รายงานฉบับสมบูรณ์
เรื่อง การวิเคราะห์ความแตกต่างของความคิดและทัศนคติทางการเมือง ระหว่างประชากรต่างช่วงวัย

โดย

ผู้ช่วยศาสตราจารย์ ดร. วสะ บูรพาเดชะ

ดร. อนันต์ โอสถศิลป์

มหาวิทยาลัยอัสสัมชัญ

เสนอ

สถาบันพระปกเกล้า

ธันวาคม 2563

สัญญาเลขที่ พป 119/2563

รายงานฉบับสมบูรณ์ (Final Report)

เรื่อง การวิเคราะห์ความแตกต่างของความคิดและทัศนคติทางการเมืองระหว่างประชากรต่างช่วงวัย

ผู้เสนอโครงการวิจัย	**สังกัด**
1. ผศ. ดร. วสะ บูรพาเดชะ (หัวหน้าโครงการวิจัย)	มหาวิทยาลัยอัสสัมชัญ
2. ดร. อนันต์ โอสถศิลป์ (นักวิจัย)	มหาวิทยาลัยอัสสัมชัญ

บทสรุปผู้บริหาร

การแบ่งขั้วความคิดทางการเมือง (Political Polarization) ถือเป็นปรากฏการณ์ที่เกิดขึ้นเป็นเรื่องปรกติในสังคมประชาธิปไตยที่มีความขัดแย้งทางการเมืองสูง ในบริบทสังคมของไทยในช่วงสิบกว่าปีที่ผ่านมานั้นเกิดปรากฏการณ์ “การเมืองสีเสื้อ (color politics)” อย่างไรก็ตามในศตวรรษที่ 21 สิ่งที่พิเศษเพิ่มเติมจากความขัดแย้งในรูปแบบเดิมคือการก่อตัวขึ้นของกลุ่มคนรุ่นใหม่ (younger generation) ที่เติบโตขึ้นมาอยู่ในวัยที่สามารถสะท้อนเสียงและความคิดของตัวเองได้ พวกเขาก้าวเข้ามามีบทบาททางการเมืองและการปกครองของประเทศอย่างมีนัยยะสำคัญ และด้วยความก้าวหน้าของการพัฒนาทางวิทยาศาสตร์และเทคโนโลยี บทบาทของกลุ่มคนรุ่นใหม่ในการเมืองไทยจึงถูกแสดงออกผ่านนวัตกรรมและเทคโนโลยีใหม่ๆ ยิ่งผู้คนโดยเฉพาะในหมู่กลุ่มคนรุ่นใหม่นั้นหันมาใช้ชีวิตในโลกโซเชียลมากยิ่งขึ้น ยิ่งก่อให้การสื่อสารที่ก่อผลกระทบในวงกว้างได้รุนแรงและรวดเร็วยิ่งขึ้น บทบาทสำคัญของอายุนั้นจึงเป็นปัจจัยที่ควรจะศึกษา ยิ่งเมื่อมันมีผลโดยตรงไม่ใช่แค่ในระดับปัจเจกบุคคลแต่ตั้งเป้าหมายไปถึงการเปลี่ยนแปลงทางการเมืองในระดับชาติด้วย

นวัตกรรมด้านวิทยาศาสตร์ข้อมูล (Data science) ได้เข้ามามีบทบาทกับองค์กรทางธุรกิจและเป็นเครื่องมือสำคัญในการเพิ่มประสิทธิภาพให้องค์กร ข้อมูลโดยเฉพาะประเภทที่ไม่มีโครงสร้างแน่นอน (Unstructured data) ซึ่งถือเป็นข้อมูลส่วนใหญ่ที่เกิดขึ้นในโลกออนไลน์กลายเป็นข้อมูลสำคัญสามารถสร้างความเข้าใจพฤติกรรมผู้บริโภคหรือกลุ่มผู้ใช้บริการได้เป็นอย่างดี สำนักนวัตกรรมเพื่อประชาธิปไตย สถาบันพระปกเกล้าได้เล็งเห็นถึงประโยชน์และความสำคัญในการใช้นวัตกรรมด้านวิทยาศาสตร์ข้อมูลเพื่อสร้างความเข้าใจเกี่ยวกับความแตกต่างทางความคิดและทัศนคติทางการเมืองของประชาชนในช่วงวัยต่างๆ โดยในเบื้องต้นจะเริ่มศึกษาและพัฒนาระบบสารสนเทศเพื่อรวบรวมข้อมูลจากหน้าแฟนเพจของสถาบันพระปกเกล้าที่นำเสนอข้อมูลความรู้เกี่ยวกับประชาธิปไตยของสถาบันพระปกเกล้า และนำมาวิเคราะห์ความแตกต่างของความคิดและทัศนคติทางการเมืองระหว่างประชาชนต่างช่วงวัยโดยกระบวนการทางวิทยาศาสตร์ข้อมูล และนำเสอนผลการวิเคราะห์ด้วยภาพข้อมูล (Data Visualization) โดยผลการวิเคราะห์ข้อมูลพบประเด็นสำคัญดังนี้

1. กลุ่มประชากรที่มีความสนใจในเนื้อหาของสถาบันฯ ส่วนใหญ่นั้นอยู่ในกลุ่มคนรุ่นใหม่ (younger generation) หรือ Generation Y รวมทั้งหมดเป็นร้อยละ 64.51 ของข้อมูลทั้งหมด กลุ่มคนที่มีความสนใจรองลงมาคือกลุ่มวัยทำงานหรือ Generation X รวมทั้งหมดเป็นร้อยละ 17.47 ของข้อมูลทั้งหมด

2. กลุ่มประชากรที่สนใจเนื้อหาของสถาบันฯ เท่าที่สามารถระบุข้อมูลส่วนตัวได้จำนวน 238 ราย ส่วนหนึ่งมีภูมิหลังในสาขาวิชาที่เกี่ยวข้องเช่นด้านรัฐศาสตร์หรือนิติศาสตร์ รวมบัญชีผู้ใช้ทั้งหมด 65 ราย คิดเป็นร้อยละ 27.3

3. กลุ่มคนที่สนใจเป็นส่วนใหญ่ประกอบอาชีพอยู่ในส่วนราชการ ซึ่งจากข้อมูลที่ระบุได้ทั้งหมด 147 ราย มีจำนวนคนในกลุ่มนี้ 65 ราย คิดเป็นร้อยละ 44.2

4. ในภาพรวมขอข้อมูลเกี่ยวกับการโพสต์ข้อความบนหน้าแฟนเพจของสถาบันฯ นั้น พบว่าไม่มีตัวแปรไหนที่มีส่วนสัมพันธ์กันอย่างมีนัยสำคัญดัง ทั้งในด้านปริมาณข้อความ (จำนวนตัวอักษร) ประเภทของสื่อที่ใช้ในข้อความการโพสต์ (media) หรือประเภทของเนื้อหา (content category) ไม่มีส่วนสัมพันธ์กับจำนวนการมีส่วนร่วมแต่อย่างใดเช่นกัน

5. ในด้านสื่อที่ใช้ในการโพสต์ข้อความ พบว่าการโพสต์ข้อความและภาพเพียงภาพเดียว (Text & Image) มีค่าเฉลี่ยการมีส่วนร่วมมากที่สุด (15.31 ครั้งต่อโพสต์) แต่ก็มีส่วนเบี่ยงเบนมาตรฐานสูงเช่นกัน (SD = 9.77) แสดงให้เห็นว่ามีการกระจายตัวของการมีส่วนร่วมสูง

ด้วยข้อจำกัดของโครงการนี้คือความสามารถในการเข้าถึงข้อมูลส่วนตัวของผู้ใช้งานที่มีส่วนร่วมกับเนื้อหาแฟนเพจของสถาบันฯ ข้อมูลดังกล่าวจึงควรใช้วิเคราะควบคู่ไปกับการเก็บข้อมูลจากแหล่งบริการเครือข่ายสังคมอื่น (source triangulation) เช่นทวิตเตอร์ (Twitter) หรืออินสตาแกรม (Instagram) และเพื่อเพิ่มโอกาสในการเชื่อมโยงข้อมูลหลายแพลตฟอร์ม ทางสถาบันฯ อาจใช้เทคนิคด้านการตลาดผ่านหน้าเพจโดยให้รางวัลกับผู้ใช้ที่ตอบแบบสอบถามความคิดเห็นต่อเนื้อหาที่ทางสถาบันเผยแพร่ และขอคำยินยอมเปิดเผยข้อมูลเพื่อทำการเชื่อมโยงไปยังแพลตฟอร์มอื่นด้วย

Executive summary

Political polarization has been a phenomenon in modern democratic societies where political conflicts present. In Thailand, the phenomenon that emerged decades ago is known as "color politics."

At the turn of the 21st century, however, a new theme emerged as the younger generation became an integral part of the political landscape, took center stage, and participated in several movements. The movement is particularly apparent in the social media arena where the new generation, who are comfortable with technology and consider technology to be a necessary part of their lives, use their skills to spread their beliefs far and wide. This phenomenon is expected to be impactful beyond individual levels and could change the course of the nation.

King Prajadhipok's Institute (KPI) realized the importance of innovation in data science and how unstructured data found in social media could inform how different generations engage in political discourse. The study supports this initiative by collecting and analyzing unstructured data from the KPI's Facebook fan page to examine and visualize data of the engagement level among generations.

The findings include:

1. Younger generation was the most engaged group of citizens in political content shared by KPI's fan page, contributing 64.51 percent of the total samples. Generation X came second with 17.47 percent of the total.

2. From the samples that the team can identify personal information regarding schools or universities, 27.3 percent of the samples (65 of 238 persons) attended or are currently studying political science or law school.

3. From the samples that the team can identify personal information regarding work experiences, 44.2 percent (65 of 147 persons) are government officers.

4. There was no significant correlation among variables related to post information and level of engagement, including post length, media used, and content category.

5. In terms of media used in the posts, text accompanying by a single image has the highest engagement (15.31 times per post), albeit the high level of dispersion among the engagement level within this type of content (SD = 9.77).

With the limitation of scraping personal information from the Facebook fan page, it is suggested to triangulate information gleaned from Facebook with other social media platforms such as Twitter or Instagram. Using a lucky draw campaign as in the marketing industry would also increase the chance to triangulate user information across platforms.

สารบัญ

บทสรุปผู้บริหาร 1

Executive summary 3

สารบัญตาราง 7

สารบัญรูปภาพ 9

บทคัดย่อ 11

Abstract 12

บทที่ 1 บทนำ 13

1.1 หลักการและเหตุผล 13

1.2 วัตถุประสงค์ 14

1.3. ขอบเขตของการวิจัย 14

1.4. วิธีดำเนินการ 15

1.5 ระยะเวลาดำเนินงานและการส่งมอบงาน 16

1.6 คณะผู้วิจัย 16

บทที่ 2 ขั้นตอนของการวิจัย 17

2.1 การเก็บข้อมูล 17

2.2 ทำความสะอาดข้อมูล 27

2.3 การเตรียมข้อมูลเพื่อการวิเคราะห์ 31

2.4 สรุปขั้นตอนการเตรียมข้อมูล 37

บทที่ 3 การวิเคราะห์ข้อมูล 39

3.1 การวิเคราะห์สถิติเชิงพรรณาเกี่ยวกับการโพสต์ 39

3.2 การวิเคราะห์สถิติเชิงพรรณาเกี่ยวกับการผู้ใช้ที่มีส่วนร่วม 41

3.3 การวิเคราะห์ความสัมพันธ์ผู้ใช้งานที่มีส่วนร่วมกับโพสต์ด้วยกราฟเครือข่ายสังคม 50

บทที่ 4 สรุปผลการวิจัย 73
4.1 สรุปขั้นตอนการวิจัยและผลการวิเคราะห์ 73
4.2 ข้อจำกัดในการเก็บข้อมูลและการนำเสนอผลการวิเคราะห์ 75
4.3 ข้อเสนอแนะในอนาคต 75
บรรณานุกรม 77

สารบัญตาราง

ตาราง 1 รายการโพสต์ (project_scrape_level1) 27
ตาราง 2 รายการโปรไฟล์ของผู้มีส่วนร่วม (post_with_who_like) 28
ตาราง 3 รายละเอียดโปรไฟล์ของผู้มีส่วนร่วม (who_like_profile) 29
ตาราง 4 รายการโปรไฟล์ที่ไลก์เพจ (data_for_people_who_click_like_page) 29
ตาราง 5 รายการไลก์เพจ (fb_kpi_with_likes-all-results) 30
ตาราง 6 ประเภทของสื่อที่ใช้ในโพสต์ 34
ตาราง 7 ประเภทของเนื้อหาการโพสต์ 35
ตาราง 8 จำนวนการมีส่วนร่วมแบ่งตามประเภทเนื้อหาการโพสต์ 39
ตาราง 9 จำนวนการมีส่วนร่วมแบ่งตามประเภทประเภทสื่อที่ใช้ในโพสต์ 39
ตาราง 10 ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของสื่อที่ใช้ในโพสต์ 40
ตาราง 11 ตารางแสดงเมทริกซ์สหสัมพันธ์ของตัวแปรที่เกี่ยวกับข้อมูลการโพสต์และการมีส่วนร่วม 41
ตาราง 12 จำนวนและร้อยละของผู้ใช้งานที่มีส่วนร่วมแบ่งตามช่วงอายุ (รุ่น) 42
ตาราง 13 จำนวนและร้อยละของผู้ใช้งานที่มีส่วนร่วมแบ่งตามเพศ 42
ตาราง 14 การแจกแจงแบบไขว้ของเพศและอายุของผู้ใช้งาน 43
ตาราง 15 การศึกษาของผู้ใช้ที่อยู่ในกลุ่มรัฐศาสตร์และนิติศาสตร์แบ่งตามช่วงอายุ 44
ตาราง 16 ข้อมูลช่วงอายุและสถานที่ทำงานของผู้ใช้ที่ระบุได้ 45
ตาราง 17 ข้อมูลสถานที่ทำงานและช่วงอายุของผู้ใช้ที่ระบุได้ 46
ตาราง 18 ความถี่ของเพจที่มีคนไลก์ 20 อันดับแรก 48
ตาราง 19 ร้อยละของผู้ใช้ในในช่วงความโน้มเอียงของขั้วความคิดทางการเมือง 49
ตาราง 20 ตัวอย่างตารางโหนดในส่วนข้อมูลโพสต์ 51
ตาราง 21 ตัวอย่างตารางโหนดในส่วนข้อมูลผู้ใช้ที่มีส่วนร่วม 52
ตาราง 22 ตัวอย่างตารางเส้นความสัมพันธ์ 52
ตาราง 23 ตารางรายละเอียดข้อมูลการมีส่วนร่วมของโพสต์อันดับหนึ่ง 54
ตาราง 24 ตารางรายละเอียดข้อมูลการมีส่วนร่วมของโพสต์อันดับสอง 55
ตาราง 25 ตารางรายละเอียดข้อมูลการมีส่วนร่วมของโพสต์อันดับสาม 56
ตาราง 26 ตารางรายละเอียดข้อมูลการมีส่วนร่วมของโพสต์อันดับสี่ 57

