศ (EFL) มักประสบข้อผิดพลาดเกี่ยวกับ Collocation ซึ่งเป็นการจับคู่คำศัพท์ เช่น Make a decision แทน Do a decision โดยผู้เรียนจำนวนมากไม่สามารถเลือกใช้คำที่ถูกต้องร่วมกับคำนามหรือคำกริยาตามธรรมชาติของเจ้าของภาษาได้ ความผิดพลาดนี้สะท้อนถึงความไม่เข้าใจในคำศัพท์เชิงแนวร่วม (Collocational Knowledge) ซึ่งถือเป็นส่วนหนึ่งของ Terminology ที่ควรใช้ให้เหมาะสมกับประเภทของข้อความและบริบทการสื่อสาร ความไม่สอดคล้องของ Collocation ดังกล่าว จึงสามารถจัดอยู่ในหมวด “Incorrect Term” และ “Inconsistent Term” ภายใต้หมวดใหญ่ Terminology Consistency ตามเกณฑ์ MQM ได้อย่างชัดเจน เพราะความคลาดเคลื่อนใน collocation ส่งผลโดยตรงต่อความแม่นยำของเนื้อหาและความเป็นธรรมชาติของภาษาที่ใช้ในบทคัดย่อเชิงวิชาการ

ทั้งนี้ งานของ พุทธชาติ ลิ้มศิริเรืองโร (2565) และ Jeharsae (2024) ยังแสดงให้เห็นว่าการวิเคราะห์เชิงคุณภาพสามารถให้ข้อมูลเชิงลึกที่การวิเคราะห์เชิงปริมาณไม่สามารถอธิบายได้เพียงพอ โดยสามารถเจาะลึกถึงข้อผิดพลาดเฉพาะจุด ซึ่งสอดคล้องกับหลักการของ MQM ที่เน้นการประเมินแยกตามประเภทข้อผิดพลาด อาทิ ด้านการตีความ (Interpretation), ความสมบูรณ์ของเนื้อหา (Completeness) และความเป็นธรรมชาติของภาษา (Naturalness) การบูรณาการการวิเคราะห์เชิงคุณภาพเช่นนี้ จึงเป็นแนวทางสำคัญในการพัฒนาไม่เพียงเฉพาะทักษะผู้แปลมนุษย์ แต่รวมถึงการออกแบบและพัฒนาเครื่องมือแปลอัตโนมัติ เช่น AI ให้สามารถแปลได้อย่างถูกต้องและมีคุณภาพมากขึ้นในอนาคต

2.6 เทคนิคการควบคุมคุณภาพการแปล

การตรวจสอบคุณภาพการแปลในบทความวิทยานิพนธ์มีความซับซ้อนและต้องการความแม่นยำสูง เพื่อให้การสื่อสารองค์ความรู้วิชาการในระดับสากลเป็นไปอย่างมีประสิทธิภาพ การตรวจสอบในปัจจุบันจึงต้องพิจารณาจากหลายมิติควบคู่กัน (Multidimensional) เพื่อระบุข้อผิดพลาดทั้งในด้านความถูกต้องของเนื้อหาและความเหมาะสมของภาษา (Lommel et al., 2014) โดยผู้วิจัยได้กำหนดเทคนิคเชิงระบบเพื่อใช้ในการควบคุมคุณภาพ ดังนี้

2.6.1 เทคนิคการแปลย้อนกลับ (Back Translation) และมุมมองเชิงหน้าที่ การควบคุมคุณภาพการแปลในงานวิจัยระดับสาขานิยมใช้เทคนิคการแปลย้อนกลับเพื่อรักษาความเที่ยงตรงของสาร ควบคู่ไปกับการพิจารณาเชิงหน้าที่เพื่อให้ภาษาลื่นไหลเป็นธรรมชาติ โดยมีรายละเอียดดังนี้

1) กระบวนการแปลย้อนกลับ การแปลย้อนกลับเป็นกระบวนการตรวจสอบความเท่าเทียมเชิงเนื้อหา (Content Equivalence) ที่เข้มงวดที่สุด โดยเทคนิคที่พัฒนาโดย Brislin (1970) ยังคงถูกนำมาปรับใช้เป็นแนวทางหลักในงานวิจัยร่วมสมัย (Tsang et al., 2017) กระบวนการนี้ประกอบด้วย 3 ระยะเวลาสำคัญ ได้แก่

(1) **ระยะเวลาการแปลไปข้างหน้า (Forward Translation)** การแปลจากภาษาไทยเป็นภาษาอังกฤษโดยเน้นความครบถ้วนของเนื้อหา

(2) **ระยะเวลาการแปลย้อนกลับ (Back Translation)** การนำบทแปลภาษาอังกฤษมาแปลกลับเป็นภาษาไทยโดยนักแปลอิสระที่ไม่เห็นต้นฉบับมาก่อนเพื่อตรวจสอบความคงอยู่ของแก่นสาร

(3) **ระยะเวลาเปรียบเทียบผลลัพธ์ (Discrepancy Analysis)** การหาจุดคลาดเคลื่อนเพื่อระบุข้อผิดพลาดในด้านความถูกต้อง (Accuracy) เช่น การละเลยเนื้อหา (Omission) หรือการแปลผิดความหมาย (Mistranslation)

2) มุมมองเชิงหน้าที่และการปรับใช้ แม้การแปลย้อนกลับจะช่วยเรื่องความแม่นยำ แต่ Colina et al. (2017) กล่าวว่า การยึดติดกับความตรงกันของคำมากเกินไปอาจทำให้บทแปลไม่เป็นธรรมชาติ จึงต้องนำทฤษฎีหน้าที่นิยม (Functionalist Theory) มาปรับใช้ร่วมด้วย โดยเน้น

ที่วัตถุประสงค์ของบทแปลและความเหมาะสมต่อกลุ่มเป้าหมาย เพื่อให้บทคัดย่อมีระดับภาษาที่เป็นทางการและลื่นไหลตามมาตรฐานสากล ซึ่งจะช่วยลดข้อผิดพลาดในด้านความคล่องแคล่ว (Fluency) และระดับภาษา (Register) ตามเกณฑ์ MQM

3) การตรวจแก้หลังการประมวลผล (Post - Editing) การตรวจแก้หลังการประมวลผล (Post - Editing: PE) คือกระบวนการที่ผู้แปลที่เป็นมนุษย์ดำเนินการปรับปรุงและแก้ไขบทแปลที่ผลิตโดยระบบแปลอัตโนมัติหรือปัญญาประดิษฐ์เพื่อให้ได้คุณภาพตามมาตรฐานวิชาการ Vieira (2019) กล่าวว่า การตรวจแก้โดยมนุษย์ได้วิวัฒนาการมาเป็นส่วนกลางของกระบวนการแปลมืออาชีพ โดยเฉพาะในงานแปลเฉพาะทางที่ AI ยังคงมีข้อจำกัดด้านการเลือกใช้ศัพท์และระดับภาษาที่ละเอียดอ่อน สอดคล้องกับแนวคิดของ Koponen (2016) ที่ระบุว่ากระบวนการดังกล่าวช่วยรักษาสมดุลระหว่างความเร็วในการผลิตงานและความถูกต้องแม่นยำในระดับที่สามารถนำไปเผยแพร่ได้ (Publishable Quality)

นอกจากนี้ Herbig et al. (2019) ยังสนับสนุนว่าการบูรณาการระหว่างปัญญาประดิษฐ์และมนุษย์ผ่านกระบวนการ Post - Editing สามารถช่วยเพิ่มประสิทธิภาพในกระบวนการแปลได้ถึงร้อยละ 36 โดยยังคงรักษาคุณภาพของเนื้อหาและความสละสลวยไว้ได้อย่างเหมาะสม

2.6.2 การจัดการคลังศัพท์เฉพาะทาง (Terminology Management and Termbase)

การรักษาความสอดคล้องในการใช้คำศัพท์ (Consistency) เป็นหัวใจสำคัญของงานแปลบทคัดย่อ โดยเฉพาะในสาขาวิทยาศาสตร์และเทคโนโลยี Cabezas - García (2022) ระบุว่า การจัดทำฐานข้อมูลคำศัพท์เฉพาะ (Termbase) มีส่วนสำคัญในการลดความกำกวมและเพิ่มความแม่นยำทางเทคนิค โดยกระบวนการนี้ประกอบด้วยการจัดเก็บข้อมูลเชิงระบบ ได้แก่ รายการคำศัพท์ (Entry) นิยามความหมายในบริบทวิชาการ (Definition) และตัวอย่างประโยคการใช้งาน (Context)

การใช้ Termbase จะช่วยป้องกันปัญหาการใช้คำศัพท์ที่ไม่สม่ำเสมอ (Inconsistent Term) ซึ่งมักพบในงานแปลจากปัญญาประดิษฐ์หรือการแปลโดยมนุษย์ที่ขาดเครื่องมือควบคุมศัพท์เฉพาะทาง การจัดการคลังศัพท์อย่างเป็นระบบจึงเป็นเทคนิคสำคัญที่ช่วยให้บทคัดย่อมีความน่าเชื่อถือและเป็นไปตามมาตรฐานสากล

2.7 งานวิจัยที่เกี่ยวข้องในประเทศไทย

ในบริบทของประเทศไทย การศึกษาเปรียบเทียบคุณภาพการแปลระหว่างมนุษย์และเครื่องแปลภาษาได้รับความสนใจอย่างมาก Tongpoon - Patanasorn และ Griffith (2020) ได้ทำการศึกษา Google Translate และคุณภาพการแปล: กรณีศึกษาการแปลบทคัดย่อทางวิชาการจากภาษาไทยเป็นภาษาอังกฤษ โดยมุ่งเน้นการประเมินคุณภาพของการแปลบทคัดย่อทางวิชาการจากภาษาไทย

เป็นภาษาอังกฤษด้วย Google Translate พบว่าผลลัพธ์การแปลมีความเข้าใจและความสามารถในการใช้งานในระดับปานกลาง โดยมีข้อผิดพลาดที่พบบ่อยคือการใช้ตัวพิมพ์ใหญ่ เครื่องหมายวรรคตอน และการแบ่งส่วนที่ไม่ถูกต้อง ซึ่งแสดงให้เห็นว่าแม้เครื่องแปลภาษาจะมีประโยชน์ แต่ยังต้องการการแก้ไขจากมนุษย์เพื่อคุณภาพที่ดียิ่งขึ้น

กัญญาภัก สาสงเคราะห์ (2562) ได้ศึกษาเชิงเปรียบเทียบการแปลเว็บไซต์โรงแรมระหว่างมนุษย์และเครื่องมือแปลภาษา Google Translate จากภาษาไทยเป็นภาษาอังกฤษ เพื่อความแตกต่างของการแปลเว็บไซต์โรงแรมระหว่างมนุษย์และโปรแกรมแปลภาษา Google Translate ในการแปลข้อความประเภทบอกกล่าวและข้อความ จูงใจ และวิเคราะห์ความแตกต่างของการแปลในระดับคำ ระดับประโยค และระดับข้อความ รวมถึงเปรียบเทียบความเป็นธรรมชาติในภาษาแปลที่โรงแรมในประเทศไทยทั้งหมด 29 แห่ง โดยผลการศึกษาพบว่า มนุษย์มีการใช้กลวิธีการแปลที่หลากหลาย มีสาทลักษณ์ของภาษาแปล สามารถปรับบทแปลได้ทุกระดับ ในขณะที่โปรแกรมแปลภาษา Google Translate ยังไม่สามารถปรับบทแปลในระดับข้อความได้ และมีข้อผิดพลาดจากการแบ่งประโยคผิด และมีข้อจำกัดในการถ่ายทอดคำทางวัฒนธรรม

Duklim (2022) ได้ศึกษาวิจัยเรื่อง ข้อผิดพลาดในการแปลของนักศึกษามหาวิทยาลัยไทย: การศึกษาประเภทและสาเหตุ งานวิจัยนี้มีวัตถุประสงค์หลัก 2 ประการ ได้แก่ 1) เพื่อสำรวจและจำแนกประเภทของข้อผิดพลาดในการแปลประโยคจากภาษาไทยเป็นภาษาอังกฤษของนักศึกษาระดับปริญญาตรี และ 2) เพื่อวิเคราะห์สาเหตุที่นำไปสู่ข้อผิดพลาดในการแปลดังกล่าว โดยผู้วิจัยได้เก็บรวบรวมข้อมูลจากแบบทดสอบการแปลและนำมาวิเคราะห์ตามกรอบทฤษฎีทางภาษาศาสตร์ ผลการศึกษาตามวัตถุประสงค์ข้อแรกพบว่า ข้อผิดพลาดที่พบบ่อยที่สุดคือ ข้อผิดพลาดทางโครงสร้างไวยากรณ์ (Syntactic Errors) ซึ่งมีสัดส่วนสูงถึงร้อยละ 65.00 ของข้อผิดพลาดทั้งหมด โดยปัญหาที่เด่นชัดคือเรื่อง กาล (Tense) และความสอดคล้องระหว่างประธานและกริยา (Subject-Verb Agreement) รองลงมาคือ ข้อผิดพลาดเชิงความหมาย (Semantic Errors) คิดเป็นร้อยละ 24.60 และข้อผิดพลาดทางพันธะภาษา (Miscellaneous Errors) ร้อยละ 10.40 ตามลำดับ สำหรับผลการศึกษาตามวัตถุประสงค์ข้อที่สองเกี่ยวกับสาเหตุของข้อผิดพลาด พบว่าเกิดจาก 2 ปัจจัยหลักคือ การแทรกแซงของภาษาแม่ (Interlingual Errors) ซึ่งผู้แปลนำโครงสร้างภาษาไทยมาใช้ในภาษาอังกฤษ และ ความซับซ้อนภายในภาษาเป้าหมายเอง (Intralingual Errors) ซึ่งทำให้นักศึกษาสับสนในกฎเกณฑ์ทางไวยากรณ์

Wongranu (2017) ได้ศึกษาวิจัยเรื่องข้อผิดพลาดในการแปลของนักศึกษาวิชาเอกภาษาอังกฤษ: การศึกษาประเภทและสาเหตุ งานวิจัยนี้มีวัตถุประสงค์หลัก 2 ประการ ได้แก่ 1) เพื่อวิเคราะห์ประเภทของข้อผิดพลาดในการแปลจากภาษาไทยเป็นภาษาอังกฤษของนักศึกษาวิชาเอกภาษาอังกฤษ และ 2) เพื่อระบุสาเหตุหลักที่ก่อให้เกิดข้อผิดพลาดในการแปลเหล่านั้น ผลการศึกษาตามวัตถุประสงค์แรกพบว่า ข้อผิดพลาดที่เกิดขึ้นบ่อยที่สุดคือ ข้อผิดพลาดทางโครงสร้าง

ไวยากรณ์ (Syntactic Errors) และข้อผิดพลาดทางความหมาย (Semantic Errors) ตามลำดับ โดยข้อผิดพลาดทางไวยากรณ์ที่เด่นชัดได้แก่ การใช้ Tense ผิด การเรียงลำดับคำผิด (Word Order) และการใช้คำบุพบทไม่ถูกต้อง ส่วนผลการศึกษาตามวัตถุประสงค์ที่สองเกี่ยวกับสาเหตุของข้อผิดพลาด ผู้วิจัยพบว่าสาเหตุที่สำคัญที่สุดคือ การแทรกแซงของภาษาแม่ (Mother Tongue Interference) หรือการที่ผู้แปลนำโครงสร้างประโยคภาษาไทยไปใช้ในการสร้างประโยคภาษาอังกฤษโดยตรง (Literal Translation) ส่งผลให้ประโยคภาษาอังกฤษที่ได้มีความผิดเพี้ยนและไม่เป็นธรรมชาติ นอกจากนี้ สาเหตุรองลงมาคือความรู้ทางไวยากรณ์ภาษาอังกฤษที่ไม่เพียงพอ (Intralingual Errors) ซึ่งทำให้ผู้แปลไม่สามารถเลือกใช้โครงสร้างที่ซับซ้อนได้อย่างถูกต้อง

Pornwiryakit และ Dandee (2022) ได้ศึกษาวิจัยเรื่องการวิเคราะห์ปัญหาการแปลและการประยุกต์ใช้วิธีแปลแบบไวยากรณ์และการแปลในชั้นเรียนภาษาอังกฤษในฐานะภาษาต่างประเทศ ได้ทำการศึกษาโดยมีวัตถุประสงค์หลัก 2 ประการ คือ 1) เพื่อศึกษาปัญหาการแปลโครงสร้างประโยคและคำศัพท์จากภาษาอังกฤษเป็นภาษาไทยของนักศึกษาสาขาวิชาภาษาอังกฤษเพื่อการสื่อสารสากล และ 2) เพื่อศึกษาแนวทางการแก้ปัญหาการแปลดังกล่าวโดยใช้วิธีสอนแบบไวยากรณ์และการแปล (Grammar - Translation Method) กลุ่มตัวอย่างในการวิจัย คือ นักศึกษามหาวิทยาลัยเทคโนโลยีราชมงคลตะวันออก จำนวน 127 คน ผลการศึกษาตามวัตถุประสงค์ข้อแรกพบว่า ก่อนการทดลอง นักศึกษามีคะแนนด้านไวยากรณ์ (Grammatical Aspect) อยู่ในระดับต่ำมาก โดยมีค่าเฉลี่ยเพียงร้อยละ 16.38 ในขณะที่ด้านคำศัพท์ (Lexical Aspect) มีคะแนนสูงกว่าที่ร้อยละ 70.30 ข้อผิดพลาดทางไวยากรณ์ที่พบบ่อยที่สุด ได้แก่ เรื่องโครงสร้างกรรมวาจก (Passive Voice) ซึ่งนักศึกษามักแปลผิดโดยยึดติดกับโครงสร้างภาษาอังกฤษ (English Structural Attachment) ถึงร้อยละ 48.7 รองลงมาคือเรื่องกาล (Tense) และคำเชื่อม (Connectives) อย่างไรก็ตาม ผลการศึกษาตามวัตถุประสงค์ข้อที่สองพบว่า หลังจากใช้วิธีสอนแบบ Grammar - Translation Method คะแนนด้านไวยากรณ์ของนักศึกษาเพิ่มขึ้นอย่างมีนัยสำคัญเป็นร้อยละ 78.8 งานวิจัยนี้ชี้ให้เห็นอย่างชัดเจนว่า จุดอ่อนที่สุดของมนุษย์ (ในบริบทนักศึกษาไทย) คือความแม่นยำทางไวยากรณ์และโครงสร้างประโยค ซึ่งเป็นผลมาจากการยึดติดกับรูปแบบภาษาต้นฉบับมากเกินไป

นอกจากนี้ นรินทร์ทิพย์ ทองศรี (2567) ได้ศึกษาวิจัยเรื่องการวิเคราะห์เปรียบเทียบการแปลภาษาอังกฤษเป็นภาษาไทยระหว่างผู้แปลที่เป็นมนุษย์และโปรแกรมแปลภาษา (Artificial Intelligence: AI ChatGPT) โดยมีวัตถุประสงค์เพื่อวิเคราะห์และเปรียบเทียบข้อบกพร่องในการแปลภาษาอังกฤษเป็นภาษาไทย โดยใช้ข้อความจากนักศึกษาและโปรแกรม AI ChatGPT เป็นกรณีศึกษา การวิจัยนี้ดำเนินการโดยกลุ่มนักศึกษาหลักสูตรศิลปศาสตรบัณฑิตจำนวน 20 คน ในภาคการศึกษาที่ 2 ปีการศึกษา 2563 ซึ่งได้ทำการแปลข้อความจำนวน 11 ข้อความ ผลการศึกษาพบว่าข้อบกพร่องในการแปลเกิดขึ้นทั้งจากนักศึกษาและ AI โดยสามารถจัดกลุ่มข้อผิดพลาดได้เป็น 3 ระดับ ได้แก่ ระดับคำ ระดับวลี และระดับประโยค การแปลที่ผิดพลาดโดยนักศึกษายู่ที่ร้อยละ 30

ในขณะที่การแปลโดย AI ChatGPT อยู่ที่ร้อยละ 70 ข้อเสนอแนะจากการวิจัยคือควรมีการทำงานร่วมกันระหว่างการแปลของมนุษย์และ AI เพื่อเพิ่มความแม่นยำและคุณภาพในการแปล

จากการสังเคราะห์งานวิจัยที่เกี่ยวข้องในบริบทของประเทศไทย พบข้อค้นพบที่น่าสนใจว่าแม้เครื่องมือแปลภาษาอัตโนมัติอย่าง Google Translate และ AI ChatGPT จะมีบทบาทสำคัญในการเพิ่มความรวดเร็วและลดต้นทุนในการแปลได้อย่างมีประสิทธิภาพ แต่ยังคงปรากฏข้อจำกัดในด้านความถูกต้องของบริบทและความเป็นธรรมชาติทางภาษา ในขณะเดียวกัน การแปลโดยมนุษย์ (โดยเฉพาะกลุ่มนักศึกษาไทย) ก็ประสบปัญหาสำคัญเช่นกัน โดยงานวิจัยชี้ให้เห็นอย่างชัดเจนว่าผู้แปลมักติดกับดักของการแทรกแซงจากภาษาแม่ (Mother Tongue Interference) และการแปลแบบตรงตัว ส่งผลให้เกิดข้อผิดพลาดทางโครงสร้างไวยากรณ์ (Syntactic Errors) ในสัดส่วนที่สูง ดังนั้นสถานการณ์ปัจจุบันในประเทศไทยจึงสะท้อนให้เห็นว่า ทั้งมนุษย์และปัญญาประดิษฐ์ต่างมีจุดแข็งและจุดอ่อนที่แตกต่างกัน กล่าวคือ มนุษย์อาจเข้าใจบริบทได้ดีกว่าแต่อ่อนด้อยเรื่องความแม่นยำทางโครงสร้าง ในขณะที่ AI อาจแม่นยำเรื่องรูปแบบแต่ขาดความลึกซึ้งทางบริบท งานวิจัยในไทยจึงชี้ไปในทิศทางเดียวกันว่า การบูรณาการเทคโนโลยี AI เข้ากับทักษะการตรวจสอบและปรับแก้ของมนุษย์ (Post - Editing) เป็นแนวทางที่เหมาะสมที่สุดในการยกระดับคุณภาพงานแปลเชิงวิชาการ ให้มีความถูกต้องตามหลักไวยากรณ์ สละสลวย และเหมาะสมกับบริบทอย่างแท้จริง

2.8 งานวิจัยที่เกี่ยวข้องในต่างประเทศ

ในบริบทการศึกษาต่างประเทศ การศึกษาคุณภาพการแปลของเครื่องแปลภาษาและมนุษย์ได้รับความนิยมเช่นกัน Khoshafah (2023) ได้ศึกษาวิจัยเรื่อง การประเมินความถูกต้องของ ChatGPT ในการแปลภาษาอาหรับเป็นภาษาอังกฤษ โดยผลการศึกษาพบว่า การแปลภาษาโดย ChatGPT โดยรวมมีความแม่นยำแต่ขาดความลึกซึ้งและละเอียดอ่อนเหมือนการแปลโดยมนุษย์ การแปลโดยมนุษย์มีความแม่นยำมากกว่า โดยเฉพาะอย่างยิ่งสำหรับข้อความที่ซับซ้อน ChatGPT มีประโยชน์สำหรับการแปลแบบง่าย ๆ แต่ต้องมีการควบคุมดูแลโดยมนุษย์สำหรับเนื้อหาเฉพาะทาง

İşim และ Balcıoğlu (2023) ได้ศึกษาประสิทธิภาพของ ChatGPT ในการแปลภาษาตุรกีเป็นภาษาอังกฤษ โดยใช้ข้อความ 50 ย่อหน้าจากบทความและบทคัดย่อทางการศึกษาเป็นตัวอย่งทดสอบ ผลการวิจัยพบว่า ChatGPT สามารถแปลเนื้อหาทางการศึกษาได้ในระดับที่เทียบเคียงกับนักแปลที่เป็นมนุษย์ อย่างไรก็ตาม ยังพบข้อผิดพลาดรวม 70 จุด ซึ่งแบ่งเป็นข้อผิดพลาดทางไวยากรณ์ 38 จุด (ร้อยละ 40) และข้อผิดพลาดทางคำศัพท์ 42 จุด (ร้อยละ 60) ซึ่งบ่งชี้ว่ายังต้องมีการปรับปรุงเนื้อหาเพิ่มเติม นอกจากนี้ Lyons (2020) ศึกษาคุณภาพการแปลภาษาไทยเป็นภาษาอังกฤษด้วยเครื่องแปลภาษา โดยประเมินผลทั้งในเชิงปริมาณด้วยเมตริกอัตโนมัติ (Automatic Metrics) และเชิงคุณภาพผ่านการประเมินโดยมนุษย์ (Human - Based Evaluation) ซึ่งครอบคลุมการจำแนกข้อผิดพลาด ความเข้าใจในการอ่าน และการวิเคราะห์จากนักแปลมืออาชีพ ผลการวิจัยพบว่าการแปลภาษาไทยเป็นภาษาอังกฤษด้วยเครื่องแปลภาษายังคงมีความท้าทาย

โดยมีอัตราข้อผิดพลาดสูงถึง 47.2% และคะแนน BLEU ที่ 0.21 ซึ่งสะท้อนให้เห็นถึงความท้าทายในการแปลระหว่างสองภาษาดังกล่าว โดยทั่วไปคะแนน BLEU ที่ต่ำบ่งชี้ว่าผลลัพธ์การแปลมีความแตกต่างจากภาษาเป้าหมายที่ควรจะเป็น ทั้งในด้านความถูกต้องของคำศัพท์ โครงสร้างประโยค และความลื่นไหลของภาษา อย่างไรก็ตาม ผู้ใช้ยังสามารถเข้าใจเนื้อหาโดยรวมได้ แม้การแปลจะไม่สมบูรณ์

Wang (2024) เน้นการประเมินคุณภาพงานแปลโดยใช้ ChatGPT - 4o ผ่านกรอบแนวคิด Multidimensional Quality Metrics (MQM) โดยทดสอบกับงานแปลต้นฉบับที่ประกอบด้วยงานเขียนประเภทนิยายและงานเขียนไม่ใช่นิยาย ซึ่งนักเรียนสาขา MTI ทำเอง แล้วเปรียบเทียบผลระหว่างเวอร์ชันของมนุษย์กับเวอร์ชันที่ปรับด้วย ChatGPT - 4o โดยใช้การให้คะแนนตาม MQM ผลการวิเคราะห์คะแนนและเชิงคุณภาพแสดงให้เห็นว่า ChatGPT - 4o มีความสอดคล้องสูงกับผู้ประเมินมนุษย์ในการใช้เกณฑ์ MQM นอกจากนี้คำแนะนำจาก ChatGPT - 4o ยังช่วยปรับปรุงคุณภาพการแปล โดยเฉพาะประเด็นที่สำคัญอย่าง “ความสอดคล้องของศัพท์เฉพาะ (Terminology Consistency)” และ “ความถูกต้องตามหลักไวยากรณ์ (Grammatical Accuracy)” ได้อย่างมีนัยสำคัญ

Lau et al. (2024) ได้ศึกษาการเปรียบเทียบคุณภาพการแปลระหว่างมนุษย์และปัญญาประดิษฐ์ (Artificial Intelligence: AI) โดยมุ่งเน้นการแปลจากภาษาจีนกลาง (Mandarin) เป็นภาษามลายู (Malay) ซึ่งเป็นภาษากลุ่มทรัพยากรต่ำ (Low - Resource Language) งานวิจัยนี้พัฒนาชุดข้อมูลที่ประกอบด้วยข้อความต้นฉบับภาษาจีนกลาง และคำแปลจากทั้งนักแปลมนุษย์และระบบ AI โดยนักแปลมนุษย์มีทั้งเจ้าของภาษาจีน (Native Chinese), ผู้แปลในโครงการ Belt and Road, และเจ้าของภาษามลายู (Native Malay) ส่วนระบบ AI ใช้โมเดล ChatGPT ทั้งเวอร์ชัน 3.5 และ 4.0 โดยในเวอร์ชัน 4.0 มีการออกแบบ Prompt ที่หลากหลาย เพื่อศึกษาผลลัพธ์ที่อาจแตกต่างกันตามบริบทการใช้งานจริง ในด้านกระบวนการวิจัย ผู้วิจัยได้ออกแบบ Rubric สำหรับตรวจจับคำสำคัญ (Keyword Detection) และประเมินความแม่นยำเชิงเนื้อหา (Semantic Accuracy) รวมถึงตรวจสอบโทนภาษา (Tone) และบริบททางวัฒนธรรม (Cultural Appropriateness) ในคำแปลแต่ละฉบับ ผลการวิจัยพบว่า การแปลโดย ChatGPT มีคุณภาพเทียบเคียงหรือดีกว่า การแปลของนักแปลมนุษย์ โดยเฉพาะในประเด็นที่เกี่ยวข้องกับการถ่ายทอดน้ำเสียง การใช้ถ้อยคำที่เหมาะสมกับบริบทวัฒนธรรม และความลื่นไหลตามธรรมชาติของภาษาเป้าหมาย แม้ว่า AI จะมีข้อได้เปรียบในด้านความเร็วและต้นทุนที่ต่ำกว่านักวิจัยเสนอว่าในอนาคต หากระบบ AI ได้รับการฝึกฝนด้วยชุดข้อมูลคุณภาพสูงและมีการออกแบบ Prompt ที่มีประสิทธิภาพมากขึ้น ก็อาจสามารถพัฒนาความสามารถในการแปลให้ทัดเทียมหรือเหนือกว่านักแปลมนุษย์ได้ โดยเฉพาะอย่างยิ่งเมื่อใช้ในรูปแบบการแปลร่วมกันระหว่างมนุษย์กับ AI (Human - AI Hybrid Translation) ซึ่งจะช่วยส่งเสริมประสิทธิภาพและลดต้นทุนของงานแปลในระยะยาว (Lau et al., 2024)

Herbig et al. (2019) ได้ศึกษาแนวทางการบูรณาการระหว่างระบบการแปลอัตโนมัติด้วยปัญญาประดิษฐ์ (Machine Translation: MT) และการแปลโดยมนุษย์ (Human Translation) โดยมีวัตถุประสงค์เพื่อเพิ่มประสิทธิภาพและคุณภาพของกระบวนการแปลภาษาในการใช้งานจริง ผู้วิจัยชี้ให้เห็นว่า แม้ระบบ AI จะสามารถให้ผลลัพธ์เบื้องต้นที่รวดเร็วและมีโครงสร้างที่สมบูรณ์ในระดับหนึ่ง แต่ยังมีข้อจำกัดด้านการเข้าใจบริบท วัฒนธรรม และโทนภาษาที่เหมาะสมกับกลุ่มเป้าหมาย ในทางกลับกัน นักแปลมนุษย์สามารถตีความข้อความที่คลุมเครือ ปรับระดับภาษา และเลือกใช้ถ้อยคำได้อย่างเหมาะสมกับบริบทของเนื้อหาและวัฒนธรรมของผู้รับสาร ซึ่งเป็นทักษะที่ AI ยังไม่สามารถทดแทนได้อย่างสมบูรณ์ นอกจากนี้ การแปลที่ได้จากระบบ AI มักยังขาดความแม่นยำทางด้านความหมาย ความสละสลวยของภาษา และความสมบูรณ์ในระดับ Pragmatic หรือสำนวนที่เป็นธรรมชาติ ด้วยเหตุนี้ Herbig et al. (2019) จึงเสนอว่าแนวทางการทำงานที่มีประสิทธิภาพที่สุดคือ การใช้ระบบ MT เพื่อสร้างร่างคำแปลเบื้องต้น และให้มนุษย์เป็นผู้ดำเนินการตรวจสอบและปรับแก้ในขั้นตอนที่เรียกว่า Post - Editing ซึ่งจะช่วยให้เพิ่มประสิทธิภาพในการแปลได้มากถึงร้อยละ 36 โดยยังคงรักษาคุณภาพของเนื้อหาอย่างเหมาะสม ทั้งนี้ ผู้วิจัยเน้นว่า การทำงานร่วมกันระหว่าง AI และมนุษย์ (AI – Human Collaboration) เป็นรูปแบบที่มีศักยภาพในการพัฒนาอุตสาหกรรมการแปลอย่างยั่งยืนในอนาคต

จากการศึกษางานวิจัยในประเทศไทยและต่างประเทศเกี่ยวกับคุณภาพการแปลระหว่างมนุษย์และเครื่องแปลภาษา พบว่าเครื่องแปลภาษา เช่น Google Translate และ AI ChatGPT มีความสามารถในการแปลที่ดีขึ้นอย่างต่อเนื่อง แต่ยังคงมีข้อจำกัดและข้อผิดพลาดที่ต้องการการแก้ไขจากมนุษย์ โดยเฉพาะอย่างยิ่งสำหรับเนื้อหาที่ซับซ้อนหรือต้องใช้ความเข้าใจทางวัฒนธรรม การแปลโดยมนุษย์ยังคงมีความแม่นยำและลึกซึ้งมากกว่า แต่เครื่องแปลภาษาสามารถช่วยเร่งความเร็วและลดต้นทุนในการแปลได้

เพื่อให้เห็นความเชื่อมโยงระหว่างแนวคิดและทฤษฎีที่ได้จากการทบทวนเอกสารในบทที่ 2 กับการออกแบบระเบียบวิธีวิจัยในบทที่ 3 ผู้วิจัยได้สังเคราะห์แนวคิดสำคัญและการนำไปใช้ในการวิจัยครั้งนี้ สรุปดังแสดงในตารางที่ 2.1

ตารางที่ 2.1 การสังเคราะห์แนวคิดและการนำไปใช้ในการวิจัย

แนวคิด/ทฤษฎีจากบทที่ 2	สาระสำคัญ	การนำไปใช้ในบทที่ 3
MQM (Accuracy, Fluency, Terminology Consistency)	เกณฑ์ประเมินคุณภาพการแปลหลายมิติ	ใช้เป็นกรอบสร้างเครื่องมือประเมินคุณภาพการแปล
การแปลโดยมนุษย์	การแปลที่อาศัยความรู้และบริบท	ใช้เป็นกลุ่มเปรียบเทียบ
การแปลโดย ChatGPT	การแปลโดย AI	ใช้เป็นกลุ่มทดลอง
ความคิดเห็นผู้เชี่ยวชาญ	การประเมินเชิงคุณภาพ	ใช้สนับสนุนผลเชิงปริมาณ

บทที่ 2 ได้นำเสนอแนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้องกับการประเมินคุณภาพการแปล บทความคัดย่อจากภาษาไทยเป็นภาษาอังกฤษ โดยครอบคลุมเกณฑ์การประเมินคุณภาพการแปล Multidimensional Quality Metrics (MQM) แนวคิดเกี่ยวกับการแปลโดยมนุษย์และการแปลโดยปัญญาประดิษฐ์ ChatGPT ตลอดจนเทคนิคการควบคุมคุณภาพการแปลและงานวิจัยที่เกี่ยวข้องทั้งในประเทศและต่างประเทศ การทบทวนเอกสารดังกล่าวช่วยให้เห็นภาพรวมของปัญหา แนวโน้ม และกรอบแนวคิดที่ใช้ในการประเมินคุณภาพการแปลอย่างเป็นระบบ และเป็นพื้นฐานสำคัญในการออกแบบระเบียบวิธีวิจัย เครื่องมือวิจัย และแนวทางการวิเคราะห์ข้อมูล ซึ่งจะนำเสนอในบทที่ 3 ต่อไป