ตาราง 27 ตารางรายละเอียดข้อมูลการมีส่วนร่วมของโพสต์อันดับห้า..... 58
ตาราง 28 จำนวนผู้กดไลก์เพจอันดับแรกแบ่งตามขั้วความคิดทางการเมือง..... 60
ตาราง 29 จำนวนผู้กดไลก์เพจอันดับที่สองแบ่งตามขั้วความคิดทางการเมือง 61
ตาราง 30 จำนวนผู้กดไลก์เพจอันดับที่สามแบ่งตามขั้วความคิดทางการเมือง 62
ตาราง 31 จำนวนผู้กดไลก์เพจอันดับที่สี่แบ่งตามขั้วความคิดทางการเมือง 63
ตาราง 32 จำนวนผู้กดไลก์เพจอันดับที่ห้าแบ่งตามขั้วความคิดทางการเมือง 64
ตาราง 33 ตารางแจกแจงแบบไขว้ของจำนวนผู้ใช้และเพจแบ่งตามขั้วความคิดทางการเมือง..... 65
ตาราง 34 ตารางแจกแจงแบบไขว้ของร้อยละของผู้ใช้และเพจแบ่งตามขั้วความคิดทางการเมือง..... 65
ตาราง 35 ตารางแจกแจงแบบไขว้ของจำนวนผู้ใช้และเพจแบ่งตามเพศและขั้วความคิดทางการเมือง..... 66
ตาราง 36 ตารางแจกแจงแบบไขว้ของร้อยละของผู้ใช้และเพจแบ่งตามเพศและขั้วความคิดทางการเมือง 67
ตาราง 37 ตารางแจกแจงแบบไขว้ของจำนวนผู้ใช้และเพจแบ่งตามเพศและช่วงอายุ 69
ตาราง 38 ตารางแจกแจงแบบไขว้ของร้อยละของผู้ใช้และเพจแบ่งตามเพศและช่วงอายุ..... 70

สารบัญรูปภาพ

รูป 1 ข้อมูลที่เก็บจากโพสต์ข้อความบนหน้าแฟนเพจ 19
รูป 2 ข้อมูลที่เก็บจากหน้าการมีส่วนร่วม 20
รูป 3 ข้อมูลที่เก็บจากหน้ารายชื่อผู้ใช้ 22
รูป 4 ตัวอย่างแท็กหน้าซอร์สโค้ดจากเว็บเบราว์เซอร์ 25
รูป 5 คำถามเรื่องช่วงอายุ (รุ่น) ในระบบ Moodle LMS 32
รูป 6 คำถามการระบุเพศในระบบ Moodle LMS 32
รูป 7 การเตรียมขั้นตอนให้ทางสถาบันฯ วิเคราะห์ข้อมูล 36
รูป 8 การเตรียมข้อมูลให้ทางสถาบันฯ วิเคราะห์ข้อมูล 36
รูป 9 ข้อมูลที่ถูกส่งกลับมาจากทางสถาบันฯ 37
รูป 10 แผนภูมิทรีแมปเปรียบเทียบสัดส่วนอายุและเพศของผู้ใช้งาน 44
รูป 11 แผนภูมิทรีแมปเปรียบเทียบสัดส่วนอายุและที่ทำงานของผู้ใช้งาน 46
รูป 12 แผนภูมิทรีแมปเปรียบเทียบสัดส่วนที่ทำงานและอายุของผู้ใช้งาน 48
รูป 13 ตัวอย่างโพสต์และการเก็บข้อมูลตาราง 51
รูป 14 กราฟเครือข่ายสังคมของโพสต์และผู้ใช้งานที่มีส่วนร่วม 53
รูป 15 กราฟเครือข่ายสังคมและภาพหน้าจอของโพสต์อันดับหนึ่ง 54
รูป 16 กราฟเครือข่ายสังคมและภาพหน้าจอของโพสต์อันดับสอง 55
รูป 17 กราฟเครือข่ายสังคมและภาพหน้าจอของโพสต์อันดับสาม 56
รูป 18 กราฟเครือข่ายสังคมและภาพหน้าจอของโพสต์อันดับสี่ 57
รูป 19 กราฟเครือข่ายสังคมและภาพหน้าจอของโพสต์อันดับห้า 58
รูป 20 กราฟเครือข่ายสังคมของผู้ใช้และเพจที่กดไลก์ 59
รูป 21 กราฟเครือข่ายสังคมของผู้ใช้และเพจที่กดไลก์เพจอันดับแรก 60
รูป 22 กราฟเครือข่ายสังคมของผู้ใช้และเพจที่กดไลก์เพจอันดับที่สอง 61
รูป 23 กราฟเครือข่ายสังคมของผู้ใช้และเพจที่กดไลก์เพจอันดับที่สาม 62
รูป 24 กราฟเครือข่ายสังคมของผู้ใช้และเพจที่กดไลก์เพจอันดับที่สี่ 63
รูป 25 กราฟเครือข่ายสังคมของผู้ใช้และเพจที่กดไลก์เพจอันดับที่ห้า 64
รูป 26 ปิรามิดประชากรเพศชายแบ่งตามขั้วความคิดทางการเมืองซ้ายและขวา 68

รูป 27 ปิรามิดประชากรเพศหญิงแบ่งตามขั้วความคิดทางการเมืองซ้ายและขวา 68
รูป 28 ปิรามิดประชากรเพศชายแบ่งตามรุ่นและขั้วความคิดทางการเมืองซ้ายและขวา 71
รูป 29 ปิรามิดประชากรเพศหญิงแบ่งตามรุ่นและขั้วความคิดทางการเมืองซ้ายและขวา 71

บทคัดย่อ

ในโลกศตวรรษที่ 21 สิ่งที่พิเศษเพิ่มเติมจากความขัดแย้งในรูปแบบเก่าที่เคยพบเจอ คือการก่อตัวขึ้นของกลุ่มคนรุ่นใหม่ (younger generation) กลุ่มคนเหล่านี้เคยเป็นเด็กในสมัยของความขัดแย้งดังที่ได้กล่าวไป แต่ปัจจุบันเริ่มเติบโตขึ้นมาอยู่ในวัยที่สามารถสะท้อนเสียงและความคิดของตัวเองได้ งานวิจัยนี้มีวัตถุประสงค์ที่จะรวบรวมและจัดเก็บข้อมูลความสนใจทางการเมืองของประชาชนช่วงวัยต่างๆ จากหน้าแฟนเพจบริการเครือข่ายสังคมที่นำเสนอข้อมูลความรู้เกี่ยวกับประชาธิปไตยของสถาบันพระปกเกล้าอย่างต่อเนื่องโดยใช้กระบวนการทางวิทยาศาสตร์ข้อมูล โดยการวิเคราะห์ข้อมูลผู้ใช้งานที่มีส่วนร่วมกับเนื้อหาในเพจของสถาบันฯ พบว่ากลุ่มคนรุ่นใหม่ คือเจอเนอเรชัน Y เป็นกลุ่มประชากรที่ให้ความมีส่วนร่วมและให้ความสนใจกับเพจของสถาบันฯ มากที่สุด โดยเฉพาะกลุ่มประชากรเพศชาย

ในด้านรายละเอียดของเนื้อหาข้อมูลของเพจของสถาบันฯ ได้แก่ความยาวของข้อความโพสต์ ประเภทของสื่อที่แบ่งปัน หรือประเภทเนื้อหา ไม่พบว่ามีตัวแปรใดมีความสัมพันธ์ต่อการมีส่วนร่วมของผู้ใช้งานอย่างมีนัยสำคัญ ในด้านประเภทสื่อที่ใช้ในการนำเสนอข้อมูล พบว่าการใช้ข้อความและภาพ (เพียงภาพเดียว) ได้รับการมีส่วนร่วมมากที่สุดแต่ก็มีการกระจายตัวของการมีส่วนร่วมสูงเช่นกัน ในด้านความโน้มเอียงของขั้วความคิดทางการเมือง พบว่าโดยเฉลี่ยผู้ใช้อยู่ในกลุ่มที่เป็นกลางทางการเมือง กลุ่มผู้ใช้ที่อยู่ในขั้วความคิดแต่ละฝั่งรับข้อมูลข่าวสารจากฝั่งตรงข้าม กลุ่มผู้ใช้ที่มีอายุมาก (Boomer) มีความโน้มเอียงไปในฝั่งขวามากกว่ากลุ่มคนที่มีอายุน้อยกว่าทุกกลุ่ม

นักวิจัยสรุปข้อมูลการวิเคราะห์ด้วยภาพข้อมูล (data visualization) ในหลายรูปแบบเช่นทรีแมป และกราฟเครือข่ายสังคม

Abstract

The younger generation's formation in the 21 century is seen as an unprecedented political conflict. They were just kids when political conflicts emerged in Thailand during the last decades. This research examines how different generations, from Baby Boomer to Generation Z, engage in social media content curated by the King Prajadhipok's Institute. The findings suggested that Generation Y are among the most engaged group of citizens, significantly younger males.

The study also found that word lengths, type of media, and content category have no significant relationship with engagement level. Sharing content with a single image seemed to gain engagement the most, albeit the high level of dispersion among the engagement level within this type of content. In terms of the generation, boomers are the group that leans more toward media and public figures on the right, while other generation groups lean toward the left. There is evidence that groups from each side do subscribe to news and follow public figures from the other side too.

The research presents findings using various data visualization techniques such as treemap and social network graph.

บทที่ 1 บทนำ

1.1 หลักการและเหตุผล

การแบ่งขั้วความคิดทางการเมือง (Political Polarization) ถือเป็นปรากฏการณ์ที่เกิดขึ้นเป็นเรื่องปรกติในสังคมประชาธิปไตยที่มีความขัดแย้งทางการเมืองสูง การแบ่งขั้วความคิดทางการเมืองภายในบริบทสังคมของไทยในช่วงสิบกว่าปีที่ผ่านมานั้นเกิดขึ้นในรูปแบบที่หลายคนเรียกว่า “การเมืองสีเสื้อ (color politics)” ระหว่างผู้คนที่มีฐานะทางเศรษฐกิจสังคมที่แตกต่างกัน (เช่น คนรากหญ้าในต่างจังหวัด-คนชั้นกลางในเมือง คนจน-คนรวย) ที่มีความคิดเห็นและจุดยืนทางการเมืองที่แตกต่างกัน อย่างไรก็ตาม ในโลกศตวรรษที่ 21 สิ่งที่พิเศษเพิ่มเติมจากความขัดแย้งในรูปแบบเก่าที่เคยพบเจอ คือการก่อตัวขึ้นของกลุ่มคนรุ่นใหม่ (younger generation) กลุ่มคนเหล่านี้เคยเป็นเด็กในสมัยของความขัดแย้งดังที่ได้กล่าวไป แต่ปัจจุบันเริ่มเติบโตขึ้นมาอยู่ในวัยที่สามารถสะท้อนเสียงและความคิดของตัวเองได้ พวกเขาก้าวเข้ามามีบทบาททางการเมืองและการปกครองของประเทศอย่างมีนัยยะสำคัญ และด้วยความก้าวหน้าของการพัฒนาทางวิทยาศาสตร์และเทคโนโลยี บทบาทของกลุ่มคนรุ่นใหม่ในการเมืองไทยจึงถูกแสดงออกผ่านนวัตกรรมและเทคโนโลยีใหม่ๆ ยิ่งผู้คนโดยเฉพาะในหมู่กลุ่มคนรุ่นใหม่นั้นหันมาใช้ชีวิตในโลกโซเชียลมากยิ่งขึ้น ยิ่งก่อให้การสื่อสารที่ก่อผลกระทบในวงกว้างได้รุนแรงและรวดเร็วยิ่งขึ้น ในขณะที่คนรุ่นเก่าเองก็มีการปรับตัวเข้ากับเทคโนโลยีสมัยใหม่และมีการแสดงออกทางการเมืองผ่านทางโซเชียลมีเดียด้วยเช่นกัน บทบาทสำคัญของอายุนั้นจึงเป็นปัจจัยที่ควรจะศึกษา ยิ่งเมื่อมันมีผลโดยตรงไม่ใช่แค่ในระดับปัจเจกบุคคลแต่ตั้งเป้าหมายไปถึงการเปลี่ยนแปลงทางการเมืองในระดับชาติด้วย

ในปัจจุบัน นวัตกรรมด้านวิทยาศาสตร์ข้อมูล (Data science) ได้เข้ามามีบทบาทกับองค์กรทางธุรกิจและเป็นเครื่องมือสำคัญในการเพิ่มประสิทธิภาพให้องค์กร ข้อมูลโดยเฉพาะประเภทที่ไม่มีโครงสร้างแน่นอน (Unstructured data) ซึ่งถือเป็นข้อมูลส่วนใหญ่ที่เกิดขึ้นในโลกออนไลน์กลายเป็นข้อมูลสำคัญสามารถสร้างความเข้าใจพฤติกรรมผู้บริโภคหรือกลุ่มผู้ใช้บริการได้เป็นอย่างดี สำนักนวัตกรรมเพื่อประชาธิปไตย สถาบันพระปกเกล้าได้เล็งเห็นถึงประโยชน์และความสำคัญในการใช้นวัตกรรมด้านวิทยาศาสตร์ข้อมูลเพื่อสร้างความเข้าใจเกี่ยวกับความแตกต่างทางความคิดและทัศนคติทางการเมืองของประชาชนในช่วงวัยต่างๆ โดยในเบื้องต้นจะเริ่มศึกษาและพัฒนาระบบสารสนเทศเพื่อรวบรวมข้อมูลจากหน้าแฟนเพจของสถาบันพระปกเกล้าก่อน โดยอาจเพิ่มเติมเว็บเพจที่นำเสนอข้อมูลความรู้เกี่ยวกับประชาธิปไตยที่มีผู้ติดตามหลากหลายและเป็นจำนวนมากเพื่อการวิเคราะห์เปรียบเทียบในบางกรณี และนำข้อมูลเหล่านั้นมาผ่านกระบวนการทางวิทยาศาสตร์ข้อมูลเพื่อสร้างความเข้าใจเกี่ยวกับลักษณะประชากร (demographic) ตลอดจนความคิดเห็นและทัศนคติทางการเมืองของผู้ใช้บริการและเข้าถึงข้อมูลทางแฟนเพจของสถาบันพระปกเกล้าต่อไป

1.2 วัตถุประสงค์

1.2.1 รวบรวมและจัดเก็บข้อมูลความคิดและทัศนคติทางการเมืองของประชาชนช่วงวัยต่างๆ จากหน้าแฟนเพจบริการเครือข่ายสังคม (social networking service) ที่นำเสนอข้อมูลความรู้เกี่ยวกับประชาธิปไตยของสถาบันพระปกเกล้า

1.2.2 วิเคราะห์ความแตกต่างของความคิดและทัศนคติทางการเมืองระหว่างประชาชนต่างช่วงวัยโดยกระบวนการทางวิทยาศาสตร์ข้อมูล

1.2.3 นำเสนอข้อมูลจากการวิเคราะห์ด้วยภาพข้อมูล (Data Visualization)

1.3. ขอบเขตของการวิจัย

1.3.1 ด้านเนื้อหาและรูปแบบการดำเนินการ

พัฒนาระบบสารสนเทศเพื่อรวบรวมข้อมูลความคิดเห็นและทัศนคติทางการเมืองของประชาชนในช่วงวัยต่างๆ จากหน้าแฟนเพจบริการเครือข่ายสังคมของสถาบันพระปกเกล้าร่วมที่นำเสนอข้อมูลความรู้เกี่ยวกับประชาธิปไตยที่มีผู้ติดตามหลากหลายและเป็นจำนวนมาก และนำข้อมูลเหล่านั้นมาผ่านกระบวนการทางวิทยาศาสตร์ข้อมูล ซึ่งประกอบด้วย

(1) การเก็บข้อมูล (Obtaining the data)

(2) การทำความสะอาดข้อมูล (Cleaning the data) และเตรียมข้อมูลเพื่อการวิเคราะห์

(3) การค้นหาสิ่งที่น่าสนใจ (Exploring ฿ visualizing the data) ในประเด็นของลักษณะประชากร (demographic) และสิ่งที่ชอบ (preference)

1.3.2 ขอบเขตด้านประชากร

ศึกษาความคิดเห็นและทัศนคติทางการเมืองของประชาชนในช่วงวัยต่างๆ เปลี่ยนจาก 3 กลุ่มหลักที่ตั้งไว้เดิมได้แก่ กลุ่มผู้สูงวัย (60 ปีขึ้นไป) กลุ่มวัยกลางคน (30-59 ปี) กลุ่มเยาวชน (18-29 ปี) ซึ่งการแบ่งกลุ่มดังกล่าวนั้นอ้างอิงมาจากการแบ่งกลุ่มตามหลักประชากรศาสตร์ โดยจะทำการแบ่งกลุ่มตามปีเกิดของประชากร เป็นการแบ่งตามรุ่น (generation) ได้แค่กลุ่ม Baby Boomer (เกิดปี พ.ศ.2489-2507) เจนเนอเรชัน X (เกิดปี พ.ศ.2508-2522) เจนเนอเรชัน Y (เกิดปี พ.ศ.2523-2540) และเจนเนอเรชัน Z (เกิดหลังปี พ.ศ.2540) ตามที่ได้ปรึกษาและชี้แจงกับทางทีมงานของผู้ว่าจ้างเมื่อวันที่ 27 ตุลาคม 2563

1.4. วิธีดำเนินการ

1.4.1 รับฟังความคิดเห็นและประเด็นปัญหาจากเจ้าหน้าที่ผู้เกี่ยวข้อง

1.4.2 ศึกษาวิธีการเก็บรวบรวมข้อมูลจากแฟนเพจที่มีประประสิทธิภาพและสมบูรณ์ตามศักยภาพของเทคโนโลยีในปัจจุบัน

(1) เทคโนโลยีในการเก็บข้อมูลผ่านหน้าเว็บในปัจจุบัน มีการพัฒนาไปมาก และมีเครื่องมือหลายชนิดให้เลือกใช้ ในขั้นตอนนี้ทีมงานวิจัย ศึกษาถึงเครื่องมือที่มีอยู่ในปัจจุบัน

1.4.3 ออกแบบและพัฒนาระบบสารสนเทศเพื่อรวบรวมข้อมูลจากหน้าแฟนเพจ ขั้นตอนการออกแบบประกอบด้วยรายละเอียดดังนี้

(1) ขั้นตอนการเลือกใช้ภาษาโปรแกรมในการพัฒนา ด้วยภาษาไพธอน (Python)

(2) ขั้นตอนการวางแผนเก็บข้อมูล ใน 4 ระดับคือ

- โพสต์ข้อความในหน้าแฟนเพจ
- การมีส่วนร่วม
- รายชื่อผู้ใช้
- รูปโปรไฟล์ผู้ใช้

1.4.4 ขั้นตอนการทดสอบใช้งานระบบสารสนเทศที่พัฒนาเพื่อเก็บข้อมูลในระดับต่างๆ

1.4.5 ทำความสะอาดข้อมูลและค้นหาสิ่งที่น่าสนใจในข้อมูล

1.4.6 การหาช่วงอายุและเพศ

(1) การจัดการรูปโปรไฟล์ที่ดึงมาจากหน้าผู้ใช้

(2) การเตรียมระบบการทำงานให้ผู้ช่วยวิจัย

1.4.7 การวิเคราะห์ข้อมูลในรูปแบบ สถิติเชิงพรรณนา (central tendency and correlation coefficient) และการสร้างภาพข้อมูล (data visualization)

1.4.8 การจัดทำรายงาน

1.5 ระยะเวลาดำเนินงานและการส่งมอบงาน

	เดือนที่	1				2				3				4				5			
	สัปดาห์ที่	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1	เก็บข้อมูล	○	○	○	○	○	○	○	○												
2	ทำความสะอาดข้อมูล							○	○	○	○	○	○								
3	วิเคราะห์ข้อมูล											○	○	○	○	○	○				
	3.1 อบรมผู้ช่วยวิจัย											○	○	○							
4	จัดทำรายงาน															○	○	○	○	○	○

1.6 คณะผู้วิจัย

ผศ. ดร. วสะ บูรพาเดชะ (หัวหน้าโครงการวิจัย) – มหาวิทยาลัยอัสสัมชัญ

ดร. อนันต์ โอสถศิลป์ (นักวิจัย) – มหาวิทยาลัยอัสสัมชัญ

บทที่ 2 ขั้นตอนของการวิจัย

ในขั้นตอนการวิจัย แบ่งออกเป็น 3 ช่วงหลัก คือ การเก็บข้อมูล การทำความสะอาดข้อมูล และการเตรียมข้อมูลเพื่อการวิเคราะห์ โดยมีรายละเอียดในแต่ละช่วงหลักดังนี้

2.1 การเก็บข้อมูล

2.1.1 ศึกษาวิธีการเก็บรวบรวมข้อมูลจากแฟนเพจที่มีประประสิทธิภาพและสมบูรณ์ตามศักยภาพของเทคโนโลยีในปัจจุบัน

(1) เทคโนโลยีในการเก็บข้อมูลผ่านหน้าเว็บหรือ web scraping ในปัจจุบันมีการพัฒนาไปมาก และมีเครื่องมือหลายชนิดให้เลือกใช้ ในขั้นตอนนี้ทีมงานวิจัย ศึกษาถึงเครื่องมือที่มีอยู่ในปัจจุบัน แบ่งได้เป็น 3 กลุ่ม

1. ซอฟต์แวร์โรบอทที่เลียนแบบพฤติกรรมของมนุษย์ (Robotic Process Automation หรือ RPA) เช่น Octoparse (https://octoparse.com) หรือ UiPath (https://uipath.com)
2. โปรแกรมเสริม หรือ plugin ในเว็บเบราว์เซอร์ เช่น
    - Web Scraper – GetData.IO (https://chrome.google.com/webstore/detail/web-scraper-getdataio/ofncgcgajhgnbkbmkdhbgkoopfbemhfj?hl=en) หรือ
    - Web Scraper – Free Web Scraping (https://chrome.google.com/webstore/detail/web-scraper-free-web-scra/jnhgnonknehpejjnehehllkliplmbmhn?hl=en)
3. พัฒนาระบบสารสนเทศเก็บข้อมูลผ่านหน้าเว็บหรือ web scraping เฉพาะทางเพื่อเก็บข้อมูลจากหน้าแฟนเพจ

(2) ทีมงานตัดสินใจเลือกใช้ตัวเลือกที่ 3 เนื่องจากความยืดหยุ่นในการเก็บข้อมูล ตามที่คาดการณ์ไว้ในขั้นตอนการที่ 2.1.2 ออกแบบและพัฒนาฯ

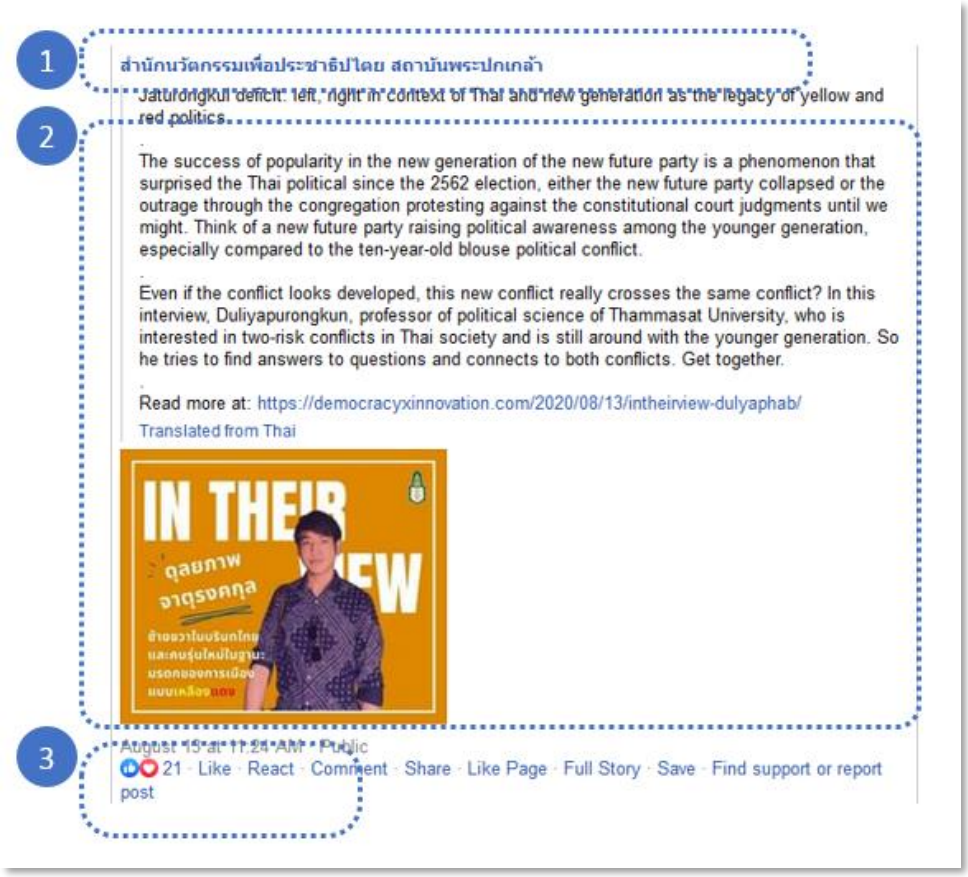
2.1.2 ออกแบบและพัฒนาระบบสารสนเทศเพื่อรวบรวมข้อมูลจากหน้าแฟนเพจ ขั้นตอนการออกแบบประกอบด้วยรายละเอียดดังนี้

(1) ขั้นตอนการเลือกใช้ภาษาโปรแกรมในการพัฒนาด้วยภาษาไพธอน (Python) เนื่องจาก

- เป็นภาษาโปรแกรมที่ได้รับความนิยมในชุมชนนักวิทยาศาสตร์ข้อมูล
- มีคลังโปรแกรม (ไลบรารี หรือ library) ที่รวบรวมฟังก์ชันย่อย (subroutine) หรือซอร์สโค้ดเฉพาะทางสำหรับการเก็บข้อมูลจากหน้าเว็บให้เลือกหลายชนิด คลังโปรแกรมที่เลือกใช้ในการวิจัยครั้งนี้คือ
  - Selenium (https://www.selenium.dev/) เป็นเครื่องมือในการช่วยควบคุมหน้าเว็บอัตโนมัติ (web-based automation tool) ซึ่งมักใช้ในการทดสอบการทำงานของซอฟท์แวร์ สำหรับการวิจัยครั้งนี้ ผู้วิจัยใช้ Selenium ในการเข้าสู่ระบบเฟซบุ๊กโดยอัตโนมัติ
  - BeautifulSoup (https://www.crummy.com/software/BeautifulSoup/) เป็นเครื่องมือที่ใช้ในการเข้าถึงซอร์สโค้ดของหน้าเว็บ (HTML parsing tool) และเลือกเอาข้อมูลที่ต้องการออกมาเก็บไว้ในรูปแบบที่ต้องการได้