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) ประเมินคุณภาพการแปลบทคัดย่อภาษาไทย เป็นภาษาอังกฤษ ตามเกณฑ์ในด้าน Accuracy (ความถูกต้องของความหมาย) Fluency (ความคล่องแคล่วทางภาษา) และ Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) ของเกณฑ์การประเมิน Multidimensional Quality Metrics (MQM) โดยพิจารณาความแตกต่างของคุณภาพระหว่างการแปลโดยมนุษย์ และการแปลโดย ChatGPT 2) เปรียบเทียบคุณภาพการแปล ระหว่างการแปลโดยมนุษย์และการแปลโดย ChatGPT จากคะแนนการประเมินของผู้เชี่ยวชาญ ด้านภาษาอังกฤษ และ 3) สังเคราะห์ความคิดเห็นของผู้เชี่ยวชาญเกี่ยวกับจุดเด่นและข้อจำกัดของการแปลทั้งสองรูปแบบ เพื่อตอบวัตถุประสงค์ดังกล่าว บทนี้นำเสนอระเบียบวิธีวิจัยในประเด็นต่อไปนี้

- 3.1 รูปแบบการวิจัย
- 3.2 แหล่งข้อมูลและผู้ให้ข้อมูล
- 3.3 ตัวแปรการวิจัย
- 3.4 เครื่องมือที่ใช้ในการวิจัย
- 3.5 การสร้างและตรวจสอบคุณภาพของเครื่องมือ
- 3.6 การเก็บรวบรวมข้อมูล
- 3.7 การวิเคราะห์ข้อมูล

3.1 รูปแบบการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเชิงปริมาณที่ผสมการวิเคราะห์เชิงคุณภาพ (Quantitative Research with Qualitative Support) โดยมีลักษณะสำคัญดังนี้

3.1.1 ส่วนเชิงปริมาณ

- 1) ใช้คะแนนการประเมินคุณภาพการแปลจากอาจารย์ผู้เชี่ยวชาญด้านภาษาอังกฤษ จำนวน 18 คน
- 2) เปรียบเทียบคะแนนคุณภาพการแปลระหว่างการแปลโดยมนุษย์ และการแปลโดย ChatGPT
- 3) ตามเกณฑ์ MQM 3 ด้าน คือ Accuracy (ความถูกต้องของความหมาย) Fluency (ความคล่องแคล่วทางภาษา) Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ)

3.1.2 ส่วนเชิงคุณภาพ

- 1) ใช้คำตอบจากคำถามปลายเปิดในแบบสอบถาม
 - 2) วิเคราะห์เนื้อหา (Content Analysis) เพื่อสังเคราะห์รูปแบบข้อผิดพลาด จุดเด่น ข้อจำกัด และมุมมองของผู้เชี่ยวชาญต่อการใช้ ChatGPT เปรียบเทียบกับมนุษย์
- ลักษณะดังกล่าวทำให้งานวิจัยนี้สามารถอธิบายคุณภาพการแปลได้ทั้งในเชิงตัวเลข (คะแนน) และเชิงอธิบาย (ความคิดเห็นและตัวอย่างข้อผิดพลาด)

3.2 แหล่งข้อมูลและผู้ให้ข้อมูล

งานวิจัยนี้กำหนดกรอบการเก็บข้อมูลโดยแยก “สิ่งที่ถูกประเมิน” ออกจาก “ผู้ให้การประเมิน” อย่างชัดเจน เพื่อความถูกต้องตามหลักระเบียบวิธีวิจัย ดังนี้

3.2.1 แหล่งข้อมูล

แหล่งข้อมูล ในการวิจัยครั้งนี้คือ บทความย่อวิทยานิพนธ์ระดับปริญญาโท ของมหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี ปีการศึกษา 2567 (มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี สำนักบัณฑิตศึกษา, 2567) ซึ่งเขียนเป็นภาษาไทย และมีฉบับแปลเป็นภาษาอังกฤษ

จากบทความย่อทั้งหมด ผู้วิจัยคัดเลือกแบบเจาะจง (Purposive Selection) จำนวน 10 เรื่อง โดยคำนึงถึง

- 1) การกระจายตัวตามกลุ่มสาขาวิชา
 - (1) กลุ่มวิทยาศาสตร์และเทคโนโลยี 5 เรื่อง
 - (2) กลุ่มมนุษยศาสตร์และสังคมศาสตร์ 5 เรื่อง
- 2) ความครบถ้วนขององค์ประกอบบทความย่อ (บทนำ วัตถุประสงค์ วิธีวิจัย ผล และข้อสรุป)
- 3) การใช้ศัพท์เฉพาะในสาขาวิชาอย่างชัดเจน ซึ่งเหมาะสมกับการประเมินด้าน Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ)
- 4) ความหลากหลายของเนื้อหาและโครงสร้างประโยค เพื่อให้การประเมินคุณภาพการแปลมีความท้าทายและสะท้อนสภาพการใช้ภาษาในงานจริง

ในการวิจัยครั้งนี้ บทความย่อ 10 เรื่องไม่ใช่ “ประชากร” หรือ “กลุ่มตัวอย่าง” ในความหมายของ “คน” แต่ทำหน้าที่เป็น “แหล่งข้อมูล (Data Source)” และ “หน่วยการวิเคราะห์ (Units of Analysis)” ซึ่งถูกจัดรูปแบบให้เป็นชุดตัวอย่างข้อความ (Thai – English) ที่แนบเป็นส่วนหนึ่งของเครื่องมือวิจัย เพื่อให้ผู้เชี่ยวชาญใช้ในการประเมินคุณภาพการแปล

3.2.2 ผู้ให้ข้อมูลสำคัญ (Key Informants) ในการวิจัยนี้ คือ อาจารย์ผู้เชี่ยวชาญด้านภาษาอังกฤษ จำนวน 18 คน ที่ปฏิบัติหน้าที่แปลและตรวจภาษาอังกฤษให้กับนักศึกษาระดับบัณฑิตศึกษาในมหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

คุณสมบัติที่ใช้ในการคัดเลือก ได้แก่

- 1) สำเร็จการศึกษาด้านภาษาอังกฤษ หรือสาขาที่เกี่ยวข้องในระดับปริญญาโทขึ้นไป
 - 2) มีประสบการณ์สอนภาษาอังกฤษ การเขียนเชิงวิชาการ หรือการแปลอย่างน้อย 5 ปี
 - 3) มีประสบการณ์ตรงในการตรวจหรือกลั่นกรองบทความภาษาอังกฤษของนักศึกษาในระดับบัณฑิตศึกษา
 - 4) ยินยอมเข้าร่วมเป็นผู้ประเมินในงานวิจัยอย่างสมัครใจ
- ทั้งนี้ ผู้วิจัยใช้รหัสแทนตัวผู้ให้ข้อมูลในการนำเสนอผลการวิจัย เพื่อรักษาความลับของผู้ให้ข้อมูลและลดอคติในการตีความผลการประเมิน

3.3 ตัวแปรการวิจัย

จากกรอบแนวคิดการวิจัย ข้อ 1.5 ในบทที่ 1 งานวิจัยนี้กำหนดตัวแปรสำหรับส่วนเชิงปริมาณ ดังนี้

3.3.1 ตัวแปรต้น (Independent Variable) ได้แก่ รูปแบบการแปล แบ่งออกเป็น การแปลโดยมนุษย์ และการแปลโดย ChatGPT

3.3.2 ตัวแปรตาม (Dependent Variables) ได้แก่ คะแนนคุณภาพการแปลในแต่ละด้านของ MQM ตามที่ผู้เชี่ยวชาญประเมิน ได้แก่ Accuracy (ความถูกต้องของความหมาย) Fluency (ความคล่องแคล่วทางภาษา) และ Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ)

3.3.3 ตัวแปรเชิงคุณภาพสนับสนุน ได้แก่ ข้อคิดเห็นและข้อเสนอแนะจากผู้เชี่ยวชาญ ซึ่งได้จากคำถามปลายเปิด นำไปใช้ในการสังเคราะห์เป็นประเด็นเกี่ยวกับ จุดเด่นของการแปลแต่ละรูปแบบ ข้อจำกัดและรูปแบบข้อผิดพลาดที่พบบ่อย และมุมมองต่อการใช้ ChatGPT ในการแปลบทความ

3.4 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือในการวิจัยประกอบด้วย 2 ส่วนที่ทำงานร่วมกัน คือ ชุดตัวอย่างบทความ (Thai – English) ที่ใช้เป็นเนื้อหาสำหรับประเมิน และแบบสอบถามประเมินคุณภาพการแปลตามเกณฑ์ MQM

3.4.1 ชุดบทความตัวอย่างสำหรับการประเมินคุณภาพการแปล ประกอบด้วย

- 1) บทความภาษาไทย 10 เรื่อง (ต้นฉบับ)
- 2) บทความภาษาอังกฤษ 20 ฉบับ ซึ่งได้จาก (1) การแปลโดยมนุษย์ 10 ฉบับ และ (2) การแปลโดย ChatGPT 10 ฉบับ โดยใช้ต้นฉบับภาษาไทยฉบับเดียวกันจับคู่ระหว่างการแปลสองรูปแบบ โดยชุดบทความเหล่านี้แนบรวมไปกับแบบสอบถาม เพื่อให้ผู้เชี่ยวชาญใช้เป็นฐานในการพิจารณาให้คะแนน ซึ่งกล่าวได้ว่า ชุดบทความทำหน้าที่เป็น “วัสดุการประเมิน (Stimulus Materials)” และเป็น “ส่วนประกอบของเครื่องมือวิจัย” มากกว่าการเป็นประชากรหรือกลุ่มตัวอย่างในความหมายแบบเดิมของการวิจัยเชิงสำรวจ

3.4.2 แบบสอบถามประเมินคุณภาพการแปลตามเกณฑ์ MQM แบ่งออกเป็น 2 ส่วนหลัก ดังนี้

1) ส่วนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม ใช้สอบถามข้อมูลพื้นฐาน เช่น วุฒิการศึกษา สาขาวิชา ประสบการณ์สอนภาษาอังกฤษ / การแปลภาษาอังกฤษ ประสบการณ์ในการตรวจบทคัดย่อ และประสบการณ์ในการใช้ ChatGPT หรือเครื่องมือ AI ด้านภาษา

2) ส่วนที่ 2 แบบประเมินคุณภาพการแปลตามเกณฑ์ Multidimensional Quality Metrics (MQM) โดยกำหนด 3 ด้านหลัก ดังนี้

(1) Accuracy (ความถูกต้องของความหมาย) องค์กรประกอบย่อยในการตรวจ ได้แก่

ก. Mistranslation (การแปลผิด)

ข. Omission (การละเลยเนื้อหา)

ค. Addition (การเพิ่มเนื้อหา)

ง. Untranslated Text (ข้อความที่ไม่ได้แปล)

ตรวจว่าข้อความภาษาอังกฤษถ่ายทอดความหมายจากต้นฉบับภาษาไทย ได้ครบถ้วน ไม่แปลผิด ไม่เพิ่ม – ลดเนื้อหาอย่างผิดบริบท

(2) Fluency (ความคล่องแคล่วทางภาษา) องค์กรประกอบย่อยในการตรวจ ได้แก่

ก. Punctuation (การใช้เครื่องหมายวรรคตอน)

ข. Spelling (การสะกดคำ)

ค. Grammar (ไวยากรณ์)

ง. Register (ระดับภาษา)

ตรวจความถูกต้องทางไวยากรณ์ การสะกดคำ การใช้เครื่องหมายวรรคตอน และความเป็นธรรมชาติของภาษาอังกฤษที่ใช้

(3) Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) องค์กรประกอบย่อยในการตรวจ ได้แก่

ก. Incorrect Term (การใช้ศัพท์ที่ไม่ตรง)

ข. Inconsistent Term (การใช้ศัพท์ไม่สอดคล้องกัน)

ตรวจว่าการใช้ศัพท์เทคนิคและคำศัพท์เชิงวิชาการถูกต้องและใช้สอดคล้องกันตลอดทั้งบทคัดย่อโดยผู้เชี่ยวชาญให้คะแนนแต่ละด้านของแต่ละฉบับแปลด้วยมาตราส่วนประมาณค่า 5 ระดับ (1 – 5) และมีช่องสำหรับ ระบุ “ลักษณะข้อผิดพลาดสำคัญ” ที่พบ โดยเขียนข้อเสนอแนะเพิ่มเติมในรูปแบบคำถามปลายเปิด

3.5 การสร้างและตรวจสอบคุณภาพของเครื่องมือวิจัย

การวิจัยครั้งนี้ใช้เครื่องมือหลัก ได้แก่ (1) ชุดแบบฟอร์มประเมินคุณภาพการแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษ ตามเกณฑ์ Multidimensional Quality Metrics (MQM) ซึ่งประกอบด้วย ด้าน Accuracy (ความถูกต้องของความหมาย) Fluency (ความคล่องแคล่วทางภาษา) และ Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) และ (2) คำถามปลายเปิดสำหรับผู้เชี่ยวชาญสะท้อนความคิดเห็นเชิงลึกเกี่ยวกับจุดเด่น ข้อจำกัด และลักษณะข้อผิดพลาดของการแปลทั้งสองรูปแบบ

การสร้างและตรวจสอบคุณภาพของเครื่องมือวิจัยในครั้งนี้จึงเน้นสองมิติสำคัญ คือ (1) ความชัดเจนและความครอบคลุมของเกณฑ์การประเมิน MQM ในบริบทของบทคัดย่อวิทยานิพนธ์ และ (2) ความตรงเชิงเนื้อหาของข้อคำถามและเกณฑ์ที่ใช้วัด โดยเฉพาะอย่างยิ่ง เนื่องจากแบบฟอร์มทำหน้าที่เป็น “แบบประเมินโดยผู้เชี่ยวชาญ (Expert Rating Form)” ไม่ใช่แบบสอบถามวัดเจตคติของประชากรจำนวนมาก การตรวจสอบคุณภาพเครื่องมือจึงมุ่งเน้นที่ความตรงเชิงเนื้อหา (Content Validity) เป็นหลัก มากกว่าการหาค่าความเชื่อมั่นภายใน (เช่น Cronbach’s Alpha)

เพื่อตอบโจทย์ดังกล่าว กระบวนการสร้างและตรวจสอบคุณภาพของเครื่องมือประกอบด้วย ขั้นตอนสำคัญดังต่อไปนี้

3.5.1 การสร้างแบบฟอร์มประเมินคุณภาพการแปลและคำถามปลายเปิด โดยเริ่มจากการกำหนด “สิ่งที่ต้องการวัด” ให้ชัดเจน โดยผู้วิจัยกำหนดว่า คุณภาพการแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษในครั้งนี้ต้องการประเมินใน 3 ด้านหลัก ตามเกณฑ์ Multidimensional Quality Metrics (MQM) ได้แก่

1) Accuracy (ความถูกต้องของความหมาย) หมายถึง ระดับความถูกต้องแม่นยำของการแปล โดยพิจารณาว่าการแปลสามารถสื่อความหมายของต้นฉบับได้อย่างครบถ้วนและถูกต้องหรือไม่

2) Fluency (ความคล่องแคล่วทางภาษา) หมายถึง ระดับความเป็นธรรมชาติและความสะดวกของการแปล โดยพิจารณาว่าการแปลมีความสละสลวยและเป็นไปตามหลักไวยากรณ์ของภาษาเป้าหมายหรือไม่

3) Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) หมายถึง ระดับความถูกต้องและสอดคล้องในการใช้ศัพท์เฉพาะทางเทคนิค โดยพิจารณาว่ามีการใช้ศัพท์เฉพาะอย่างถูกต้องและสอดคล้องกันตลอดทั้งข้อความหรือไม่

จากนิยามทั้งสามด้าน ผู้วิจัยดำเนินการสร้างเครื่องมือเป็นลำดับดังนี้

1) ออกแบบแบบฟอร์มประเมิน (Rating Form) โดย

(1) แยกส่วนประเมินเป็น 3 ด้าน ตาม MQM คือ Accuracy (ความถูกต้องของความหมาย) Fluency (ความคล่องแคล่วทางภาษา) และ Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ)

(2) กำหนดเกณฑ์การให้คะแนนในแต่ละด้านเป็นมาตราส่วนประมาณค่า 5 ระดับ (5-point rating scale) ของ Tongpoon - Patanasorn และ Griffith (2020) ดังนี้

5 คะแนน หมายถึง คุณภาพการแปลดีเยี่ยม

(ไม่มีข้อผิดพลาด)

4 คะแนน หมายถึง คุณภาพการแปลดี

(มีข้อผิดพลาดเล็กน้อย ไม่กระทบต่อความเข้าใจ)

3 คะแนน หมายถึง คุณภาพการแปลปานกลาง

(มีข้อผิดพลาดบางส่วนที่กระทบความเข้าใจเล็กน้อย)

2 คะแนน หมายถึง คุณภาพการแปลต่ำ

(ข้อผิดพลาดหลายแห่ง เข้าใจได้ยาก)

1 คะแนน หมายถึง คุณภาพการแปลต่ำมาก

(ข้อผิดพลาดร้ายแรง ทำให้เข้าใจผิดหรือไม่เข้าใจ)

(3) กำหนดคำอธิบายประกอบแต่ละระดับคะแนนอย่างย่อ เพื่อให้ผู้เชี่ยวชาญเข้าใจมาตรฐานการให้คะแนนตรงกัน (เช่น “มีข้อผิดพลาดเล็กน้อย ไม่กระทบต่อความเข้าใจ” “มีข้อผิดพลาดหลายแห่ง เข้าใจได้ยาก” เป็นต้น)

2) เชื่อมโยงกับชุดบทคัดย่อที่ใช้เป็นวัสดุการประเมิน

(1) นำบทคัดย่อภาษาไทยจำนวน 10 เรื่อง และฉบับแปลภาษาอังกฤษจำนวน 20 ฉบับ (แปลโดยมนุษย์ 10 ฉบับ และแปลโดย ChatGPT 10 ฉบับ) มาจัดเรียงในชุดเอกสารสำหรับผู้เชี่ยวชาญ

(2) สำหรับบทคัดย่อแต่ละเรื่อง ผู้วิจัยจัดแบบฟอร์มให้คะแนน 2 ชุดคู่กัน (แต่ปิดบังไม่บอกว่าเป็นฉบับที่แปลโดยมนุษย์หรือโดย ChatGPT) เพื่อให้ผู้เชี่ยวชาญประเมินคุณภาพการแปลโดยไม่เกิดอคติจากแหล่งที่มา (Blind Evaluation)

3) สร้างคำถามปลายเปิดเพื่อเก็บข้อมูลเชิงคุณภาพ ซึ่งนอกจากคะแนนเชิงปริมาณ ผู้วิจัยได้เพิ่มคำถามปลายเปิดในตอนท้ายของแบบฟอร์มแต่ละชุดและส่วนรวมของแบบสอบถาม เพื่อให้ผู้เชี่ยวชาญสะท้อนความคิดเห็นเชิงลึก เช่น

- (1) ข้อดีที่เห็นได้ชัดของการแปลโดยมนุษย์และการแปลโดย ChatGPT
- (2) ลักษณะข้อผิดพลาดที่พบบ่อย ๆ ในแต่ละด้านของ MQM
- (3) ตัวอย่างประโยคที่ สะท้อนข้อบกพร่องด้าน Accuracy, Fluency หรือ Terminology Consistency
- (4) มุมมองต่อความเหมาะสมในการใช้ ChatGPT เป็นเครื่องมือช่วยแปล บทความอภินิพนธ์ในบริบทระดับบัณฑิตศึกษา

การออกแบบคำถามปลายเปิดดังกล่าวช่วยให้ข้อมูลเชิงคุณภาพที่ได้สามารถนำไปวิเคราะห์เนื้อหา (Content Analysis) เพื่อสังเคราะห์ภาพรวม จุดเด่น ข้อจำกัด และข้อเสนอแนะเชิงลึกเกี่ยวกับการแปลทั้งสองรูปแบบได้อย่างรอบด้าน สนับสนุนการตีความผลเชิงปริมาณอย่างมีนัยสำคัญ

3.5.2 การตรวจสอบความตรงเชิงเนื้อหาด้วยดัชนี IOC ซึ่งเมื่อได้ร่างแบบฟอร์มประเมิน และคำถามปลายเปิด เบื้องต้นผู้วิจัยได้นำเครื่องมือดังกล่าวเสนอให้ ผู้ทรงคุณวุฒิ 3 ท่าน ซึ่งมีความเชี่ยวชาญด้านภาษาอังกฤษ การแปล และการเขียนเชิงวิชาการระดับบัณฑิตศึกษา ตรวจสอบความตรงเชิงเนื้อหา (Content Validity) ของเครื่องมือ โดยใช้ ดัชนีความสอดคล้องระหว่างข้อคำถามกับวัตถุประสงค์ (Index of Item – Objective Congruence: IOC) โดยมีขั้นตอนการตรวจสอบมีดังนี้

1) ผู้ทรงคุณวุฒิแต่ละท่านประเมินว่า ข้อคำถาม/เกณฑ์ย่อยในแบบฟอร์มประเมิน MQM และคำถามปลายเปิดแต่ละข้อ “สอดคล้องกับวัตถุประสงค์ของการวิจัย” มากน้อยเพียงใด โดยให้คะแนน 3 ระดับ คือ +1 = สอดคล้อง 0 = ไม่แน่ใจ และ -1 = ไม่สอดคล้อง

2) ผู้วิจัยนำคะแนนจากผู้ทรงคุณวุฒิมาคำนวณเป็นค่า IOC ของแต่ละข้อ โดยใช้สูตรมาตรฐาน

$$IOC = \frac{\sum \square \square}{\square}$$

เมื่อ

$\square \square$ คือคะแนนที่ผู้ทรงคุณวุฒิแต่ละท่านให้ (+1, 0 หรือ -1)

\square คือจำนวนผู้ทรงคุณวุฒิ (ในที่นี้เท่ากับ 3 ท่าน)

3) พิจารณาค่าที่ได้ตามเกณฑ์ว่า ค่า IOC ตั้งแต่ 0.50 ขึ้นไป ถือว่าข้อคำถาม/เกณฑ์ย่อยนั้นมีความสอดคล้องกับวัตถุประสงค์ของการวิจัย สามารถนำไปใช้ได้ สำหรับค่า IOC ต่ำกว่า 0.50 ควรได้รับการพิจารณาปรับปรุงหรือตัดออก ซึ่งจากการตรวจสอบในครั้งนี้ พบว่า ข้อคำถามทั้งหมดมีค่า IOC = 1.00 ซึ่งเป็นค่าที่สูงที่สุด แสดงว่าผู้ทรงคุณวุฒิทั้งสามท่านเห็นว่าเนื้อหาของแบบฟอร์มมีความเหมาะสมและสอดคล้องกับวัตถุประสงค์ของการวิจัยอย่างสมบูรณ์ ผู้ทรงคุณวุฒิไม่ได้เสนอให้ปรับปรุงแก้ไขแบบฟอร์มใด ๆ เนื่องจากเกณฑ์ MQM และคำอธิบายประกอบมีความครอบคลุม ชัดเจน และสามารถใช้เป็นเครื่องมือประเมินคุณภาพการแปลได้ตรงตามมาตรฐานของสาขาภาษาศาสตร์และเทคโนโลยีการแปล

แนวทางการตรวจสอบความตรงเชิงเนื้อหาโดยใช้ IOC ในลักษณะนี้สอดคล้องกับแนวคิดของ พิสิษฐ ตัณฑวณิช และพนา จินดาศรี (2561) ซึ่งเสนอว่า IOC เป็นเครื่องมือสำคัญที่ใช้ยืนยันความสอดคล้องของข้อคำถามกับตัวแปรที่ต้องการวัด และเหมาะสมสำหรับแบบสอบถามหรือแบบประเมินที่อาศัยความคิดเห็นของผู้เชี่ยวชาญโดยเฉพาะ

3.5.3 เหตุผลเชิงระเบียบวิธีที่ไม่ตรวจสอบค่าความเชื่อมั่นภายในของเครื่องมือซึ่งแม้งานวิจัยเชิงสำรวจจำนวนมากมักตรวจสอบ “ความเชื่อมั่นภายใน (Internal Consistency Reliability)” ของแบบสอบถาม เช่น การคำนวณค่าสัมประสิทธิ์ Cronbach’s Alpha แต่ในงานวิจัยครั้งนี้ ผู้วิจัยไม่ได้คำนวณค่า Cronbach’s Alpha สำหรับแบบฟอร์มประเมิน MQM โดยมีเหตุผลเชิงระเบียบวิธีดังนี้

1) ลักษณะเครื่องมือเป็นแบบประเมินโดยผู้เชี่ยวชาญ (Expert Rating Form) ไม่ได้เป็น “สเกลวัดเจตคติ” หรือ “แบบวัดคุณลักษณะเชิงจิตพิสัย” ของผู้ตอบ แต่เป็นแบบฟอร์มที่ใช้ให้ผู้เชี่ยวชาญ “ใช้เกณฑ์ MQM ไปประเมินคุณภาพของข้อความแปล” การใช้ IOC เพื่อตรวจสอบว่าข้อคำถามและเกณฑ์ที่กำหนด “ตรงกับวัตถุประสงค์การวัด” และครอบคลุมด้านของคุณภาพการแปลที่ต้องการ ถือว่าเป็นหัวใจสำคัญมากกว่าการตรวจสอบความสอดคล้องภายในของข้อคำถามในลักษณะเดียวกับสเกลทางจิตวิทยาหรือสังคมศาสตร์ทั่วไป

2) ขนาดและลักษณะของกลุ่มผู้ประเมิน ซึ่งเป็นกลุ่มผู้ตอบแบบฟอร์มในการวิจัยครั้งนี้คืออาจารย์ผู้เชี่ยวชาญด้านภาษาอังกฤษจำนวน 18 คน ซึ่งนับเป็น “กลุ่มผู้เชี่ยวชาญเฉพาะทาง” มากกว่าจะเป็น “กลุ่มตัวอย่างขนาดใหญ่” ที่สุ่มมาจากประชากร โดยผู้วิจัยใช้ผู้เชี่ยวชาญทั้ง 18 คนทั้งหมดในฐานะผู้ให้ข้อมูลสำคัญ มิได้มีเจตนาจะอนุมานคุณลักษณะของประชากรในวงกว้างจึงไม่จำเป็นต้องใช้สถิติความเชื่อมั่นภายในเพื่อรองรับการใช้เครื่องมือในกลุ่มตัวอย่างจำนวนมาก

3) ธรรมชาติของข้อมูลที่ต้องการวิเคราะห์ที่ได้จากแบบฟอร์มคือ “คะแนนคุณภาพการแปลของข้อความแต่ละฉบับ” ภายใต้เกณฑ์ MQM ต่าง ๆ มิใช่ “คะแนนคุณลักษณะของผู้ตอบแบบสอบถาม” ดังนั้น ความสำคัญจึงอยู่ที่การกำหนดเกณฑ์ให้ชัดเจนและเหมาะสม และการทำให้ผู้เชี่ยวชาญเข้าใจเกณฑ์ร่วมกัน มากกว่าการตรวจสอบว่าข้อคำถามแต่ละข้อในสเกลทดสอบสอดคล้องกันเพียงใดในเชิงสถิติ

4) บทบาทหลักของเครื่องมือนี้ คือ “มาตรฐานการประเมิน” ไม่ใช่การ “วัดตัวบุคคล” โดยจุดมุ่งหมายของการสร้างเครื่องมือ คือ การกำหนดมาตรฐานในการประเมินคุณภาพการแปลบทความให้มีเกณฑ์เดียวกันสำหรับผู้เชี่ยวชาญทุกคน การตรวจสอบความตรงเชิงเนื้อหาและการปรับเกณฑ์ให้ชัดเจนจึงมีความสำคัญมากที่สุด เมื่อเครื่องมือมีความตรงเชิงเนื้อหาที่ดีและผู้ประเมินมีความเชี่ยวชาญในสาขาเดียวกัน การใช้เครื่องมือจะให้ข้อมูลที่เพียงพอสำหรับการเปรียบเทียบคุณภาพการแปลของมนุษย์และ ChatGPT ตามวัตถุประสงค์ของงานวิจัย

กล่าวโดยสรุป ในบริบทของงานวิจัยที่ใช้ผู้เชี่ยวชาญ 18 คนเป็นผู้ประเมินคุณภาพข้อความแปลตามเกณฑ์ MQM การตรวจสอบ ความตรงเชิงเนื้อหาด้วย IOC ถือว่าเป็นวิธีที่เหมาะสมและเพียงพอสำหรับการยืนยันคุณภาพของเครื่องมือวิจัย โดยไม่จำเป็นต้องตรวจสอบค่าความเชื่อมั่นภายในเช่น Cronbach's Alpha ซึ่งถูกออกแบบมาเพื่อใช้กับสเกลวัดเจตคติหรือคุณลักษณะทางจิตพิสัยในกลุ่มตัวอย่างจำนวนมากเป็นหลัก

3.6 การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลดำเนินการตามขั้นตอนต่อไปนี้

3.6.1 จัดเตรียมชุดบทคัดย่อและแบบสอบถาม โดย

1) รวบรวมบทคัดย่อภาษาไทย 10 เรื่อง และจัดเตรียมฉบับแปลภาษาอังกฤษ 20 ฉบับ (การแปลโดยมนุษย์ และการแปลโดย ChatGPT)

2) แนวนฉบับแปลเหล่านี้เข้าในชุดแบบสอบถาม โดยจัดรูปแบบให้ผู้เชี่ยวชาญสามารถอ่านและประเมินได้สะดวก

3.6.2 จัดเรียงลำดับบทคัดย่อแบบปกปิดแหล่งที่มา โดย

1) สลับลำดับและไม่เปิดเผยว่าแต่ละฉบับแปลโดยใคร เพื่อป้องกันอคติล่วงหน้า
2) ใช้รหัสแทน เช่น H1 – H10 สำหรับฉบับแปลโดยมนุษย์ และ C1 – C10 สำหรับฉบับแปลโดย ChatGPT

3.6.3 แจกแบบสอบถามให้ผู้เชี่ยวชาญ 18 คน พร้อมคำชี้แจงวิธีการให้คะแนนใน 3 ด้านของ MQM คือ Accuracy (ความถูกต้องของความหมาย) Fluency (ความคล่องแคล่วทางภาษา) Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) และเน้นว่าการประเมินต้องอิงจากข้อความแปลเท่านั้น ไม่คำนึงว่าเป็นผลงานของมนุษย์หรือ AI

3.6.4 การตอบแบบสอบถาม โดย

1) ผู้เชี่ยวชาญอ่านต้นฉบับภาษาไทยและฉบับแปลภาษาอังกฤษแต่ละคู่
2) ให้คะแนน MQM ทั้ง 3 ด้านในแบบฟอร์ม
3) ระบุลักษณะข้อผิดพลาดหลัก (ถ้าพบ) และตอบคำถามปลายเปิดตามประสบการณ์และมุมมองของตน

3.6.5 ตรวจสอบและจัดเก็บข้อมูล

1) วิจัยตรวจสอบแบบสอบถามที่ได้รับคืนว่ากรอกข้อมูลครบถ้วน
2) เข้ารหัสข้อมูลและจัดเตรียมสำหรับการวิเคราะห์

3.7 การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลแบ่งเป็น 2 ส่วน คือ ข้อมูลเชิงปริมาณ และข้อมูลเชิงคุณภาพ

3.7.1 การวิเคราะห์ข้อมูลเชิงปริมาณ เพื่อตอบวัตถุประสงค์ข้อที่ 1 และ 2

1) สถิติเชิงพรรณนา (Descriptive Statistics) หาค่าเฉลี่ย (Mean) และส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) ของคะแนน MQM ในแต่ละด้าน แยกตามรูปแบบการแปล (มนุษย์ / ChatGPT) โดยใช้มาตราส่วนประมาณค่า 5 ระดับ มีเกณฑ์การแปลความหมายค่าคะแนน (Tongpoon - Patanasorn และ Griffith, 2020) ดังนี้

4.51-5.00	หมายถึง คุณภาพการแปลอยู่ในระดับดีเยี่ยม
3.51-4.50	หมายถึง คุณภาพการแปลอยู่ในระดับดี
2.51-3.50	หมายถึง คุณภาพการแปลอยู่ในระดับปานกลาง
1.51-2.50	หมายถึง คุณภาพการแปลอยู่ในระดับต่ำ
1.00-1.50	หมายถึง คุณภาพการแปลอยู่ในระดับต่ำมาก

2) สถิติเชิงอนุมาน (Inferential Statistics) การวิจัยครั้งนี้ทดสอบสมมติฐานเพื่อเปรียบเทียบคุณภาพการแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษระหว่างการแปลโดยมนุษย์และการแปลโดย ChatGPT โดยใช้คะแนนการประเมินตามเกณฑ์ MQM ซึ่งเป็นข้อมูลจากสเกลประมาณค่าและเป็นการเปรียบเทียบคะแนนจากต้นฉบับเดียวกันที่ถูกแปลด้วยสองวิธี จึงใช้สถิติไม่เป็นพารามิเตอร์ Wilcoxon Signed - Rank Test ในการทดสอบความแตกต่างของคะแนนแบบจับคู่ (Paired Samples) ที่ระดับนัยสำคัญทางสถิติ .05 การทดสอบสมมติฐานดำเนินการแยกตามแต่ละด้านของ MQM ได้แก่ Accuracy (ความถูกต้องของความหมาย) Fluency (ความคล่องแคล่วทางภาษา) และ Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) โดยกำหนดสมมติฐานทางสถิติ ดังนี้

สมมติฐานศูนย์ (Null Hypothesis)

H_0 : คะแนนการแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษโดยมนุษย์ และโดย ChatGPT ไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ในด้าน Accuracy/ Fluency/ Terminology Consistency

สมมติฐานทางเลือก (Alternative Hypothesis)

H_1 : คะแนนการแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษโดยมนุษย์ และโดย ChatGPT มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ในด้าน Accuracy/ Fluency/ Terminology Consistency

ผลการทดสอบสมมติฐานจะนำเสนอในรูปค่ามัธยฐาน (Median) ค่าอันดับเฉลี่ย (Mean Rank) ค่าสถิติ Z และค่า Sig. (p-value) โดยใช้เกณฑ์การตัดสินใจที่ระดับนัยสำคัญทางสถิติ .05 หากค่า p-value น้อยกว่า .05 จะปฏิเสธสมมติฐานศูนย์

3.7.2 การวิเคราะห์ข้อมูลเชิงคุณภาพ จากข้อมูลจากคำถามปลายเปิดและช่องความคิดเห็นที่ผู้เชี่ยวชาญกรอก จะนำมาวิเคราะห์ด้วยวิธี การวิเคราะห์เนื้อหา (Content Analysis) โดยมีขั้นตอนดังนี้