(2) ขั้นตอนการวางแผนเก็บข้อมูล มีดังนี้

- โพสต์ข้อความบนหน้าแฟนเพจ (scraper.py)



รูป 1 ข้อมูลที่เก็บจากโพสต์ข้อความบนหน้าแฟนเพจ

โดยในระดับแรกนี้ จะทำการเก็บข้อมูล

(1) หัวข้อโพสต์ เลขรหัสของโพสต์ (post ID) และลิงก์ตรงของโพสต์

(2) ข้อความโพสต์ และชนิดของโพสต์ (เช่น รูปภาพ แชร์ลิงก์)

(3) การมีส่วนร่วม (engagement) ต่างๆ เช่น ไลก์ แชร์

- การมีส่วนร่วม (engagement)

รูป 2 ข้อมูลที่เก็บจากหน้าการมีส่วนร่วม

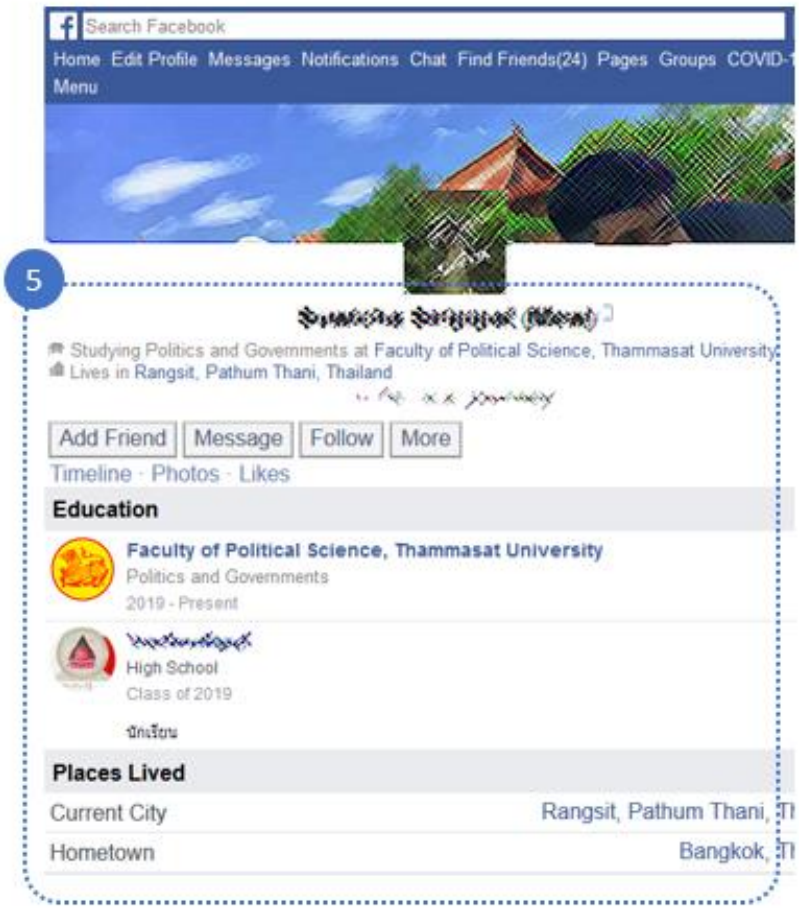
จากขั้นตอนที่ (3) เมื่อกดลิงก์ก็จะปรากฏหน้ารายละเอียดที่สามารถเห็นรายชื่อผู้ใช้ (profiles) ที่เข้ามามีส่วนร่วมกับโพสต์ได้ ทีมงานสามารถเก็บข้อมูลดังนี้

(4) ลิงก์ตรงเพื่อเข้าถึงรายชื่อผู้ใช้ ซึ่งข้อมูลในส่วนนี้จะถูกใช้ในการสร้างรายชื่อโปรไฟล์ผู้ใช้งานที่เข้ามามีส่วนร่วมและเก็บไว้ที่ big data list ในรูปแบบ JSON (JavaScript Object Notation) โดยเก็บข้อมูลแยกไว้ 2 ไฟล์คือ posts.json และ users.json (และทำการคัดลอกไว้ใน users2.json เพื่อดำเนินการขั้นต่อไป)

คำอธิบายซูโดโค้ด (Pseudocode) ของ scraper.py มีดังนี้

```
Open target web page
Login with authorize user account
Scroll to the end of web page
Find all comments tags and expand each comments tree
Parse html content into a tag tree
Find posts in the tag tree and save the data as a json file( post_with_no_likers.json )
Create an empty Big data list
For each post in posts:-
  detect post text
  detect post id
  detect post image link
  detect #shares
  detect reactions information
  detect comments
  save each new post with detected information as a dictionary in big data list
For each post in big data list:-
  detect the list of persons who clicked like in the post and save url and name of those persons
Save the big data list as  two files including posts.json and users.json
//note that a user is the person is defined as the one who click like to any post form the web page
//users.json contains the list of all the user
```

- รายชื่อผู้ใช้ (get_person_info.py)



รูป 3 ข้อมูลที่เก็บจากหน้ารายชื่อผู้ใช้

ในกรณีที่เข้าถึงหน้ารายชื่อผู้ใช้แล้วในขั้นตอนที่ (4) จะทำการเก็บข้อมูลดังนี้

(5) ใช้รายการข้อมูลจาก users2.json รายละเอียดต่างๆ ที่ผู้ใช้เปิดเป็นสาธารณะ ได้แก่

- ชื่อโปรไฟล์
- สถานศึกษา (ทุกรายการที่เข้าถึงข้อมูลได้)
- สนานที่ทำงานปัจจุบันและในอดีต (ทุกรายการที่เข้าถึงข้อมูลได้)
- ที่อยู่ (อำเภอหรือเขต และจังหวัด)
- ลิงก์ของรูปโปรไฟล์

คำอธิบายซูโดโค้ด (Pseudocode) ของ get_person_info.py มีดังนี้

```
Open web browser
Login with authorized user account
Open users2.json and dump all the data into a list of users
Foreach user object in users:-
  get user name and user link
  open new web page using url from the user link
  detect user's about data from the page
  open user's follower link
  scroll to the end of the page
  detect all followers/friends information including name and
link
  append user's about data and user's followers data into each
user object
Save a list of users back into users2.json
```

- การบันทึกรายละเอียดข้อมูลเพื่อการวิเคราะห์ในขั้นตอนต่อไป (make_csv.py)

  ขั้นตอนนี้เป็นการแปลงไฟล์ JSON ที่เก็บข้อมูลไว้ให้เป็นไฟล์ CSV (Comma-separated values) ซึ่งเป็นไฟล์ที่สามารถใช้งานได้ในรูปแบบตาราง เหมาะกับโปรแกรมการคำนวณแบบสเปรดชีททั่วไป

คำอธิบายซูโดโค้ด (Pseudocode) ของ make_csv.py มีดังนี้

```
Open posts.json file
For each post object in posts:-
   save post information including post_id, post_text, image link
into posts.csv
   uniquely save who like profiles in who_like_profile.csv
   save the list of who like for each post in posts_with_like.csv
   save a list who like friends in who_like_friends.csv
```

- รูปโปรไฟล์ผู้ใช้ (get_who_like_photo.py)

  ขั้นตอนการบันทึกรูปโปรไฟล์ของผู้ใช้ โดยการดึงข้อมูลจาก user2.json กลับมาเพื่อเข้าถึงหน้าโปรไฟล์ผู้ใช้แต่ละคน ตรวจหารูปโปรไฟล์และบันทึกรูปมาเก็บไว้ใน

คำอธิบายซูโดโค้ด (Pseudocode) ของ get_who_like_photo.py มีดังนี้

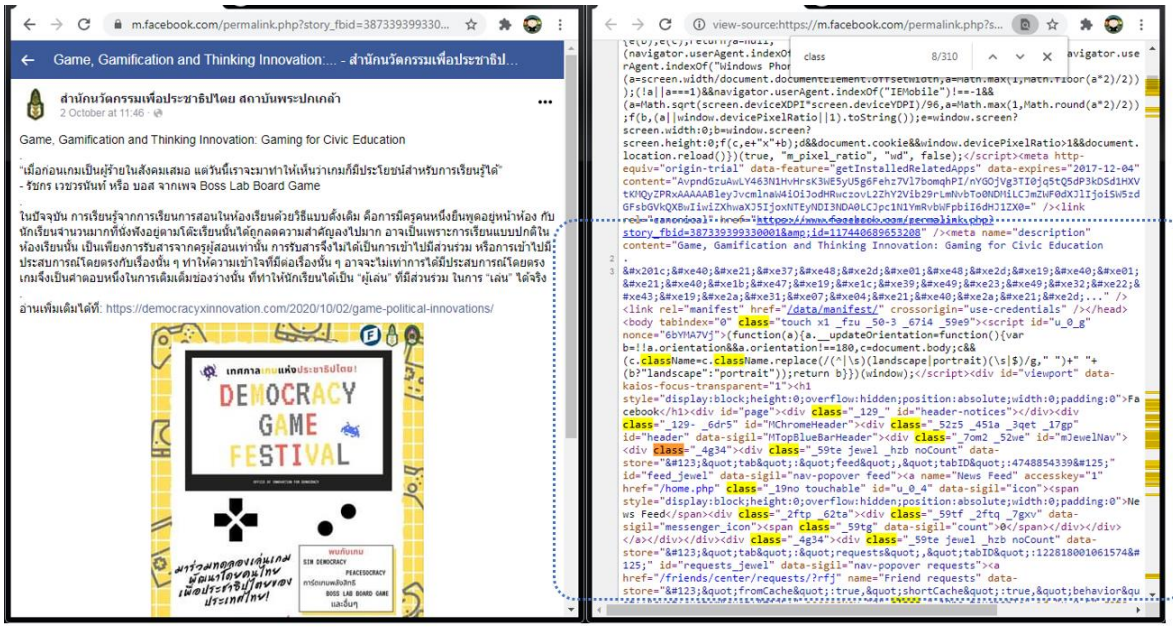
```
Open web browser
Login with authorized user account
Open users2.json and dump all the data into a list of users
Foreach user object in users:-
  get user name and user link
  open new web page using url from the user link
  detect user's photo from the photo tag in the page
  save each user's photo as a file with the username
```

2.1.3 ขั้นตอนการทดสอบใช้งานระบบสารสนเทศที่พัฒนาเพื่อเก็บข้อมูลในระดับต่างๆ แพลตฟอร์มโซเชียลมีเดียของเฟซบุ๊ก มีการปรับเปลี่ยนการออกแบบการติดต่อประสานระหว่างเครื่องคอมพิวเตอร์กับผู้ใช้ (หรืออินเตอร์เฟซ) มาตลอด[1] และได้ปรับเปลี่ยนนโยบายความเป็นส่วนตัวของผู้ใช้มาตลอดเช่นเดียวกัน ในช่วงหลังเหตุการณ์ “เคมบริดจ์แอนะลิติกา” ทางเฟซบุ๊กเพิ่มความเข้มงวดในการเข้าถึงข้อมูลผู้ใช้ รวมถึงการเพิ่มความซับซ้อนของซอร์สโค้ดหลังบ้านของแพลตฟอร์ม มาตรการเหล่านี้ทำให้การเก็บข้อมูลจากหน้าเว็บของเฟซบุ๊ก มีขั้นตอนที่ซับซ้อนมากขึ้น โดยการทดสอบระบบมีแนวทางดังนี้

- พัฒนาซอร์สโค้ดในการเก็บข้อมูลจากหน้าเพจเฟซบุ๊กในแต่ละขั้นตอน จาก 2.1.2 โดยการเลือกคุณสมบัติของแท็กจากซอร์สโค้ดหลังบ้านของแพลตฟอร์มที่เกี่ยวข้องกับข้อมูลที่ต้องการ โดยทดสอบว่าแท็กนั้นมีความเฉพาะตัว (unique) เพียงพอหรือไม่ที่จะใช้ระบุชุดข้อมูลที่ต้องการ

[1] ในระหว่างโครงการ เฟซบุ๊กได้ทำการปรับอินเตอร์เฟซใหม่ โดยประกาศประมาณเดือนพฤษภาคมในงาน Facebook F8 2019 และบังคับให้ผู้ใช้งานปรับมาเป็นอินเตอร์เฟซใหม่ภายในสิ้นเดือนกันยายน อินเตอร์เฟซใหม่นี้มีการปรับโครงสร้างซอร์สโค้สใหม่ทั้งหมด ทำให้ระบบที่พัฒนาขึ้นไม่สามารถเข้าถึงข้อมูลได้ ทีมงานจึงต้องเร่งเก็บข้อมูลทั้งหมดภายในสิ้นเดือนกันยายน และปรับไปใช้หน้าอินเตอร์เฟซอุปกรณ์พกพา (โดยการเพิ่ม m. ด้านหน้าชื่อ facebook.com) โดยอินเตอร์เฟซอุปกรณ์พกพานี้ ไม่สามารถสั่งการได้อัตโนมัติเหมือนระบบเดสก์ทอป) ทำให้ต้องปรับการทำงานเป็นแบบกึ่งอัตโนมัติ โดยผู้เก็บข้อมูลต้องเป็นผู้เลื่อนหน้าอินเตอร์เฟซเพื่อโหลดข้อมูลเพิ่มเอง

- ตัวอย่างการค้นหาแท็กจากซอร์สโค้ดว่ามีความเฉพาะตัว โดยดูจากค่าของ class



รูป 4 ตัวอย่างแท็กหน้าซอร์สโค้ดจากเว็บเบราว์เซอร์

จากตัวอย่างข้างต้นจะเห็นว่าในซอร์สโค้ดมีค่าของ class อยู่หลายค่า เช่น

```
<div class="_129_" …
<div class="_129- _6dr5" …
<div class="_52z5 _451a _3qet _17gp" …
<div class="_7om2 _52we" …
<div class="_4g34">
```

การพัฒนาระบบสารสนเทศเพื่อดึงค่าที่ถูกต้อง จึงต้องใช้การค้นหาแท็ก (ซึ่งโดยส่วนใหญ่แล้วเป็นแท็ก) และหาค่า class ที่มีความเฉพาะตัว