- 1) อ่านคำตอบทั้งหมดซ้ำหลายรอบเพื่อทำความเข้าใจภาพรวม
- 2) กำหนดหน่วยวิเคราะห์ (เช่น ประโยค/ย่อหน้าที่มีความหมายสมบูรณ์)
- 3) เข้ารหัส (Coding) แยกตามประเด็น เช่น ข้อดีของการแปลโดยมนุษย์ การแปลโดย ChatGPT รูปแบบข้อผิดพลาดด้าน Accuracy, Fluency, Terminology Consistency และข้อเสนอแนะต่อการใช้ ChatGPT ในบริบทบัณฑิตศึกษา

- 4) จัดกลุ่มรหัสเป็นหมวดหมู่ (Categories) และธีมหลัก (Themes)
- 5) สังเคราะห์และสรุปเป็นข้อความบรรยาย พร้อมยกตัวอย่างคำพูดของผู้เชี่ยวชาญประกอบ

ผลการวิเคราะห์เชิงคุณภาพจะใช้ตอบวัตถุประสงค์ข้อที่ 3 และช่วยอธิบายเชิงลึกประกอบผลการวิเคราะห์เชิงปริมาณในบทที่ 4 และบทที่ 5

บทที่ 3 ได้นำเสนอระเบียบวิธีวิจัยที่ใช้ในการศึกษาคั้งนี้อย่างเป็นระบบ โดยครอบคลุมรูปแบบการวิจัย แหล่งข้อมูลและผู้ให้ข้อมูล เครื่องมือที่ใช้ในการวิจัย รวมถึงกระบวนการสร้างและตรวจสอบคุณภาพของเครื่องมือ ตลอดจนขั้นตอนการเก็บรวบรวมและการวิเคราะห์ข้อมูล ทั้งนี้ เพื่อให้การดำเนินการวิจัยเป็นไปอย่างมีความเหมาะสม สอดคล้องกับวัตถุประสงค์และสมมติฐานการวิจัยที่กำหนดไว้ ในบทถัดไปจะนำเสนอผลการวิเคราะห์ข้อมูลตามแนวทางและวิธีการที่ได้กำหนดไว้ในบทนี้

บทที่ 4

ผลการวิจัย

บทนี้นำเสนอผลการวิจัยตามวัตถุประสงค์ของการวิจัยทั้งสามประการ โดยใช้ข้อมูลจากบทคัดย่อวิทยานิพนธ์ภาษาไทยจำนวน 10 เรื่อง และฉบับแปลภาษาอังกฤษจำนวน 20 ฉบับ ซึ่งประกอบด้วย การแปลโดยมนุษย์และการแปลโดย ChatGPT จากต้นฉบับภาษาไทยเดียวกัน ทั้งนี้ บทคัดย่อดังกล่าวใช้เป็นหน่วยการวิเคราะห์ในการประเมินคุณภาพการแปลตามเกณฑ์ Multidimensional Quality Metrics (MQM) โดยการนำเสนอผลการวิจัยในบทนี้แบ่งออกเป็นสามส่วนหลัก ตามลำดับของวัตถุประสงค์การวิจัย ได้แก่ 1) ประเมินคุณภาพการแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษ ตามเกณฑ์ในด้าน Accuracy (ความถูกต้องของความหมาย) Fluency (ความคล่องแคล่วทางภาษา) และ Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) ของเกณฑ์การประเมิน Multidimensional Quality Metrics (MQM) โดยพิจารณาความแตกต่างของคุณภาพ ระหว่างการแปลโดยมนุษย์ และการแปลโดย ChatGPT 2) เปรียบเทียบคุณภาพการแปล ระหว่างการแปลโดยมนุษย์และการแปลโดย ChatGPT จากคะแนนการประเมินของผู้เชี่ยวชาญด้านภาษาอังกฤษ และ 3) สังเคราะห์ความคิดเห็นของผู้เชี่ยวชาญเกี่ยวกับจุดเด่นและข้อจำกัดของการแปล ทั้งสองรูปแบบ ผลการวิจัยเชิงปริมาณในส่วนแรกนำเสนอในลักษณะของสถิติเชิงพรรณนาและสถิติเชิงอนุมาน โดยรายงานค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และผลการทดสอบความแตกต่างทางสถิติ เพื่อสะท้อนระดับคุณภาพและความแตกต่างของการแปลบทคัดย่อในแต่ละมิติของเกณฑ์ MQM ขณะที่ผลการวิจัยเชิงคุณภาพในส่วนที่สามได้นำมาวิเคราะห์ด้วยวิธีการวิเคราะห์เนื้อหา (Content Analysis) เพื่อสังเคราะห์รูปแบบข้อผิดพลาด จุดเด่น และข้อเสนอแนะจากมุมมองของผู้เชี่ยวชาญ โดยเรียงลำดับดังนี้

- 4.1 ผลการประเมินคุณภาพการแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษตามเกณฑ์ MQM
- 4.2 ผลการเปรียบเทียบคุณภาพการแปลระหว่างการแปลโดยมนุษย์และการแปลโดย ChatGPT
- 4.3 ผลการสังเคราะห์ความคิดเห็นของผู้เชี่ยวชาญเกี่ยวกับการแปลบทคัดย่อโดยมนุษย์และการแปลโดย ChatGPT

4.1 ผลการประเมินคุณภาพการแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษตามเกณฑ์ MQM

ผู้วิจัยนำเสนอผลการประเมินคุณภาพการแปลบทคัดย่อวิทยานิพนธ์จากภาษาไทยเป็นภาษาอังกฤษโดยมนุษย์และ ChatGPT ตามเกณฑ์ MQM ใน 3 ด้าน โดยใช้สถิติพื้นฐาน ได้แก่ ค่าเฉลี่ย (μ) ส่วนเบี่ยงเบนมาตรฐาน (σ) และระดับคุณภาพ ดังปรากฏในตารางที่ 4.1

ตารางที่ 4.1 แสดงค่าเฉลี่ย (μ) ส่วนเบี่ยงเบนมาตรฐาน (σ) และระดับคุณภาพการแปลโดยภาพรวม และรายด้าน ของมนุษย์และ ChatGPT (N = 10 บทคัดย่อ)

รายการประเมิน (MQM)	รูปแบบการแปล	□	□	ระดับคุณภาพ
1. Accuracy (ความถูกต้องของ ของความหมาย)	มนุษย์	3.59	0.30	ดี
	ChatGPT	4.53	0.08	ดีเยี่ยม
2. Fluency (ความคล่องแคล่ว ทางภาษา)	มนุษย์	3.58	0.29	ดี
	ChatGPT	4.48	0.11	ดี
3. Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ)	มนุษย์	3.72	0.30	ดี
	ChatGPT	4.58	0.11	ดีเยี่ยม
รวมเฉลี่ยทุกด้าน	มนุษย์	3.63	0.29	ดี
	ChatGPT	4.53	0.11	ดีเยี่ยม

จากตารางที่ 4.1 ผลการประเมินคุณภาพการแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษตามเกณฑ์ Multidimensional Quality Metrics (MQM) พบว่า การแปลโดย ChatGPT มีคุณภาพโดยภาพรวมอยู่ในระดับดีเยี่ยม ($\mu = 4.53$, $\sigma = 0.11$) ในขณะที่การแปลโดยมนุษย์มีคุณภาพอยู่ในระดับดี ($\mu = 3.63$, $\sigma = 0.29$) เมื่อพิจารณาตามเกณฑ์แปลผลคะแนนแบบมาตราส่วน 5 ระดับแสดงให้เห็นถึงความแตกต่างของระดับคุณภาพการแปลโดยรวมระหว่างทั้งสองรูปแบบอย่างชัดเจน

เมื่อพิจารณาเป็นรายด้าน พบว่า ด้าน Accuracy (ความถูกต้องของความหมาย) การแปลโดย ChatGPT มีค่าเฉลี่ยอยู่ในระดับดีเยี่ยม ($\mu = 4.53$, $\sigma = 0.08$) สูงกว่าการแปลโดยมนุษย์ซึ่งอยู่ในระดับดี ($\mu = 3.59$, $\sigma = 0.30$) และเป็นด้านที่มีความแตกต่างของค่าเฉลี่ยมากที่สุดเมื่อเปรียบเทียบระหว่างทั้งสองรูปแบบ สะท้อนให้เห็นว่าคุณภาพการถ่ายทอดความหมายจากต้นฉบับภาษาไทยเป็นภาษาอังกฤษมีระดับแตกต่างกันอย่างชัดเจนในมิตินี้

สำหรับ ด้าน Fluency (ความคล่องแคล่วทางภาษา) พบว่า ChatGPT มีค่าเฉลี่ยอยู่ในระดับดี ($\mu = 4.48$, $\sigma = 0.11$) ใกล้เคียงกับระดับดีเยี่ยม ขณะที่การแปลโดยมนุษย์มีค่าเฉลี่ยอยู่ในระดับดีเช่นเดียวกัน ($\mu = 3.58$, $\sigma = 0.29$) ทั้งนี้ ด้าน Fluency (ความคล่องแคล่วทางภาษา) เป็นด้านที่ทั้งสองรูปแบบการแปลมีค่าเฉลี่ยต่ำที่สุดเมื่อเปรียบเทียบกับด้านอื่น สะท้อนให้เห็นว่าความคล่องแคล่วของภาษาเป็นมิติที่มีความแตกต่างของระดับคุณภาพน้อยกว่าด้านอื่นและยังเป็นด้านที่มีความแปรปรวนของคะแนนค่อนข้างสูงในกรณีของการแปลโดยมนุษย์

ในด้าน Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) พบว่า ChatGPT มีค่าเฉลี่ยสูงที่สุดในบรรดาทุกด้าน อยู่ในระดับดีเยี่ยม ($\mu = 4.58, \sigma = 0.11$) ขณะที่การแปลโดยมนุษย์มีค่าเฉลี่ยอยู่ในระดับดี ($\mu = 3.72, \sigma = 0.30$) แม้ว่าความแตกต่างของค่าเฉลี่ยในด้านนี้จะไม่มากที่สุดเมื่อเทียบกับด้าน Accuracy แต่ผลการประเมินแสดงให้เห็นถึงระดับคุณภาพที่แตกต่างกันอย่างสม่ำเสมอระหว่างทั้งสองรูปแบบการแปล

นอกจากนี้ เมื่อพิจารณาค่าความแปรปรวนของคะแนนจากส่วนเบี่ยงเบนมาตรฐาน พบว่า ChatGPT มีค่า σ ต่ำในทุกด้าน แสดงถึงความสม่ำเสมอของคุณภาพการแปลในชุดข้อมูลที่ศึกษา ขณะที่การแปลโดยมนุษย์มีค่า σ สูงกว่าในทุกด้าน สะท้อนถึงความแตกต่างของคุณภาพการแปลในแต่ละฉบับ ทั้งนี้ การจัดลำดับค่าเฉลี่ยรายด้านของทั้งสองรูปแบบการแปลมีแนวโน้มสอดคล้องกัน โดยมีค่าเฉลี่ยสูงสุดในด้าน Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ และต่ำสุดในด้าน Fluency (ความคล่องแคล่วทางภาษา)

4.2 ผลการเปรียบเทียบคุณภาพการแปลระหว่างการแปลโดยมนุษย์และการแปลโดย ChatGPT

ภายหลังจากการประเมินคุณภาพการแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษตามเกณฑ์ Multidimensional Quality Metrics (MQM) ในภาพรวมแล้ว ในหัวข้อนี้ผู้วิจัยนำเสนอผลการเปรียบเทียบความแตกต่างของคุณภาพการแปลบทคัดย่อวิทยานิพนธ์จากภาษาไทยเป็นภาษาอังกฤษ ระหว่างฉบับแปลโดยมนุษย์และฉบับแปลโดย ChatGPT ซึ่งเป็นข้อมูลสองชุดที่ได้จากต้นฉบับเดียวกัน และมีความสัมพันธ์กันเป็นรายคู่ (Paired Samples)

การวิเคราะห์ข้อมูลใช้สถิติทดสอบแบบไม่เป็นพารามิเตอร์ คือ Wilcoxon Signed - Rank Test เพื่อทดสอบความแตกต่างของคะแนนคุณภาพการแปล เนื่องจากลักษณะข้อมูลเป็นคะแนนเชิงอันดับ และมีขนาดกลุ่มตัวอย่างจำกัด การเปรียบเทียบครอบคลุมองค์ประกอบของเกณฑ์ MQM ทั้ง 3 ด้าน ได้แก่ Accuracy (ความถูกต้องของความหมาย) Fluency (ความคล่องแคล่วทางภาษา) และ Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) รวมถึงคะแนนเฉลี่ยรวม โดยกำหนดระดับนัยสำคัญทางสถิติที่ .05 ดังผลการวิเคราะห์ที่แสดงในตารางที่ 4.2

ตารางที่ 4.2 ผลการเปรียบเทียบความแตกต่างของคุณภาพการแปลระหว่างมนุษย์และ ChatGPT

ด้านการประเมิน (MQM)	Mean Rank (Negative) ^a	Mean Rank (Positive) ^b	Z	p - value	ผลการทดสอบ สมมติฐาน
1. Accuracy (ความถูกต้องของ ความหมาย)	5.50	.00	-2.80	.002*	แตกต่างกัน
2. Fluency (ความคล่องแคล่วทาง ภาษา)	5.50	.00	-2.80	.002*	แตกต่างกัน
3. Terminology Consistency (ความสอดคล้องของ ศัพท์เฉพาะ)	5.50	.00	-2.80	.002*	แตกต่างกัน
รวมเฉลี่ย	5.50	.00	-2.80	.002*	แตกต่างกัน

^aNegative Ranks = คะแนน ChatGPT สูงกว่า มนุษย์ (Human < ChatGPT)

^bPositive Ranks = คะแนน ChatGPT ต่ำกว่า มนุษย์ (Human > ChatGPT)

*มีนัยสำคัญทางสถิติที่ระดับ .05

จากตารางที่ 4.2 ผลการเปรียบเทียบความแตกต่างของคุณภาพการแปลบทคัดย่อภาษาไทย เป็นภาษาอังกฤษระหว่างการแปลโดยมนุษย์และการแปลโดย ChatGPT ด้วยสถิติ Wilcoxon Signed - Rank Test พบว่า คะแนนคุณภาพการแปลของทั้งสองรูปแบบมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ในทุกองค์ประกอบของเกณฑ์ MQM ได้แก่ Accuracy (ความถูกต้องของ ความหมาย) Fluency (ความคล่องแคล่วทาง ภาษา) และ Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) รวมถึงคะแนนเฉลี่ยรวม

เมื่อพิจารณาผลการทดสอบรายด้าน พบว่า ด้าน Accuracy (ความถูกต้องของ ความหมาย) ให้ค่า Z = -2.80 และค่า p = .002 แสดงให้เห็นว่าคะแนนคุณภาพการแปลระหว่างการแปลโดยมนุษย์ และการแปลโดย ChatGPT แตกต่างกันอย่างมีนัยสำคัญทางสถิติในด้านดังกล่าว เช่นเดียวกับด้าน Fluency (ความคล่องแคล่วทาง ภาษา) และด้าน Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) ซึ่งให้ผลการทดสอบในทิศทางเดียวกัน โดยมีค่า Z และค่า p เท่ากัน (Z = -2.80, p = .002) ทั้งสองด้าน

นอกจากนี้ เมื่อพิจารณาทิศทางของความแตกต่างจากลำดับผลต่างคะแนน (Ranks) พบว่าคะแนนทั้งหมดอยู่ในกลุ่ม Negative Ranks และไม่ปรากฏ Positive Ranks ในทุกองค์ประกอบของการประเมิน แสดงให้เห็นว่าคะแนนคุณภาพการแปลโดย ChatGPT สูงกว่าการแปลโดยมนุษย์อย่างสม่ำเสมอในทุกคู่การเปรียบเทียบ ภายใต้เกณฑ์ MQM ทั้งสามด้าน ทั้งนี้ การวิเคราะห์มุ่งเปรียบเทียบความแตกต่างของคุณภาพการแปลในระดับองค์ประกอบของเกณฑ์ MQM โดยไม่ได้พิจารณาความแตกต่างเชิงรายการณของบทคัดย่อแต่ละคู่

เมื่อพิจารณาคะแนนคุณภาพการแปลโดยรวม พบว่ามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 เช่นเดียวกัน ($Z = -2.80, p = .002$) ซึ่งสะท้อนให้เห็นถึงความแตกต่างของระดับคุณภาพการแปลโดยภาพรวมระหว่างการแปลโดยมนุษย์และการแปลโดย ChatGPT อย่างชัดเจน

ผลการวิเคราะห์ดังกล่าวนำไปสู่การปฏิเสศสมมติฐานศูนย์ และสนับสนุนสมมติฐานการวิจัยที่ตั้งไว้ว่า คุณภาพการแปลบทคัดย่อวิทยานิพนธ์จากภาษาไทยเป็นภาษาอังกฤษที่แปลโดยมนุษย์และที่แปลโดย ChatGPT มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 เมื่อพิจารณาตามองค์ประกอบของเกณฑ์ MQM ทั้งสามด้าน ได้แก่ Accuracy, Fluency และ Terminology Consistency

4.3 ผลการสังเคราะห์ความคิดเห็นของผู้เชี่ยวชาญเกี่ยวกับการแปลบทคัดย่อโดยมนุษย์และการแปลโดย ChatGPT

ผลการวิเคราะห์ข้อมูลเชิงคุณภาพในส่วนนี้มีวัตถุประสงค์เพื่อสังเคราะห์ความคิดเห็นและมุมมองของผู้เชี่ยวชาญเกี่ยวกับจุดเด่น ข้อจำกัด และลักษณะการใช้การแปลโดยมนุษย์และการแปลโดย ChatGPT โดยใช้ข้อมูลจากคำถามปลายเปิดในแบบสอบถามและความคิดเห็นเพิ่มเติมที่ผู้เชี่ยวชาญระบุไว้หลังการประเมินบทคัดย่อแต่ละฉบับ การวิเคราะห์ดำเนินการด้วยวิธีการวิเคราะห์เนื้อหา (Content Analysis) ตามที่กำหนดไว้ในบทที่ 3 ทั้งนี้ ผลการวิเคราะห์เชิงคุณภาพในหัวข้อนี้มีได้มีวัตถุประสงค์เพื่อประเมินหรือให้คะแนนคุณภาพการแปลซ้ำ แต่มีบทบาทในการอธิบายและสนับสนุนผลการวิเคราะห์เชิงปริมาณตามเกณฑ์ Multidimensional Quality Metrics (MQM) ที่นำเสนอในหัวข้อ 4.1

4.3.1 ความแตกต่างระหว่างการแปลโดยมนุษย์และ ChatGPT ในภาพรวม

จากการสังเคราะห์ความคิดเห็นของผู้เชี่ยวชาญด้านภาษาอังกฤษและการแปล พบว่าการแปลโดยมนุษย์และการแปลโดย ChatGPT มีลักษณะความแตกต่างเชิงคุณภาพอย่างชัดเจนในทุกองค์ประกอบของเกณฑ์ MQM ได้แก่ Accuracy (ความถูกต้องของความหมาย) Fluency (ความคล่องแคล่วทางภาษา) และ Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) โดยผู้เชี่ยวชาญส่วนใหญ่มองว่าการแปลทั้งสองรูปแบบมีจุดเด่นและข้อจำกัดที่แตกต่างกัน และควรได้รับการพิจารณาให้เหมาะสมกับบริบทการใช้งาน

ในด้าน Accuracy (ความถูกต้องของความหมาย) ผู้เชี่ยวชาญจำนวนมากเห็นว่าการแปลโดยมนุษย์มีความได้เปรียบในด้านความสามารถในการเข้าใจบริบทเชิงลึกของเนื้อหาต้นฉบับ โดยเฉพาะบริบททางวัฒนธรรม สังคม และระบบการศึกษาไทย เช่น ชื่อหน่วยงานราชการ ชื่อสถาบันการศึกษา หรือแนวคิดเฉพาะที่ไม่สามารถถ่ายทอดได้ตรงตัวเป็นภาษาอังกฤษ อย่างไรก็ตาม ผู้เชี่ยวชาญยังสะท้อนว่าการแปลโดยมนุษย์มีความเสี่ยงต่อการเกิดข้อผิดพลาดจากความไม่รอบคอบหรือความเคยชิน เช่น การแปลข้อมูลไม่ครบถ้วน การใช้คำไม่สม่าเสมอ หรือความคลาดเคลื่อนด้านตัวเลข หน่วยวัด และข้อมูลเชิงสถิติ ซึ่งเป็นข้อผิดพลาดที่แม้จะเกิดขึ้นในรายละเอียด แต่ส่งผลกระทบต่อความน่าเชื่อถือของบทคัดย่อโดยรวม

ในทางตรงกันข้าม ผู้เชี่ยวชาญส่วนใหญ่เห็นว่า ChatGPT มีจุดเด่นด้านการรักษาโครงสร้างและความครบถ้วนของข้อมูล โดยเฉพาะการถ่ายทอดข้อมูลเชิงตัวเลข ลำดับขั้นตอนการวิจัย และองค์ประกอบหลักของบทคัดย่อ อย่างไรก็ตาม ผู้เชี่ยวชาญบางรายตั้งข้อสังเกตว่า ChatGPT อาจเกิดการตีความคลาดเคลื่อนหรือการขยายความเกินต้นฉบับ (Addition) โดยเฉพาะในกรณีที่ต้นฉบับภาษาไทยมีความคลุมเครือหรือใช้ถ้อยคำเชิงนัย ซึ่งจำเป็นต้องอาศัยการตรวจสอบและปรับแก้โดยมนุษย์

ในด้าน Fluency (ความคล่องแคล่วทางภาษา) ผู้เชี่ยวชาญมีความเห็นสอดคล้องกันว่า ChatGPT มีความโดดเด่นในการสร้างประโยคภาษาอังกฤษที่ลื่นไหล เป็นธรรมชาติ และสอดคล้องกับรูปแบบการเขียนเชิงวิชาการ โดยเฉพาะการใช้โครงสร้างประโยคภาษาอังกฤษ การเลือกใช้คำเชื่อมและการใช้ Passive Voice อย่างเหมาะสม ส่งผลให้บทคัดย่อมีลักษณะเป็นภาษาอังกฤษเชิงวิชาการมากขึ้น ขณะที่การแปลโดยมนุษย์บางฉบับยังคงปรากฏลักษณะของการยึดโครงสร้างภาษาไทย (Translationese) ทำให้ภาษาอังกฤษขาดความลื่นไหลหรือมีโครงสร้างประโยคซับซ้อนเกินความจำเป็น

อย่างไรก็ตาม ผู้เชี่ยวชาญบางรายเห็นว่าความลื่นไหลของภาษาที่ได้จาก ChatGPT อาจทำให้ข้อความภาษาอังกฤษมีลักษณะ “ราบรื่นมาก” เมื่อเทียบกับต้นฉบับภาษาไทย โดยเฉพาะในกรณีที่บทคัดย่อภาษาไทยมีลักษณะกระชับหรือเป็นทางการน้อยกว่า ซึ่งอาจทำให้น้ำเสียงของข้อความภาษาอังกฤษเปลี่ยนไปจากต้นฉบับ หากไม่มีการตรวจทานโดยมนุษย์

ในด้าน Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) ผู้เชี่ยวชาญเห็นว่ามนุษย์ที่มีความเชี่ยวชาญเฉพาะสาขาสามารถเลือกใช้คำศัพท์เฉพาะได้อย่างเหมาะสม และสอดคล้องกับบริบทมากกว่า โดยเฉพาะศัพท์เชิงเทคนิค ชื่อทฤษฎี หรือคำศัพท์ที่มีความหมายเฉพาะในสาขาวิชานั้น ๆ อย่างไรก็ตาม หากผู้แปลไม่มีความเชี่ยวชาญเฉพาะด้าน ก็มีโอกาสเกิดความไม่สม่าเสมอในการใช้คำศัพท์หรือเลือกใช้คำที่ไม่เป็นมาตรฐานทางวิชาการ

สำหรับ ChatGPT ผู้เชี่ยวชาญเห็นว่าสามารถรักษาความสม่ำเสมอของคำศัพท์ตลอดทั้งบทความได้ดีในระดับหนึ่ง อย่างไรก็ตาม ยังพบข้อจำกัดในกรณีคำศัพท์ที่เกี่ยวข้องกับบริบทไทยหรือคำศัพท์ราชการ ซึ่ง ChatGPT อาจเลือกใช้คำที่ถูกต้องในเชิงภาษา แต่ไม่ตรงกับการใช้จริงในบริบททางวิชาการของประเทศไทย

โดยสรุป ผู้เชี่ยวชาญส่วนใหญ่เห็นว่าการแปลโดยมนุษย์และการแปลโดย ChatGPT ไม่ได้มีลักษณะทดแทนกันโดยสมบูรณ์ แต่มีบทบาทที่แตกต่างและสามารถนำมาใช้ร่วมกันได้อย่างเหมาะสม โดย ChatGPT สามารถใช้เป็นเครื่องมือช่วยร่างและปรับสำนวนภาษา ขณะที่มนุษย์มีบทบาทสำคัญในการตรวจสอบความถูกต้องของเนื้อหา การเลือกใช้คำศัพท์เฉพาะ และการรักษาความสอดคล้องกับบริบททางวิชาการและวัฒนธรรมของต้นฉบับภาษาไทย

4.3.2 ข้อเสนอแนะในการพัฒนาคุณภาพการแปลบทความภาษาอังกฤษ

จากการสังเคราะห์ความคิดเห็นของผู้เชี่ยวชาญ พบว่าการพัฒนาคุณภาพการแปลบทความภาษาอังกฤษควรดำเนินการในลักษณะของกระบวนการแบบบูรณาการ โดยผสานจุดแข็งของมนุษย์และปัญญาประดิษฐ์ (Artificial Intelligence: AI) เข้าด้วยกัน ข้อเสนอแนะของผู้เชี่ยวชาญสามารถสรุปเป็นแนวทางสำคัญได้ดังนี้

1) ผู้เชี่ยวชาญเห็นตรงกันว่า AI ควรใช้เป็นเครื่องมือช่วยแปลในขั้นต้น ไม่ใช่เครื่องมือทดแทนมนุษย์โดยสมบูรณ์ โดยเฉพาะในงานวิชาการระดับบัณฑิตศึกษา AI มีศักยภาพในการช่วยจัดทำร่างการแปล (Draft Translation) ที่มีความรวดเร็วและความสับสนไหลทางภาษาแต่ยังจำเป็นต้องมีการตรวจทานและตัดสินใจขั้นสุดท้ายโดยมนุษย์ (Post - Editing) เพื่อควบคุมความถูกต้องของเนื้อหาและความเหมาะสมเชิงวิชาการ

2) ผู้เชี่ยวชาญเสนอให้พัฒนาทักษะการตรวจทานงานแปลของผู้แปลและนักศึกษา โดยเฉพาะการตรวจสอบสาระสำคัญของบทความได้แก่ วัตถุประสงค์การวิจัย วิธีดำเนินการวิจัย ผลการวิจัย และข้อสรุป รวมถึงข้อมูลเชิงตัวเลข หน่วยวัด และสถิติ ซึ่งเป็นประเด็นที่ AI อาจเกิดความคลาดเคลื่อนได้

3) ผู้เชี่ยวชาญเสนอให้จัดทำฐานข้อมูลคำศัพท์เฉพาะทาง (Terminology Database / Termbase) สำหรับแต่ละสาขาวิชา โดยเฉพาะสาขาที่มีศัพท์เชิงเทคนิคหรือศัพท์เฉพาะบริบทไทย เพื่อให้ทั้งมนุษย์และ AI ใช้เป็นแหล่งอ้างอิงเดียวกัน ลดปัญหาความไม่สม่ำเสมอในการเลือกใช้คำศัพท์

4) ผู้เชี่ยวชาญเน้นความสำคัญของการพัฒนาทักษะการเขียนเชิงวิชาการภาษาอังกฤษควบคู่กับการแปล โดยเฉพาะความเข้าใจเรื่องระดับภาษา โครงสร้างประโยคเชิงวิชาการ และการใช้ Passive Voice อย่างเหมาะสม เพื่อให้สามารถตรวจสอบและปรับแก้งานแปลได้อย่างมีประสิทธิภาพ

5) ผู้เชี่ยวชาญเสนอให้จัดอบรมหรือกิจกรรมพัฒนาศักยภาพด้านการใช้ AI เพื่อการแปลเชิงวิชาการ โดยเนื้อหาควรครอบคลุมการออกแบบคำสั่ง (Prompting), การตรวจสอบความถูกต้องของผลลัพธ์จาก AI และเทคนิคการปรับแก้งานแปลให้สอดคล้องกับมาตรฐานทางวิชาการ

6) ผู้เชี่ยวชาญเสนอให้ใช้กระบวนการควบคุมคุณภาพงานแปลหลายขั้นตอน เช่น การตรวจทานโดยผู้เชี่ยวชาญด้านภาษาอังกฤษหรือผู้เชี่ยวชาญเฉพาะสาขา และการใช้เทคนิค Back Translation (การแปลย้อนกลับ) ในกรณีที่บทความมีเนื้อหาเชิงเทคนิคสูง ทั้งนี้ ผู้เชี่ยวชาญมิได้เสนอให้ใช้ Back Translation เป็นขั้นตอนหลักในทุกกรณี แต่เป็นเครื่องมือเสริมในการตรวจสอบคุณภาพงานแปล

4.3.3 ลักษณะคำสั่ง (Prompt) ที่ผู้เชี่ยวชาญใช้กับ ChatGPT

จากการวิเคราะห์ข้อมูลเชิงคุณภาพ พบว่าผู้เชี่ยวชาญให้ความสำคัญกับการออกแบบคำสั่ง (Prompt Engineering) โดยเห็นว่าคุณภาพของผลลัพธ์จาก ChatGPT ขึ้นอยู่กับความชัดเจน ความเฉพาะเจาะจง และระดับรายละเอียดของคำสั่งที่ใช้ ผู้เชี่ยวชาญมีการปรับระดับคำสั่งตามลักษณะของบทความและสาขาวิชา ซึ่งสามารถสังเคราะห์ลักษณะคำสั่งได้เป็น 3 ระดับ ได้แก่ คำสั่งพื้นฐาน (Basic Prompt), คำสั่งเฉพาะเจาะจง (Specific Prompt) และคำสั่งเพื่อปรับปรุงคุณภาพและการกลั่นภาษา (Refinement Prompt)

1) คำสั่งพื้นฐาน (Basic Prompt) ผู้เชี่ยวชาญมักใช้คำสั่งที่ระบุเป้าหมายของการแปลอย่างชัดเจน โดยเน้นให้ ChatGPT แปลบทความจากภาษาไทยเป็นภาษาอังกฤษเชิงวิชาการที่เหมาะสมกับงานวิจัย ตัวอย่างคำสั่งที่พบ เช่น

“Translate the following Thai abstract into academic English suitable for a thesis abstract.”

คำสั่งในลักษณะนี้ช่วยให้ได้ผลลัพธ์ที่มีโครงสร้างภาษาอังกฤษที่เป็นทางการ และอ่านเข้าใจได้ง่าย อย่างไรก็ตาม ผู้เชี่ยวชาญเห็นว่าคำสั่งระดับพื้นฐานเพียงอย่างเดียวอาจยังไม่เพียงพอสำหรับงานวิจัยที่มีเนื้อหาเชิงเทคนิคหรือมีข้อมูลเชิงตัวเลขจำนวนมาก

2) คำสั่งเฉพาะเจาะจง (Specific Prompt) ซึ่งผู้เชี่ยวชาญจะเพิ่มเงื่อนไขเพื่อควบคุมความถูกต้องของเนื้อหา (Accuracy) และการใช้คำศัพท์เฉพาะทาง (Terminology) โดยเฉพาะ ตัวอย่างคำสั่งที่ผู้เชี่ยวชาญใช้ เช่น

“Translate the following Thai abstract into concise academic English suitable for journal publication, strictly adhering to the source text’s intent and data. Do not change or interpret any numerical data, percentages, or statistical results.”

นอกจากนี้ ผู้เชี่ยวชาญบางรายยังระบุคำศัพท์เฉพาะหรือชื่อเฉพาะที่ต้องการให้ใช้แบบตายตัว เพื่อหลีกเลี่ยงความคลาดเคลื่อน เช่น

“Ensure that the term ‘Priority Needs Index (PNI)’ is used consistently throughout the abstract.”

“Ensure the name ‘Davies’ is accurately translated and not confused with ‘Dewey’.”

คำสั่งลักษณะนี้สะท้อนให้เห็นว่าผู้เชี่ยวชาญตระหนักถึงข้อจำกัดของ ChatGPT ในการจัดการคำศัพท์เฉพาะ และพยายามลดความเสี่ยงของข้อผิดพลาดผ่านการกำหนดเงื่อนไขอย่างละเอียด

นอกจากนี้ จากข้อมูลเชิงคุณภาพยังพบว่าผู้เชี่ยวชาญจำนวนหนึ่งให้ความสำคัญกับการกำหนดบทบาท (Role) ให้กับ ChatGPT ก่อนการแปล โดยกำหนดให้ทำหน้าที่เสมือนผู้เชี่ยวชาญด้านการแปลเชิงวิชาการหรือผู้เชี่ยวชาญเฉพาะสาขา ผู้เชี่ยวชาญเห็นว่าการกำหนด Role ช่วยควบคุมระดับภาษา ความถูกต้องของความหมาย และการเลือกใช้ศัพท์เฉพาะให้สอดคล้องกับบริบทของงานวิจัย ส่งผลให้ผลลัพธ์จาก ChatGPT มีคุณภาพและสอดคล้องกับมาตรฐานทางวิชาการมากยิ่งขึ้น ซึ่งสามารถจำแนกบทบาทที่พบได้เป็น 3 ลักษณะ ดังนี้

(1) ผู้เชี่ยวชาญด้านการแปลเชิงวิชาการ

ตัวอย่างคำสั่งจากผู้เชี่ยวชาญ ได้แก่

“You are an expert academic translator. Translate the following Thai abstract into formal academic English.”

(2) ผู้เชี่ยวชาญเฉพาะสาขา

ผู้เชี่ยวชาญบางรายกำหนดบทบาทให้สอดคล้องกับสาขาวิชา เช่น วิศวกรรม สังคมศาสตร์ หรือการศึกษา เช่น

“You are an expert in educational research. Translate the following abstract while maintaining accurate educational terminology.”

(3) บทบาทผู้ช่วยตรวจสอบความถูกต้องของงานแปล

ในบางกรณี ผู้เชี่ยวชาญใช้การกำหนดบทบาทเพื่อเน้นความถูกต้องของเนื้อหา มากกว่าความสละสลวยของภาษา โดยใช้คำสั่ง เช่น

“You are a professional proofreader and academic editor. Check the translation for accuracy, numbers, and consistency.”

3) คำสั่งเพื่อปรับปรุงคุณภาพและการกลั่นภาษา (Refinement Prompt) ซึ่งมักใช้หลังจากได้ผลการแปลครั้งแรกแล้ว ผู้เชี่ยวชาญจะใช้คำสั่งเพิ่มเติมเพื่อปรับสำนวนให้มีความสื่อนไหวเหมาะสมกับการตีพิมพ์ และคงลักษณะภาษาเชิงวิชาการ ตัวอย่างคำสั่งที่พบ ได้แก่

“Revise the previous translation to improve fluency and ensure the use of formal academic register. Use passive voice where appropriate and maintain consistency in terminology.”

หรือในบางกรณี ผู้เชี่ยวชาญจะขอให้ ChatGPT ปรับภาษาให้เหมาะกับผู้อ่านเฉพาะกลุ่ม เช่น

“Please refine the abstract for an international academic audience, ensuring clarity, conciseness, and adherence to academic writing conventions.”

จากข้อมูลทั้งหมด ผู้เชี่ยวชาญเห็นตรงกันว่า การใช้ ChatGPT อย่างมีประสิทธิภาพไม่ใช่เพียงการสั่งให้แปล แต่เป็นกระบวนการเชิงขั้นตอน ที่เริ่มจากการให้คำสั่งพื้นฐาน ตามด้วยการกำหนดเงื่อนไขเฉพาะ และจบด้วยการปรับแก้สำนวนและตรวจสอบความถูกต้อง ผู้เชี่ยวชาญจึงมองว่า Prompt ไม่ได้เป็นเพียงคำสั่ง แต่เป็นเครื่องมือควบคุมคุณภาพการแปล ซึ่งช่วยลดข้อผิดพลาดด้านความหมาย คำศัพท์ และความไม่สม่ำเสมอของภาษาในบทความภาษาอังกฤษ

4.3.4 ประสบการณ์ของผู้เชี่ยวชาญในการใช้ AI เครื่องมืออื่น

จากการสังเคราะห์ความคิดเห็นของผู้เชี่ยวชาญเกี่ยวกับประสบการณ์ในการใช้ปัญญาประดิษฐ์ (Artificial Intelligence: AI) เครื่องมืออื่น ๆ นอกเหนือจาก ChatGPT พบว่า ผู้เชี่ยวชาญส่วนใหญ่มีประสบการณ์ในการใช้เครื่องมือแปลภาษาอัตโนมัติหลากหลายประเภท ทั้งเครื่องมือแปลแบบดั้งเดิมที่อาศัยกฎหรือสถิติ และเครื่องมือแปลที่ใช้เทคโนโลยีปัญญาประดิษฐ์ขั้นสูง โดยเครื่องมือที่ถูกกล่าวถึงบ่อย ได้แก่ Google Translate, DeepL, Microsoft Translator รวมถึงเครื่องมือในกลุ่ม Generative AI ที่สามารถสร้างข้อความภาษาอังกฤษเชิงวิชาการได้อย่างเป็นธรรมชาติ

ผู้เชี่ยวชาญส่วนใหญ่เห็นว่า Google Translate เป็นเครื่องมือที่มีความสะดวกและรวดเร็ว เหมาะสำหรับการทำความเข้าใจเนื้อหาในระดับเบื้องต้น อย่างไรก็ตาม เมื่อใช้ในบริบทการแปลบทความวิทยานิพนธ์ พบข้อจำกัดด้าน Accuracy (ความถูกต้องของความหมาย) และ Fluency (ความคล่องแคล่วทางภาษา) โดยเฉพาะในกรณีประโยคที่มีโครงสร้างซับซ้อนหรือมีคำศัพท์เฉพาะทางวิชาการ ผู้เชี่ยวชาญบางรายสะท้อนว่า Google Translate มีแนวโน้มแปลแบบตรงตัว ส่งผลให้ภาษาอังกฤษที่ได้ขาดความสื่อนไหวและไม่สอดคล้องกับรูปแบบการเขียนเชิงวิชาการ

ในกรณีของ DeepL ผู้เชี่ยวชาญบางรายให้ความเห็นว่าเครื่องมือนี้มีจุดเด่นด้านความลื่นไหลของภาษาและโครงสร้างประโยคที่ใกล้เคียงภาษาเจ้าของภาษา อย่างไรก็ตาม ยังมีข้อจำกัดด้านการรองรับภาษาไทยและการจัดการคำศัพท์เฉพาะในบริบทไทย ทำให้การใช้งานในงานแปลบทความภาษาไทยเป็นภาษาอังกฤษยังไม่แพร่หลายเท่าที่ควร

สำหรับ Microsoft Translator และเครื่องมือแปลภาษาอื่น ๆ ผู้เชี่ยวชาญเห็นว่าให้ผลลัพธ์ในระดับปานกลาง เหมาะสำหรับการใช้งานทั่วไปหรือใช้เป็นเครื่องมือเสริมในการตรวจสอบความหมาย แต่ยังไม่สามารถตอบโจทย์ด้านความถูกต้องของความหมายและความเหมาะสมเชิงวิชาการในงานวิจัยระดับบัณฑิตศึกษาได้อย่างเต็มที่

เมื่อเปรียบเทียบกับเครื่องมือแปลภาษาอื่น ผู้เชี่ยวชาญส่วนใหญ่เห็นว่า ChatGPT มีความยืดหยุ่นสูงกว่า เนื่องจากสามารถปรับผลลัพธ์ได้ตามคำสั่ง (Prompt) ที่ผู้ใช้กำหนด ทั้งในด้านระดับภาษา ความกระชับ และรูปแบบการเขียนเชิงวิชาการ อย่างไรก็ตาม ผู้เชี่ยวชาญยังเน้นย้ำว่าไม่ว่าเครื่องมือ AI ใดจะถูกนำมาใช้ การแปลบทความวิทยานิพนธ์ยังคงจำเป็นต้องมีการตรวจทานโดยมนุษย์ (Post – Editing) เพื่อควบคุมความถูกต้องของความหมาย การใช้คำศัพท์เฉพาะและความสอดคล้องกับบริบททางวิชาการ

นอกจากนี้ ผู้เชี่ยวชาญบางรายสะท้อนแนวคิดการใช้ AI มากกว่าหนึ่งเครื่องมือร่วมกัน เพื่อเปรียบเทียบผลลัพธ์จากหลายแหล่งและช่วยลดความเสี่ยงของข้อผิดพลาด อย่างไรก็ตาม แนวทางดังกล่าวยังคงต้องอาศัยการตัดสินใจของมนุษย์เป็นหลักในการเลือกผลลัพธ์ที่เหมาะสมที่สุดสำหรับบริบทของงานวิจัย

โดยสรุป ประสบการณ์ของผู้เชี่ยวชาญสะท้อนให้เห็นว่า AI เครื่องมือแปลภาษาแต่ละประเภทมีจุดเด่นและข้อจำกัดที่แตกต่างกัน และยังไม่สามารถทดแทนบทบาทของมนุษย์ได้อย่างสมบูรณ์ การเลือกใช้เครื่องมือ AI จึงควรพิจารณาลักษณะของงาน ระดับความซับซ้อนของเนื้อหา และการจัดให้มีกระบวนการควบคุมคุณภาพที่เหมาะสม เพื่อให้บทความแปลเป็นภาษาอังกฤษมีคุณภาพและสอดคล้องกับมาตรฐานทางวิชาการ

4.3.5 ข้อเสนอแนะของผู้เชี่ยวชาญต่อการแปลบทความรายฉบับ

จากการวิเคราะห์ข้อเสนอแนะเชิงคุณภาพที่ผู้เชี่ยวชาญให้ไว้หลังการประเมินคุณภาพการแปลบทความในเชิงปริมาณ พบว่าผู้เชี่ยวชาญได้ให้ข้อเสนอแนะที่สะท้อนลักษณะเฉพาะของงานแปลในแต่ละรูปแบบอย่างชัดเจน ทั้งในกรณีของการแปลโดยมนุษย์และการแปลโดย ChatGPT ทั้งนี้ ผู้เชี่ยวชาญมิได้รับทราบข้อมูลเกี่ยวกับผู้แปลของแต่ละฉบับในระหว่างการให้ข้อเสนอแนะเพื่อหลีกเลี่ยงอคติ (Bias) ที่อาจเกิดจากการรับรู้แหล่งที่มาของงานแปล ส่งผลให้ข้อเสนอแนะที่ได้รับสะท้อนการประเมินจากคุณภาพของตัวบทแปลโดยตรง

ข้อเสนอแนะดังกล่าวมิได้มีวัตถุประสงค์เพื่อให้คะแนนคุณภาพการแปลซ้ำ แต่มีบทบาทในการอธิบายลักษณะข้อผิดพลาด จุดเด่น และประเด็นที่ควรปรับปรุงของงานแปลในแต่ละรูปแบบ ซึ่งช่วยสนับสนุนและอธิบายผลการวิเคราะห์เชิงปริมาณตามกรอบ Multidimensional Quality Metrics (MQM) ที่นำเสนอในหัวข้อ 4.1 โดยสามารถสังเคราะห์ข้อเสนอแนะออกเป็น 2 ประเด็นหลัก ได้แก่ ข้อเสนอแนะต่อการแปลโดยมนุษย์ และข้อเสนอแนะต่อการแปลโดย ChatGPT ดังรายละเอียดต่อไปนี้

1) ข้อเสนอแนะต่อการแปลโดยมนุษย์

ผู้เชี่ยวชาญส่วนใหญ่ให้ข้อเสนอแนะว่าการแปลโดยมนุษย์มีแนวโน้มเกิดข้อผิดพลาดด้าน Accuracy (ความถูกต้องของความหมาย) และความครบถ้วนของเนื้อหา (Content Completeness) โดยเฉพาะในลักษณะของการละเลยข้อความสำคัญ (Omission), การแปลไม่ครบถ้วน และการแปลคลาดเคลื่อนในเชิงเทคนิค (Mistranslation) ซึ่งในบางกรณีส่งผลต่อความน่าเชื่อถือของบทคัดย่อ

ตัวอย่างข้อผิดพลาดที่พบซ้ำในหลายบทคัดย่อ ได้แก่ ความคลาดเคลื่อนด้านข้อมูลเชิงตัวเลขและหน่วยวัด เช่น การแปลคำร้อยละหรือหน่วยทางวิทยาศาสตร์ไม่ตรงกับต้นฉบับ ซึ่งผู้เชี่ยวชาญเห็นว่าเป็นประเด็นที่ควรให้ความสำคัญเป็นพิเศษในงานวิจัยเชิงทดลองและงานวิจัยเชิงปริมาณ

นอกจากนี้ ผู้เชี่ยวชาญยังสะท้อนว่าการแปลโดยมนุษย์บางฉบับมีลักษณะของการแปลตรงตัวจากภาษาไทย (Literal Translation) ส่งผลให้โครงสร้างประโยคภาษาอังกฤษไม่เป็นธรรมชาติ และไม่สอดคล้องกับสำนวนการเขียนเชิงวิชาการ เช่น การเลือกใช้คำกริยาหรือโครงสร้างประโยคที่ไม่เหมาะสมกับบริบท

ในด้าน Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) ผู้เชี่ยวชาญพบว่าการแปลโดยมนุษย์อาจเกิดความไม่สม่ำเสมอในการใช้คำศัพท์ โดยเฉพาะชื่อทฤษฎี ชื่อบุคคล หรือศัพท์เชิงเทคนิค ซึ่งในบางกรณีสะท้อนถึงความจำเป็นในการตรวจสอบศัพท์เฉพาะอย่างเป็นระบบมากขึ้น

โดยสรุป ผู้เชี่ยวชาญเสนอแนะว่าการแปลโดยมนุษย์ควรให้ความสำคัญกับการตรวจสอบข้อมูลเชิงตัวเลขและศัพท์เฉพาะอย่างรอบคอบ ลดการแปลแบบตรงตัว และปรับสำนวนให้สอดคล้องกับรูปแบบการเขียนเชิงวิชาการสากล รวมถึงควรมีการตรวจทานเพิ่มเติมก่อนนำไปใช้จริง

2) ข้อเสนอแนะต่อการแปลโดย ChatGPT

ในภาพรวม ผู้เชี่ยวชาญส่วนใหญ่มีความเห็นว่าการแปลโดย ChatGPT มีจุดเด่นด้านความครบถ้วนของเนื้อหา (Content Completeness) และ Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) โดยพบว่าสามารถรักษารายละเอียดจากต้นฉบับภาษาไทยได้ค่อนข้างครบถ้วน และมีการใช้คำศัพท์เชิงวิชาการอย่างสม่ำเสมอตลอดทั้งบทความ

อย่างไรก็ตาม ผู้เชี่ยวชาญยังให้ข้อสังเกตว่า ChatGPT มีข้อจำกัดในบางประการ โดยเฉพาะการตีความเพิ่มเติมหรือการขยายความเกินต้นฉบับ (Over – Interpretation/ Addition) ซึ่งอาจส่งผลให้ความหมายของข้อความภาษาอังกฤษคลาดเคลื่อนไปจากเจตนาเดิมของต้นฉบับ นอกจากนี้ ยังพบข้อสังเกตเกี่ยวกับความยาวของประโยคและรูปแบบการจัดข้อความที่อาจไม่สอดคล้องกับแนวทางของสถาบันในบางกรณี

ผู้เชี่ยวชาญจึงเสนอแนะว่า แม้ ChatGPT จะมีศักยภาพในการช่วยแปลบทความเชิงวิชาการได้อย่างมีประสิทธิภาพ แต่ควรใช้ควบคู่กับการตรวจทานโดยมนุษย์ (Post - Editing) และในบางกรณีอาจใช้เทคนิคเสริม เช่น Back Translation (การแปลย้อนกลับ) หรือการใช้ฐานข้อมูลคำศัพท์เฉพาะ เพื่อช่วยควบคุมความถูกต้องของคำแปลและความสอดคล้องของศัพท์เฉพาะ

โดยสรุป ผู้เชี่ยวชาญเห็นว่า ChatGPT เหมาะสมสำหรับการใช้เป็นเครื่องมือช่วยแปลในขั้นต้น ขณะที่บทบาทของมนุษย์ยังคงมีความสำคัญในการตรวจสอบ ปรับแก้ และควบคุมความเหมาะสมของเนื้อหาและบริบททางวิชาการก่อนการนำไปใช้หรือเผยแพร่

4.4 สรุปผลการวิจัยในภาพรวม

จากผลการวิเคราะห์ข้อมูลเชิงปริมาณตามเกณฑ์ Multidimensional Quality Metrics (MQM) และผลการวิเคราะห์ข้อมูลเชิงคุณภาพจากความคิดเห็นของผู้เชี่ยวชาญ สามารถสรุปผลการวิจัยในภาพรวมได้ดังนี้

4.4.1 ภาพรวมคุณภาพการแปลบทความภาษาไทยเป็นภาษาอังกฤษ โดยผลการประเมินคุณภาพการแปลตามเกณฑ์ MQM พบว่า การแปลโดย ChatGPT มีระดับคุณภาพโดยรวมอยู่ในระดับดีเยี่ยม ขณะที่การแปลโดยมนุษย์มีคุณภาพอยู่ในระดับดี โดย ChatGPT ได้คะแนนเฉลี่ยสูงกว่าการแปลโดยมนุษย์อย่างสม่ำเสมอในทุกองค์ประกอบ ได้แก่ Accuracy (ความถูกต้องของคำแปล), Fluency (ความคล่องแคล่วทางภาษา) และ Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) นอกจากนี้ การแปลโดย ChatGPT ยังมีค่าความแปรปรวนของคะแนนต่ำกว่า สะท้อนถึงความสม่ำเสมอของคุณภาพการแปลในชุดข้อมูลที่ศึกษา

4.4.2 ความแตกต่างของคุณภาพการแปลระหว่างมนุษย์และ ChatGPT โดยผลการทดสอบความแตกต่างของคะแนนคุณภาพการแปลด้วยสถิติ Wilcoxon Signed - Rank Test แสดงให้เห็นว่าคะแนนคุณภาพการแปลระหว่างการแปลโดยมนุษย์และการแปลโดย ChatGPT แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ในทุกองค์ประกอบของเกณฑ์ MQM รวมถึงคะแนนเฉลี่ยรวมโดยทิศทางของความแตกต่างเป็นไปในแนวทางเดียวกันทั้งหมด กล่าวคือ การแปลโดย ChatGPT ได้คะแนนสูงกว่าการแปลโดยมนุษย์ในทุกคู่การเปรียบเทียบ ทั้งนี้ การเปรียบเทียบดังกล่าวเป็นการสรุปผลในระดับองค์ประกอบของเกณฑ์และคะแนนรวม มิได้มุ่งพิจารณาความแตกต่างเชิงรายการณีของบทคัดย่อแต่ละฉบับ

4.4.3 บทบาท จุดเด่น และข้อจำกัดของการแปลโดยมนุษย์และ ChatGPT โดยผลการวิเคราะห์ความคิดเห็นของผู้เชี่ยวชาญสะท้อนว่า การแปลโดย ChatGPT มีจุดเด่นด้านความเป็นระบบของภาษา ความสั้นไหลของสำนวน และความสม่ำเสมอในการใช้ศัพท์เฉพาะ ขณะที่การแปลโดยมนุษย์มีจุดเด่นด้านความเข้าใจบริบทเชิงลึกและความเหมาะสมของสำนวนทางวิชาการ อย่างไรก็ตาม การแปลทั้งสองรูปแบบยังมีข้อจำกัดที่แตกต่างกัน ส่งผลให้ผู้เชี่ยวชาญเห็นว่า การแปลโดย ChatGPT และการแปลโดยมนุษย์ควรถูกนำมาใช้ในลักษณะเสริมกัน โดยมนุษย์ยังมีบทบาทสำคัญในการตรวจทานและปรับแก้ก่อนนำบทคัดย่อไปใช้งานจริง

โดยสรุป ผลการวิจัยในบทนี้สะท้อนให้เห็นภาพรวมของคุณภาพ ความแตกต่าง และบทบาทของการแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษโดยมนุษย์และ ChatGPT อย่างชัดเจน โดยสามารถแสดงได้ดังภาพที่ 4.1 ดังนี้



ภาพที่ 4.1 แสดงแผนภาพสรุปผลการวิจัยในภาพรวม

ที่มา : ผู้วิจัยสังเคราะห์จากผลการวิเคราะห์ในบทที่ 4

จากภาพที่ 4.1 แสดงผลการสังเคราะห์ข้อมูลเชิงปริมาณจากการประเมินคุณภาพการแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษตามเกณฑ์ Multidimensional Quality Metrics (MQM) ร่วมกับข้อมูลเชิงคุณภาพจากความคิดเห็นของผู้เชี่ยวชาญ เพื่อสะท้อนภาพรวมของคุณภาพ ความแตกต่าง และบทบาทของการแปลโดยมนุษย์และ ChatGPT

จากผลการวิเคราะห์ข้อมูลในบทนี้ สามารถสรุปได้ว่า การแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษโดยมนุษย์และโดย ChatGPT มีระดับคุณภาพและลักษณะเด่นที่แตกต่างกันอย่างชัดเจน ทั้งในมิติของความถูกต้องของความหมาย ความคล่องแคล่วทางภาษา และความสอดคล้องของศัพท์เฉพาะ ภายใต้กรอบการประเมิน Multidimensional Quality Metrics (MQM) ทั้งนี้ ผลการวิจัยในบทที่ 4 เป็นการรายงานผลตามข้อมูลเชิงประจักษ์ทั้งเชิงปริมาณและเชิงคุณภาพ โดยยังมีได้ทำการตีความหรือเชื่อมโยงกับแนวคิด ทฤษฎี หรือผลงานวิจัยที่เกี่ยวข้อง ซึ่งจะนำเสนอและอภิปรายอย่างเป็นระบบในบทที่ 5 ต่อไป



บทที่ 5

สรุปผลการวิจัย อภิปรายผลและข้อเสนอแนะ

งานวิจัยนี้มุ่งประเมินและเปรียบเทียบคุณภาพการแปลบทคัดย่อวิทยานิพนธ์จากภาษาไทยเป็นภาษาอังกฤษ ระหว่างการแปลโดยมนุษย์และการแปลโดย ChatGPT ภายใต้กรอบการประเมิน Multidimensional Quality Metrics (MQM) ซึ่งครอบคลุม 3 องค์ประกอบ ได้แก่ Accuracy (ความถูกต้องของความหมาย) Fluency (ความคล่องแคล่วทางภาษา) และ Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) โดยใช้ข้อมูลบทคัดย่อภาษาไทยจำนวน 10 เรื่อง และฉบับแปลภาษาอังกฤษจำนวน 20 ฉบับ (ฉบับแปลโดยมนุษย์ และฉบับแปลโดย ChatGPT จากต้นฉบับเดียวกัน) จากนั้นให้ผู้เชี่ยวชาญด้านภาษาอังกฤษประเมินคุณภาพตามเกณฑ์ MQM และนำคะแนนมาวิเคราะห์ด้วยสถิติเชิงพรรณนา รวมถึงทดสอบความแตกต่างของคะแนนแบบข้อมูลคู่สัมพันธ์ด้วยสถิติ Wilcoxon Signed - Rank Test ที่ระดับนัยสำคัญ .05 ทั้งนี้ ข้อมูลเชิงคุณภาพได้จากคำถามปลายเปิดในแบบสอบถาม และวิเคราะห์ด้วยการวิเคราะห์เนื้อหา (Content Analysis) เพื่อนำไปสังเคราะห์จุดเด่น ข้อจำกัด และข้อเสนอแนะเชิงปฏิบัติในการใช้ AI เพื่อสนับสนุนการแปลเชิงวิชาการให้เหมาะสมกับบริบทระดับบัณฑิตศึกษา โดยนำเสนอ ดังนี้

- 5.1 สรุปผลการวิจัย
- 5.2 อภิปรายผลการวิจัย
- 5.3 ข้อเสนอแนะในการวิจัย
- 5.4 สรุปประโยชน์ที่ได้จากการวิจัย

5.1 สรุปผลการวิจัย

หัวข้อนี้นำเสนอการสรุปผลการวิจัยตามวัตถุประสงค์ที่กำหนดไว้ โดยมุ่งรายงานผลการประเมินและการเปรียบเทียบคุณภาพการแปลบทคัดย่อวิทยานิพนธ์จากภาษาไทยเป็นภาษาอังกฤษ ระหว่างการแปลโดยมนุษย์และการแปลโดย ChatGPT ภายใต้กรอบการประเมิน Multidimensional Quality Metrics (MQM) ซึ่งประกอบด้วยองค์ประกอบหลัก 3 ด้าน ได้แก่ Accuracy (ความถูกต้องของความหมาย), Fluency (ความคล่องแคล่วทางภาษา) และ Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) โดยการสรุปผลในหัวข้อนี้เป็นการรายงานผลตามข้อมูลเชิงประจักษ์ (Empirical Evidence) ที่ได้จากการประเมินโดยผู้เชี่ยวชาญด้านภาษาอังกฤษ โดยนำเสนอทั้งในลักษณะเชิงพรรณนาเพื่ออธิบายภาพรวมของคุณภาพการแปล และเชิงเปรียบเทียบเพื่อแสดงแนวโน้มความแตกต่างของคะแนนระหว่างการแปลโดยมนุษย์และการแปลโดย ChatGPT ในระดับองค์ประกอบของเกณฑ์ MQM และคะแนนรวม ทั้งนี้ การรายงานผลดังกล่าวเป็นการสรุปในระดับภาพรวมของข้อมูลคู่สัมพันธ์ มิได้มุ่งอธิบายเชิงรายกรณีของบทคัดย่อแต่ละคู่

นอกจากนี้ หัวข้อนี้ยังสรุปผลการวิเคราะห์ข้อมูลเชิงคุณภาพจากคำถามปลายเปิด โดยใช้การวิเคราะห์เนื้อหา (Content Analysis) เพื่อสังเคราะห์ความคิดเห็นของผู้เชี่ยวชาญเกี่ยวกับ จุดเด่น ข้อจำกัด และบทบาทที่เหมาะสมของการใช้ ChatGPT ในกระบวนการแปลเชิงวิชาการ ซึ่งนำเสนอในลักษณะของประเด็นร่วม (Common Themes) เพื่อเสริมความเข้าใจต่อผลเชิงปริมาณ โดยมีได้ใช้ข้อมูลเชิงคุณภาพเพื่อสนับสนุนหรือหักล้างผลการทดสอบทางสถิติ

5.1.1 ผลการประเมินคุณภาพการแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษตามเกณฑ์

MQM

ผลการประเมินคุณภาพการแปลตามเกณฑ์ MQM พบว่า การแปลโดย ChatGPT มีคะแนนเฉลี่ยสูงกว่าการแปลโดยมนุษย์ในทุกองค์ประกอบ ได้แก่ Accuracy (ความถูกต้องของ ความหมาย) Fluency (ความคล่องแคล่วทางภาษา) และ Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) รวมถึงคะแนนเฉลี่ยรวมทุกด้าน โดยคะแนนเฉลี่ยรวมของ ChatGPT อยู่ในระดับดีเยี่ยม ขณะที่คะแนนเฉลี่ยรวมของมนุษย์อยู่ในระดับดี ตามเกณฑ์การแปลผลคะแนน แบบมาตราส่วน 5 ระดับ

เมื่อพิจารณาเชิงลักษณะคะแนน (จากค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน) พบว่า ChatGPT มีค่า σ ต่ำกว่าในทุกด้าน สะท้อนความสม่ำเสมอของคุณภาพการแปลมากกว่า ขณะที่ การแปลโดยมนุษย์มีค่า σ สูงกว่า แสดงถึง ความแปรปรวนของคุณภาพระหว่างฉบับแปลที่มากกว่า นอกจากนี้ เมื่อจัดลำดับค่าเฉลี่ยรายด้าน พบว่า ChatGPT มีค่าเฉลี่ยสูงสุดในด้าน Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) และมีค่าเฉลี่ยต่ำสุดในด้าน Fluency (ความคล่องแคล่วทางภาษา) เช่นเดียวกับแนวโน้มของการแปลโดยมนุษย์ ที่คะแนนต่ำสุด อยู่ในด้าน Fluency (ความคล่องแคล่วทางภาษา)

5.1.2 ผลการเปรียบเทียบคุณภาพการแปลระหว่างการแปลโดยมนุษย์และการแปล โดย ChatGPT

ผลการทดสอบความแตกต่างของคะแนนคุณภาพการแปลระหว่างการแปลโดยมนุษย์ และการแปลโดย ChatGPT ด้วยสถิติ Wilcoxon Signed - Rank Test พบว่า คะแนนคุณภาพการแปล แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ในทุกองค์ประกอบของเกณฑ์ MQM และคะแนน รวมเฉลี่ย ($p = .002$)

เมื่อพิจารณาทิศทางของความแตกต่างจากค่าอันดับของผลต่างคะแนน (Ranks) พบว่าผลอยู่ในกลุ่ม Negative Ranks ตามคำอธิบายใต้ตาราง (Human < ChatGPT) แสดงว่า ChatGPT ได้คะแนนสูงกว่ามนุษย์อย่างสม่ำเสมอในทุกคู่การเปรียบเทียบ ภายใต้กรอบ MQM

ทั้งนี้ การเปรียบเทียบในส่วนนี้เป็นการสรุปผลในระดับ “องค์ประกอบของเกณฑ์ MQM” และ “คะแนนรวม” ของข้อมูลคู่สัมพันธ์ มิได้มุ่งชี้ผลต่างเชิงรายการณ์ของบทคัดย่อแต่ละคู่ เป็นรายชิ้นงาน

5.1.3 ผลการสังเคราะห์ความคิดเห็นของผู้เชี่ยวชาญเกี่ยวกับจุดเด่น ข้อจำกัด และแนวทางการพัฒนาคุณภาพการแปล

ผลการวิเคราะห์เชิงคุณภาพจากคำถามปลายเปิดสะท้อนว่า ผู้เชี่ยวชาญให้ข้อสังเกตเชิงร่วมใน 3 ประเด็นหลัก ได้แก่

- 1) จุดเด่นของ ChatGPT คือการสร้างข้อความภาษาอังกฤษที่มีรูปแบบภาษาเป็นระบบและค่อนข้างสม่ำเสมอ รวมถึงความสอดคล้องของศัพท์เฉพาะ
- 2) ข้อจำกัดของ ChatGPT คือความเสี่ยงต่อความคลาดเคลื่อนเชิงความหมายในบางบริบท และความจำเป็นต้องตรวจทานรูปแบบ/ความเหมาะสมตามข้อกำหนดของสถาบัน
- 3) บทบาทของมนุษย์ ยังคงมีความสำคัญในขั้นตอนตรวจทานและปรับแก้ (Post - Editing) เพื่อควบคุมคุณภาพเชิงความหมาย ความเหมาะสมของสำนวน และความสอดคล้องกับรูปแบบการเขียนบทคัดย่อเชิงวิชาการ

5.2 อภิปรายผลการวิจัย

จากการศึกษาวิจัยเรื่อง การเปรียบเทียบคุณภาพการแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษระหว่างการแปลโดยมนุษย์และการแปลโดย ChatGPT ตามเกณฑ์ Multidimensional Quality Metrics (MQM) สามารถอภิปรายผลการวิจัยตามวัตถุประสงค์ที่กำหนดไว้ได้ ดังนี้

5.2.1 ผลการประเมินคุณภาพการแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษตามเกณฑ์ MQM

จากผลการประเมินคุณภาพการแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษตามเกณฑ์ Multidimensional Quality Metrics (MQM) พบว่า การแปลโดย ChatGPT ได้คะแนนเฉลี่ยสูงกว่าการแปลโดยมนุษย์ในทุกองค์ประกอบ ได้แก่ Accuracy (ความถูกต้องของความหมาย) Fluency (ความคล่องแคล่วทางภาษา) และ Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) สะท้อนว่าในชุดข้อมูลบทคัดย่อที่ศึกษา ChatGPT สามารถสร้างข้อความภาษาอังกฤษที่คงสาระสำคัญของต้นฉบับได้ครบถ้วน มีความสละสลวยของภาษา และมีความสม่ำเสมอในการใช้ศัพท์เฉพาะในภาพรวม ทั้งนี้อาจเป็นเพราะ บทคัดย่อวิทยานิพนธ์มีโครงสร้างการเขียนที่ค่อนข้างเป็นแบบแผนและใช้รูปแบบภาษาเชิงวิชาการซ้ำรูป (Patterned Academic Language) ทำให้โมเดลภาษาอย่าง ChatGPT สามารถสร้างประโยคภาษาอังกฤษที่มีโครงสร้างทางไวยากรณ์ชัดเจนและมีความต่อเนื่องของถ้อยคำได้ดี นอกจากนี้ เกณฑ์การประเมินที่ใช้ในงานวิจัยนี้เน้นองค์ประกอบหลัก 3 ด้าน ได้แก่ Accuracy, Fluency และ Terminology Consistency ซึ่งเป็นมิติของคุณภาพภาษาในระดับโครงสร้างและการใช้คำศัพท์ที่ระบบปัญญาประดิษฐ์มักสามารถประมวลผลได้อย่างมีประสิทธิภาพ โดยเฉพาะเมื่อข้อความต้นฉบับมีรูปแบบและบริบททางวิชาการที่ชัดเจน ผลการวิจัยดังกล่าว สอดคล้องกับ Wang (2024) ที่พบว่า ChatGPT เมื่อประเมินด้วยเกณฑ์ MQM มีแนวโน้มให้คุณภาพด้าน Terminology

Consistency และ Grammatical Accuracy อยู่ในระดับสูง โดยเฉพาะเมื่อใช้เป็นเครื่องมือช่วยปรับปรุงร่างงานแปล นอกจากนี้ยังสอดคล้องกับแนวคิดเกี่ยวกับ Post-editing ที่เสนอว่า AI สามารถช่วยสร้างร่างงานแปลที่มีโครงสร้างภาษาและความลื่นไหลของข้อความได้ดี ขณะที่มนุษย์ทำหน้าที่ตรวจทานและปรับแก้เพื่อให้ได้คุณภาพงานแปลที่พร้อมเผยแพร่ (Koponen, 2016; Vieira, 2019)

5.2.2 ผลการเปรียบเทียบคุณภาพการแปลระหว่างการแปลโดยมนุษย์และการแปลโดย ChatGPT

จากผลการทดสอบความแตกต่างของคะแนนคุณภาพการแปลด้วยสถิติ Wilcoxon Signed-Rank Test พบว่า คะแนนคุณภาพการแปลของ ChatGPT สูงกว่าการแปลโดยมนุษย์อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ในทุกองค์ประกอบของเกณฑ์ Multidimensional Quality Metrics (MQM) ได้แก่ Accuracy (ความถูกต้องของความหมาย), Fluency (ความคล่องแคล่วทางภาษา) และ Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) แสดงให้เห็นว่าความแตกต่างของคุณภาพการแปลไม่ได้เกิดขึ้นเฉพาะบางด้าน แต่เกิดขึ้นอย่างสอดคล้องกันในทุกองค์ประกอบของกรอบการประเมินคุณภาพการแปล ทั้งนี้อาจเป็นเพราะ ChatGPT เป็นโมเดลภาษาประเภท Generative AI ที่ได้รับการฝึกจากข้อมูลขนาดใหญ่ ทำให้มีความสามารถในการสร้างประโยคภาษาอังกฤษที่มีโครงสร้างทางไวยากรณ์ชัดเจนและมีความลื่นไหลของภาษา นอกจากนี้การออกแบบคำสั่ง (Prompt Engineering) ที่กำหนดบทบาทและบริบทของการแปลอย่างชัดเจนยังสามารถช่วยยกระดับความถูกต้องของความหมาย ความลื่นไหลของภาษา และความสม่ำเสมอของการใช้ศัพท์เฉพาะได้ อีกทั้งข้อผิดพลาดของมนุษย์บางประเภท เช่น Mother Tongue Interference (การแทรกแซงของภาษาแม่) หรือความสับสนในกฎไวยากรณ์ของภาษาอังกฤษ อาจทำให้คุณภาพของงานแปลมีความแปรปรวนระหว่างผู้แปลแต่ละคนได้ (Wongranu, 2017; Duklim, 2022) ผลการวิจัยดังกล่าว สอดคล้องกับ Freitag et al. (2021) ที่เสนอว่าเกณฑ์ MQM สามารถให้เห็นความแตกต่างของคุณภาพการแปลในแต่ละมิติได้อย่างชัดเจนมากกว่าการใช้คะแนนรวมเพียงค่าเดียว และยังสอดคล้องกับ Park and Padó (2024) ที่ระบุว่า MQM ช่วยเพิ่มความสามารถในการตีความผลลัพธ์ของการประเมินคุณภาพการแปล เนื่องจากสามารถแยกองค์ประกอบของคุณภาพออกจากกันได้อย่างเป็นระบบ อย่างไรก็ตาม ผลการวิจัยนี้แตกต่างจาก Tongpooon-Patanasorn and Griffith (2020) ที่พบว่า Google Translate ในการแปลบทคัดย่อไทย-อังกฤษยังมีข้อผิดพลาดด้าน Fluency เช่น Punctuation และ Grammar อยู่จำนวนมาก ซึ่งความแตกต่างดังกล่าวอาจสะท้อนถึงพัฒนาการของเทคโนโลยีการแปลอัตโนมัติที่มีความก้าวหน้าขึ้นในช่วงเวลาที่แตกต่างกัน

5.2.3 ผลการสังเคราะห์ความคิดเห็นของผู้เชี่ยวชาญเกี่ยวกับการแปลบทคัดย่อโดยมนุษย์และการแปลโดย ChatGPT

ผลการสังเคราะห์ความคิดเห็นของผู้เชี่ยวชาญ สะท้อนว่า การแปลบทคัดย่อโดยมนุษย์และ ChatGPT มีลักษณะจุดเด่นที่แตกต่างกัน โดยมนุษย์มีความได้เปรียบด้านการตีความบริบทของเนื้อหาและความเหมาะสมเชิงวัฒนธรรม (Contextual and Cultural Appropriateness)