(1) เมื่อคัดเลือกแท็กที่มีความเฉพาะตัวที่สุดและสามารถระบุชุดข้อมูลที่ต้องการได้แล้ว ทำการทดลองการเก็บข้อมูลแบบอัตโนมัติผ่านคลังโปรแกรม Selenium

(2) เก็บข้อมูลที่ดึงลงมาได้ในรูปแบบ JavaScript Object Notation (JSON)

(3) แปลงข้อมูลจากรูปแบบ JSON ให้อยู่ในรูปแบบไฟล์ Comma Separated Value (CSV) เพื่อนำเข้าโปรแกรม Microsoft Excel และเข้าสู่กระบวนการ 4.4 ทำความสะอาดข้อมูลและค้นหาสิ่งที่น่าสนใจในข้อมูล (finding insights)

2.1.4 **การเก็บข้อมูลเพิ่มเติมที่เกี่ยวกับความคิดเห็นทางการเมือง**

(1) เพิ่มคอลัมน์ ตาราง 4 *รายการโปรไฟล์ที่ไลก์เพจ* (data_for_people_who_click_like_page) โดยเพิ่ม “/likes_all” ท้ายลิงก์ของโปรไฟล์ เช่น https://www.facebook.com/vasa.buraphadeja/likes_all โดยลิงก์นี้จะเปิดรายการเพจต่างๆ ที่ผู้ใช้ได้กดไลก์เอาไว้

(2) ทดสอบเครื่องมือในการเก็บข้อมูลจากหน้าเฟซบุ๊ก เพื่อเป็นการลดระยะเวลาในการพัฒนาระบบเก็บข้อมูล ทีมงานจึงตัดสินใจใช้โปรแกรม UiPath (https://www.uipath.com/product/studio) เพื่อทำการเก็บข้อมูลเพจต่างๆ ที่ผู้ใช้ได้กดไลก์เอาไว้

UiPath คือ Robotic Process Automation (RPA) หรือหุ่นยนต์ซอฟต์แวร์ (robot) สำหรับงานออฟฟิศที่ใช้การผสมผสานเทคโนโลยีหลายแขนง เช่น Image Recognition, Machine Learning และ AI เพื่อให้การทำงานเป็นแบบอัตโนมัติและเพิ่มประสิทธิภาพมากขึ้น โดยในการใช้งานของโครงการนี้ ทางทีมงานต้องการให้พัฒนาหุ่นยนต์ซอฟต์แวร์เพื่อช่วยในขั้นตอนต่อไปนี้

- เปิดหน้าลิงก์ผู้ใช้งานผ่านเบราเซอร์
- เลื่อนหน้าจอลง (scroll) มาให้ได้มากที่สุด (เป็นขั้นตอนที่ใช้เวลานาน เนื่องจากเฟซบุ๊กจะทำการแสดงผลข้อมูลเพียงบางส่วน เช่นจะแสดงรายการเพจทีละชุด อาจจะเพียง 8 รายการ หรือ 10 รายการ)

(3) ดำเนินการเก็บข้อมูลบนเฟซบุ๊ก

- ทีมงานประเมินว่าการเลื่อนหน้าจอลงอย่างรวดเร็วเกินไป และการเปิดเบราเซอร์หลายหน้าพร้อมกัน มีความเสี่ยงในการถูกตรวจจับได้จากระบบ และอาจเข้าข่ายการละเมิดข้อตกลงการให้บริการ[2]
- เพื่อเป็นการลดความเสี่ยง ทีมงานจึงตัดสินใจใช้การเก็บข้อมูลแบบกึ่งอัตโนมัติ ดังนี้
    i. ขั้นตอนที่ใช้ทำด้วยมือ คือการเปิดหน้าเบราเซอร์และเลื่อนหน้าจอลง
    ii. ขั้นตอนที่ใช้ RPA คือการกวาดข้อมูลต่างๆ ที่เกี่ยวข้องกับเพจ ได้แก่ ชื่อบัญชีผู้ใช้งาน ชื่อเพจ ลิงก์ โดยเก็บข้อมูลของแต่ละโปรไฟล์ไม่เกิน 500 เพจ
    iii. ข้อมูลที่ได้ถูกเก็บไว้โดยแต่ละโปรไฟล์ จะมีข้อมูล 3 คอลัมน์แรกตามที่แสดงในตาราง 5 *รายการไลก์เพจ* ซึ่งได้รวบรวมไว้ทั้งสิ้น 269 ไฟล์

[2] Facebook Takes Legal Action Against Data Scrapers in Latest Enforcement Effort
https://www.socialmediatoday.com/news/facebook-takes-legal-action-against-data-scrapers-in-latest-enforcement-eff/586288/

iv. จากนั้น ทำการรวมไฟล์ทั้งหมดให้เป็นไฟล์เดียว โดยใช้ Pandas (library สำหรับวิเคราะห์ข้อมูลด้วย Python) โดยมีคำอธิบายโค้ดดังนี้

คำอธิบายซูโดโค้ด (Pseudocode) ของ add_column_with_username.py.py มีดังนี้

```
import pandas library
import file management library (Glob)
set path to "csv"
collect all .csv files

set empty data frame (dfs = [])
for each file name in the .csv folder
  read all rows (max = 500)
  add new column "USER" set all rows to username (i.e., file name)
  attach (append) the new column to the data frame dfs
  write the data frame out as a .csv file to the same file

assign all dfs data to big_frame variable
write all data from big_frame to a all_like.csv
```

2.2 ทำความสะอาดข้อมูล

2.2.1 นำเข้าข้อมูล CSV มาที่โปรแกรม Microsoft Excel และจัดเรียงตามรูปแบบฐานข้อมูลที่ออกแบบไว้ โดยออกแบบเบื้องต้นไว้ดังนี้

ตาราง 1 รายการโพสต์ (project_scrape_level1)

Field	Data Type	Remark
postNo	number	หมายเลขโพสต์ (เลขอัตโนมัติ)
id	text	ลิงก์ของโพสต์
postTitle	text	ข้อความของโพสต์
image	text	ลิงก์ของภาพ (ปัจจุบันไม่สามารถเข้าถึงได้)
shares	text	จำนวนการแชร์โพสต์
likes	text	จำนวนการไลก์โพสต์

linkLink	text	ลิงก์ของข้อมูลไลก์ (ปัจจุบันไม่สามารถเข้าถึงได้)

ตาราง 2 รายการโปรไฟล์ของผู้มีส่วนร่วม (post_with_who_like)

Field	**Data Type**	**Remark**
postNo	number	หมายเลขโพสต์ (เลขอัตโนมัติ)
whoLike	text	ชื่อโปรไฟล์
whoLikeLink	text	ลิงก์ของโปรไฟล์

ตาราง 3 รายละเอียดโปรไฟล์ของผู้มีส่วนร่วม (who_like_profile)

Field	Data Type	Remark
Who Like	text	ชื่อโปรไฟล์
Studies at	text	สถานศึกษาปัจจุบัน
Studied at	text	ประวัติสถานศึกษา (มหาวิทยาลัย)
Went to	text	ประวัติสถานศึกษา (โรงเรียน)
Lives in	text	เมือง จังหวัดที่อยู่ปัจจุบัน
From	Text	ภูมิลำเนาเดิม
Works at	Text	สถานที่ทำงานปัจจุบัน
Worked at	Text	ประวัติสถานที่ทำงาน

ตาราง 4 รายการโปรไฟล์ที่ไลก์เพจ (data_for_people_who_click_like_page)

Field	Data Type	Remark
No	number	หมายเลข (เลขอัตโนมัติ)
Name	text	ชื่อโปรไฟล์
Link	text	ลิงก์ของโปรไฟล์

(1) ยุบรวมรายชื่อผู้มีส่วนร่วม (post_with_who_like) กับหน้าเพจที่มีความซ้ำซ้อนกันให้เหลือเพียงชื่อละหนึ่งเดียว (unique records) และเพิ่มเติมรายชื่อที่ขาดหายไปในขั้นตอน 2.1.2 (2) หัวข้อย่อย รายชื่อผู้ใช้ (get_person_info.py) (รายละเอียดบันทึกไว้ในไฟล์ moodle_GIFT_gen_quiz.xlsx ชีท likes)

(2) แปลงข้อมูลให้อยู่ในรูปแบบ GIFT-ready (พร้อมสำหรับการนำเข้าเป็นแบบทดสอบใน Moodle LMS) (moodle_GIFT_gen_quiz.xlsx ชีท GIFT)

ตาราง 5 รายการไลก์เพจ (fb_kpi_with_likes-all-results)

Field	Data Type	Remark
PAGE	Text	ชื่อเพจ
URL	Text	ลิงก์ของเพจ
CAT	Text	ประเภทของเพจ (เช่น นักการเมือง นักแสดง)
USER	Text	ชื่อโปรไฟล์
[BIAS]	Number	ความโน้มเอียงของขั้วความคิดทางการเมือง (1-5) โดย 1 คือซ้ายจัด และ 5 คือขวาจัด

2.2.2 ข้อมูลในส่วนความโน้มเอียงของขั้วความคิดทางการเมือง

(1) รวบประเภทเพจ ข้อมูลที่ได้จาก ตาราง 5 *รายการไลก์เพจ* นั้น ในส่วนของประเภท (Category หรือ CAT) มีหลากหลายประเภท ซึ่งบางประเภทสะกดต่างกัน เช่น

- Media, Media/news company, และ News and media website
- Blogger และ Personal blog
- Jewellery & watches shop และ Jewellery/watches

ซึ่งในหลายกรณี เป็นการสะกดที่ต่างกันของ American English และ British English โดยทีมงานได้ทำการเปลี่ยนข้อมูลประเภทเพจให้อยู่ในกลุ่มเดียวกัน สะกดเหมือนกันตามแบบ American English

(2) กำหนดค่าความโน้มเอียงของขั้วการเมือง

จากตาราง 5 *รายการไลก์เพจ* หลังจากขั้นตอนที่ (4) จะมีประเภทเพจที่ถูกต้อง หลังจากนั้นหาความถี่ของเพจที่คนนิยม (กดไลก์ไว้และแสดงในโปรไฟล์มากที่สุด) ด้วย PivotTable ใน Excel

ทีมงานได้ทำการเลือกเพจที่มีโปรไฟล์ผู้ใช้ติดตามรวมมากกว่า 10 คน และเป็นประเภทดังต่อไปนี้ เพื่อให้ผู้เชี่ยวชาญของสำนักนวัตกรรมเพื่อประชาธิปไตย สถาบันพระปกเกล้ากำหนดค่าความโน้มเอียงของขั้วความคิดทางการเมืองต่อไป

- Media/news company
- Public figure
- News and media website

- Politician
- Political organization
- Political party
- Newspaper
- Political candidate

2.3 การเตรียมข้อมูลเพื่อการวิเคราะห์

2.3.1 การหาช่วงอายุและเพศ

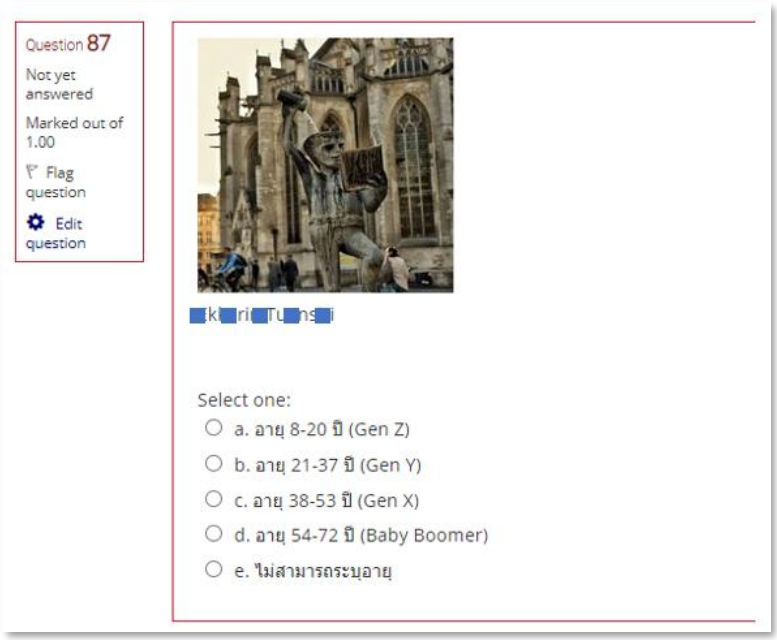
ในขณะที่การใช้งานปัญญาประดิษฐ์ (Artificial Intelligence หรือ AI) เริ่มเป็นที่แพร่หลาย เช่นการวิเคราะห์รูปภาพ หรือสิ่งของ ยังมีงานอีกหลายประเภทที่ AI ยังทำได้ไม่เที่ยงตรง และมีข้อจำกัดอีกหลายประการ ทางทีมงานจึงตัดสินใจใช้ผู้ช่วยวิจัยช่วยตัดสินข้อมูลเชิงคุณภาพ โดยแบ่งงานเป็นชิ้นเล็กและให้ผู้ช่วยวิจัย 3 ท่านได้ตัดสินใจจากระบบที่วางไว้ดังนี้

(1) การจัดการรูปโปรไฟล์ที่ดึงมาจากหน้าผู้ใช้ 2.1.2 (2) หัวข้อย่อย รายชื่อผู้ใช้ (get_person_info.py)

(2) การเตรียมระบบการทำงานให้ผู้ช่วยวิจัย โดยใช้ระบบทดสอบ (quiz) ในระบบจัดการเรียนการสอน Moodle LMS เป็นเครื่องมือในการประเมิน โดยในแต่ละโปรไฟล์จะถูกใช้ในการตั้งคำถาม 2 คำถาม