ขณะที่ ChatGPT มีความโดดเด่นด้านความต่อเนื่องของภาษา การรักษาโครงสร้างภาษาเชิงวิชาการ และความสม่ำเสมอของการใช้ศัพท์เฉพาะ ซึ่งสัมพันธ์กับมิติของ Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) และ Fluency (ความคล่องแคล่วทางภาษา) ทั้งนี้อาจเป็นเพราะระบบปัญญาประดิษฐ์ประเภทโมเดลภาษา (Large Language Models) ถูกพัฒนาจากการเรียนรู้ข้อมูลข้อความจำนวนมาก จึงมีความสามารถในการสร้างประโยคที่มีความต่อเนื่องทางภาษาและรักษารูปแบบการใช้ศัพท์ได้อย่างสม่ำเสมอ ขณะที่มนุษย์มีศักยภาพในการตีความความหมายเชิงบริบท ความหมายแฝง และความเหมาะสมทางวัฒนธรรมได้ดีกว่า ซึ่งเป็นองค์ประกอบสำคัญของงานแปลเชิงวิชาการ ผลการวิเคราะห์ดังกล่าวสอดคล้องกับแนวคิดการทำงานร่วมกันระหว่างมนุษย์และปัญญาประดิษฐ์ (Human-AI Collaboration) ที่เสนอให้ใช้ระบบ AI ในการสร้างร่างงานแปลเบื้องต้น และให้มนุษย์ทำหน้าที่ตรวจแก้ภายหลัง (Post-editing) เพื่อเพิ่มความถูกต้องและความเหมาะสมของภาษา (Herbig et al., 2019) อีกทั้งยังสอดคล้องกับแนวคิดเชิงหน้าที่ของการแปล (Functionalist Perspective) ซึ่งชี้ว่างานแปลควรพิจารณาตามวัตถุประสงค์ของการสื่อสารและกลุ่มผู้อ่าน โดยผู้แปลยังคงมีบทบาทสำคัญในการปรับระดับภาษาและรูปแบบการนำเสนอให้เหมาะสมกับบริบทของงานวิชาการ (Colina et al., 2017) อย่างไรก็ตาม ผู้เชี่ยวชาญบางส่วนเห็นว่า ระบบ AI ยังมีข้อจำกัดในการตีความบริบทเฉพาะทาง และอาจเกิดการขยายความหรือใช้คำที่คลาดเคลื่อนจากต้นฉบับในบางกรณี ซึ่งสอดคล้องกับงานวิจัยของ Altakhaine et al. (2025) และ Lau et al. (2024) ที่พบว่าเครื่องมือแปลอัตโนมัติยังมีความอ่อนไหวต่อคำศัพท์เฉพาะและบริบททางวัฒนธรรม โดยเฉพาะในเนื้อหาทางวิชาการหรือเนื้อหาเฉพาะสาขา ดังนั้น แนวทางที่เหมาะสมในการประยุกต์ใช้เทคโนโลยีปัญญาประดิษฐ์ในการแปลบทคัดย่อเชิงวิชาการ คือการใช้ AI เพื่อช่วยสร้างร่างการแปลในขั้นต้น และให้มนุษย์ทำหน้าที่ตรวจแก้ ปรับปรุง และตรวจสอบความถูกต้องของเนื้อหาและรูปแบบภาษา ก่อนการเผยแพร่ผลงานวิชาการ เพื่อให้ได้คุณภาพของงานแปลที่มีความถูกต้อง เหมาะสม และสอดคล้องกับมาตรฐานทางวิชาการ (Koponen, 2016; Vieira, 2019)

5.3 ข้อเสนอแนะในการวิจัย

5.3.1 ข้อเสนอแนะสำหรับการนำผลการวิจัยไปใช้

1) การนำ ChatGPT มาใช้เป็นเครื่องมือสนับสนุนการแปลบทคัดย่อเชิงวิชาการ ซึ่งจากผลการวิจัย หน่วยงานและผู้ใช้งานสามารถนำ ChatGPT มาใช้ในบทบาทของ เครื่องมือสนับสนุนการแปล (Translation Support Tool) โดยกำหนดให้ ChatGPT ทำหน้าที่สร้างร่างคำแปลเบื้องต้น (Draft Translation) เพื่อช่วยจัดโครงสร้างภาษาและเพิ่มความคล่องแคล่วของสำนวนภาษาอังกฤษเชิงวิชาการ ขณะที่มนุษย์ยังคงมีบทบาทสำคัญในการตรวจสอบและควบคุมคุณภาพของบทแปลก่อนนำไปใช้งานจริง โดยในการนำไปใช้เชิงปฏิบัติ สามารถกำหนดแนวทางดำเนินงานได้ดังนี้

(1) ควรกำหนดขอบเขตการใช้ให้ชัดเจนว่า ChatGPT ใช้เพื่อช่วยร่างภาษามิใช่ใช้แทนการตัดสินใจเชิงความหมาย โดยเฉพาะส่วนที่เกี่ยวข้องกับวัตถุประสงค์ วิธีดำเนินการ และผลการวิจัย ซึ่งต้องถ่ายทอดสาระสำคัญให้ตรงกับต้นฉบับภาษาไทยมากที่สุด

(2) ควรกำหนดรูปแบบผลลัพธ์ที่ต้องการ เช่น การใช้ภาษาอังกฤษเชิงวิชาการ การจัดโครงสร้างบทคัดย่อให้ครบองค์ประกอบ และการหลีกเลี่ยงการขยายความหรือเพิ่มเติมเนื้อหาที่ไม่มีในต้นฉบับ

(3) ควรกำหนดให้มีการตรวจทานโดยมนุษย์เป็นขั้นตอนมาตรฐาน โดยตรวจทั้งด้านความถูกต้องของความหมาย (Accuracy) ความคล่องแคล่วทางภาษา (Fluency) และความสอดคล้องของศัพท์เฉพาะ (Terminology Consistency) ก่อนนำบทคัดย่อไปใช้ในการสอบหรือการเผยแพร่ผลงาน

แนวทางดังกล่าวช่วยให้การใช้ ChatGPT สนับสนุนงานแปลเชิงวิชาการเป็นไปอย่างเหมาะสม ลดความเสี่ยงด้านความคลาดเคลื่อนของเนื้อหา และยังคงรักษามาตรฐานทางวิชาการของบทคัดย่อในระดับบัณฑิตศึกษา

2) การใช้เกณฑ์ Multidimensional Quality Metrics (MQM) เป็นแนวทางตรวจสอบคุณภาพบทคัดย่อภาษาอังกฤษ ซึ่งจากผลการวิจัย สามารถนำเกณฑ์ Multidimensional Quality Metrics (MQM) ไปใช้เป็นกรอบมาตรฐานในการตรวจสอบคุณภาพบทคัดย่อภาษาอังกฤษในระดับหน่วยงานได้อย่างเป็นรูปธรรม โดยเฉพาะการใช้เป็นแนวทางกลางในการพิจารณาคุณภาพงานแปลที่มีความชัดเจนและตรวจสอบได้ สำหรับการนำไปใช้โดยหน่วยงานสามารถจัดทำรายการตรวจ (Checklist) ที่ครอบคลุมองค์ประกอบหลักทั้งสามด้าน ได้แก่ Accuracy (ความถูกต้องของความหมาย) Fluency (ความคล่องแคล่วทางภาษา) และ Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) เพื่อใช้ประกอบการตรวจบทคัดย่อก่อนการสอบหรือก่อนการเผยแพร่ผลงาน ทั้งนี้ การใช้ MQM ช่วยลดความคลาดเคลื่อนจากดุลยพินิจส่วนบุคคล ทำให้การตรวจมีมาตรฐานเดียวกัน และสามารถชี้จุดที่ต้องปรับปรุงได้อย่างชัดเจน นอกจากนี้ ผลการตรวจตามเกณฑ์ MQM ยังสามารถนำมาใช้เป็นข้อมูลย้อนกลับเพื่อพัฒนาทักษะการเขียนและการแปลของนักศึกษา โดยช่วยให้นักศึกษามองเห็นข้อบกพร่องของบทคัดย่อในเชิงโครงสร้างภาษา ความหมาย และการใช้ศัพท์เฉพาะอย่างเป็นระบบ

3) การพัฒนาทักษะการออกแบบคำสั่ง (Prompt) สำหรับการใช้ ChatGPT ในงานแปล ซึ่งผลการวิจัยสะท้อนให้เห็นว่า คุณภาพของบทแปลที่ได้จาก ChatGPT มีความสัมพันธ์กับความชัดเจนของคำสั่งที่ผู้ใช้กำหนด ดังนั้น ในการนำผลการวิจัยไปใช้ ควรส่งเสริมให้ผู้ใช้งานมีความรู้และทักษะในการออกแบบคำสั่ง (Prompt) ที่เหมาะสมกับงานแปลเชิงวิชาการ ดังนั้นในเชิงปฏิบัติ คำสั่งควรระบุบทบาทของ ChatGPT ให้ชัดเจน ระบุบริบทของงานแปล เช่น สาขาวิชาและกลุ่มผู้อ่าน และระบุเกณฑ์คุณภาพที่ต้องการ เช่น ความถูกต้องของความหมาย ความสอดคล้องของศัพท์เฉพาะ

และระดับภาษาที่เป็นทางการ หน่วยงานสามารถนำผลการวิจัยไปจัดทำตัวอย่างคำสั่งมาตรฐาน หรือจัดกิจกรรมอบรมเพื่อให้ผู้ใช้งานสามารถใช้ ChatGPT ได้อย่างมีประสิทธิภาพและสม่ำเสมอ

5.3.2 ข้อเสนอแนะสำหรับการวิจัยครั้งต่อไป

แม้งานวิจัยฉบับนี้จะบรรลุวัตถุประสงค์ที่กำหนดไว้ในการประเมินและเปรียบเทียบคุณภาพการแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษระหว่างการแปลโดยมนุษย์และการแปลโดย ChatGPT ภายใต้กรอบ Multidimensional Quality Metrics (MQM) อย่างเป็นระบบแล้ว อย่างไรก็ตาม เพื่อขยายขอบเขตองค์ความรู้และเพิ่มความหลากหลายของมุมมองในการศึกษาครั้งต่อไป สามารถนำเสนอแนวทางการวิจัยเพิ่มเติมเป็นประเด็นสำคัญดังต่อไปนี้

1) การขยายประเภทของข้อความเชิงวิชาการที่ใช้ในการศึกษา โดยการวิจัยครั้งต่อไปอาจขยายขอบเขตการศึกษาไปยังข้อความเชิงวิชาการประเภทอื่นนอกเหนือจากบทคัดย่อ เช่น บทนำ (Introduction) วิธีดำเนินการวิจัย (Methodology) หรืออภิปรายผล (Discussion) เพื่อเปรียบเทียบคุณภาพการแปลในข้อความที่มีระดับความซับซ้อน โครงสร้าง และหน้าที่ของข้อความแตกต่างกัน ซึ่งจะช่วยให้เห็นลักษณะข้อจำกัดและศักยภาพของการแปลในบริบทที่หลากหลายมากขึ้น

2) การเปรียบเทียบรูปแบบการออกแบบคำสั่ง (Prompt) ที่แตกต่างกัน โดยการศึกษาครั้งต่อไปอาจมุ่งเปรียบเทียบผลของการใช้คำสั่งในรูปแบบที่หลากหลาย เช่น คำสั่งทั่วไป คำสั่งที่มีการกำหนดบทบาทของผู้แปล หรือคำสั่งที่ระบุบริบทและเกณฑ์คุณภาพอย่างชัดเจน เพื่อวิเคราะห์ว่าลักษณะของคำสั่งมีผลต่อคุณภาพการแปลตามเกณฑ์ MQM ในแต่ละด้านมากน้อยเพียงใด

3) การศึกษากระบวนการตรวจแก้หลังการแปล (Post - Editing) อย่างเป็นระบบ ซึ่งการวิจัยในอนาคตอาจศึกษากระบวนการตรวจแก้บทแปลที่ได้จาก ChatGPT โดยมนุษย์ในเชิงปฏิบัติการ เช่น ระยะเวลาในการตรวจแก้ ประเภทของข้อผิดพลาดที่ถูกแก้ไข และระดับคุณภาพของบทแปลหลังการปรับปรุงตามเกณฑ์ MQM เพื่อพัฒนาแนวทางการทำงานร่วมกันระหว่างมนุษย์และ AI ให้มีประสิทธิภาพยิ่งขึ้น

4) การประยุกต์ใช้การแปลด้วย AI ในบริบทการเรียนการสอนระดับบัณฑิตศึกษา ซึ่งการวิจัยครั้งต่อไปอาจนำ ChatGPT มาใช้เป็นเครื่องมือประกอบการเรียนการสอนด้านการเขียนและการแปลเชิงวิชาการ และศึกษาผลต่อการพัฒนาทักษะของผู้เรียน โดยใช้เกณฑ์ MQM เป็นเครื่องมือประเมินคุณภาพ เพื่อสะท้อนผลการเรียนรู้ด้านการใช้ภาษาและความตระหนักรู้ด้านคุณภาพงานแปลในเชิงวิชาการ

5.4 สรุปประโยชน์ที่ได้จากการวิจัย

การวิจัยเรื่อง การเปรียบเทียบคุณภาพการแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษ ระหว่างการแปลโดยมนุษย์และการแปลโดย ChatGPT ตามเกณฑ์ Multidimensional Quality Metrics (MQM) ก่อให้เกิดประโยชน์ที่สำคัญ โดยสอดคล้องกับวัตถุประสงค์ของการวิจัย ดังนี้

5.4.1 ประโยชน์ด้านองค์ความรู้เกี่ยวกับการประเมินคุณภาพการแปล โดยผลการวิจัย ช่วยสร้างองค์ความรู้เกี่ยวกับการประเมินคุณภาพการแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษตาม เกณฑ์ Multidimensional Quality Metrics (MQM) ซึ่งประกอบด้วย Accuracy, Fluency และ Terminology Consistency ทำให้เกิดความเข้าใจเชิงระบบเกี่ยวกับลักษณะข้อผิดพลาด จุดเด่น และระดับคุณภาพของการแปลในแต่ละมิติ อันสามารถใช้เป็นพื้นฐานในการศึกษาหรือพัฒนางานวิจัย ด้านการแปลเชิงวิชาการต่อไป

5.4.2 ประโยชน์ด้านข้อมูลเชิงประจักษ์เกี่ยวกับบทบาทของปัญญาประดิษฐ์ในงานแปล โดย ผลการเปรียบเทียบคุณภาพการแปลระหว่างมนุษย์และ ChatGPT ซึ่งวิเคราะห์ด้วยสถิติ Wilcoxon Signed-Rank Test ช่วยให้ได้ข้อมูลเชิงประจักษ์เกี่ยวกับศักยภาพและข้อจำกัดของการใช้เทคโนโลยี ปัญญาประดิษฐ์ในงานแปลเชิงวิชาการ ข้อมูลดังกล่าวสามารถใช้ประกอบการพิจารณาการนำเครื่องมือ AI มาใช้สนับสนุนการจัดทำบทคัดย่อภาษาอังกฤษในระดับบัณฑิตศึกษาได้อย่างเหมาะสม

5.4.3 ประโยชน์ด้านแนวทางการประยุกต์ใช้ ChatGPT ในการแปลบทคัดย่อเชิงวิชาการ โดย ผลการสังเคราะห์ความคิดเห็นของผู้เชี่ยวชาญช่วยเสนอแนวทางการใช้ ChatGPT ในกระบวนการ แปลบทคัดย่อเชิงวิชาการ โดยใช้ AI เป็นเครื่องมือช่วยสร้างร่างการแปลเบื้องต้น และให้มนุษย์ทำหน้าที่ ตรวจสอบแก้ (Post - Editing) เพื่อควบคุมความถูกต้องของความหมาย ความเหมาะสมของภาษา และความสอดคล้องของศัพท์เฉพาะ ซึ่งสามารถนำไปประยุกต์ใช้ในการพัฒนาคุณภาพบทคัดย่อ ภาษาอังกฤษของนักศึกษาระดับบัณฑิตศึกษาได้



ภาพที่ 5.1 แสดงแผนภาพสรุปประโยชน์จากการวิจัยเรื่องการประเมินคุณภาพการแปลบทคัดย่อภาษาไทย – ภาษาอังกฤษด้วย MQM

ที่มา : ผู้วิจัยสังเคราะห์จากผลการวิจัย

โดยสรุป งานวิจัยฉบับนี้แสดงให้เห็นว่า การใช้ปัญญาประดิษฐ์ โดยเฉพาะ ChatGPT ในการแปลบทคัดย่อเชิงวิชาการ สามารถช่วยสนับสนุนกระบวนการจัดทำบทคัดย่อภาษาอังกฤษได้อย่างมีประสิทธิภาพ หากมีการกำหนดบทบาทการใช้งานที่ชัดเจนและมีการตรวจทานโดยมนุษย์ควบคู่กัน ผลการวิจัยช่วยให้การประเมินคุณภาพงานแปลมีกรอบที่เป็นระบบมากขึ้น ผ่านเกณฑ์ Multidimensional Quality Metrics (MQM) และช่วยสะท้อนแนวทางการทำงานร่วมกันระหว่างมนุษย์และ AI ที่เหมาะสมกับบริบทระดับบัณฑิตศึกษา อันสามารถนำไปต่อยอดสู่การพัฒนามาตรฐานงานวิชาการและการประกันคุณภาพด้านภาษาอังกฤษของมหาวิทยาลัยได้อย่างเป็นรูปธรรม

บรรณานุกรม

- กัญญาภัค สาสงเคราะห์. (2562). การศึกษาเชิงเปรียบเทียบการแปลเว็บไซต์โรงแรมระหว่างมนุษย์และเครื่องมือแปลภาษา Google Translate จากภาษาไทยเป็นภาษาอังกฤษ. มหาวิทยาลัยธรรมศาสตร์.
- ธัญญภัคภรณ์ สารชาติ. (2566). การวิเคราะห์ข้อผิดพลาดที่พบในบทคัดย่อภาษาอังกฤษของนักศึกษา ระดับปริญญาโท มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี. สืบค้นจาก <https://online.anyflip.com/tpjzm/hvxs/mobile/index.html>
- นรินทร์ทิพย์ ทองศรี. (2567). การวิเคราะห์เปรียบเทียบการแปลภาษาอังกฤษเป็นภาษาไทยระหว่างผู้แปลที่เป็นมนุษย์และโปรแกรมแปลภาษา (Artificial Intelligence: AI ChatGPT). *วารสารรวมคำแห่ง ฉบับมนุษยศาสตร์*, 43(2), 99–122. สืบค้นจาก <https://so05.tci-thaijo.org/index.php/huru/article/view/277709>
- พิสิฐ ตัณฑวณิช, และ พนา จินดาศรี. (2561). ดัชนีความสอดคล้องระหว่างข้อความกับวัตถุประสงค์ (IOC) และการนำไปใช้ในการวิจัยทางสังคมศาสตร์. สืบค้นจาก <https://so03.tci-thaijo.org/index.php/jiqs/article/view/162831>
- พุทธชาติ ลิ้มศิริเรืองโร. (2565). ปัญหาและข้อผิดพลาดในการแปลจากภาษาไทยเป็นภาษาอังกฤษของนักศึกษาสาขาวิชาภาษาอังกฤษธุรกิจชั้นปีที่ 2 มหาวิทยาลัยราชภัฏบุรีรัมย์. *วารสารมนุษยศาสตร์ มหาวิทยาลัยนครสวรรค์*, 19(2), 95–108. สืบค้นจาก <https://so03.tci-thaijo.org/index.php/jhnu/article/view/244154>
- มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี สำนักบัณฑิตศึกษา . (2565). รายงานจำนวนบทคัดย่อที่แก้ไขของนักศึกษาระดับปริญญาโท ปีการศึกษา 2565. สำนักบัณฑิตศึกษา.
- _____. (2567). รายงานจำนวนบทคัดย่อที่แก้ไขของนักศึกษาระดับปริญญาโท ปีการศึกษา 2567. สำนักบัณฑิตศึกษา.
- มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี. (2567). คำสั่งมหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี ที่ 1204/2567 เรื่อง แต่งตั้งคณะกรรมการกลั่นกรองบทคัดย่อและบทความวิจัยเพื่อสำเร็จการศึกษาระดับบัณฑิตศึกษา สั่ง ณ วันที่ 23 ธันวาคม พ.ศ. 2567.
- สภาวิชาการมหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี. (2559). รายงานการประชุมสภาวิชาการ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี ครั้งที่ 4 วันที่ 7 เมษายน 2559. สำนักส่งเสริมวิชาการและงานทะเบียน.
- Alspach, G. J. (2017). Writing for publication 101: Why the abstract is so important. *Critical Care Nurse*, 37(4), 12–15. <https://doi.org/10.4037/ccn2017466>

บรรณานุกรม (ต่อ)

- Altakhaineh, A. R. K., Alghathian, R., & Jarrah, M. A. (2025). Comparing ChatGPT and human translation of legal texts from English into Arabic: A case study. *International Journal of Language & Law*, 14(1), 39–64. Retrieved from <https://www.languageandlaw.eu/jll/article/view/190>
- Boonraksa, T., & Naisena, S. (2022). A study on English collocation errors of Thai EFL students. *English Language Teaching*, 15(1), 164–172. <https://doi.org/10.5539/elt.v15n1p164>
- Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, 1(3), 185–216. <https://doi.org/10.1177/135910457000100301>
- Cabezas-García, M. (2022). *Terminology management in specialized translation*. Retrieved from https://www.researchgate.net/publication/359579450_Terminology_Management_in_Specialized_Translation
- Chuanpipatpong, A. (2025). Thai EFL university students' writing in the digital age: Error analysis revisited. *PASAA: Journal of Language Teaching and Learning in Thailand*, 70, 308–339. Retrieved from <https://files.eric.ed.gov/fulltext/EJ1478073.pdf>
- Duklim, B. (2022). Translation errors made by Thai university students: A study on types and causes. *rEFLections*, 29(2), 344–360. <https://doi.org/10.61508/refl.v29i2.259858>
- Freitag, M., Foster, G., Grangier, D., Ratnakar, V., Tan, Q., & Macherey, W. (2021). Experts, errors, and context: A large-scale study of human evaluation for machine translation. *Transactions of the Association for Computational Linguistics*, 9, 1460–1474. https://doi.org/10.1162/tacl_a_00437
- Gao, Y., Wang, R., & Hou, F. (2023). How to design translation prompts for ChatGPT: An empirical study. *arXiv*. <https://doi.org/10.48550/arXiv.2304.02182>

บรรณานุกรม (ต่อ)

- He, S. (2024). Prompting ChatGPT for translation: A comparative analysis of translation brief and persona prompts. In C. Scarton et al. (Eds.), *Proceedings of the 25th Annual Conference of the European Association for Machine Translation* (Vol. 1, pp. 316–326). European Association for Machine Translation. Retrieved from <https://aclanthology.org/2024.eamt-1.27/>
- Herbig, N., Pal, S., van Genabith, J., & Krüger, A. (2019). Integrating artificial and human intelligence for efficient translation. *arXiv*. Retrieved from <https://arxiv.org/abs/1903.02978>
- Human Science Co., Ltd. (2024). How to use ChatGPT for English translation? Retrieved from <https://www.science.co.jp/en/nmt/blog/37037/>
- Işim, Ç., & Balcioğlu, Y. S. (2023). ChatGPT: Performance of translation. In *Proceedings of the 3rd International ACHARAKA Congress on Humanities and Social Sciences* (pp. 47–51).
- Jeharsae, F., & Boonsuk, Y. (2024). “I’ll find it... or I’ll look it up...?”: Decoding Thai EFL students’ translation errors in the Thai–English language pair in multi-linguacultural higher education. *rEFLECTIONS*, 31(2), 670–691. <https://doi.org/10.61508/refl.v31i2.275131>
- Kampookaew, P. (2020). An analysis of grammatical errors made by Thai EFL university students in an EAP writing class: Issues and recommendations. *rEFLECTIONS*, 27(2), 246–273. <https://doi.org/10.61508/refl.v27i2.248862>
- Khoshafah, F. (2023). ChatGPT for Arabic-English translation: Evaluating the accuracy. *Preprint*. <https://doi.org/10.21203/rs.3.rs-2814154/v1>
- Koponen, M. (2016). Is machine translation post-editing worth the effort? A survey of research into post-editing and effort. *The Journal of Specialised Translation*, (25), 131–148.
- Lau, Y. L., Chee, S. P., Chua, R. H. H., Yong, Z. H., Yong, I. K., Tan, J. C., Yong, H. W., & Bakar, A. L. A. (2024). The datasets of human and AI translation. *Journal of Open Humanities Data*, 10(45), 1–10. <https://doi.org/10.5334/johd.212>

บรรณานุกรม (ต่อ)

- Lommel, A. R., Burchardt, A., & Uszkoreit, H. (2014). Multidimensional quality metrics (MQM): A framework for declaring and describing translation quality metrics. *Tradumàtica: Tecnologies de la Traducció*. <https://doi.org/10.5565/rev/tradumatica.77>
- Lyons, S. (2020). Quality of Thai to English machine translation. *Machine Translation*. <https://doi.org/10.1007/s10590-020-09248-8>
- Mariana, V., Cox, T., & Melby, A. (2025). The Multidimensional Quality Metrics (MQM) framework: A new framework for translation quality assessment. *Journal of Translation Studies*, 12(1), 45–67.
- Melby, A., & Foster, G. (2014). The development and use of MQM: A master error typology for translation quality assessment. *The Journal of Specialised Translation*, 23, 134–150.
- Park, D., & Padó, S. (2024). Multi-dimensional machine translation evaluation: Model evaluation and resource for Korean. *arXiv preprint*. Retrieved from <https://arxiv.org/abs/2403.12666v1>
- Patanasorn, A. T. (2017). Analysis of the quality of translated texts produced by Google Translate from Thai to English. In *Proceedings of the 13th International Conference on Humanities & Social Sciences 2017 (IC-HUSO 2017)* (pp. 237–238). Faculty of Humanities and Social Sciences, Khon Kaen University.
- Pewnim, P. (2020). *Assessing the quality of English-to-Thai machine interpreting: Translating Google Translate* [Unpublished master's special research]. Chulalongkorn University. Retrieved from <https://digital.car.chula.ac.th/cgi/viewcontent.cgi?article=7751&context=chulaetd>
- Pornwiriyaakit, P., & Dandee, W. (2022). Translation problems analysis and application of the grammar-translation method in EFL class. *Journal of Educational Issues*, 8(1), 572–591. <https://doi.org/10.5296/jei.v8i1.19825>
- Stockemer, D., & Wigginton, M. J. (2019). Publishing in English or another language: An inclusive study of scholars' language publication preferences in the natural, social, and interdisciplinary sciences. *Scientometrics*, 118, 645–652. <https://doi.org/10.1007/s11192-018-2987-0>

บรรณานุกรม (ต่อ)

- Tongpoon-Patanasorn, A., & Griffith, P. (2020). Google Translate and translation quality: A case of translating academic abstracts from Thai to English. *Manuscript available on ResearchGate*. Retrieved from <https://www.researchgate.net/publication/344770726>
- Tsang, S., Royse, C. F., & Terkawi, A. S. (2017). Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. *Saudi Journal of Anaesthesia*, 11(Suppl 1), S80–S89. https://doi.org/10.4103/sja.SJA_203_17
- Vieira, L. N. (2019). Post-editing of machine translation. In M. O'Hagan (Ed.), *The Routledge Handbook of Translation and Technology* (pp. 319–336). Routledge.
- Wang, J. (2024). Exploring the potential of ChatGPT-4o in translation quality assessment. *Journal of Theory and Practice in Humanities and Social Sciences*, 1(3), 18–30. Retrieved from <https://woodyinternational.com/index.php/jtphss/article/view/27>
- Wongranu, P. (2017). Errors in translation made by English major students: A study on types and causes. *Kasetsart Journal of Social Sciences*, 38(2), 117–122. <https://doi.org/10.1016/j.kjss.2016.11.003>
- Yamada, M. (2023). Optimizing machine translation through prompt engineering: An investigation into ChatGPT's customizability. In *Proceedings of Machine Translation Summit XIX (Vol. 2: Users Track)* (pp. 195–204). Macau SAR, China. Retrieved from <https://arxiv.org/abs/2308.01391>

ภาคผนวก



ภาคผนวก ก
เครื่องมือที่ใช้ในการวิจัย



ชุดบทความตัวอย่างสำหรับการประเมินคุณภาพการแปล

หมายเหตุเกี่ยวกับภาคผนวก: บทความที่ปรากฏในภาคผนวกนี้ถูกนำเสนอในลักษณะตัดตอนและไม่เปิดเผยแหล่งที่มา เพื่อใช้เป็นตัวอย่างของเครื่องมือในการประเมินคุณภาพการแปลตามเกณฑ์ Multidimensional Quality Metrics (MQM) เท่านั้น ทั้งนี้เพื่อคำนึงถึงจริยธรรมการวิจัยและการคุ้มครองผลงานทางวิชาการของเจ้าของผลงาน การเก็บข้อมูลจริง ผู้ประเมินไม่ได้รับทราบว่าเป็นการแปลโดยมนุษย์หรือโดยระบบปัญญาประดิษฐ์ เพื่อป้องกันอคติในการประเมินคุณภาพการแปล

โครงสร้างมาตรฐานของแต่ละชุด (ใช้เหมือนกันทุกฉบับ)

บทความฉบับที่

ส่วนที่ 1 ต้นฉบับภาษาไทย

[ข้อความบทความภาษาไทย ความยาวประมาณ 150–200 คำ ไม่ระบุชื่อเรื่อง ชื่อผู้วิจัย หรือคณะสังกัด]

ส่วนที่ 2 ฉบับแปลภาษาอังกฤษ (ฉบับที่ A)

[ข้อความบทความภาษาอังกฤษ แปลจากต้นฉบับเดียวกัน ไม่ระบุผู้แปล]

ส่วนที่ 3 ฉบับแปลภาษาอังกฤษ (ฉบับที่ B)

[ข้อความบทความภาษาอังกฤษ แปลจากต้นฉบับเดียวกัน ไม่ระบุผู้แปล]

ชุดบทความทั้ง 10 ฉบับ

บทความฉบับที่ 1 (กลุ่มวิทยาศาสตร์และเทคโนโลยี)

(ใช้โครงสร้างมาตรฐานตามด้านบน)

บทความฉบับที่ 2 (กลุ่มวิทยาศาสตร์และเทคโนโลยี)

(ใช้โครงสร้างมาตรฐานตามด้านบน)

บทความฉบับที่ 3 (กลุ่มวิทยาศาสตร์และเทคโนโลยี)

(ใช้โครงสร้างมาตรฐานตามด้านบน)

บทความฉบับที่ 4 (กลุ่มวิทยาศาสตร์และเทคโนโลยี)

(ใช้โครงสร้างมาตรฐานตามด้านบน)

บทความฉบับที่ 5 (กลุ่มวิทยาศาสตร์และเทคโนโลยี)

(ใช้โครงสร้างมาตรฐานตามด้านบน)

บทความฉบับที่ 6 (กลุ่มมนุษยศาสตร์และสังคมศาสตร์)

(ใช้โครงสร้างมาตรฐานตามด้านบน)

บทความฉบับที่ 7 (กลุ่มมนุษยศาสตร์และสังคมศาสตร์)

(ใช้โครงสร้างมาตรฐานตามด้านบน)

บทความฉบับที่ 8 (กลุ่มมนุษยศาสตร์และสังคมศาสตร์)

(ใช้โครงสร้างมาตรฐานตามด้านบน)

บทความฉบับที่ 9 (กลุ่มมนุษยศาสตร์และสังคมศาสตร์)

(ใช้โครงสร้างมาตรฐานตามด้านบน)

บทความฉบับที่ 10 (กลุ่มมนุษยศาสตร์และสังคมศาสตร์)

(ใช้โครงสร้างมาตรฐานตามด้านบน)

แบบสอบถามประเมินคุณภาพการแปลตามเกณฑ์ MQM

แบบสอบถามโครงการวิจัย (Research Project)

งานประจำสู่งานวิจัย (R2R) มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

ประจำปีงบประมาณ 2568

ชื่อโครงการวิจัย: การประเมินคุณภาพการแปลบทคัดย่อวิทยานิพนธ์จากภาษาไทยเป็นภาษาอังกฤษโดยมนุษย์และ ChatGPT ที่มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

Evaluating the Quality of Human and ChatGPT Thai – to - English Translations of Thesis Abstracts at Rajamangala University of Technology Thanyaburi

วัตถุประสงค์การวิจัย

1. เพื่อประเมินคุณภาพการแปลบทคัดย่อภาษาไทยเป็นภาษาอังกฤษ ตามเกณฑ์ในด้าน Accuracy (ความถูกต้องของความหมาย) Fluency (ความคล่องแคล่วทางภาษา) และ Terminology Consistency (ความสอดคล้องของศัพท์เฉพาะ) ของเกณฑ์การประเมิน Multidimensional Quality Metrics (MQM) โดยพิจารณาความแตกต่างของคุณภาพระหว่างการแปลโดยมนุษย์ และการแปลโดย ChatGPT

2. เพื่อเปรียบเทียบคุณภาพการแปลระหว่างการแปลโดยมนุษย์และการแปลโดย ChatGPT จากคะแนนการประเมินของผู้เชี่ยวชาญด้านภาษาอังกฤษ

3. เพื่อสังเคราะห์ความคิดเห็นของผู้เชี่ยวชาญเกี่ยวกับจุดเด่นและข้อจำกัดของการแปลทั้งสองรูปแบบ โดยใช้ข้อมูลจากคำถามปลายเปิดในแบบสอบถาม

คำชี้แจง : แบบสอบถามฉบับนี้จัดทำขึ้นเพื่อใช้ในการวิจัยเรื่อง การประเมินคุณภาพการแปลบทคัดย่อวิทยานิพนธ์จากภาษาไทยเป็นภาษาอังกฤษ โดยมนุษย์และ ChatGPT ตามเกณฑ์ Multidimensional Quality Metrics (MQM) ผู้ตอบแบบสอบถามจะอ่านบทคัดย่อต้นฉบับภาษาไทย จากนั้นอ่านบทคัดย่อภาษาอังกฤษจำนวน 2 ฉบับที่แปลจากต้นฉบับเดียวกัน และประเมินคุณภาพการแปลตามเกณฑ์ที่กำหนด รวมทั้งให้ข้อเสนอแนะเพิ่มเติมตามความเหมาะสม ข้อมูลที่ได้รับจะถูกเก็บเป็นความลับและใช้เพื่อการวิจัยเท่านั้น

ส่วนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบสอบถาม

1. วุฒิต่างการศึกษา

ระดับปริญญาเอก ระดับปริญญาโท อื่น ๆ (โปรดระบุ)

2. สาขาวิชาที่สำเร็จการศึกษา

.....