(3) คำถามแรกคือเรื่องช่วงอายุ (รุ่น) โดยมีส่วนประกอบดังนี้

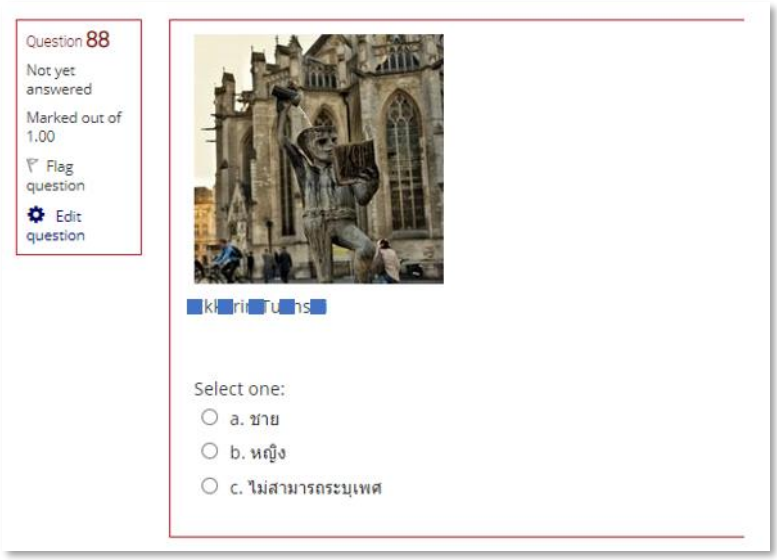
- รูปโปรไฟล์
- ลิงก์โปรไฟล์
- ตัวเลือกแบบปุ่มวิทยุ (Radio button) สองชุด ที่ผู้ประเมินสามารถเลือกได้เพียงตัวเลือกเดียว โดยแสดง 5 ตัวเลือกช่วงอายุ ดังนี้



รูป 5 คำถามเรื่องช่วงอายุ (รุ่น) ในระบบ Moodle LMS

(4) คำถามที่สองคือเรื่องเพศ โดยมีส่วนประกอบดังนี้

- รูปโปรไฟล์
- ลิงก์โปรไฟล์
- ตัวเลือกแบบปุ่มวิทยุ (Radio button) สองชุด ที่ผู้ประเมินสามารถเลือกได้เพียงตัวเลือกเดียว โดยแสดง 3 ตัวเลือกระบุเพศ ดังนี้



รูป 6 คำถามการระบุเพศในระบบ Moodle LMS

(5) ทีมงานจัดเรียงข้อคำถามไว้ทั้งสิ้น 710 คำถาม โดยแบ่งเป็น 4 ช่วง ช่วงละ 200 ข้อ (หรือ 100 โปรไฟล์) และช่วงสุดท้ายมีเพียง 110 ข้อ (หรือ 55 โปรไฟล์) ระบบ Moodle LMS ทั้งนี้เพื่อให้การทำงานของผู้ช่วยวิจัยสะดวกขึ้นโดยสามารถส่งงานได้ทีละช่วง

(6) การจัดหาผู้ช่วยวิจัยและการอบรม
ทีมงานประกาศหาผู้ช่วยวิจัยสัญชาติไทยที่มีความชำนาญการใช้งานคอมพิวเตอร์และระบบสื่อสังคมออนไลน์ เมื่อได้ผู้สมัครแล้วทีมงานได้จัดอบรมการใช้งานระบบทดสอบ Moodle LMS และให้ทดลองทำแบบทดสอบชุดแรกจำนวน 200 ข้อเป็นชุดแรกเพื่อทดลองระบบและรับฟังข้อแนะนำจากผู้ช่วยวิจัย กรณีที่อาจมีความเข้าใจคลาดเคลื่อนหรือมีข้อสงสัย

(7) หลังจากผู้ช่วยวิจัยทำการทดสอบชุดแรกเป็นที่เรียบร้อย ทีมงานได้นำข้อมูลมาวิเคราะห์และสอบถามความคิดเห็นการใช้งาน จากนั้นจึงสร้างแบบทดสอบชุดที่เหลือและให้ผู้ช่วยวิจัยทำงานต่อจนแล้วเสร็จ

(8) การตรวจสอบความเที่ยงระหว่างผู้ประเมิน (inter-rater reliability) (Randolph, J. J., 2005)

- <u>คำถามเรื่องช่วงอายุ</u> ผู้ช่วยวิจัย 3 ท่านทำการประเมินผู้ใช้งานเพจที่มีส่วนร่วมในเนื้อหาของเพจ (engagement) ด้านอายุแบ่งเป็นรุ่น (generation) ตั้งแต่ Baby Boomer ถึง Generation Y และระบุรุ่นไม่ได้ ในการประเมินรอบแรกพบว่ามีร้อยละของความเหมือนของรุ่น (percent overall agreement of 3 raters) อยู่ที่ 65.35% ซึ่งต่ำกว่าค่าที่รับได้ (70%) ผู้วิจัยจึงคัดรายชื่อผู้ใช้ที่มีผลคะแนนความเหมือนต่ำออกมาและให้ผู้ช่วยวิจัยทำการประเมินอีกครั้งหนึ่งด้วยวิธีการต่อรองไปพร้อมกัน (percent overall agreement after negotiation) โดยได้ค่าร้อยละของความเหมือนที่ 98.78% (fixed-marginal kappa = 0.98)
- <u>คำถามเรื่องการระบุเพศ</u> ผู้ช่วยวิจัย 3 ท่านทำการประเมินผู้ใช้งานเพจที่มีส่วนร่วมในเนื้อหาของเพจ (engagement) ด้านการระบุเพศแบ่งเพศหญิง เพศชาย หรือระบุเพศไม่ได้ ในการประเมินรอบแรกพบว่ามีร้อยละของความเหมือนของเพศอยู่ที่ 93.71% (fixed-marginal kappa = 0.85)

2.3.2 การสรุปคุณลักษณะของโพสต์ จากข้อมูลในขั้นตอนที่ 2.2.1 ได้จัดเรียงไว้ใน Microsoft Excel แล้วนั้น ในตารางที่ 1 รายละเอียดของโพสต์ถูกนำมาจัดหมวดหมู่ในของกลุ่มคือ

(1) ประเภทของสื่อที่ใช้ในโพสต์ (media) เช่นข้อความ วิดีโอคลิป รูปภาพ หรือรูปภาพชุด

ตาราง 6 ประเภทของสื่อที่ใช้ในโพสต์

Code	**Count**	**Remark**
Event	2	แบ่งปันกิจกรรม
Image	4	รูปภาพ
Live	3	แบ่งปันกิจกรรมถ่ายทอดสด
Text	1	ข้อความ
Text & Image	48	ข้อความและรูปภาพ
Text & Images	10	ข้อความและรูปภาพชุด
Text & Share Poll	1	ข้อความและแบ่งปันโพล
Text & Shared Link	1	ข้อความและแบ่งปันลิ้งก์
Text & Shared Post	7	ข้อความและแบ่งปันโพสต์อื่น
Text & YouTube	7	ข้อความและแบ่งปันวิดีโอยูทูบ
รวมทั้งสิ้น	**84**	

(2) ประเภทของเนื้อหาการโพสต์ แบ่งเป็นโพสต์ที่อ้างอิงบทความของสำนักฯ (democracy x innovation) โพสต์แบ่งปันการสัมมนา (google meet) โพสต์แบ่งปันพอดแคสต์ (podcast) โพสต์แบ่งปันรายการ THE STANDARD Daily X สถาบันพระปกเกล้า และโพสต์อื่นๆ

ตาราง 7 ประเภทของเนื้อหาการโพสต์

Code	Count	Remark
Others	46	โพสต์อื่นๆ
DemxInno	25	โพสต์ที่อ้างอิงบทความของสำนักฯ
meet.gle	5	โพสต์แบ่งปันการสัมมนา (google meet)
Podcast	3	โพสต์แบ่งปันพอดแคสต์
STD Daily X	5	โพสต์แบ่งปันรายการ THE STANDARD Daily X สถาบันพระปกเกล้า
รวมทั้งสิ้น	**84**	

2.3.3 การสรุปค่าคุณลักษณะอื่นๆ ของผู้ใช้งานเพจที่มีส่วนร่วมในเนื้อหาของเพจ (engagement) จากข้อมูลในขั้นตอนที่ 2.2.1 ที่ได้จัดเรียงไว้ใน Microsoft Excel แล้วนั้น ในตารางที่ 3 รายละเอียดโปรไฟล์ของผู้มีส่วนร่วม ถูกนำมาวิเคราะห์ดังนี้

(1) สถานศึกษา แบ่งหมวดหมู่เท่าที่เก็บข้อมูลได้ 238 รายชื่อจากทั้งหมด 355 รายชื่อคิดเป็นร้อยละ 67.0

(2) สถานที่ทำงาน แบ่งหมวดหมู่เท่าที่เก็บข้อมูลได้ 147 รายชื่อจากทั้งสิ้น 355 รายชื่อคิดเป็นร้อยละ 41.4

2.3.4 การสรุปข้อมูลโดยใช้กราฟสังคม (social network graph) เพื่อสรุปภาพรวมความสัมพันธ์ของโพสต์และผู้ใช้งานที่มีส่วนร่วมในเนื้อหาของเพจ โดยการแปลงข้อมูลให้อยู่ในรูปของตารางกราฟสังคมที่มีสองตารางประกอบกันคือ โหนด (node) และเส้นความสัมพันธ์ (edge) โดยสร้างเป็นกราฟแบบเดินได้สองทาง (undirected graph)

(1) ตารางโหนดมีข้อมูลทั้งสิ้น 442 แถว แบ่งเป็นข้อมูลโพสต์ 84 แถว ข้อมูลผู้ใช้งานที่มีส่วนร่วม 355 แถว และข้อมูลเพจที่มีส่วนร่วม 3 แถว

(2) ตารางเส้นความสัมพันธ์มีข้อมูลทั้งสิ้น 1,078 แถว เป็นข้อมูลความสัมพันธ์แบบเดินได้สองทาง (undirected graph)

2.3.5 การวิเคราะห์ความโน้มเอียงของขั้วความคิดทางการเมือง

(1) นำข้อมูลที่ได้สร้างเป็นไฟล์ Excel เพื่อทำการวิเคราะห์ความโน้มเอียงของขั้วความคิดทางการเมือง โดยเตรียมชีตไว้ 2 ชีต

- ชีตแรกเป็นขั้นตอนการให้คะแนน โดยมี 5 ลำดับคือ 1 คือซ้ายจัด 2 คือเอียงซ้าย 3 คือกลาง 4 คือเอียงขวา และ 5 คือขวาจัด

	A	B
1		
2	ใน Sheet 1 มีตาราง 4 คอลัมน์ เป็นเพจที่ผู้ใช้งานเฟซบุ๊กที่ติดตาม democracyxinnovation	
3	อยากให้ทางทีมงาน democracyxinnovation	
4		
5	1. เลือกเพจที่คิดว่าสะท้อนมุมมองทางการเมืองได้ ในคอลัมน์ E	
6	ถ้าสะท้อนได้ ใส่ 1	
7	ถ้าสะท้อนไม่ได้ ไม่เกี่ยวกับการเมือง ใส่ 0	
8		
9	2. เลือกทัศนคติของเพจนั้นๆ	
10	ส่วนนี้คิดว่าอาจจะทำคล้ายๆ กับเพจนี้ครับ	
11	https://www.allsides.com/media-bias/media-bias-chart	
12	โดยเราอาจจะให้ตัวเลข 1 - 5 ด้วยค่าดังนี้	
13	1 = ซ้ายจัด	
14	2 = ซ้าย	
15	3 = กลาง	
16	4 = ขวา	
17	5 = ขวาจัด	
18		
19		
20		

รูป 7 การเตรียมขั้นตอนให้ทางสถาบันฯ วิเคราะห์ข้อมูล

- และชีตที่สองเป็นข้อมูลเพจ

	C	D	E	F	G
1	**Count**	**URL**	**INCULDE?**	**BIAS**	
2	75	https://www.facebook.com/thestandardth/			
3	69	https://www.facebook.com/TheReportersTH/			
4	64	https://www.facebook.com/thematterco/			
5	55	https://www.facebook.com/themomentumco/			
6	51	https://www.facebook.com/the101.world/			
7	41	https://www.facebook.com/brandthink.me/			
8	39	https://www.facebook.com/VoiceOnlineTH/			
9	34	https://www.facebook.com/ejan2016/			
10	31	https://www.facebook.com/punpromotion/			
11	31	https://www.facebook.com/ThePoliticsByMatichon/			
12	31	https://www.facebook.com/salehere/			
13	29	https://www.facebook.com/NetflixTH/			
14	28	https://www.facebook.com/thestandardpop/			
15	27	https://www.facebook.com/readthecloud/			
16	27	https://www.facebook.com/AgendaPlatform/			
17	26	https://www.facebook.com/bbcnews/			
18	24	https://www.facebook.com/isranewsfanpage/			
19	23	https://www.facebook.com/futuretrends.th/			
20	23	https://www.facebook.com/XinhuaNewsAgency.th/			
21	21	https://www.facebook.com/cafestorythailand/			

รูป 8 การเตรียมข้อมูลให้ทางสถาบันฯ วิเคราะห์ข้อมูล

(2) เมื่อทางทีมงานของสถาบันฯ ส่งข้อมูลกลับมา โดยมีการกรองเพจที่ต้องการใช้ (INCLUDE) และให้คะแนนความโน้มเอียงของขั้วความคิดทางการเมือง ดังภาพด้านล่าง

	B	C	D	E	F
1	PAGE	Cou	URL	INCULDE?	BIAS
2	THE STANDARD	75	https://www.facebook.com/thestandardth/	1	4
3	The Reporters	69	https://www.facebook.com/TheReportersTH/	1	3
4	The MATTER	64	https://www.facebook.com/thematterco/	1	2
5	The Momentum	55	https://www.facebook.com/themomentumco/	1	4
6	The101.world	51	https://www.facebook.com/the101.world/	1	2
7	BrandThink	41	https://www.facebook.com/brandthink.me/	1	2
8	Voice TV	39	https://www.facebook.com/VoiceOnlineTH/	1	1
9	อีจัน	34	https://www.facebook.com/ejan2016/	0	4
10	บันโปร - Punpromotion	31	https://www.facebook.com/punpromotion/	0	3
11	The Politics ข่าวบ้าน การเมือง	31	https://www.facebook.com/ThePoliticsByMatichon/	1	1
12	SALE HERE	31	https://www.facebook.com/salehere/	0	3
13	Netflix	29	https://www.facebook.com/NetflixTH/	0	3
14	THE STANDARD POP	28	https://www.facebook.com/thestandardpop/	1	3
15	The Cloud	27	https://www.facebook.com/readthecloud/	0	3
16	Agenda	27	https://www.facebook.com/AgendaPlatform/	1	2
17	BBC News	26	https://www.facebook.com/bbcnews/	1	2
18	สำนักข่าวอิศรา	24	https://www.facebook.com/isranewsfanpage/	1	4
19	Future Trends	23	https://www.facebook.com/futuretrends.th/	0	3
20	China Xinhua News	23	https://www.facebook.com/XinhuaNewsAgency.th/	1	4
21	Cafe Story : รีวิวคาเฟ่ ทุกวัน	21	https://www.facebook.com/cafestorythailand/	0	3

รูป 9 ข้อมูลที่ถูกส่งกลับมาจากทางสถาบันฯ

(3) ทีมงานวิจัย ได้นำข้อมูลดังกล่าวกลับมาเชื่อมต่อกับตาราง 5 คือการเชื่อมต่อกลับไปที่โปรไฟล์ของผู้ใช้ที่ได้กดไลก์เพจเหล่านี้ไว้ใน Excel ด้วยคำสั่ง VLOOKUP โดยใช้ค่าเฉลี่ยของเพจทั้งหมดที่ทางสถาบันฯ ได้เลือกไว้ เช่น

- นายเอ กดไลก์ THE STANDARD (4), The Reporters (3), The101.world (2) และ Agenda (2)
- ใช้คำสั่ง PivotTable คิดคะแนนความโน้มเอียงของขั้วความคิดทางการเมืองเฉลี่ยของนายเอ คือ

  4 + 3 + 2 + 2 = 2.75 ปัดเศษทศนิยมเป็น 3 ดังนั้นนายเอ จัดอยู่ในกลุ่มที่เป็นกลางทางการเมือง

2.4 สรุปขั้นตอนการเตรียมข้อมูล

ข้อมูลที่เก็บได้ผ่านหน้าเว็บหรือ web scraping นั้น ถูกนำมาทำความสะอาด และจัดเรียงไว้เป็นหมวดหมู่คือกลุ่มข้อมูลเกี่ยวกับโพสต์ และกลุ่มข้อมูลเกี่ยวกับผู้ใช้ โดยข้อมูลสองกลุ่มนี้ถูกปรับปรุงให้มีความสมบูรณ์ขึ้นด้วยเทคนิควิธีต่างกันไป

ข้อมูลโพสต์ได้ถูกจัดหมวดหมู่ออกเป็นสองกลุ่มคือ หมวดหมู่ของสื่อที่ใช้ในการโพสต์ และหมวดหมู่ประเภทเนื้อหาของโพสต์ ทั้งหมดผ่านการเขียนเงื่อนไขตรวจสอบใน Microsoft Excel ข้อมูลผู้ใช้ได้ถูกจัดหมวดหมู่ไว้หลายด้าน คือด้านประวัติการศึกษาและสถานที่ทำงาน ซึ่งใช้การเขียนเงื่อนไขตรวจสอบใน

Microsoft Excel เช่นเดียวกัน และด้านอายุ (รุ่น) และเพศ โดยใช้ผลการประเมินของผู้ช่วยวิจัย ผ่านการตรวจสอบความเที่ยงระหว่างผู้ประเมิน เพื่อให้ได้ข้อมูลที่น่าเชื่อถือ

ในขั้นสุดท้ายของการเตรียมข้อมูล ทีมงานได้นำข้อมูลทั้งหมดมาปรับให้อยู่ในรูปแบบของตารางกราฟสังคมแบบเดินได้สองทาง (undirected network graph) เพื่อทำการวิเคราะห์ความสัมพันธ์ของผู้ใช้ในแง่มุมต่างๆ และโพสต์ของสถาบันฯ ต่อไป

บทที่ 3 การวิเคราะห์ข้อมูล

ในขั้นตอนการวิเคราห์ข้อมูล แบ่งเป็นการวิเคราะห์สถิติเชิงพรรณาและการแสดงภาพข้อมูล (data visualization) ทั้งในรูปแบบกราฟทั่วไปด้วยโปรแกรม Microsoft Excel เช่น ตาราง (table) ทรีแมป (treemap) ตัวหนังสือภาพ (pictogram) และกราฟเครือข่ายสังคม (social network graph) ด้วยโปรแกรม Gephi 0.9.2

3.1 การวิเคราะห์สถิติเชิงพรรณาเกี่ยวกับการโพสต์

ข้อมูลการโพสต์ในเพจสถาบัน มีความหลากหลายทั้งทางด้านเนื้อหาและสื่อที่ใช้ ซึ่งได้แสดงไว้ในตารางที่ 6 และตารางที่ 7 ในบทที่ 2 แล้ว ในด้านความสัมพันธ์ของเนื้อหาที่ใช้ในโพสต์ต่อจำนวนการมีส่วนร่วม (engagement) ของผู้ใช้งานนั้น มีปริมาณดังนี้

ตาราง 8 จำนวนการมีส่วนร่วมแบ่งตามประเภทเนื้อหาการโพสต์

Code	Count of Posts	Count of engagement	% of engagement
Others	46	573	53%
DemxInno	25	349	32%
meet.gle	5	51	5%
Podcast	3	70	6%
STD Daily X	5	35	3%
รวมทั้งสิ้น	**84**	**1078**	**100%**

ความสัมพันธ์ของสื่อที่ใช้ในโพสต์ต่อจำนวนการมีส่วนร่วม (engagement) ของผู้ใช้งานนั้นมีปริมาณดังนี้

ตาราง 9 จำนวนการมีส่วนร่วมแบ่งตามประเภทประเภทสื่อที่ใช้ในโพสต์

Media	Count of Posts	Count of engagement	% of engagement
Text & Shared Link	1	17	1.58%
Text & Image	48	735	68.18%
Event	2	30	2.78%

Media	Count of Posts	Count of engagement	% of engagement
Text & YouTube	7	87	8.07%
Text & Images	10	114	10.58%
Image	4	29	2.69%
Text & Shared Post	7	41	3.80%
Live	3	17	1.58%
Text	1	5	0.46%
Text & Share Poll	1	3	0.28%
รวมทั้งสิ้น	**84**	**1078**	**100%**

ค่าเฉลี่ย (mean) และส่วนเบี่ยงเบนมาตรฐาน (standard deviation หรือ SD) ของสื่อที่ใช้ในโพสต์ต่อจำนวนการมีส่วนร่วม (engagement) ของผู้ใช้งานนั้นมีดังนี้

ตาราง 10 ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของสื่อที่ใช้ในโพสต์

Media	Count of Posts	Count of engagement	Mean	SD
Text & Shared Link	1	17	17.00	0.00
Text & Image	48	735	15.31	9.77
Event	2	30	15.00	8.00
Text & YouTube	7	87	12.43	7.07
Text & Images	10	114	11.40	6.86
Image	4	29	7.25	4.71
Text & Shared Post	7	41	5.86	3.36
Live	3	17	5.67	0.47
Text	1	5	5.00	0.00

Media	Count of Posts	Count of engagement	Mean	SD
Text & Share Poll	1	3	3.00	0.00
รวมทั้งสิ้น	**84**	**1078**	**12.83**	**9.01**

จากตารางที่ 9 จะเห็นว่าแม้ข้อความและภาพ (Text & Image) มีจำนวนการมีส่วนร่วมมากที่สุด (735) และมีค่าเฉลี่ยในการมีส่วนร่วมสูงที่สุดเช่นกัน (Mean = 15.31) แต่ก็มีส่วนเบี่ยงเบนมาตรฐานสูงเช่นกัน (SD = 9.77) แสดงให้เห็นว่ามีการกระจายตัวของการมีส่วนร่วมสูง

ความยาวของข้อความมีความหลากหลายอยู่สูงมากในจำนวนข้อมูลโพสต์ที่เก็บได้ ตั้งแต่ 41 ตัวอักษร ถึง 2773 ตัวอักษรและขนาดของข้อความ (word length) ไม่มีส่วนสัมพันธ์กับจำนวนการมีส่วนร่วม (engagement) ประเภทของสื่อ (media) หรือประเภทของเนื้อหา (content category) แต่อย่างใดดังแสดงในตารางเมทริกซ์สหสัมพันธ์ (correlation matrix)

ตาราง 11 ตารางแสดงเมทริกซ์สหสัมพันธ์ของตัวแปรที่เกี่ยวกับข้อมูลการโพสต์และการมีส่วนร่วม

	Word Length	Engagement	Media	Category
Word Length	1			
Engagement	0.182798564	1		
Media	0.142486852	-0.083270749	1	
Category	0.156999583	-0.024229438	0.223777346	1

3.2 การวิเคราะห์สถิติเชิงพรรณาเกี่ยวกับการผู้ใช้ที่มีส่วนร่วม

ข้อมูลผู้ใช้งานที่มีส่วนร่วม (engagement) กับเนื้อหาของเพจที่สามารถเก็บได้ผ่านการเก็บข้อมูลจากหน้าเว็บ (web scraping) นั้นพบว่าการเข้าถึงข้อมูลที่เป็นสาธารณะในระดับผู้ใช้งานหรือหน้าข้อมูลส่วนตัว (profile page) นั้นค่อนข้างจำกัดทั้งข้อมูลที่ระบุโดยผู้ใช้งานเองเช่นประวัติการทำงานหรือการศึกษาและข้อมูลที่ต้องนำมาให้ความหมายต่อโดยทีมผู้วิจัยเช่นข้อมูลเพศหรือช่วงอายุ

โดยทีมผู้วิจัยพบว่าข้อมูลที่ดึงมาได้และต้องนำมาให้ความหมายสามารถทำได้มากกว่าแก่ข้อมูลรูปโปรไฟล์ จึงขออธิบายในส่วนที่มีข้อมูลมากกว่าเป็นส่วนแรกดังนี้ ข้อมูลผู้ใช้งานทั้งหมดแบ่งตามช่วงอายุ (รุ่น) ตามการ

ประเมินของผู้ช่วยวิจัย ได้ผลดังตารางที่ 12 ข้างล่างนี้ จะเห็นว่ามีข้อมูลรูปภาพโปรไฟล์ที่ไม่สามารถระบุช่วงอายุได้ร้อยละ 5.35

ตาราง 12 จำนวนและร้อยละของผู้ใช้งานที่มีส่วนร่วมแบ่งตามช่วงอายุ (รุ่น)

Generation	No. of Users	% of Users
Boomer	10	2.82%
Gen X	62	17.46%
Gen Y	230	64.79%
Gen Z	34	9.58%
N/A (ไม่สามารถระบุได้)	19	5.35%

ข้อมูลรูปภาพเหล่าบางส่วนยังไม่สามารถใช้ระบุเพศได้เช่นกัน ดังแสดงในตารางข้างล่างนี้ โดยมีข้อมูลรูปโปรไฟล์ที่ไม่สามารถระบุเพศได้ร้อยละ 1.41

ตาราง 13 จำนวนและร้อยละของผู้ใช้งานที่มีส่วนร่วมแบ่งตามเพศ

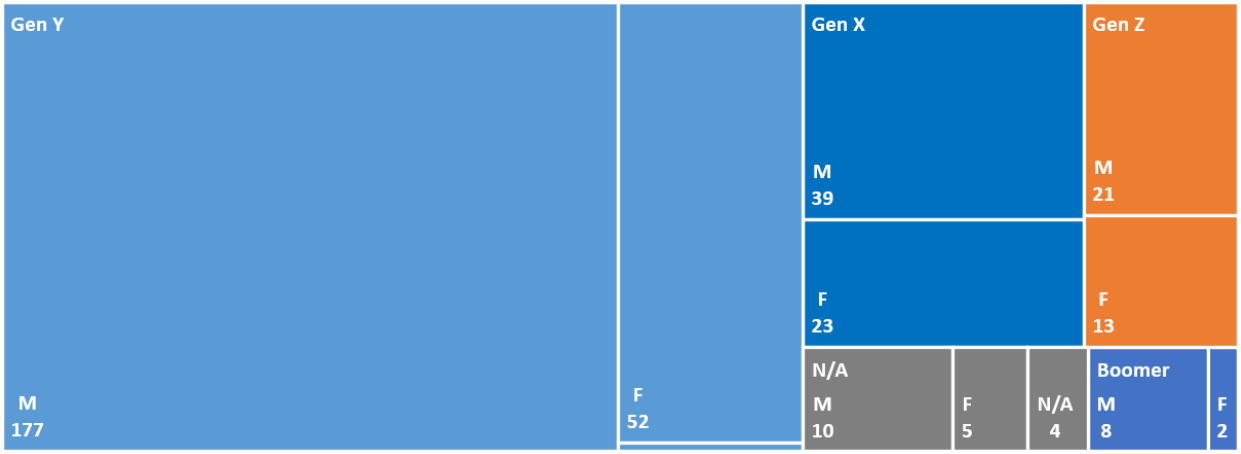
Gender	No. of Users	% of Users
Female	95	26.76%
Male	255	71.83%
N/A (ไม่สามารถระบุได้)	5	1.41%

เมื่อนำข้อมูลในส่วนของอายุและเพศมาแสดงแจกแจงแบบไขว้ (crosstabs) จะมีผลดังนี้

ตาราง 14 การแจกแจงแบบไขว้ของเพศและอายุของผู้ใช้งาน

Generation	No. of Users	% of Users
Boomer		
Female	2	0.56%
Male	8	2.25%
Gen X		
Female	23	6.48%
Male	39	10.99%
Gen Y		
Female	52	14.65%
Male	177	49.86%
N/A	1	0.28%
Gen Z		
Female	13	3.66%
Male	21	5.92%
N/A (ไม่สามารถระบุได้)		
Female	5	1.41%
Male	10	2.82%
N/A	4	1.13%
รวมทั้งสิ้น	**355**	**100.0%**

ข้อมูลชุดเดียวกันสามารถแสดงเป็นทรีแมป (treemap) ซึ่งเหมาะสำหรับการเปรียบเทียบสัดส่วนภายในลำดับชั้นจะได้ภาพดังนี้