3. ประสบการณ์ในการสอนภาษาอังกฤษ (ปี)

น้อยกว่า 1 ปี 1-3 ปี 4-6 ปี 7-10 ปี มากกว่า 10 ปี

4. ประสบการณ์เกี่ยวกับการแปลภาษา (ปี)

- น้อยกว่า 1 ปี 1-3 ปี 4-6 ปี 7-10 ปี มากกว่า 10 ปี

5. ความคุ้นเคยกับการใช้ ChatGPT

- ไม่เคยใช้ เคยใช้บ้างเป็นครั้งคราว ใช้เป็นประจำ

ส่วนที่ 2 คำชี้แจงการประเมินบทคัดย่อ

โปรดประเมินคุณภาพของบทคัดย่อภาษาอังกฤษทั้ง 2 แบบ โดยใช้มาตราส่วนประมาณค่า (Rating Scale) 5 ระดับ สำหรับแต่ละด้านตามเกณฑ์ MQM (Multidimensional Quality Metrics: Accuracy, Fluency, Terminology) ตามเกณฑ์ของ Lommel et al. (2014) หากพบข้อผิดพลาดใด ๆ โปรดทำเครื่องหมาย “√” ในช่องเกณฑ์ย่อยที่ตรงกับข้อผิดพลาดนั้น โดยสามารถเลือกได้มากกว่า 1 ข้อ

MQM Criteria	(บทคัดย่อฉบับที่ 1) แบบแปลที่ 1	เกณฑ์ย่อยที่พบ	(บทคัดย่อฉบับที่ 1) แบบแปลที่ 2	เกณฑ์ย่อยที่พบ
1) Accuracy (ความถูกต้อง)	<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5	<input type="checkbox"/> Mistranslation <input type="checkbox"/> Omission <input type="checkbox"/> Addition <input type="checkbox"/> Untranslated Text	<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5	<input type="checkbox"/> Mistranslation <input type="checkbox"/> Omission <input type="checkbox"/> Addition <input type="checkbox"/> Untranslated Text
2) Fluency (ความคล่องแคล่ว)	<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5	<input type="checkbox"/> Punctuation <input type="checkbox"/> Spelling <input type="checkbox"/> Grammar <input type="checkbox"/> Register	<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5	<input type="checkbox"/> Punctuation <input type="checkbox"/> Spelling <input type="checkbox"/> Grammar <input type="checkbox"/> Register
3) Terminology (ศัพท์เฉพาะ)	<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5	<input type="checkbox"/> Incorrect Term <input type="checkbox"/> Inconsistent Term	<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5	<input type="checkbox"/> Incorrect Term <input type="checkbox"/> Inconsistent Term

ส่วนที่ 4 ช่องแสดงความคิดเห็นรายบทคัดย่อ

ความคิดเห็นเพิ่มเติมเกี่ยวกับการแปลบทคัดย่อแบบที่ 1

.....

ความคิดเห็นเพิ่มเติมเกี่ยวกับการแปลบทคัดย่อแบบที่ 2

.....

ส่วนที่ 5 คำถามปลายเปิด (4 ข้อ)

1. ท่านเห็นความแตกต่างระหว่างการแปลโดยมนุษย์และ ChatGPT อย่างไรในภาพรวม
2. ท่านมีข้อเสนอแนะในการพัฒนาคุณภาพการแปลบทคัดย่อวิทยานิพนธ์ในอนาคตอย่างไร
3. คำสั่งที่ท่านใช้สั่ง (Prompt) ChatGPT ในการแปลภาษาไทยเป็นภาษาอังกฤษมีอะไรบ้าง
4. นอกจาก ChatGPT ท่านมีประสบการณ์ในการใช้ AI ตัวอื่นในการแปลภาษาไทยเป็นภาษาอังกฤษหรือไม่

ภาคผนวก ข

ผลการวิเคราะห์ค่าดัชนีความสอดคล้อง IOC



ตารางสรุปค่าดัชนีความสอดคล้อง IOC
แบบบันทึกข้อผิดพลาดที่พบในบทคัดย่อภาษาอังกฤษ

ผู้เชี่ยวชาญ

- 1) รองศาสตราจารย์ ดร.พิมพ์นภัส ภูมิภิตติพิชญ์
สังกัดคณะศิลปศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
- 2) ผู้ช่วยศาสตราจารย์ ดร.สุรศักดิ์ โจถาวร
สังกัดคณะศิลปศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
- 3) ดร.เบญจวรรณ รุ่งเรืองศุภรัตน์
สังกัดคณะศิลปศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

ลำดับ ที่	รายการประเมิน (องค์ประกอบของ แบบฟอร์ม)	ผู้เชี่ยวชาญ			รวม	แปล ผล (IOC)
		1	2	3		
1	ส่วนที่ 1 ข้อมูลทั่วไปของ ผู้ตอบแบบสอบถาม: ระบุ วุฒิทางการศึกษา (ระดับ ปริญญาเอก/ระดับปริญญา โทและอื่น ๆ)	1	1	1	3.00	ผ่าน
2	ข้อมูลทั่วไปของผู้ตอบ แบบสอบถาม: ระบุ สาขาวิชาที่สำเร็จการศึกษา	1	1	1	3.00	ผ่าน
3	ข้อมูลทั่วไปของผู้ตอบ แบบสอบถาม: ระบุ ประสบการณ์ในการสอน ภาษาอังกฤษ (ปี) (น้อยกว่า 1 ปี/1-3 ปี/4-6 ปี/7-10 ปี/ มากกว่า 10 ปี)	1	1	1	3.00	ผ่าน
4	ข้อมูลทั่วไปของผู้ตอบ แบบสอบถาม: ระบุ ประสบการณ์เกี่ยวกับการ แปลภาษา (ปี) (น้อยกว่า 1 ปี/1-3 ปี/4-6 ปี/7-10 ปี/ มากกว่า 10 ปี)	1	1	1	3.00	ผ่าน

ลำดับ ที่	รายการประเมิน (องค์ประกอบของ แบบฟอร์ม)	ผู้เชี่ยวชาญ 1	ผู้เชี่ยวชาญ 2	ผู้เชี่ยวชาญ 3	รวม	แปล ผล (IOC)
5	ข้อมูลทั่วไปของผู้ตอบ แบบสอบถาม:ระบุ ความคุ้นเคยกับการใช้ ChatGPT (ไม่เคยใช้/เคยใช้ บ้างเป็นครั้งคราว/ใช้เป็น ประจำ)	1	1	1	3.00	ผ่าน
6	ส่วนที่ 2 การประเมิน บทคัดย่อ : เกณฑ์หลัก MQM 3 ด้าน ประกอบด้วย 1) ความถูกต้อง (Accuracy 2) ความคล่องแคล่ว (Fluency) 3) ความสอดคล้องของศัพท์ เฉพาะ (Terminology Consistency)	1	1	1	3.00	ผ่าน
7	ส่วนที่ 2 การประเมิน บทคัดย่อ : เกณฑ์ย่อย MQM ประกอบไปด้วย 1) ความถูกต้อง (Accuracy) เกณฑ์ย่อย ได้แก่ 1.1 การแปลผิด (Mistranslation) 1.2 การละเลยเนื้อหา (Omission) 1.3 การเพิ่มเนื้อหา (Addition) 1.4 ข้อความที่ไม่ได้แปล (Untranslated Text)	1	1	1	3.00	ผ่าน

ลำดับ ที่	รายการประเมิน (องค์ประกอบของ แบบฟอร์ม)	ผู้เชี่ยวชาญ 1	ผู้เชี่ยวชาญ 2	ผู้เชี่ยวชาญ 3	รวม	แปล ผล (IOC)
	ส่วนที่ 2 การประเมิน บทคัดย่อ : เกณฑ์ย่อย MQM ประกอบไปด้วย 2) ความคล่องแคล่ว (Fluency) เกณฑ์ย่อยได้แก่					
8	2.1 การใช้เครื่องหมายวรรค ตอน (Punctuation) 2.2 การสะกดคำ (Spelling) 2.3 ไวยากรณ์ (Grammar) 2.4 ระดับภาษา (Register)	1	1	1	3.00	ผ่าน
	ส่วนที่ 2 การประเมิน บทคัดย่อ : เกณฑ์ย่อย MQM ประกอบไปด้วย 3) ความสอดคล้องของศัพท์ เฉพาะ (Terminology Consistency) เกณฑ์ย่อย ได้แก่					
9	3.1 การใช้ศัพท์ไม่ตรง (Incorrect Term) 3.2 การใช้ศัพท์ไม่ สอดคล้องกัน (Inconsistent Term)	1	1	1	3.00	ผ่าน

ลำดับ ที่	รายการประเมิน (องค์ประกอบของ แบบฟอร์ม)	ผู้เชี่ยวชาญ 1	ผู้เชี่ยวชาญ 2	ผู้เชี่ยวชาญ 3	รวม	แปล ผล (IOC)
10	<p>มาตราส่วนประมาณค่า (Rating scale) 5 ระดับ ซึ่งแต่ละระดับมีความหมาย คุณภาพการแปล ดังนี้: 5 = คุณภาพดีเยี่ยม (ไม่มีข้อผิดพลาด)</p> <p>4 = คุณภาพดี (มีข้อผิดพลาดเล็กน้อย แต่ ไม่ส่งผลต่อความเข้าใจ โดยรวม)</p> <p>3 = คุณภาพปานกลาง (มีข้อผิดพลาดบางส่วน ที่ ส่งผลต่อความเข้าใจ เล็กน้อย)</p> <p>2 = คุณภาพน้อย (มีข้อผิดพลาดหลายแห่ง ทำ ให้เข้าใจได้ยาก)</p> <p>1 = คุณภาพน้อยที่สุด (มี ข้อผิดพลาดร้ายแรง ทำให้ เข้าใจผิดหรือไม่สามารถ เข้าใจได้)</p>	1	1	1	3.00	ผ่าน

ลำดับ ที่	รายการประเมิน (องค์ประกอบของ แบบฟอร์ม)	ผู้เชี่ยวชาญ 1	ผู้เชี่ยวชาญ 2	ผู้เชี่ยวชาญ 3	รวม	แปล ผล (IOC)
11	<p>คำถามปลายเปิด 4 ข้อ</p> <p>หลังจากตอบแบบสอบถาม:</p> <p>1. ท่านเห็นความแตกต่าง ระหว่างการแปลโดยมนุษย์ และ ChatGPT อย่างไรใน ภาพรวม</p> <p>2. ท่านมีข้อเสนอแนะใน การพัฒนาคุณภาพการแปล บทความวิทยานิพนธ์ใน อนาคตอย่างไร</p> <p>3. คำสั่งที่ท่านใช้สั่ง ChatGPT ในการแปลภาษา ไทยเป็นภาษาอังกฤษมี อะไรบ้าง</p> <p>4. นอกจาก ChatGPT ท่าน มีประสบการณ์ในการใช้ AI ตัวอื่นในการแปลภาษาไทย เป็นภาษาอังกฤษหรือไม่</p>	1	1	1	3.00	ผ่าน

สรุป: ค่า IOC ทุกข้อ = 1.00 (ผ่านเกณฑ์) และไม่มีข้อเสนอแนะเพิ่มเติมจากผู้เชี่ยวชาญ

หมายเหตุ: รายการประเมินอ้างอิงจากแบบบันทึกข้อผิดพลาด ในภาคผนวก ก

ภาคผนวก ค


ตารางข้อมูลดิบข้อเสนอแนะผู้เชี่ยวชาญตอบคำถามปลายเปิด 4 ข้อ



ข้อเสนอแนะผู้เชี่ยวชาญตอบคำถามปลายเปิด 4 ข้อ

ผช. (รายที่)	1. ความแตกต่าง ระหว่างมนุษย์ และ ChatGPT ในภาพรวม	2. ข้อเสนอแนะใน การพัฒนาคุณภาพ การแปล	3. คำสั่ง (Prompt) ที่ใช้	4. ประสบการณ์ ใช้ AI อื่น
1	ในบทแปลไม่ได้ระบุว่าขั้นตอนเป็นการแปลจาก Chat GPT หรือแปลโดยมนุษย์ทำให้วิเคราะห์ไม่ได้ อย่างชัดเจน	สามารถมาใช้ AI ร่วมกับมนุษย์ได้เป็นการ Recheck ทำให้การแปลมีประสิทธิภาพมากขึ้น	ท่านเป็นผู้เชี่ยวชาญในการแปลภาษาไทยและภาษาอังกฤษ ช่วยแปลบทคัดย่อต่อไป โดยคงภาษาที่ถูกต้องตามหลักไวยากรณ์ และเป็นภาษาทางการที่ใช้ในการเขียนงานวิจัย	มี เช่น Gemini Google Translate ใช้หลายแอปร่วมกัน
2	มนุษย์มีความแม่นยำสูงในการถ่ายทอดตัวเลข สถิติ และเจตนาของข้อความ โดยเฉพาะการตีความบริบทและศัพท์เฉพาะทางราชการหรือสังคมส่วน AI แม้จะรักษาความถูกต้องของตัวเลขได้ดีและมีความคล่องแคล่วในการสร้างประโยค แต่มีแนวโน้มผิดพลาดจากการตีความหรือเพิ่มหรือละเว้นคำ ทำให้ความหมายเปลี่ยนไป	เน้นการกำหนด Role/Persona ให้ชัดเจน และระบุข้อจำกัดด้านสไตล์ (Style Constraints) ให้ชัดเจน เช่น “The translation must maintain an Academic Register”. ควรเน้นย้ำเรื่องความถูกต้องของตัวเลขและศัพท์เฉพาะทางเทคนิค	ตัวอย่างคำสั่ง Role/Persona: “You are an expert academic translator specializing in scientific/social science thesis abstracts. Your goal is to translate the provided Thai abstract into formal, clear, and concise English suitable for a journal	Google Translate: เร็วแต่คุณภาพศัพท์เทคนิคยังไม่สูงเท่า LLM. Gemini: จัดการ Prompt ซับซ้อนได้ดี รักษา Accuracy ของข้อมูล และปรับรูปแบบตามเกณฑ์ MQM ได้ยืดหยุ่นกว่า

ข้อเสนอแนะผู้เชี่ยวชาญตอบคำถามปลายเปิด 4 ข้อ

ผช. (รายที่)	1. ความแตกต่าง ระหว่างมนุษย์ และ ChatGPT ในภาพรวม	2. ข้อเสนอแนะใน การพัฒนาคุณภาพ การแปล	3. คำสั่ง (Prompt) ที่ใช้	4. ประสบการณ์ ใช้ AI อื่น
	<p>ความคล่องแคล่วของ AI เหนือกว่าในด้านโครงสร้างภาษาอังกฤษเชิงวิชาการและการจัดรูปแบบที่เป็นระเบียบ อย่างไรก็ตาม AI มักพลาดในรายละเอียดเชิงบริบทและการใช้ชื่อเฉพาะที่ซับซ้อน มนุษย์จึงยังจำเป็นสำหรับการตรวจสอบศัพท์เฉพาะที่ละเอียดอ่อนและไม่ชัดเจน</p> <p>ในต้นฉบับภาษาไทย เช่น ชื่อทฤษฎีหรือคำเฉพาะทางเทคนิค</p>		<p>publication, adhering strictly to the source text's intent and data.”</p> <p>Style/Format Constraints:</p> <p>“The translation must maintain an **Academic Register**. Use formal vocabulary and structure. Ensure that all **numbers, statistics (e.g., p<.05, mean, SD), and technical terms (e.g., ANOVA, t-test, specific chemical names/plant parts)** are translated accurately and consistently.”</p>	


ข้อเสนอแนะผู้เชี่ยวชาญตอบคำถามปลายเปิด 4 ข้อ

ผช. (รายที่)	1. ความแตกต่าง ระหว่างมนุษย์ และ ChatGPT ในภาพรวม	2. ข้อเสนอแนะใน การพัฒนาคุณภาพ การแปล	3. คำสั่ง (Prompt) ที่ใช้	4. ประสบการณ์ ใช้ AI อื่น
			<p>**Specific Instructions (ถ้ามี):**</p> <p>“The thesis subject is (สาขา). Please pay special attention to the consistency of the following key terms:</p> <ul style="list-style-type: none"> - (คำไทย 1) must be translated as (คำอังกฤษ 1) - (คำไทย 2) must be translated as (คำอังกฤษ 2)” <p>**Input:**</p> <p>(บทคัดย่อภาษาไทยฉบับเต็ม)</p> <p>ตัวอย่างการปรับแต่งคำสั่ง (Modification)</p>	

ข้อเสนอแนะผู้เชี่ยวชาญตอบคำถามปลายเปิด 4 ข้อ

ผช. (รายที่)	1. ความแตกต่าง ระหว่างมนุษย์ และ ChatGPT ในภาพรวม	2. ข้อเสนอแนะใน การพัฒนาคุณภาพ การแปล	3. คำสั่ง (Prompt) ที่ใช้	4. ประสบการณ์ ใช้ AI อื่น
			<p>ที่สำคัญ: เน้นความถูกต้อง ของข้อมูล (Accuracy Focus): หากเป็นงานวิจัย เชิงปริมาณ (เช่น บทความย่อ 3, 4, 10) ผมจะเพิ่มคำสั่ง: “DO NOT change or interpret any numerical data, percentages, or statistical results. Translate them exactly as provided in the Thai text.”</p> <p>เน้นความสอดคล้อง ของชื่อเฉพาะ (Terminology Focus): หากเป็น งานวิจัยที่มีชื่อทฤษฎี (เช่น บทความย่อ 7) อาจารย์จะเน้นย้ำ: “Ensure the name 'เดวีส์' is accurately</p>	

ข้อเสนอแนะผู้เชี่ยวชาญตอบคำถามปลายเปิด 4 ข้อ

ผช. (รายที่)	1. ความแตกต่าง ระหว่างมนุษย์ และ ChatGPT ในภาพรวม	2. ข้อเสนอแนะใน การพัฒนาคุณภาพ การแปล	3. คำสั่ง (Prompt) ที่ใช้	4. ประสบการณ์ ใช้ AI อื่น
			translated as 'Davies' and not confused with 'Dewey'."	
3	หากพิจารณา ในด้านความถูกต้อง (Accuracy) การแปล โดยมนุษย์อาจมี บางกรณีที่เลือกใช้ คำศัพท์หรือประโยค ที่ไม่ตรงกับบริบท ภาษาอังกฤษที่ควรใช้ โดยเฉพาะอย่างยิ่ง ในด้านเชิงวิชาการ หรือเฉพาะทางที่ต้อง ใช้เทคนิคเฉพาะ หากผู้แปลไม่คุ้นเคย กับบริบทนั้นอาจทำให้ การเลือกคำไม่แม่นยำ เท่าที่ควร ในขณะที่ ChatGPT มักจะมี ข้อมูลที่ทำให้แปล ได้ตรงตัวและ สอดคล้องกับบริบท ได้มากกว่า	สำหรับข้อเสนอแนะ ในการพัฒนา คุณภาพการแปล ในอนาคต เห็นว่า ควรดำเนินการใน หลายด้านควบคู่กัน ได้แก่ 1.การใช้เครื่องมือ AI ร่วมกับมนุษย์ แม้ AI จะมีศักยภาพสูง ในการช่วยแปลและ ให้ข้อมูลที่ถูกต้อง แต่ยังมีข้อจำกัด ในด้านบริบทภาษา และวัฒนธรรมไทย เฉพาะอย่าง เมื่อแปลจาก ภาษาไทยเป็น ภาษาอังกฤษ ซึ่งอาจทำให้ ความหมายหรือ หลักการเชิงวิชาการ	คำสั่งที่ใช้ เช่น ท่าน เป็นผู้เชี่ยวชาญ ด้านภาษาอังกฤษ และเป็นผู้เชี่ยวชาญ ในสาขาการศึกษา โปรดช่วยแปล บทความย่อภาษาไทย ต่อไปนี้เป็น ภาษาอังกฤษโดยให้ เป็นแบบทาง การใช้งานใน งานวิจัยทาง การศึกษา ขอให้ ใช้งานเป็นทางการ ถูกต้อง ตามหลัก วิชาการ และเลือกใช้ ศัพท์เฉพาะ ที่ เหมาะสมกับ สาขาวิชาการนี้ [รวม บทความย่อภาษาไทย]	นอกจากการใช้ ChatGPT อาจารย์ มีประสบการณ์ ในการใช้ AI ตัวอื่น สำหรับการแปล ภาษาไทย เป็นภาษาอังกฤษ เช่น Claude และ Gemini Claude จาก ประสบการณ์พบว่า ภาษาที่ได้ค่อนข้าง ฟุ่มเฟือยคือมี การขยายความ หรือใช้ถ้อยคำ ยืดเยื้อเกิน ความจำเป็น แม้จะอ่านได้สั้น แต่อาจไม่ค่อย ชัดเจนและกระชับ เพียงพอต่อการใช้

ข้อเสนอแนะผู้เชี่ยวชาญตอบคำถามปลายเปิด 4 ข้อ

ผช. (รายที่)	1. ความแตกต่าง ระหว่างมนุษย์ และ ChatGPT ในภาพรวม	2. ข้อเสนอแนะใน การพัฒนาคุณภาพ การแปล	3. คำสั่ง (Prompt) ที่ใช้	4. ประสบการณ์ ใช้ AI อื่น
	<p>ในด้าน ความคล่องแคล่ว (Fluency) มนุษย์อาจ มีข้อจำกัดจากความรู้ และการเข้าถึงข้อมูล ในบางครั้ง การแปล อาจติดขัด หรือไม่สละสลวย เท่าที่ควร ในขณะที่ ChatGPT สามารถ สร้างประโยคที่อ่าน แล้วลื่นไหลและ ใกล้เคียงกับ ภาษาธรรมชาติ ได้ดีกว่า สุดท้าย ในด้านการใช้ศัพท์ เฉพาะ (Terminology) มักพบว่าการแปล ของมนุษย์อาจมี ความไม่สม่ำเสมอ ากผู้แปลไม่ชำนาญ ในสาขานั้น ๆ แต่ ChatGPT มีแนวโน้มที่จะใช้ศัพท์ เฉพาะได้อย่างถูกต้อง และสม่ำเสมอกว่า</p>	<p>ในบริบทไทย ถูกถ่ายทอด ไม่ครบถ้วน ดังนั้น การผสมผสาน ระหว่าง AI และ ผู้แปลมนุษย์ จึงเป็นแนวทาง ที่เหมาะสมที่สุด</p> <p>2.การอบรม เชิงปฏิบัติการ ด้านการแปล: ควรมี การจัดอบรมเพื่อให้ ผู้แปล นักศึกษา และบุคลากร ได้เรียนรู้การใช้ เครื่องมือ AI อย่างมีประสิทธิภาพ รวมถึงเทคนิค การตรวจทาน งานแปล (Post - Editing) เพื่อ ยกระดับคุณภาพ การแปลในภาพรวม</p> <p>3.การแลกเปลี่ยน เรียนรู้ระหว่าง ผู้เชี่ยวชาญ: ควรมี เวทีหรือลักษณะ</p>	<p>Prompt ภาษาอังกฤษ: > “You are an expert in English language and also a specialist in the field of Education. Please translate the following Thai abstract into English, ensuring that the style is academic, the terminology is appropriate for the field of Education, and the meaning is conveyed accurately and fluently. [Insert Thai abstract]”</p>	<p>สำหรับบทคัดย่อ เชิงวิชาการ Gemini ลักษณะ การแปลคล้ายกับ Claude คือแม้ ข้อความ จะสละสลวย แต่การเลือกใช้ศัพท์ และการถ่ายทอด ความหมาย อาจไม่ค่อยตรงกับ บริบททางวิชาการ มากนัก ทำให้ ความแม่นยำ ในเชิงภาษาเฉพาะ ยังไม่ได้เท่าที่ควร</p>

ข้อเสนอแนะผู้เชี่ยวชาญตอบคำถามปลายเปิด 4 ข้อ

ผช. (รายที่)	1. ความแตกต่าง ระหว่างมนุษย์ และ ChatGPT ในภาพรวม	2. ข้อเสนอแนะใน การพัฒนาคุณภาพ การแปล	3. คำสั่ง (Prompt) ที่ใช้	4. ประสบการณ์ ใช้ AI อื่น
	ซึ่งช่วยให้ข้อความ ที่แปลมีความ น่าเชื่อถือและเข้าถึงได้ ในเชิงวิชาการ หรือวิชาชีพ	การทำงานร่วมกัน ที่เปิดโอกาส ให้สถาบัน ผู้แปล และผู้เชี่ยวชาญ สาขาต่าง ๆ ได้แลกเปลี่ยน ความรู้และ ประสบการณ์ เพื่อให้การแปล บทความสอดคล้องกับ ความถูกต้อง เชิงวิชาการและ ความเหมาะสม เชิงภาษาและ วัฒนธรรม		
4	-	-		-
5	จากการประเมินตาม กรอบ MQM พบว่า การแปลโดย AI มีความแม่นยำ เชิงโครงสร้างและ รายละเอียดสถิติ แต่บางครั้งแปลตรงตัว เกินไปจนทำให้ข้อมูล คลาดเคลื่อน และมี	ไม่ควรใช้ AI แบบ “แปลแล้วจบ” แต่ควรใช้แบบ “AI - Powered Translation with Human Post - Editing”. ควร ตรวจสอบศัพท์ เฉพาะ โดยใช้	“โปรดรับบทบาท เป็นผู้เชี่ยวชาญ ด้าน การแปล วิชาการ (Academic Translator/Expert) ที่มีความเข้มงวด ในการประเมิน ความถูกต้อง ตรงตามต้นฉบับ	DeepL Translator เด่นเรื่องความเป็น ธรรมชาติ แปลสละสลวย ไวยากรณ์ถูกต้อง เหมาะกับ ภาษาอังกฤษ เข้าใจบริบทดี เหมาะกับประโยค


ข้อเสนอแนะผู้เชี่ยวชาญตอบคำถามปลายเปิด 4 ข้อ

ผช. (รายที่)	1. ความแตกต่าง ระหว่างมนุษย์ และ ChatGPT ในภาพรวม	2. ข้อเสนอแนะใน การพัฒนาคุณภาพ การแปล	3. คำสั่ง (Prompt) ที่ใช้	4. ประสบการณ์ ใช้ AI อื่น
	<p>ข้อผิดพลาดในการจัดการตัวเลขหรือหน่วยวัด ส่วนการแปลโดยมนุษย์ปรับสำนวนให้เหมาะกับบริบทวิชาการ ใช้ศัพท์เฉพาะได้ถูกต้อง และรักษาความแม่นยำของข้อมูลเชิงตัวเลขดีกว่าในด้านความคล่องแคล่ว AI ใช้ศัพท์วิชาการได้ดี แต่บางประโยคแข็งทื่อและซ้ำซ้อน ขณะที่มนุษย์ใช้โครงสร้างประโยคหลากหลายและเป็นธรรมชาติ ส่วนการใช้ศัพท์เฉพาะ AI มักผิดพลาดในชื่อเทคนิคหรือบุคคล และมีความไม่สม่ำเสมอ ขณะที่มนุษย์มีความแม่นยำและสม่ำเสมอมากกว่า</p>	<p>Glossary ตรวจสอบก่อนการรวม AI หลายตัว ช่วยให้เลือกผลลัพธ์ที่ดีที่สุดได้</p>	<p>ภาษาไทย (Literal Accuracy) สูงสุด</p> <p>1.ภารกิจ: ประเมินบทความย่อภาษาไทยฉบับที่ [หมายเลข] โดยเทียบกับแบบแปลที่ 1 และ 2 ตามเกณฑ์ MQM.</p> <p>2.เกณฑ์ MQM: ความถูกต้อง (Accuracy): เน้นความแม่นยำของข้อมูล ตัวเลข ชื่อเฉพาะ และความครบถ้วนของเนื้อหา.</p> <p>ความคล่องแคล่ว (Fluency): เน้นความสละสลวยและระดับความเป็นวิชาการ (Academic Register) ที่เหมาะสม</p> <p>ศัพท์เฉพาะ (Terminology): เน้นความสอดคล้องและการใช้</p>	<p>ซับซ้อนในงานวิชาการ</p>

ข้อเสนอแนะผู้เชี่ยวชาญตอบคำถามปลายเปิด 4 ข้อ

ผช. (รายที่)	1. ความแตกต่าง ระหว่างมนุษย์ และ ChatGPT ในภาพรวม	2. ข้อเสนอแนะใน การพัฒนาคุณภาพ การแปล	3. คำสั่ง (Prompt) ที่ใช้	4. ประสบการณ์ ใช้ AI อื่น
			ศัพท์เทคนิคที่เป็น มาตรฐานสากล. 3.ข้อจำกัด: ห้ามมี การแก้ไขการแปล แต่ให้ระบุ ข้อผิดพลาดและ ให้คะแนน (Scale 1 - 5). 4.ผลลัพธ์: สรุป คุณภาพโดยรวม ของแต่ละแบบแปล พร้อมระบุแบบแปล ที่มีคุณภาพสูงกว่า.”	
6	โดยภาพรวม การแปล โดยมนุษย์ มีความโดดเด่น ในด้าน Accuracy เนื่องจากผู้แปล สามารถทำความเข้าใจ เข้าใจเนื้อหาต้นฉบับ เชิงลึก และถ่ายทอด ความหมายให้ตรง กับบริบทเชิงวิชาการ ได้อย่างครบถ้วน พร้อมทั้งสามารถปรับ สำนวนให้สอดคล้อง กับแนวทางการเขียน	ควรใช้แนวทาง ผสมผสาน โดยใช้ AI แปลฉบับร่าง เพื่อความรวดเร็ว แต่ต้องมีผู้เชี่ยวชาญ ตรวจสอบ. ควรจัด อบรมเชิงปฏิบัติการ ด้านการแปล วิชาการให้นักศึกษา เพื่อพัฒนาทักษะ การเลือกคำและ โครงสร้างประโยค	“Translate the following Thai abstract into academic English suitable for a thesis/ dissertation.” “Please ensure accuracy, fluency, and consistency in terminology.”	นอกจาก ChatGPT เคยมีประสบการณ์ ใช้ Google Translate และ DeepL ในการแปลภาษา ไทยเป็น ภาษาอังกฤษ โดย Google Translate มักเหมาะ กับการแปล ข้อความทั่วไป แต่ยังมีข้อจำกัด

ข้อเสนอแนะผู้เชี่ยวชาญตอบคำถามปลายเปิด 4 ข้อ

ผช. (รายที่)	1. ความแตกต่าง ระหว่างมนุษย์ และ ChatGPT ในภาพรวม	2. ข้อเสนอแนะใน การพัฒนาคุณภาพ การแปล	3. คำสั่ง (Prompt) ที่ใช้	4. ประสบการณ์ ใช้ AI อื่น
	<p>บทความทางวิชาการ ที่เหมาะสมกับผู้อ่าน เป้าหมาย ขณะที่ การแปลโดย ChatGPT มีจุดแข็ง ในด้าน Fluency และ ความคล่องตัว ในการเรียบเรียง ทำให้ประโยค อ่านสั้นไหล ในเชิงภาษาศาสตร์ แต่บางครั้งอาจเกิด ข้อผิดพลาด ด้าน Terminology โดยเฉพาะคำศัพท์ วิชาการเฉพาะสาขา หรือคำที่มีความหมาย เฉพาะทาง ซึ่งอาจ ไม่สอดคล้องกับ มาตรฐานวิชาการ หรือไม่ตรงตามเนื้อหา ต้นฉบับ</p>		<p>“Use formal academic style and appropriate terminology for research abstracts.”</p> <p>บางครั้งจะมีการระบุ เพิ่มเติม เช่น “retain statistical terms as in the original,” หรือ “translate technical terms using standard English terminology.”</p> <p>เพื่อให้การแปลมี ความถูกต้องทาง วิชาการมากขึ้น</p>	<p>ด้านสำนวนวิชาการ ส่วน DeepL สามารถให้ผลลัพธ์ ที่เป็นธรรมชาติ มากขึ้น แต่ยังไม่ รองรับรับภาษาไทย ได้อย่างสมบูรณ์ ดังนั้น ChatGPT จึงเป็นทางเลือก ที่มีประสิทธิภาพ มากกว่าในด้าน การแปลบทความ เชิงวิชาการ</p>
7	ChatGPT ความ ถูกต้องและ คล่องแคล่วมากกว่า แต่ก็ต้องตรวจสอบ	-	-	-

ข้อเสนอแนะผู้เชี่ยวชาญตอบคำถามปลายเปิด 4 ข้อ

ผช. (รายที่)	1. ความแตกต่าง ระหว่างมนุษย์ และ ChatGPT ในภาพรวม	2. ข้อเสนอแนะใน การพัฒนาคุณภาพ การแปล	3. คำสั่ง (Prompt) ที่ใช้	4. ประสบการณ์ ใช้ AI อื่น
	การใช้เฉพาะให้ตี เพราะอาจมีความ ผิดพลาดในบางครั้ง			
8	มนุษย์เด่น ด้าน Accuracy และ Terminology ส่วน ChatGPT เต้น ด้าน Fluency แต่ยังต้องมีการ ตรวจสอบ ความถูกต้องของคำ และความหมาย	ควรจัดทำฐานข้อมูล ศัพท์เฉพาะ (terminology database) ของแต่ละสาขาวิชา เพื่อให้นักศึกษา และผู้แปล (ทั้งมนุษย์และ AI) ใช้คำศัพท์ ได้อย่างถูกต้อง การแปลควรผ่าน การตรวจทาน โดยผู้เชี่ยวชาญ ภาษาอังกฤษ	ใช้คำสั่งตรงไปตรงมา เช่น “ช่วยแปล บทความนี้ เป็นภาษาอังกฤษ โดยใช้ภาษาทางการ เชิงวิชาการ” กำหนดบริบท เพิ่มเติม เช่น “ให้แปล โดยคง รูปแบบบทความย่อ วิทยานิพนธ์ ใช้โครงสร้าง ภาษาอังกฤษ เชิงวิชาการ และรักษา ความหมายต้นฉบับ ให้ครบถ้วน” บางครั้งมีการปรับ แต่ง Prompt เช่น “ใช้คำศัพท์ทางการ, หลีกเลี่ยงการทับ ศัพท์, แยก วัตถุประสงค์	เคยใช้ Google Translate ซึ่งให้ผลลัพธ์ รวดเร็ว แต่มี ข้อจำกัดเรื่อง ความแม่นยำ ในเชิงวิชาการ โดยเฉพาะ การเลือกคำศัพท์ และโครงสร้าง ประโยคที่ไม่เป็น ธรรมชาติ บางครั้งใช้ DeepL Translator สำหรับ การตรวจสอบ เพิ่มเติม พบว่า ภาษาอังกฤษ ที่ได้อ่านสั้นไหล และใกล้เคียง ภาษาธรรมชาติมาก แต่ยังไม่มี การรองรับ

ข้อเสนอแนะผู้เชี่ยวชาญตอบคำถามปลายเปิด 4 ข้อ

ผช. (รายที่)	1. ความแตกต่าง ระหว่างมนุษย์ และ ChatGPT ในภาพรวม	2. ข้อเสนอแนะใน การพัฒนาคุณภาพ การแปล	3. คำสั่ง (Prompt) ที่ใช้	4. ประสบการณ์ ใช้ AI อื่น
			<p>เป็นข้อ ๆ ตามต้นฉบับ”</p> <p>ในกรณีที่เป็นศัพท์ เฉพาะ จะใช้คำสั่ง เพิ่มเติม เช่น “ช่วยตรวจสอบ การใช้คำว่า juvenile correctional institution ว่าตรงกับ ทัณฑสถานวัยหนุ่ม กลางหรือไม่”</p>	<p>ภาษาไทยอย่าง สมบูรณ์แบบ</p> <p>โดยรวม ChatGPT เหนือกว่า ด้านความยืดหยุ่น ในการปรับคำสั่ง และ การรักษา รูปแบบวิชาการ ทำให้เหมาะสมกว่า ในการใช้กับ บทความย่อ วิทยานิพนธ์</p>
9	<p>การแปล โดย ChatGPT มีความแม่นยำและ สม่ำเสมอ ในด้านไวยากรณ์ และการใช้ศัพท์เฉพาะ มากกว่ามนุษย์ทั่วไป แต่ยังขาด ความเป็นธรรมชาติ และ ความละเอียดอ่อน ที่นักแปลมนุษย์ ผู้เชี่ยวชาญสามารถ</p>	<p>ควรใช้ AI เป็นเครื่องมือแปล เบื้องต้น แล้วให้ ผู้เชี่ยวชาญ ตรวจทานและ ปรับปรุง เพื่อความสละสลวย และแม่นยำ จัด อบรมการแปล เชิงวิชาการ ให้ผู้เกี่ยวข้อง สร้างคลังคำศัพท์ เฉพาะทาง สำหรับ</p>	<p>รูปแบบคำสั่งพื้นฐาน เช่น “Translate the following abstract from Thai to English. Ensure the tone is formal and academic. The translation should be accurate, concise, and clear. Maintain the</p>	<p>Google Translate, DeepL และ Microsoft Translator</p>

ข้อเสนอแนะผู้เชี่ยวชาญตอบคำถามปลายเปิด 4 ข้อ

ผช. (รายที่)	1. ความแตกต่าง ระหว่างมนุษย์ และ ChatGPT ในภาพรวม	2. ข้อเสนอแนะใน การพัฒนาคุณภาพ การแปล	3. คำสั่ง (Prompt) ที่ใช้	4. ประสบการณ์ ใช้ AI อื่น
	<p>ทำได้ โดยมนุษย์ มีข้อได้เปรียบ ในด้านการเข้าใจ บริบทและเลือกใช้คำ ที่เหมาะสมที่สุด ขณะที่ ChatGPT เต็ม เรื่องความถูกต้อง และความสั้นไหล ของโครงสร้างประโยค แต่บางครั้งใช้ภาษา ที่เป็นแบบแผนเกินไป และไม่แพรวพราว</p>	<p>แต่ละสาขา และ กำหนดให้มีการ การตรวจทาน โดยผู้เชี่ยวชาญ เพื่อยกระดับ คุณภาพงานแปล</p>	<p>original structure and formatting." รูปแบบคำสั่ง ที่เจาะจงมากขึ้น (เพื่อเพิ่ม ความแม่นยำ): "Translate the following thesis abstract from Thai to English. Pay close attention to the research methodology and findings sections. For technical terms like 'ค่าดัชนี การเรียงลำดับ ความสำคัญ ของความต้องการ จำเป็น,' please use the standard academic term 'Priority Needs Index (PNI).' The subject is related to [ระบุสาขา]. The</p>	

ข้อเสนอแนะผู้เชี่ยวชาญตอบคำถามปลายเปิด 4 ข้อ

ผช. (รายที่)	1. ความแตกต่าง ระหว่างมนุษย์ และ ChatGPT ในภาพรวม	2. ข้อเสนอแนะใน การพัฒนาคุณภาพ การแปล	3. คำสั่ง (Prompt) ที่ใช้	4. ประสบการณ์ ใช้ AI อื่น
			<p>tone must be highly formal and suitable for a scientific journal."</p> <p>รูปแบบคำสั่งที่มีการปรับปรุงแก้ไข (หลังจากการแปลครั้งแรก): "The previous translation of 'ทัณฑสถานวัยหนุ่มกลาง' was 'Central Juvenile Correctional Institution.' This is incorrect. The correct term is 'Central Youth Correctional Institution.'</p> <p>Please regenerate the full abstract with this correction and ensure all instances of the</p>	

ข้อเสนอแนะผู้เชี่ยวชาญตอบคำถามปลายเปิด 4 ข้อ

ผช. (รายที่)	1. ความแตกต่าง ระหว่างมนุษย์ และ ChatGPT ในภาพรวม	2. ข้อเสนอแนะใน การพัฒนาคุณภาพ การแปล	3. คำสั่ง (Prompt) ที่ใช้	4. ประสบการณ์ ใช้ AI อื่น
			term are consistent."	
10	<p>เป็นการยากที่จะตอบว่าอันใดดีกว่ากันระหว่างมนุษย์หรือแปลโดย ChatGPT เพราะในความเป็นจริงงานที่แปลโดย ChatGPT จะมีการ Edit เพิ่มเติมซึ่งทำให้ดีขึ้น และกลายเป็นว่า GPT แปลได้ดีกว่าแต่จริง ๆ แล้วมีการแก้ไขเพิ่มเติมโดยมนุษย์</p> <p>มนุษย์ แปลได้ถูกต้องตามลักษณะของเอกสารนั้น ๆ แต่อาจมีข้อผิดพลาดมาก</p> <p>AI ทำได้เร็ว ผิดพลาดน้อย แต่บางที่ผิดพลาดที่บางที่ไม่สามารถแก้ไขเองได้</p>	<p>ต้องพัฒนาไปตามสถานการณ์ปัจจุบันให้เลย เพราะ AI มีการพัฒนาอยู่เสมอ ๆ อย่างรวดเร็ว</p>	<p>ตรวจตัวสะกด ไวยากรณ์ นำคำแปลที่ได้มาอ่านอีกครั้งเพื่อดูความถูกต้องเหมาะสมหรือไม่</p> <p>ในรายละเอียดเฉพาะ</p>	<p>AI ใช้ได้ทุกตัว เพราะมี Large Language Model ในกลุ่มการแปลใช้ได้ดีหลายตัว เช่น Gemini, Perplexity, DeepSeek, ChatGPT, ฯลฯ</p>

ข้อเสนอแนะผู้เชี่ยวชาญตอบคำถามปลายเปิด 4 ข้อ

ผช. (รายที่)	1. ความแตกต่าง ระหว่างมนุษย์ และ ChatGPT ในภาพรวม	2. ข้อเสนอแนะใน การพัฒนาคุณภาพ การแปล	3. คำสั่ง (Prompt) ที่ใช้	4. ประสบการณ์ ใช้ AI อื่น
	ต้องมีมนุษย์ช่วยแก้ไข เปรียบเทียบพาทหนะ ที่เราขับซี รถจักรยาน รถยนต์ หรือเครื่องบิน -> AI มนุษย์ คือผู้กำหนดว่าจะให้ ไปไหน ทิศทางใด			
11	เนื่องจาก ChatGPT ถูกสร้างจากมนุษย์ นับล้าน คลังคำศัพท์ ถูกสะสมมาเป็น ศตวรรษ การเลือกใช้ คำหรือการใช้คำศัพท์ ย่อมดีกว่าคน 1 คน แปล ในเรื่อง ความ หลากหลาย อย่างไรก็ตาม เมื่อต้อง ใช้ Register ในบาง Context ต้อง มีการคัดกรอง โดยผู้เชี่ยวชาญ ด้านนั้น ๆ เพราะ หลายครั้งศัพท์ไม่ตรง แบบ 100%	เห็นด้วยกับการใช้ AI ร่วมกับมนุษย์ แต่ขึ้นอยู่กับนโยบาย องค์กร + ผู้บริหาร ต้องเปิดใจยอมรับ ในการนำมาปรับใช้	ส่วนใหญ่จะใช้ ChatGPT แปล อย่างเดียว ดังที่กล่าว รายละเอียดในข้อ 2 ส่วนที่สั่งให้เขียน ภาษาอังกฤษเลย ยังไม่เคยใช้	ยังไม่มี ประสบการณ์ ในการใช้ AI ตัวอื่น ในการแปล ภาษาไทย เป็นภาษาอังกฤษ
12	-	-	-	-

ข้อเสนอแนะผู้เชี่ยวชาญตอบคำถามปลายเปิด 4 ข้อ

ผช. (รายที่)	1. ความแตกต่าง ระหว่างมนุษย์ และ ChatGPT ในภาพรวม	2. ข้อเสนอแนะใน การพัฒนาคุณภาพ การแปล	3. คำสั่ง (Prompt) ที่ใช้	4. ประสบการณ์ ใช้ AI อื่น
13	1. มนุษย์แปล 1.1 แปลข้อมูลไม่ครบ 1.2 ภาษาสละสลวย 1.3 Register (ระดับ ภาษา) ดีกว่ามาก 1.4 Punctuation (เครื่องหมายวรรค ตอน) ถูกต้องและมี ความเป็นสากล 2. AI แปล 2.1 คำเฉพาะบาง ศัพท์ที่ใช้ไม่ Consist (ไม่สม่ำเสมอ) 2.2 บางครั้งแปล วกวนไปมา 2.3 บางครั้งภาษาที่ใช้ ประหลาดเกินบริบท ที่นักวิจัยใช้จริง 2.4 Add up ข้อมูล เกินมา 2.5 ลื่นไหลสู้มนุษย์ ไม่ได้ ณ ปัจจุบัน	ใช้ AI ช่วยแปล อย่างน้อย 2-3 AI และ Cross - Check นักวิชาการ หรือ ผู้แปล ตรวจสอบ ความถูกต้อง ดูข้อกำหนด ของวารสาร ปลายทาง ที่จะตีพิมพ์ Prompt ต้องมี ความเหมาะสม สามารถดู Prompt ตัวอย่างได้ ใน YouTube ของนักวิชาการ อื่น ๆ ได้ สำคัญ: มนุษย์ ต้องตรวจสอบ	"ช่วยแปลข้อความ จากภาษาไทย ต่อไปนี้ เป็นภาษาอังกฤษ วิชาการ // (แล้วใส่ ข้อความ) ยกตัวอย่าง Abstract ให้แปล ทั้งหมด และนำไป ทดสอบกับ AI ตัวอื่น ๆ แล้วนำมา เปรียบเรียงใหม่"	Gemini AI Copilot Grammarly
14	โดยรวมแล้ว การแปล โดย มนุษย์ มีข้อดี ด้าน ความแม่นยำ (Accuracy) ที่ยึดเกาะ ต้นฉบับได้ค่อนข้างดี	ควรมีการใช้ เครื่องมือ AI ร่วมกับมนุษย์ โดยเฉพาะการให้ มนุษย์	"Translate this Thai abstract into academic English suitable for a thesis abstract."	มีประสบการณ์ใช้ AI แปลภาษาอื่น ๆ ได้แก่:

ข้อเสนอแนะผู้เชี่ยวชาญตอบคำถามปลายเปิด 4 ข้อ

ผช. (รายที่)	1. ความแตกต่าง ระหว่างมนุษย์ และ ChatGPT ในภาพรวม	2. ข้อเสนอแนะใน การพัฒนาคุณภาพ การแปล	3. คำสั่ง (Prompt) ที่ใช้	4. ประสบการณ์ ใช้ AI อื่น
	<p>มักหลีกเลี่ยง การตีความเกิน ต้นฉบับ และใช้ศัพท์ เฉพาะได้ตรงบริบท แต่ข้อเสียที่พบ คือ ความคล่องแคล่ว (Fluency) โดยเฉพาะ เรื่องโครงสร้าง ประโยคและระดับ ภาษาที่บางครั้งยังคง มีลักษณะตรงตัว จากภาษาไทย ไม่สลับไหลตามลักษณะ ภาษาอังกฤษ เชิงวิชาการ</p> <p>ส่วนการแปล โดย ChatGPT มีความโดดเด่นเรื่อง ความคล่องแคล่ว (Fluency) ใช้ภาษาอังกฤษ ได้สลับไหล เป็น ธรรมชาติ ในบริบท วิชาการ อ่านเข้าใจ ง่าย โดยเฉพาะในฉบับ ที่เป็นสาขา มนุษยศาสตร์และ</p>	<p>เป็นผู้ตรวจทาน ขั้นสุดท้าย เพื่อความถูกต้อง และเหมาะสม เชิงบริบท</p> <p>อบรมเชิงปฏิบัติการ ด้านการแปล โดยเฉพาะ การพัฒนาทักษะ การเขียน ภาษาอังกฤษ เชิงวิชาการสำหรับ ผู้เขียนบทความ จัดให้มี กระบวนการ ตรวจสอบ โดยผู้เชี่ยวชาญ ด้านการแปล เพื่อประเมิน ความถูกต้อง ของคำศัพท์เฉพาะ การแปลคำราชการ และบริบท ทางวัฒนธรรม ส่งเสริมการใช้ ฐานข้อมูลศัพท์</p>	<p>Keep the content faithful to the original and use proper terminology for scientific/academi c context.”</p> <p>หรือปรับรูปแบบ ดังนี้:</p> <p>“Please translate the following Thai research abstract into formal academic English, ensuring fluency, accuracy, and proper use of domain-specific terminology.”</p> <p>โดยมีการปรับคำสั่ง เพิ่มเติมตามบริบท ของสาขาวิชา เช่น วิทยาศาสตร์ เทคโนโลยี หรือ การศึกษา</p>	<p>DeepL – ให้ผลลัพธ์ใกล้เคียง ภาษาอังกฤษ แบบเจ้าของภาษา แต่ยังรองรับ ภาษาไทยจำกัด</p> <p>Google Translate – พัฒนาได้ดีขึ้น ในระยะหลัง แต่ยังมีข้อจำกัด ในการแปลเชิง วิชาการ</p> <p>Microsoft Translator – ให้ผลลัพธ์ในระดับ ปานกลาง ดีสำหรับ ศัพท์เทคนิค บางกลุ่ม</p> <p>อย่างไรก็ตาม ChatGPT ยังเป็น เครื่องมือที่โดดเด่น ที่สุดในด้าน ความสลับไหล ของสำนวนและ การปรับรูปแบบ</p>

ข้อเสนอแนะผู้เชี่ยวชาญตอบคำถามปลายเปิด 4 ข้อ

ผช. (รายที่)	1. ความแตกต่าง ระหว่างมนุษย์ และ ChatGPT ในภาพรวม	2. ข้อเสนอแนะใน การพัฒนาคุณภาพ การแปล	3. คำสั่ง (Prompt) ที่ใช้	4. ประสบการณ์ ใช้ AI อื่น
	<p>สังคมศาสตร์ อย่างไร ก็ตาม พบว่า มีบางกรณีที่ดีความ เกินจากต้นฉบับ (Addition) และการใช้ศัพท์เฉพาะ (Terminology) ยังมี โอกาสใช้ผิด โดยเฉพาะกรณีที่มี บริบทเฉพาะ ของระบบราชการหรือ การศึกษาไทย เช่น การแปลชื่อหน่วยงาน ราชการ หรือชื่อ สถานศึกษา</p>	<p>เฉพาะ เฉพาะสาขาวิชา เพื่อความสอดคล้อง ในการใช้คำ ในแต่ละสาขา</p>		<p>เป็นภาษา เชิงวิชาการ</p>
15	<p>1. ความถูกต้อง (Accuracy) การแปลโดยมนุษย์: มีความเข้าใจบริบท ทางวิชาการ อย่างลึกซึ้ง สามารถ แปลศัพท์เฉพาะ และแนวคิดซับซ้อน ได้อย่างแม่นยำ ChatGPT: พบ ข้อผิดพลาดบ่อย</p>	<p>การใช้ AI ร่วมกับ มนุษย์: ใช้ AI สำหรับแปลร่างแรก แล้วให้ผู้เชี่ยวชาญ ตรวจทานและ ปรับแก้ พร้อม พัฒนา AI ให้เข้าใจ ศัพท์เฉพาะของแต่ละ สาขา การอบรม: จัด อบรมการแปล</p>	<p>แนะนำให้ระบุบริบท ทางวิชาการ, การ รักษาความถูกต้อง ของตัวเลข, การใช้ ศัพท์เฉพาะที่ สอดคล้อง, และการ ตรวจสอบไวยากรณ์. แนะนำให้ใช้ prompt ที่ระบุ: บริบททางวิชาการ และสาขาวิชา</p>	<p>มีความรู้เกี่ยวกับ Google Translate, DeepL, Microsoft Translator แต่เห็น ว่างานวิชาการ ระดับสูงยังต้อง อาศัยการตรวจทาน โดยมนุษย์เป็นหลัก</p>

ข้อเสนอแนะผู้เชี่ยวชาญตอบคำถามปลายเปิด 4 ข้อ

ผช. (รายที่)	1. ความแตกต่าง ระหว่างมนุษย์ และ ChatGPT ในภาพรวม	2. ข้อเสนอแนะใน การพัฒนาคุณภาพ การแปล	3. คำสั่ง (Prompt) ที่ใช้	4. ประสบการณ์ ใช้ AI อื่น
	<p>ในส่วนข้อมูล เชิงตัวเลข การผสม ภาษา และการตีความ บริบททางวิชาการ</p> <p>2. ความคล่องแคล่ว (Fluency) การแปลโดยมนุษย์: ใช้ภาษาที่เป็น ธรรมชาติ โครงสร้าง ประโยคถูกต้อง ตามหลักไวยากรณ์ ภาษาอังกฤษ ChatGPT: บางครั้ง สร้างประโยคที่ยาว เกินไปหรือมีโครงสร้าง ที่ไม่เป็นธรรมชาติ</p> <p>3. การใช้ศัพท์เฉพาะ (Terminology) การแปลโดยมนุษย์: ใช้ศัพท์เทคนิค อย่างสม่ำเสมอและ เลือกคำที่เหมาะสม กับสาขาวิชา ChatGPT: มักใช้ศัพท์ ไม่สอดคล้องกัน หรือ แปลศัพท์เฉพาะผิด</p>	<p>เอกสารวิชาการ สำหรับนักแปล และ สอน การใช้ เครื่องมือ CAT ร่วมกับ AI</p> <p>การควบคุม คุณภาพ: มีการ ตรวจทานโดย ผู้เชี่ยวชาญ และ สร้างฐานข้อมูลศัพท์ เฉพาะของแต่ละ สาขา</p>	<p>การรักษาความ ถูกต้องของข้อมูลเชิง ตัวเลข การใช้ศัพท์เฉพาะที่ สอดคล้อง การตรวจสอบ ไวยากรณ์และความ เป็นธรรมชาติ</p>	

ข้อเสนอแนะผู้เชี่ยวชาญตอบคำถามปลายเปิด 4 ข้อ

ผช. (รายที่)	1. ความแตกต่าง ระหว่างมนุษย์ และ ChatGPT ในภาพรวม	2. ข้อเสนอแนะใน การพัฒนาคุณภาพ การแปล	3. คำสั่ง (Prompt) ที่ใช้	4. ประสบการณ์ ใช้ AI อื่น
16	<p>การแปลของมนุษย์มีความใกล้เคียงต้นฉบับมากกว่าในด้าน Accuracy โดยเฉพาะประเด็นเชิงเทคนิคหรือศัพท์เฉพาะในสาขาวิทยาศาสตร์ อย่างไรก็ตาม บางฉบับมีการแปลแบบตรงตัว ทำให้ Fluency ไม่เป็นธรรมชาติในภาษาอังกฤษ ChatGPT ให้ผลลัพธ์ที่มีความ Fluency สูงกว่าชัดเจนในหลายฉบับ โดยเฉพาะในสาขาสังคมศาสตร์และมนุษยศาสตร์ แต่บางครั้งมีการตีความ หรือขยายความเกินกว่าต้นฉบับ ทำให้ส่งผลต่อ Accuracy ของ Terminology ของทั้งสองวิธี</p>	<p>ควรใช้ AI ร่วมกับมนุษย์ โดยเฉพาะการให้ ChatGPT เป็นผู้แปลร่างเบื้องต้น และให้ผู้เชี่ยวชาญปรับปรุงเพิ่มเติม</p> <p>จัดอบรมเชิงปฏิบัติการด้านการแปล โดยเน้นกลยุทธ์การแปลเพื่อเลี่ยงการแปลตรงตัวและให้ความสำคัญกับผู้อ่านภาษาเป้าหมาย</p> <p>ใช้ Back Translation เป็นเครื่องมือช่วยตรวจสอบคุณภาพการแปล โดยเฉพาะในหัวข้อที่เป็นศัพท์เฉพาะหรือเชิงเทคนิค</p> <p>ส่งเสริมการ Proofread โดยผู้เชี่ยวชาญ</p>	<p>“ช่วยแปลบทคัดย่อวิทยานิพนธ์ฉบับนี้จากภาษาไทยเป็นภาษาอังกฤษ โดยคงความหมายให้ถูกต้อง และใช้สำนวนแบบภาษาอังกฤษเชิงวิชาการ”</p> <p>“Translate this Thai academic abstract into English with high accuracy and fluency for an international academic audience”</p> <p>“แปลบทคัดย่อจากวิทยานิพนธ์ไทย โดยใช้สไตล์การเขียนแบบ APA abstract”</p> <p>มีการปรับแต่งเพิ่มเติม เช่น การระบุว่าการ “ไม่ต้องเพิ่มเนื้อหา</p>	<p>ประสบการณ์ที่ใช้ เช่น DeepL Translator: ดีด้าน Fluency และมีแนวโน้มแปลได้สั้นไหลกว่า Google Translate แต่ยังมีปัญหาเรื่อง Terminology ในบริบทไทย</p> <p>Google Translate: ใช้งานง่าย แปลเร็ว แต่มีปัญหาความคลาดเคลื่อนในระดับข้อความเชิงวิชาการ</p> <p>Microsoft Translator: ผลลัพธ์ใกล้เคียงกับ Google Translate แต่ใช้น้อยกว่า</p> <p>จากการเปรียบเทียบคุณภาพ</p>

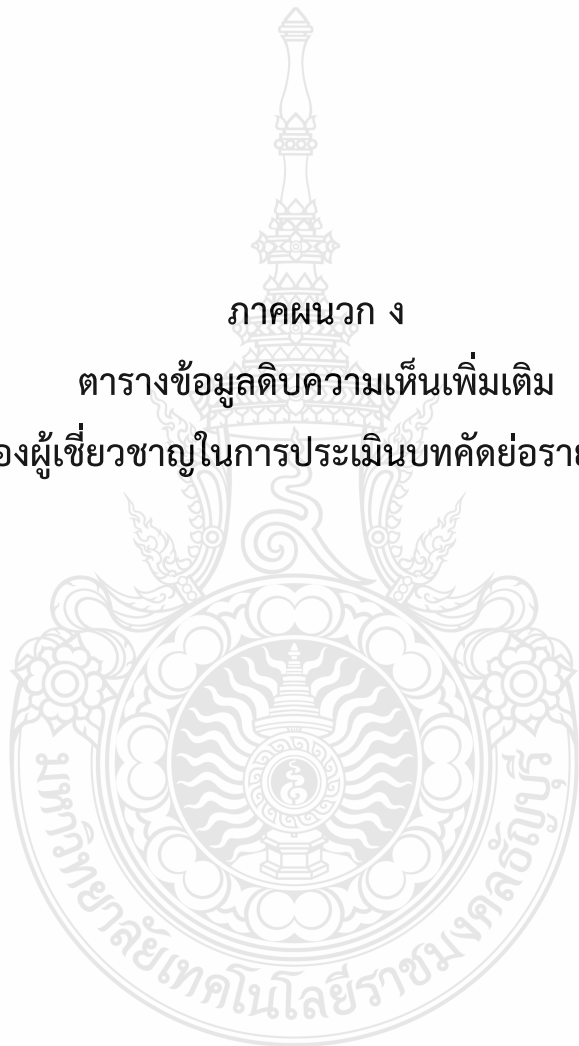
ข้อเสนอแนะผู้เชี่ยวชาญตอบคำถามปลายเปิด 4 ข้อ

ผช. (รายที่)	1. ความแตกต่าง ระหว่างมนุษย์ และ ChatGPT ในภาพรวม	2. ข้อเสนอแนะใน การพัฒนาคุณภาพ การแปล	3. คำสั่ง (Prompt) ที่ใช้	4. ประสบการณ์ ใช้ AI อื่น
	มีความถูกต้อง ในระดับที่ใกล้เคียงกัน แต่ ChatGPT อาจใช้ ศัพท์ที่เหมาะสมกับ Academic Writing มากกว่าในบางฉบับ ขณะที่มนุษย์มักใช้ คำซ้ำหรือตรงตัว จากต้นฉบับภาษาไทย	ด้านภาษาอังกฤษ ที่เข้าใจเนื้อหา ทางวิชาการ	เกินต้นฉบับ” หรือ “เน้นความกระชับ แต่ครบถ้วน”	เปรียบเทียบแล้ว ChatGPT ให้ผลลัพธ์ที่มี ความสม่ำเสมอ และสามารถ Customize ได้ผ่านคำสั่ง Prompt
17	ในภาพรวม ChatGPT จะแปลได้ดีกว่ามนุษย์ ในด้านความถูกต้อง ความคล่องแคล่ว และการใช้ศัพท์เฉพาะ จะมีเฉพาะบางศาสตร์ เช่น ชื่อหน่วยงาน ในไทย ชื่อเฉพาะทาง วิทยาศาสตร์ ที่ ChatGPT ยังแปล ไม่ถูกต้องแม่นยำ	บทคัดย่อควรเขียน เป็นภาษาไทยก่อน แล้วใช้ ChatGPT แปลเป็น ภาษาอังกฤษ จากนั้นนำ Abstract ภาษาอังกฤษ ที่ ChatGPT แปล มาตรวจสอบ ความถูกต้อง โดยเฉพาะชื่อเฉพาะ ทางศาสตร์นั้น ๆ เช่น ชื่อเครื่องพิมพ์ ชื่อทางวิทยาศาสตร์ ทางเคมี ทางชีววิทยา และ ชื่อคน เป็นต้น	ส่วนใหญ่จะใช้ ChatGPT แปล อย่างเดียว ดังที่กล่าว รายละเอียดในข้อ 2 ส่วนที่สั่งให้เขียน ภาษาอังกฤษเลย ยังไม่เคยใช้	-

ข้อเสนอแนะผู้เชี่ยวชาญตอบคำถามปลายเปิด 4 ข้อ

ผช. (รายที่)	1. ความแตกต่าง ระหว่างมนุษย์ และ ChatGPT ในภาพรวม	2. ข้อเสนอแนะใน การพัฒนาคุณภาพ การแปล	3. คำสั่ง (Prompt) ที่ใช้	4. ประสบการณ์ ใช้ AI อื่น
		หลังการตรวจสอบ ในข้อ 3 ก็จะได้ Abstract ที่สมบูรณ์ ที่ไม่น่าจะมี ส่วนใดผิดพลาด		
18	มนุษย์มีความติดขัด ในสำนวน ไวยากรณ์ ไม่สมบูรณ์ และใช้ ศัพท์ไม่สม่ำเสมอ ChatGPT ภาษา สั่นไหวแบบเจ้าของ ภาษาใช้ Passive Voice ได้เหมาะสม กับงานวิจัย และมี ความแม่นยำ ของศัพท์สูง	เน้นเรียนรู้เทคนิค Prompting และ Post - Editing การตรวจทาน ขั้นสุดท้าย โดยผู้เชี่ยวชาญ การพัฒนา ฐานข้อมูลศัพท์ เพื่อลด ความไม่สม่ำเสมอ	Basic: “ช่วยแปล บทความนี้... โดยใช้ สำนวนทางวิชาการ” Specific: “คง โครงสร้าง ตามต้นฉบับ และใช้ ศัพท์ทางวิชาการ” Refine: “เน้นการ ใช้ Passive Voice”, “ปรับประโยค ให้อ่านสั่นไหว แบบ Native Speaker”	Google: แปล คร่าว ๆ ได้ความหมายเร็ว DeepL: เด่นเรื่อง โครงสร้างประโยค ธรรมชาติ Microsoft: ใช้ง่าย ใน Word แต่คุณภาพ ใกล้เคียง Google

ภาคผนวก ง
ตารางข้อมูลตีความเห็นเพิ่มเติม
ของผู้เชี่ยวชาญในการประเมินบทความคัดย่อรายฉบับ



ความเห็นเพิ่มเติมของผู้เชี่ยวชาญในการประเมินบทความฉบับ

บทความ (ฉบับที่)	ผู้เชี่ยวชาญ (รายที่)	ความเห็นเพิ่มเติม การแปลบทความโดยมนุษย์	ความเห็นเพิ่มเติม การแปลบทความโดย ChatGPT
1	1	เป็นการแปลเอาความ ไม่ได้แปล ทุกคำ ใช้ภาษาสั้นและกระชับกว่า หัวข้อวิจัยของฉบับที่ 1 เป็นไป ตามมาตรฐานการแปลของหัวข้อวิจัย ที่ใช้กันทั่วไป คือขึ้นต้นเป็นชื่อเรื่อง เลย ไม่ต้องมี “The” พวก Article นำ	แปลละเอียด แปลทุกคำจากภาษาไทย
	2	<ul style="list-style-type: none"> - พบข้อความที่ไม่ได้แปลจากต้นฉบับ ภาษาไทยคือ “ซึ่งสามารถอยู่ในพื้นที่ ได้ตลอดเวลาโดยไม่ส่งผลกระทบต่อ” ทำให้ การแปลจากต้นฉบับไม่สมบูรณ์ - จากประโยคต้นฉบับภาษาไทย “เป็นไปในแนวทางเดียวกัน” แปลเป็น อังกฤษ “related in the same direction” แม้จะเข้าใจได้ แต่ในทาง สถิติหรือวิทยาศาสตร์ถือว่าไม่แม่นยำ เท่าที่ควร - มีลักษณะการแปลตรงตัวเกินไป จากประโยค “to find safety distance” ควรใช้คำที่เป็นวิชาการ กว่านี้ เช่น “to determine the safe distance” - ใช้คำศัพท์ทั่วไป ไม่เป็นศัพท์เฉพาะ “high voltage electrical tower” ซึ่งต้นฉบับเน้นที่สายส่ง (Transmission lines) มากกว่า ตัวเสา 	<ul style="list-style-type: none"> - แปลได้ถูกต้อง เก็บรายละเอียดครบถ้วน แปลคำว่า “ผลจากการวัดจริง” เป็น “actual field measurements” ซึ่งสื่อความหมายทางวิทยาศาสตร์ - ใช้ภาษาในระดับสูง ใช้คำว่า “impact” แทน “effects” หรือ “safety thresholds” แทน “standard”
	3	ตำแหน่งของข้อความไม่ตรง กับต้นฉบับ และข้อความไม่ค่อย สมบูรณ์	ตรงและสมบูรณ์

ความเห็นเพิ่มเติมของผู้เชี่ยวชาญในการประเมินบทความฉบับ

บทความ (ฉบับที่)	ผู้เชี่ยวชาญ (รายที่)	ความเห็นเพิ่มเติม การแปลบทความโดยมนุษย์	ความเห็นเพิ่มเติม การแปลบทความโดย ChatGPT
	4	-	-
	5	<ul style="list-style-type: none"> - มีข้อความที่ไม่ได้แปลในส่วน ผลการศึกษา ส่วนท้ายของประโยค ที่ว่า “ซึ่งสามารถอยู่ในพื้นที่ได้ ตลอดเวลาโดยไม่ส่งผลกระทบต่อ” - ใช้ศัพท์ไม่สอดคล้องและไม่ตรง บริบททางเทคนิค ใช้ high voltage electrical tower แทน Transmission lines สายส่งจ่าย ไฟฟ้าแรงสูง 	<p>คุณภาพดี มีความแม่นยำ ภาษาเป็น ธรรมชาติ แต่มีการเติมคำว่า “power” ในหัวข้อ high - voltage power transmission Line Areas ซึ่งไม่มี ในต้นฉบับภาษาไทย</p>
	6	<ul style="list-style-type: none"> - แปลได้ใกล้เคียงต้นฉบับ แต่บางส่วนคลาดเคลื่อน เช่น ประโยค “find safety distance” แปลไม่เป็นธรรมชาติ ควรเป็น “determine the safe distance” - ไวยากรณ์บางส่วนไม่สมบูรณ์ เช่น “It indicated that at the distance...” 	<ul style="list-style-type: none"> - แปลได้ตรงกับต้นฉบับ ๐ ภาษาเป็นธรรมชาติ สั้นไหล และ สอดคล้องกับการเขียนเชิงวิชาการ - ใช้คำศัพท์วิศวกรรมตรงและสม่ำเสมอ เช่น “electric and magnetic field strengths,” “high - voltage power transmission lines,” “safety thresholds.” เป็นต้น
	7	-	-
	8	มีประโยคที่แปลไม่ชัดเจนและไม่เป็น ธรรมชาติ และมีการละเลยการแปล เนื้อหาสำคัญ	แปลได้ถูกต้อง สั้นไหลกว่าแบบที่ 1 คำว่า high - voltage power transmission line areas ตรงกับภาษาทางวิศวกรรม
	9	มีข้อผิดพลาดเล็กน้อยในการรวม ประโยคทำให้ซับซ้อน โครงสร้าง ประโยค ยังไม่เป็นธรรมชาติ เท่าที่ควร	ไม่พบข้อผิดพลาด ถ่ายทอดความหมายได้ ครบถ้วน ภาษาเป็นธรรมชาติ
	10	สั้นกว่าฉบับที่ 2	ยาวกว่าฉบับที่ 1 และคำแปลตรงต้นฉบับ ภาษาไทย และเป็นไปได้ว่ามีกรอแก้ไข

ความเห็นเพิ่มเติมของผู้เชี่ยวชาญในการประเมินบทความฉบับ

บทความ (ฉบับที่)	ผู้เชี่ยวชาญ (รายชื่อ)	ความเห็นเพิ่มเติม การแปลบทความโดยมนุษย์	ความเห็นเพิ่มเติม การแปลบทความโดย ChatGPT
			เพิ่มเติมโดยมนุษย์ แต่ส่วนใหญ่จะเป็น word for word translation
	11	Omission และ Untranslated Text มีความใกล้เคียงกันในการตีความ และในทางทฤษฎีการแปล (Mona Baker) Omission อาจถือเป็นกลยุทธ์หนึ่ง แนะนำให้พิจารณาการให้คะแนนในส่วนนี้ อย่างระมัดระวัง	-
	12	-	-
	13	-	-
	14	<p>- ต้นฉบับไทยระบุว่า “ด้วยโปรแกรมจำลองคอมพิวเตอร์ และผลจากการวัดจริง...” ในแบบแปล “with a computer simulation program and measurements.” คำว่า “ผลจากการวัดจริง” ควรถูกระบุอย่างชัดเจนว่าเป็น actual field measurements ไม่ใช่แค่ “measurements”</p> <p>- บางประโยคมีลักษณะการแปลตรงตัวจากไทย เช่น “to find safety distance of the residents...” ควรปรับเป็น “to determine the safe distance for residents...” เพื่อความเป็นธรรมชาติและวิชาการ</p> <p>- “the voltage of 24 and 69 kV” ไม่เป็นธรรมชาติในบริบท</p>	ภาษาสั่นไหวตีความในเชิง Academic ครอบคลุมเนื้อหาได้ครบถ้วน อย่างไรก็ตาม บางจุดมีลักษณะของการตีความเพิ่มเติมเล็กน้อย เช่น การสรุปผลกระทบที่ไม่ได้ปรากฏตรง ๆ ในต้นฉบับ

ความเห็นเพิ่มเติมของผู้เชี่ยวชาญในการประเมินบทคัดย่อรายฉบับ

บทคัดย่อ (ฉบับที่)	ผู้เชี่ยวชาญ (รายที่)	ความเห็นเพิ่มเติม การแปลบทคัดย่อโดยมนุษย์	ความเห็นเพิ่มเติม การแปลบทคัดย่อโดย ChatGPT
		วิชาการ ควรใช้รูปแบบที่เป็นมาตรฐาน เช่น “24 and 69 kV voltage levels” หรือ “voltages of 24 kV and 69 kV”	
	15	<ul style="list-style-type: none"> - แปลผิด “เป็นไปในแนวทางเดียวกัน” แปลเป็น "related in the same direction" ซึ่งไม่ถูกต้อง ควรเป็น “consistent” หรือ “aligned” - บางประโยคยาวเกินไปและไม่ค่อยเป็นธรรมชาติ - ศัพท์เฉพาะส่วนใหญ่ถูกต้อง 	แปลได้ดี ถูกต้องครบถ้วน ใช้ภาษาธรรมชาติ อ่านง่าย ใช้ศัพท์ สอดคล้องตลอดทั้งบทคัดย่อ
	16	<p>ประโยค “the measurement results of the electric and magnetic fields of the voltage of 24 and 69 kV were also related in the same direction” สื่อความหมายไม่ชัดเจน และแปลคลาดเคลื่อนจากต้นฉบับที่ต้องการจะสื่อว่า “ผลการวัดและการจำลองเป็นไปในทิศทางเดียวกัน” ไม่ใช่ “related in the same direction” ควรแก้ไขเป็น “The measurement results... were consistent with the computer simulation results.”</p>	จากประโยค “residents can safely remain in the area without adverse effects” มีการขยายความเกินจากต้นฉบับเล็กน้อย ซึ่งต้นฉบับไม่ได้กล่าวถึง “without adverse effects” อย่างชัดเจน
	17	-	-
	18	-	-

ความเห็นเพิ่มเติมของผู้เชี่ยวชาญในการประเมินบทความฉบับ

บทความ (ฉบับที่)	ผู้เชี่ยวชาญ (รายที่)	ความเห็นเพิ่มเติม การแปลบทความโดยมนุษย์	ความเห็นเพิ่มเติม การแปลบทความโดย ChatGPT
2	1	แปลได้สั้น และกระชับ ครอบคลุม	แปลแบบได้รายละเอียดเยอะครบถ้วน ตามเวอร์ชันภาษาไทยแต่ควรมีการวงเล็บ คำที่เป็น Technical Term ด้วย
	2	แปลได้ดี แปลศัพท์เทคนิคซับซ้อน ได้ถูกต้อง เช่น ชื่อสารเคมี ภาษา เป็นวิชาการ	- แปลได้ดีแต่มีข้อบกพร่องเรื่อง ศัพท์เทคนิค ละเลยคำว่า micro ใน microsuspension - ในหัวข้อวิทยานิพนธ์ใช้ “Thermal Energy Storage” แต่ในส่วนบทความ และ Keywords ใช้ “thermal energy storage material” และ “thermal energy storage”
	3	-	-
	4	-	-
	5	-	-
	6	-	-
	7	-	-
	8	-	-
	9	-	-
	10	แปลได้ตรงกับฉบับภาษาไทย และ ถูกต้องตามหลักไวยากรณ์	-
	11	-	-
	12	-	-
	13	-	-

ความเห็นเพิ่มเติมของผู้เชี่ยวชาญในการประเมินบทความฉบับ

บทความ (ฉบับที่)	ผู้เชี่ยวชาญ (รายที่)	ความเห็นเพิ่มเติม การแปลบทความโดยมนุษย์	ความเห็นเพิ่มเติม การแปลบทความโดย ChatGPT
	14	จากประโยคต้นฉบับภาษาไทย "... ด้วยกระบวนการสังเคราะห์แบบ แวนลอยโดยกลไกโยกย้ายสายโซ่ ไอโอดีน..." แปลเป็น "...produced by micro- suspension iodine transfer polymerization..." แม้จะกล่าวถึง "iodine transfer polymerization" แต่ไม่สื่อชัดเจนถึง "chain transfer mechanism"	ไม่มีการแปลผิดหรือละเลยข้อมูลสำคัญ ภาษามีความเป็นวิชาการ แต่ประโยค "...show promise for applications in thermal energy storage..." คำว่า "show promise" ค่อนข้างไม่เป็นกลาง และให้ความรู้สึก "คาดการณ์" มากกว่า รายงานข้อเท็จจริง
	15	- คำว่า "วัสดุเก็บความร้อนชีวภาพ" จากที่แปลเป็น bio- heat storage material ควรแปลเป็น "bio- based thermal energy storage material จะถูกต้องมากกว่า - ขาดคำว่า chain mechanism ในกระบวนการสังเคราะห์ - High percentages of EE, 70% โครงสร้างประโยคไม่เป็นธรรมชาติ ควรเป็น encapsulation efficiency exceeding 70%	-
	16	การแปลโดยรวมเข้าใจได้ แต่ควรปรับปรุงการใช้ศัพท์เทคนิคให้แม่นยำมากขึ้น	แนะนำให้ใช้ Back Translation เพื่อตรวจสอบความหมายย้อนกลับอีกครั้ง เพื่อให้แน่ใจว่าข้อความแปลมีความถูกต้องตามต้นฉบับ และใช้ Termbase หรือ Translation Memory เฉพาะทาง เพื่อให้การใช้ศัพท์สอดคล้องกัน
	17	-	-
	18	-	-

ความเห็นเพิ่มเติมของผู้เชี่ยวชาญในการประเมินบทความรายฉบับ

บทความ (ฉบับที่)	ผู้เชี่ยวชาญ (รายที่)	ความเห็นเพิ่มเติม การแปลบทความโดยมนุษย์	ความเห็นเพิ่มเติม การแปลบทความโดย ChatGPT
3	1	ใช้คำฟุ่มเฟือยและไม่กระชับ มีแปลผิดตรง 0.3% แปลเป็น 3%	- ชื่อเรื่อง มีบางคำอาจตัดทิ้งได้ ไม่ควรใช้ในชื่อเรื่อง เช่นคำว่า “Culture” - เวอร์ชันนี้แปลได้กระชับกว่า เช่น การใช้คำ
	2	- แปลต้นฉบับภาษาไทย “ปริมาณ 0.3 เปอร์เซ็นต์ผสมกับอาหาร” เป็น “at the concentration of 3%” (สามเปอร์เซ็นต์) ซึ่งผิดจากต้นฉบับคือ 0.3% (ศูนย์จุดสามเปอร์เซ็นต์) การผิดพลาดของตัวเลขนี้มีผลกระทบอย่างมาก ต่อความน่าเชื่อถือของงานวิจัย - แปลต้นฉบับภาษาไทย “น้ำล้างการแปรรูปปลา” แปลเป็น “Leftover water from Fish Processing Production” ควรใช้ Fish Processing Wastewater หรือ Fish processing wash water” มากกว่า	แปลได้แม่นยำสูง รายละเอียดเชิงชีววิทยาและสถิติถูกต้อง ภาษาเป็นธรรมชาติ
	3	-	-
	4	-	-
	5	การแปลตรงตัวแต่ทำให้ความหมายผิด เช่นคำว่า Leftover water	-
	6	ความเข้มข้นของโปรไบโอติกในอาหารระบุว่า 3% ผิดจากต้นฉบับที่ 0.3%	-
	7	-	-

ความเห็นเพิ่มเติมของผู้เชี่ยวชาญในการประเมินบทความฉบับ

บทความ (ฉบับที่)	ผู้เชี่ยวชาญ (รายที่)	ความเห็นเพิ่มเติม การแปลบทความโดยมนุษย์	ความเห็นเพิ่มเติม การแปลบทความโดย ChatGPT
	8	-	-
	9	<p>- คำว่า “น้ำล้างการแปรรูปปลา” แปลเป็น “Leftover water from fish processing production” ซึ่งมีความหมายที่ใกล้เคียง แต่คำว่า “leftover water” อาจสื่อถึงน้ำทิ้งที่เหลือใช้มากกว่า “น้ำล้าง” ที่ใช้ในการแปรรูป</p> <p>- มีการแปลผิดที่สำคัญ “ความเป็นกรดต่างตั้งแต่ 2 - 10% (NaCl) 0-10 (w/v)” ถูกแปลว่า "acid - base properties of 2- 10% (NaCl) 0-10 (w/v)" ซึ่งผิดความหมายอย่างชัดเจน เพราะ 2 - 10 คือค่า pH แต่ไม่ได้ถูกแปล</p>	<p>- โดยรวมแล้วมีความถูกต้องดีมาก แต่มีจุดที่ขาดความแม่นยำเล็กน้อย เช่น น้ำล้างการแปรรูปปลา (fish processing wastewater) เป็นการแปลที่ถูกต้อง แต่ยังไม่ครอบคลุม คำว่า leftover water from fish processing production ที่อาจจะมีความหมายกว้างกว่า</p> <p>- มีการละเว้นเนื้อหาบางส่วน เช่น ต้นฉบับระบุ “เจริญเติบโต ในความเป็นกรดต่างตั้งแต่ 2 - 10% (NaCl) 0- 10 (w/v) ในน้ำที่สวดปลอดเชื้อของปลา” แต่ในฉบับแปล มีการแปลที่ยังไม่ครบถ้วน ในเรื่องของตัวแปร โดยเฉพาะส่วนของ “pH values” ที่ในต้นฉบับไม่ได้ระบุเป็นเปอร์เซ็นต์</p>
	10	แปลคำต่อคำมากกว่า ระบุข้อมูลผิดพลาด เช่น 3%	ใช้คำเชิงวิชาการได้ดีกว่า แม้ผิดพลาดเรื่องการเว้นวรรค เครื่องหมายต่าง ๆ
	11	-	-
	12	-	-
	13	-	-
	14	-	-
	15	-	-
	16	-	-

ความเห็นเพิ่มเติมของผู้เชี่ยวชาญในการประเมินบทความฉบับ

บทความ (ฉบับที่)	ผู้เชี่ยวชาญ (รายที่)	ความเห็นเพิ่มเติม การแปลบทความโดยมนุษย์	ความเห็นเพิ่มเติม การแปลบทความโดย ChatGPT
	17	เนื่องจากมีศัพท์เฉพาะศัพท์เทคนิค ตลอดทั้งบทความทำให้ผู้แปลแปลผิด ในทุกหัวข้อ	-
	18	-	-
4	1	ชื่อเรื่องภาษาอังกฤษไม่ตรงกับภาษาไทย คำว่า เปลือกหุ้มฝัก มะรุม และ morning pods	มีแปลผิดเช่น design (ออกแบบ) ในภาษาไทยใช้ ประดิษฐ์
	2	- ต้นฉบับคือ เปลือกฝักมะรุม แปล เป็น Moringa Pods (ฝักมะรุม) ละเลยส่วนของเปลือกไป - ใช้คำว่า artisan products และ customer satisfaction แทน งานประดิษฐ์ และ ความพึงพอใจ ของผู้บริโภค ถือว่ารับได้ แต่การละเลย คำว่า shell/husk ใน หัวข้อและเนื้อหาเป็นข้อบกพร่องด้าน ศัพท์เฉพาะที่สำคัญ	แปลได้ดี มีความแม่นยำทางศัพท์สูง แปล Moringa Pod Shells ถูกต้อง ภาษาเป็นธรรมชาติ
	3	-	-
	4	-	-
	5	- เปลือกฝักมะรุม แปลเป็น Moringa Pods (ฝักมะรุม) แทนที่จะเป็น Moringa Pod Shells (เปลือกฝัก มะรุม) - มีโครงสร้างประโยคที่ดูเหมือนแปล ตรงตัวและไม่สั่นไหวเท่าที่ควร เช่น The factorial in CRD method was applied to the experiments	ใช้คำว่า Handicraft Applications ตรงกับคำว่า งานประดิษฐ์

ความเห็นเพิ่มเติมของผู้เชี่ยวชาญในการประเมินบทความฉบับ

บทความ (ฉบับที่)	ผู้เชี่ยวชาญ (รายที่)	ความเห็นเพิ่มเติม การแปลบทความโดยมนุษย์	ความเห็นเพิ่มเติม การแปลบทความโดย ChatGPT
		of natural mosaics' physical properties...	
	6	มีบางประโยคไม่เป็นธรรมชาติ เช่น “the highest value of gloss depicts”	-
	7	-	-
	8	มีบางจุดที่แปลตรงเกินไป จนความหมายไม่ชัด เช่น “the highest value of gloss depicts was 2.43” มีคำกริยาแท้ซ้อนกัน สองตัวคือ depicts และ was	แปลได้ดี
	9	คุณภาพปานกลาง แปลคลาดเคลื่อน ในผลการวิจัย และใช้คำที่ไม่เป็นธรรมชาติ	คุณภาพดีเยี่ยม สื่อความหมายถูกต้อง ครบถ้วน ใช้ภาษาอังกฤษเชิงวิชาการได้ดี
	10	เวอร์ชันนี้แปลยังไม่สอดคล้องกับ ภาษาไทยต้นฉบับ ตอนท้าย ๆ ประโยคสุดท้ายแปลได้ตรงกว่า	หัวข้อดีกว่าอีกเวอร์ชัน เนื้อหาชัดเจนกว่า รายละเอียดถูกต้องตามต้นฉบับมากกว่า
	11	-	-
	12	-	-
	13	-	-
	14	จากการแปล “The Natural mosaic prototypes ... were wooden boxes ...” คำว่า “depicts” ใน “The highest value of gloss depicts”	ไม่พบข้อผิดพลาด สื่อสารข้อมูลตรงต้นฉบับ ภาษาสโลวไทล เป็นธรรมชาติ

ความเห็นเพิ่มเติมของผู้เชี่ยวชาญในการประเมินบทความฉบับ

บทความ (ฉบับที่)	ผู้เชี่ยวชาญ (รายที่)	ความเห็นเพิ่มเติม การแปลบทความโดยมนุษย์	ความเห็นเพิ่มเติม การแปลบทความโดย ChatGPT
		ใช้ไม่เหมาะสมกับบริบท เป็นการใช้ กริยาแทนคำนามโดยไม่จำเป็น	
	15	แปลผิดวัตถุดิบ (Moringa pods) และโครงสร้างประโยคผิด	-
	16	ให้ความหมายครบถ้วนแต่ใช้ โครงสร้างประโยคซ้ำซ้อน และเลือกคำที่อาจทำให้สับสน	สื่อสารได้ชัดเจน ลื่นไหล เป็นธรรมชาติ เหมาะกับการตีพิมพ์
	17	-	-
	18	เนื้อหาถูกต้องแต่ภาษาไม่เป็น ธรรมชาติ เช่นคำว่า gloss depicts และใช้ศัพท์ไม่มาตรฐาน เช่น softener, craft products	แปลถูกต้อง ลื่นไหล สอดคล้องกับ มาตรฐานการเขียนวิชาการ ใช้ศัพท์ถูกต้อง handicraft applications
5	1	- ชื่องานวิจัย ไม่ควรใส่ชื่อ ทางวิทยาศาสตร์ - บางข้อความในย่อหน้าที่ 2 ไม่ได้แปลตามฉบับภาษาไทย - บางข้อความในย่อหน้าที่ 3 ไม่ได้แปลตามฉบับภาษาไทย - ใน Keywords ไม่ควรใช้คำศัพท์ ทางวิทยาศาสตร์ (ตัวแรก) และใช้ อักษรตัวเล็กตามเกณฑ์การเขียน บทความของ RMUTT, ใช้ - ภาษาไทยใช้ “ผลกระทบ” ภาษาอังกฤษใช้ “effects” ผิดความหมาย สำนวนภาษา มีความกระชับกว่าอีกรเวอร์ชัน	วงเล็บผิด ไม่เป็นไปตามเงื่อนไขการเขียน บทความของ RMUTT เช่น (1) ควรเป็น 1)
	2	ละเลยหน่วย % ใน “ผลผลิตที่ผลิต ได้เท่ากับ 0.14 และ 18.45%”	แปลได้ดี ใช้คำศัพท์สอดคล้องทั้งเนื้อหา

ความเห็นเพิ่มเติมของผู้เชี่ยวชาญในการประเมินบทความฉบับ

บทความ (ฉบับที่)	ผู้เชี่ยวชาญ (รายที่)	ความเห็นเพิ่มเติม การแปลบทความโดยมนุษย์	ความเห็นเพิ่มเติม การแปลบทความโดย ChatGPT
		แปลเป็น “yield at 0.14 and 18.45, respectively.”	
	3	-	-
	4	-	-
	5	“น้ำมันสกัดรกจันทน์เทศ” แปลเป็น “Extracted Mace (Myristica Fragrans Houtt.) Oil” คำว่า Mace (ดอกจันทน์เทศ) เป็นส่วนที่ต่างจาก Aril (รกจันทน์เทศ)	คำว่า “น้ำมันสกัดรกจันทน์เทศ” แปลเป็น “Nutmeg Aril Oil Encapsulate” แปลตรงต้นฉบับ
	6	ใช้ mace oil กว้างไป ภาษาไม่เป็นธรรมชาติเท่าแบบที่ 1 มีความซ้ำซ้อน	-
	7	-	-
	8	- “extracted- mace oil” ซึ่งอาจสร้างความสับสนเพราะ mace = รกจันทน์เทศ แต่คำว่า nutmeg aril oil (แบบแปลที่ 1) ตรงกว่า - ใช้ศัพท์บางส่วนไม่สม่ำเสมอ เช่น gum Arabic (ตัว A ใหญ่) ไม่คงที่กับ gum arabic ที่ใช้ทั่วไป	- ตัวเลขตรงกับต้นฉบับครบถ้วน - ใช้คำว่า “nutmeg aril oil” ตรงกับ “น้ำมันสกัดรกจันทน์เทศ”
	9	-	-
	10	ฉบับนี้มีเนื้อหาที่ตรงกับภาษาไทยมากกว่า	ฉบับนี้ไม่มีข้อผิดพลาดด้านข้อมูลหลัก แต่อาจปรับปรุงด้านสำนวนการเรียงลำดับข้อมูล
	11	-	-
	12	-	-

ความเห็นเพิ่มเติมของผู้เชี่ยวชาญในการประเมินบทความฉบับ

บทความ (ฉบับที่)	ผู้เชี่ยวชาญ (รายที่)	ความเห็นเพิ่มเติม การแปลบทความโดยมนุษย์	ความเห็นเพิ่มเติม การแปลบทความโดย ChatGPT
	13	-	-
	14	มีความพยายามใช้โครงสร้างประโยค และศัพท์เทคนิคที่ดี แต่ยังมี ข้อผิดพลาดเล็กน้อยด้านการแปล ความหมายเฉพาะทาง และความไม่สอดคล้องของโครงสร้าง ไวยากรณ์ ควรตรวจสอบการใช้คำ ทางวิทยาศาสตร์ให้แม่นยำมากขึ้น	แปลถูกต้องครบถ้วน ไม่ละเว้นเนื้อหา ภาษาเป็นทางการ แต่บางประโยค อาจดูตรงตัวไปบ้าง
	15	-	-
	16	-	-
	17	-	-
	18	-	-
6	1	ชื่อบทความวิจัย “The Influence” ควรให้ตรงกับภาษาไทย	-
	2	-	-
	3	-	-
	4	-	-
	5	-	-
	6	โครงสร้างประโยคไม่เป็นธรรมชาติ เช่น “the results of the study were found that ...”	ใช้ The findings revealed that... จะเป็นธรรมชาติมากกว่า
	7	-	-
	8	-	-

ความเห็นเพิ่มเติมของผู้เชี่ยวชาญในการประเมินบทความฉบับ

บทความ (ฉบับที่)	ผู้เชี่ยวชาญ (รายที่)	ความเห็นเพิ่มเติม การแปลบทความโดยมนุษย์	ความเห็นเพิ่มเติม การแปลบทความโดย ChatGPT
	9	-	-
	10	สามารถปรับสำนวนได้ เป็นทางวิชาการให้มากขึ้น	-
	11	-	-
	12	-	-
	13	-	-
	14	-	-
	15	-	-
	16	-	-
	17	-	-
	18	ภาพรวมแปลได้ดี แต่มีข้อสังเกต เล็กน้อย คำว่า “สายใยแก้วนำแสง” แปลเป็น LAN fiber optic cable systems แต่อีกแบบแปลเป็น fiber optic communication systems ซึ่งถ้าวิเคราะห์จากภาษาไทย พูดในภาพรวมถึงสายใยแก้วนำแสง ไม่ได้เจาะจงว่าเป็น LAN ดังนั้น fiber optic communication systems น่าจะตรงกว่า	แปลได้ตรงต้นฉบับ
7	1	- ชื่องานวิจัย ไม่กระชับเหมือน เวอร์ชันที่ 2 และการใช้คำไม่เป็น มาตรฐาน เช่น "ผลสัมฤทธิ์ทางการ เรียน" ควรใช้ “learning achievement”, รายวิชา ควรใช้	คำกระชับกว่าเวอร์ชันที่ 2

ความเห็นเพิ่มเติมของผู้เชี่ยวชาญในการประเมินบทความฉบับ

บทความ (ฉบับที่)	ผู้เชี่ยวชาญ (รายชื่อ)	ความเห็นเพิ่มเติม การแปลบทความโดยมนุษย์	ความเห็นเพิ่มเติม การแปลบทความโดย ChatGPT
		<p>“course” และไม่ควรรู้ใช้ “who” ในชื่อบทความวิจัย</p> <p>- แปลคำต่อคำตามเวอร์ชันภาษาไทย ใส่คำเยอะโดยไม่มีการละ ให้ข้อความกระชับ</p>	
	2	เปลี่ยนชื่อผู้สร้างทฤษฎีผิด Davies เป็น Dewey	เปลี่ยนชื่อผู้สร้างทฤษฎีถูกต้อง Davies
	3	-	-
	4	-	-
	5	<p>- เดวิส แปลเป็น Dewey ไม่ถูกต้อง</p> <p>- เครื่องมือข้อ 4 กลับกลายเป็น ประโยชน์ผลการวิจัย (ละเลยเครื่องมือ และแทรกผลการวิจัยเข้ามาแทนที่)</p> <p>- ตรงคำสำคัญใช้คำว่า Davies’ practical skill แต่ในเนื้อหาของบทความ กลับใช้คำว่า Dewey's practical skills</p>	-
	6	เปลี่ยนผิดพลาดสำคัญ Davies เป็น Dewey เนื้อหาสับสน เครื่องมือปณผลวิจัย รายละเอียด ไม่ครบ	แปลครบถ้วน
	7	-	-
	8	พบทั้ง Davies และ Dewey ทำให้ไม่สอดคล้องกันทั้งเนื้อหา	-
	9	-	-

ความเห็นเพิ่มเติมของผู้เชี่ยวชาญในการประเมินบทความฉบับ

บทความ (ฉบับที่)	ผู้เชี่ยวชาญ (รายที่)	ความเห็นเพิ่มเติม การแปลบทความโดยมนุษย์	ความเห็นเพิ่มเติม การแปลบทความโดย ChatGPT
	10	ผิดพลาดมากกว่าแปล 1	แปลได้ดีกว่าแปล 2 ข้อผิดพลาดน้อยกว่าแปล 2
	11	-	-
	12	-	-
	13	-	-
	14	โครงสร้างประโยคสับสนและซ้ำซ้อน	เนื้อหาครบถ้วน
	15	คุณภาพน้อย แปลผิด (Dewey), โครงสร้างเนื้อหาผิดพลาด (ผลวิจัย ปนในเครื่องมือ), สับสน	คุณภาพดีเยี่ยม แปลความหมายถูกต้อง ครบถ้วน ไม่พบข้อผิดพลาดสำคัญ
	16	มีข้อผิดพลาดสำคัญ Omission ข้อมูลเครื่องมือ Addition ผลวิจัย ผิดที่ และแปลศัพท์ผิด (Dewey)	-
	17	-	-
	18	มีข้อผิดพลาดในชื่อบุคคล Dewey ข้อมูลสับสน เครื่องมือปนผลลัพธ์ ภาษาไม่ลื่นไหล	แปลได้ดี
8	1	บางข้อความแปลไม่เหมือนเวอร์ชัน ภาษาไทย การใช้เครื่องหมายวงเล็บ ไม่ตรงกับเกณฑ์ของ RMUTT	-
	2	- คำว่า interviewees อาจสื่อได้ว่า ทำการวิจัยกับตัวบุคคลมากกว่า จะหมายถึงการสัมภาษณ์ ควรใช้ key informants เพื่อความชัดเจน และเป็นทางการมากขึ้น	เขียน priority needs index (PNI) แบบสากล แต่ยึดตามรูปแบบมหาวิทยาลัย จะเหมาะสมมากกว่า

ความเห็นเพิ่มเติมของผู้เชี่ยวชาญในการประเมินบทความฉบับ

บทความ (ฉบับที่)	ผู้เชี่ยวชาญ (รายที่)	ความเห็นเพิ่มเติม การแปลบทความโดยมนุษย์	ความเห็นเพิ่มเติม การแปลบทความโดย ChatGPT
		<p>- ความต้องการจำเป็น เวอร์ชันนี้ เขียนแบบนี้ priority needs index (PNI_{modified}) แต่อีกเวอร์ชัน เขียนแบบนี้ priority needs index (PNI) ซึ่งจากการตรวจสอบบทความที่ผ่านมาจะพบ priority needs index (PNI_{modified}) การเขียนรูปแบบนี้มากกว่าและยึดว่ามหาวิทยาลัยเราจะเขียนรูปแบบนี้ จะมีความเป็นสากลมากกว่า</p>	
	3	-	-
	4	-	-
	5	<p>- การแปลเครื่องมือว่า The research tool utilized a dual-response format questionnaire with a 5- point rating scale ซึ่งในต้นฉบับคือ แบบสอบถามที่มีลักษณะแบบตอบสนองคู่ (Paired-Response) แต่การเพิ่มคำว่า with a 5 - point rating scale เข้ามาถือเป็นการเพิ่มข้อมูลที่ไม่มีในต้นฉบับ</p> <p>- priority needs index (PNI_{modified}) ใช้รูปแบบนี้ถูกต้อง</p>	
	6	<p>แปลตรงตัวเกินไป เช่น “was good and ... was excellent” ฟังไม่เป็นภาษาวิชาการเท่า high level / highest level</p>	
	7	-	-

ความเห็นเพิ่มเติมของผู้เชี่ยวชาญในการประเมินบทความฉบับ

บทความ (ฉบับที่)	ผู้เชี่ยวชาญ (รายที่)	ความเห็นเพิ่มเติม การแปลบทความโดยมนุษย์	ความเห็นเพิ่มเติม การแปลบทความโดย ChatGPT
	8	มีการเพิ่มข้อความ with a 5 - point rating scale ซึ่งไม่ปรากฏในต้นฉบับ	-
	9	-	-
	10	แปล 1 มีข้อผิดพลาดน้อยกว่า แปล 2	-
	11	-	-
	12	-	-
	13	-	-
	14	ประโยค interviewees กลุ่มเครือ ไวยากรณ์ซับซ้อน ใช้ศัพท์ ไม่สม่ำเสมอ	แปลข้อมูลครบถ้วนและตรงตามเนื้อหา ใช้ภาษาสละสลวยในเชิงวิชาการ
	15	-	-
	16	ประโยคกลุ่มเครือเรื่องผู้ให้สัมภาษณ์ ไวยากรณ์ซับซ้อนเกินไป ใช้คำไม่สม่ำเสมอ	-
	17	-	-
	18	มีความหมายครบ แต่คลาดเคลื่อน ทางข้อมูล 5- point scale จำนวนไม่ เป็นธรรมชาติ และใช้ศัพท์ไม่ สม่ำเสมอ	ดีเยี่ยม ถูกต้อง ครบถ้วน สั้นไหล ใช้ศัพท์ มาตรฐาน เหมาะสมสำหรับงานวิชาการ
9	1	หัวข้อวิจัยเป็นไปตามมาตรฐาน RMUTT	หัวข้อวิจัยใช้คำว่า standard affairs Administration จะกระชับกว่าเวอร์ชัน ที่ 1
	2	PNI _{Modified} ถูกต้องตามมหาวิทยาลัย	priority needs index แปลตามสากล

ความเห็นเพิ่มเติมของผู้เชี่ยวชาญในการประเมินบทคัดย่อรายฉบับ

บทคัดย่อ (ฉบับที่)	ผู้เชี่ยวชาญ (รายที่)	ความเห็นเพิ่มเติม การแปลบทคัดย่อโดยมนุษย์	ความเห็นเพิ่มเติม การแปลบทคัดย่อโดย ChatGPT
	3	-	-
	4	- แปลชื่อหน่วยงานถูกต้อง - การใช้ศัพท์อย่าง “educational institutions” แทน “schools” ในบริบทของประเทศไทยควร หลีกเลี่ยง เพราะอาจก่อให้เกิด ความสับสน	- ใช้ชื่อหน่วยงานว่า Pathum Thani Secondary Educational Service Area Office ซึ่งไม่ถูกต้องตามการเรียงลำดับชื่อ หน่วยงานในประเทศไทย ต้องเป็น Secondary Educational Service Area Office Pathum Thani - มีการใช้ประโยคที่ดีและสั้นไหล แต่บางช่วงอาจยาวเกินไป
	5	-	-
	6	แปลชื่อเฉพาะได้ถูกต้อง	แปลชื่อเฉพาะยังไม่ถูก
	7	-	-
	8		
	9		
	10	แปล 1 แปลได้ดีกว่า	-
	11	-	-
	12	-	-
	13	-	-
	14	-	-
	15	-	-
	16	-	-
	17	-	-

ความเห็นเพิ่มเติมของผู้เชี่ยวชาญในการประเมินบทความฉบับ

บทความ (ฉบับที่)	ผู้เชี่ยวชาญ (รายที่)	ความเห็นเพิ่มเติม การแปลบทความโดยมนุษย์	ความเห็นเพิ่มเติม การแปลบทความโดย ChatGPT
	18	ภาษาอ่านเข้าใจได้ดี ลื่นไหล แต่บางช่วงยาวไป แปลชื่อ หน่วยงานไทยได้ถูกต้อง	แปลได้ถูกต้องครบถ้วน แต่ชื่อ หน่วยงานไทยอาจจะแปลไม่ถูกต้อง
10	1	- เป็นไปตามเกณฑ์การเขียนบทความ ของ RMUTT - ใน Keywords ควรใช้คำ ให้เหมือนกับในชื่องานวิจัย “Inmate visitation” กับ “Prison visit” - หัวข้องานวิจัยของเวอร์ชันที่ 2 จะสั้นและกระชับกว่าเวอร์ชันที่ 1 - ไม่มีการใส่วงเล็บในหัวข้อย่อ เช่น 1)...2) - คำว่า Province ตัว p ต้องเป็น ตัวเล็ก	-
	2	- ในวัตถุประสงค์ข้อ 2 มีการแปลว่า personal factors of inmates, relatives (ปัจจัยส่วนบุคคลของ ผู้ต้องขัง,ญาติ) ทั้งที่ต้นฉบับระบุ เพียง “ปัจจัยส่วนบุคคล” (โดยเจาะจงที่ญาติผู้ต้องขัง) - การเรียบเรียงประโยค มีความสละสลวยและเป็นธรรมชาติ เชิงวิชาการสูงในภาพรวม	คุณภาพดีเยี่ยม ใช้ชื่อสถาบันถูกต้อง (Central Youth Correctional Institution) แปลมิติการบริการสมบูรณ์ ภาษาชัดเจน
	3	-	-
	4	-	-
	5	ชื่อหน่วยงาน ทัณฑสถานวัยหนุ่ม กลาง แปลเป็น Central Youth Correctional Institution ในหัวข้อ	ใช้ Central Youth Correctional Institution สอดคล้องและถูกต้อง ตลอดทั้งบทความ

ความเห็นเพิ่มเติมของผู้เชี่ยวชาญในการประเมินบทความฉบับ

บทความ (ฉบับที่)	ผู้เชี่ยวชาญ (รายชื่อ)	ความเห็นเพิ่มเติม การแปลบทความโดยมนุษย์	ความเห็นเพิ่มเติม การแปลบทความโดย ChatGPT
		วิทยานิพนธ์ แต่มีการเปลี่ยน เป็น Central Young Offender Institution ในส่วนผลการวิจัย	
	6		
	7	-	-
	8		
	9		
	10	แปล 2 แปลในภาพรวมได้ดีกว่า	-
	11	-	-
	12	-	-
	13	-	-
	14	“ของทัณฑสถานวัยหนุ่มกลาง จังหวัดปทุมธานี” แปลว่า at the Central Youth Correctional Institution in Pathum Thani Province แต่ไปใช้สลับเป็น “Central Juvenile Correctional Institution” และ “Central Young Offender Institution” ซึ่งไม่สอดคล้องกับชื่อสถาบันจริง ที่ใช้ในประเทศไทย	-
	15	-	-
	16	Central Juvenile Correctional Institution แปลคลาดจาก	ไม่มีข้อผิดพลาดสำคัญด้านการแปล ความหมายตรงกับต้นฉบับ

ความเห็นเพิ่มเติมของผู้เชี่ยวชาญในการประเมินบทความฉบับ

บทความ (ฉบับที่)	ผู้เชี่ยวชาญ (รายที่)	ความเห็นเพิ่มเติม การแปลบทความโดยมนุษย์	ความเห็นเพิ่มเติม การแปลบทความโดย ChatGPT
		“ทัศนสถานวัยหนุ่มกลาง” ซึ่งใช้คำว่า Youth จะเหมาะสมกว่า Juvenile	
	17	-	-
	18	<p>- สำนวนดี แต่มีความไม่สม่ำเสมอในการใช้ศัพท์ชื่อหน่วยงาน ในวัตถุประสงค์ข้อ 2 เขียนว่า “personal factors of inmates, relatives” ทำให้ความหมายเพี้ยน เพราะต้นฉบับหมายถึง “ญาติผู้ต้องขัง” ไม่ใช่รวมตัวผู้ต้องขังด้วย</p> <p>- โครงสร้างประโยคชัดเจน เป็นทางการ ใช้ Passive อย่างเหมาะสม</p>	<p>ระบุ “relatives of inmates” ชัดเจน ตรงกับต้นฉบับ</p> <p>ภาษาเป็นธรรมชาติ กระชับกว่าฉบับที่ 1 เล็กน้อย</p>



ประวัติผู้วิจัย

ชื่อ-นามสกุล	นางสาวธัญภักษณ์ สารชาติ
ประวัติการศึกษา	วิทยาศาสตรบัณฑิต (วิทยาการคอมพิวเตอร์) มหาวิทยาลัยขอนแก่น ศิลปศาสตรมหาบัณฑิต (ภาษาอังกฤษ) มหาวิทยาลัยราชภัฏบุรีรัมย์
ตำแหน่ง	นักวิชาการศึกษา ระดับปฏิบัติการ
สถานที่ทำงาน	ฝ่ายวิชาการและวิจัย สำนักบัณฑิตศึกษา มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี โทรศัพท์ 0-2549-3618
Email Address	thanyaphak_s@rmutt.ac.th