รูป 10 แผนภูมิทรีแมปเปรียบเทียบสัดส่วนอายุและเพศของผู้ใช้งาน

ข้อมูลการศึกษาทั้งในอดีตและปัจจุบันสามารถระบุได้รวม 65 บัญชีผู้ใช้จากทั้งสิ้น 355 ราย คิดเป็นร้อยละ 18.3 หรือจากบัญชีผู้ใช้ทั้งสิ้น 238 รายที่แสดงข้อมูลการศึกษา คิดเป็นร้อยละ 27

โดยในจำนวนทั้งหมด 65 บัญชีแบ่งตามช่วงอายุที่ระบุว่าศึกษาทางรัฐศาสตร์และนิติศาสตร์ดังนี้

ตาราง 15 การศึกษาของผู้ใช้ที่อยู่ในกลุ่มรัฐศาสตร์และนิติศาสตร์แบ่งตามช่วงอายุ

Generation	No. of Users studied Political Science	No. of Users studied Law
Boomer	0	0
Gen X	5	2
Gen Y	32	11
Gen Z	11	2
N/A (ไม่สามารถระบุได้)	2	0
รวมทั้งสิ้น	**50**	**15**

ข้อมูลสถานที่ทำงานทั้งในอดีตและปัจจุบันสามารถระบุได้รวม 147 บัญชีผู้ใช้จากทั้งสิ้น 355 ราย คิดเป็นร้อยละ 41.4 โดยข้อมูลสถานที่ทำงานทั้งในอดีตและปัจจุบันของผู้ใช้ที่ระบุข้อมูลในหน้าข้อมูลส่วนตัว (profile

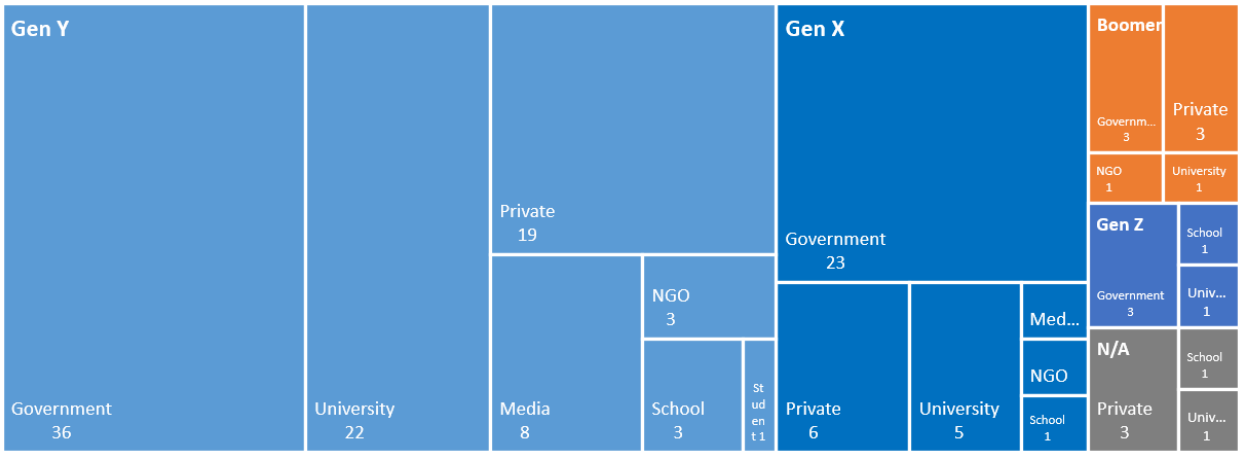
page) แบ่งตามประเภทสถานที่ทำงานคือภาครัฐ (Government) ภาคเอกชน (Private) องค์กรไม่แสวงหาผลกำไร (NGO) โรงเรียน (School) มหาวิทยาลัย (University) สื่อ (Media) และนักเรียน สามารถแสดงได้ดังนี้

ตาราง 16 ข้อมูลช่วงอายุและสถานที่ทำงานของผู้ใช้ที่ระบุได้

GENERATION	WORKS	No. of Users	% of Users
Boomer			
	Government	3	2.0%
	NGO	1	0.7%
	Private	3	2.0%
	University	1	0.7%
Gen X			
	Government	23	15.6%
	Media	1	0.7%
	NGO	1	0.7%
	Private	6	4.1%
	School	1	0.7%
	University	5	3.4%
Gen Y			
	Government	36	24.5%
	Media	8	5.4%
	NGO	3	2.0%
	Private	19	12.9%
	School	3	2.0%
	Student	1	0.7%
	University	22	15.0%
Gen Z			
	Government	3	2.0%
	School	1	0.7%
	University	1	0.7%
N/A			
	Private	3	2.0%
	School	1	0.7%
	University	1	0.7%

GENERATION	WORKS	No. of Users	% of Users
	รวมทั้งสิ้น	**147**	

ข้อมูลชุดเดียวกันสามารถแสดงเป็นทรีแมป (treemap) ซึ่งเหมาะสำหรับการเปรียบเทียบสัดส่วนภายในลำดับชั้น จะได้ภาพดังนี้



รูป 11 แผนภูมิทรีแมปเปรียบเทียบสัดส่วนอายุและที่ทำงานของผู้ใช้งาน

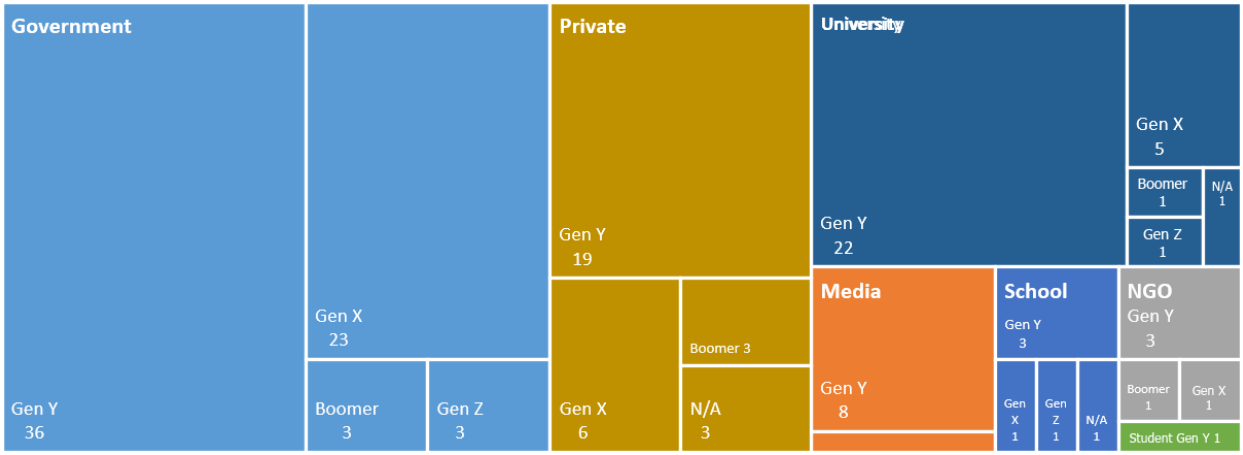
เพื่อเปรียบเทียบโดยยึดสถานที่ทำงานเป็นข้อมูลหลัก ข้อมูลชุดเดียวกันสามารถแสดงตารางได้ดังนี้

ตาราง 17 ข้อมูลสถานที่ทำงานและช่วงอายุของผู้ใช้ที่ระบุได้

WORKS	GENERATION	No. of Users	% of Users
Government			
	Boomer	3	2.0%
	Gen X	23	15.6%
	Gen Y	36	24.5%
	Gen Z	3	2.0%
Media			
	Gen X	1	0.7%
	Gen Y	8	5.4%
NGO			
	Boomer	1	0.7%
	Gen X	1	0.7%
	Gen Y	3	2.0%

WORKS	GENERATION	No. of Users	% of Users
Private			
	Boomer	3	2.0%
	Gen X	6	4.1%
	Gen Y	19	12.9%
	N/A	3	2.0%
School			
	Gen X	1	0.7%
	Gen Y	3	2.0%
	Gen Z	1	0.7%
	N/A	1	0.7%
Student			
	Gen Y	1	0.7%
University			
	Boomer	1	0.7%
	Gen X	5	3.4%
	Gen Y	22	15.0%
	Gen Z	1	0.7%
	N/A	1	0.7%
	รวมทั้งสิ้น	**147**	

และข้อมูลข้างต้นสามารถแสดงเป็นทรีแมป (treemap) ซึ่งเหมาะสำหรับการเปรียบเทียบสัดส่วนภายในลำดับชั้น จะได้ภาพดังนี้



รูป 12 แผนภูมิทรีแมปเปรียบเทียบสัดส่วนที่ทำงานและอายุของผู้ใช้งาน

เปรียบเทียบสัดส่วนของผู้ใช้ที่มีความโน้มเอียงของขั้วความคิดทางการเมืองแบบต่างๆ ข้อมูลความถี่ของการกดไลก์เพจ (Count) และข้อมูลคะแนนความโน้มเอียงของขั้วความคิดทางการเมือง (Bias) แสดงได้รูปแบบตารางได้ดังนี้

ตาราง 18 ความถี่ของเพจที่มีคนไลก์ 20 อันดับแรก

Page	Count	Bias
THE STANDARD (080)	75	4
The Reporters (079)	69	3
The MATTER (074)	64	2
The Momentum (075)	55	4
The101.world (083)	51	2
BrandThink	41	2
Voice TV	39	1
The Politics ข่าวบ้าน การเมือง	31	1
THE STANDARD POP	28	3
Agenda	27	2

BBC News	26	2
สำนักข่าวอิศรา	24	4
China Xinhua News	23	4
CNN	20	1
ThaiPublica	19	4
GM Live	18	2
สำนักข่าวไทย	17	4
Prachatai English	17	1
China Report ASEAN - Thailand	16	4
Prachachat - ประชาชาติ	15	4

เมื่อทำการคำนวณคะแนนความโน้มเอียงของขั้วความคิดทางการเมืองของผู้ใช้งานทั้งสิ้น 355 คน สามารถเข้าถึงหน้าไลก์ได้ 250 คน (70.42%) สามารถแสดงสัดส่วนผู้ใช้ที่มีคะแนนอยู่ในช่วงความโน้มเอียงของขั้วความคิดทางการเมืองได้ดังนี้ จะเห็นว่าผู้ใช้ส่วนใหญ่อยู่ในกลุ่มเอียงซ้าย (46.00%) และกลุ่มที่เป็นกลาง (38.40%)

ตาราง 19 ร้อยละของผู้ใช้ในในช่วงความโน้มเอียงของขั้วความคิดทางการเมือง

ขั้วความคิดทางการเมือง	**จำนวนเพจ**	**ร้อยละ**
ซ้ายจัด (1)	7	2.80%
เอียงซ้าย (2)	115	46.00%
กลาง (3)	96	38.40%
เอียงขวา (4)	29	11.60%
ขวาจัด (5)	3	1.20%
	250	**100.00%**

3.3 การวิเคราะห์ความสัมพันธ์ผู้ใช้งานที่มีส่วนร่วมกับโพสต์ด้วยกราฟเครือข่ายสังคม

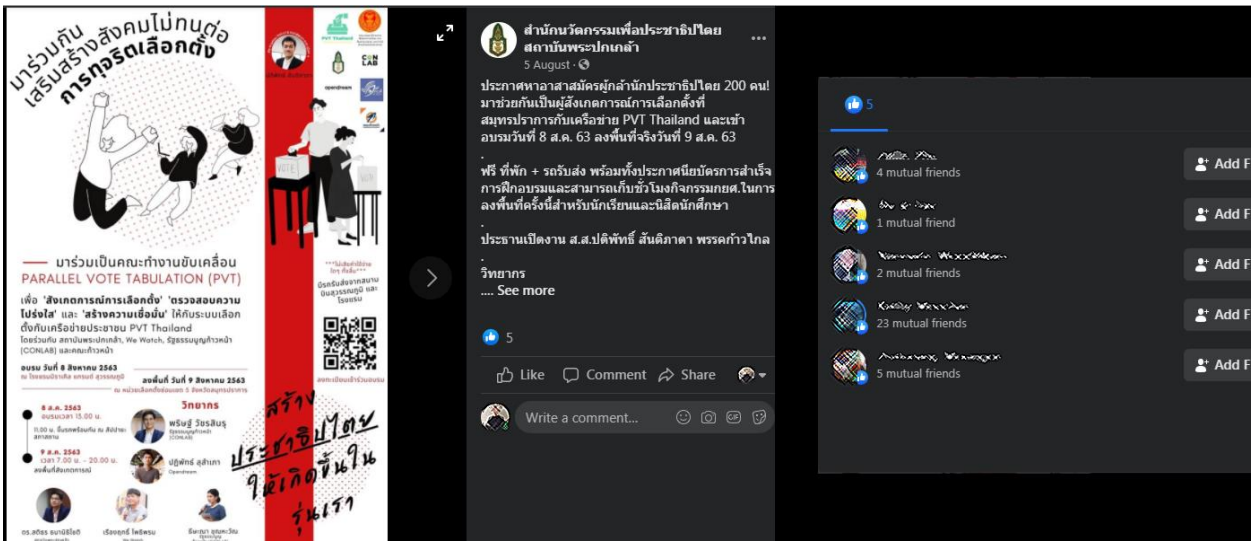
กราฟเครือข่ายสังคมคือกราฟแสดงความสัมพันธ์ระหว่างโหนด (node) และเส้นความสัมพันธ์ (edge) โดยสามารถสร้างเป็นกราฟแบบเดินได้สองทาง (undirected graph) (Wasserman & Faust, 1994) ทีมวิจัยทำการเตรียมข้อมูลเพื่อนำเข้าโปรแกรม Gephi 0.9.2 ในการสร้างภาพเครือข่ายสังคมจำเป็นต้องใช้ข้อมูลตารางสองชุดตารางโหนด (Node Table) และตารางเส้นความสัมพันธ์ (Edge Table)

ข้อมูลในตารางโหนดมีรายละเอียดดังนี้

(1) ข้อมูลโพสต์ 84 แถว
(2) ข้อมูลผู้ใช้งานที่มีส่วนร่วม 355 แถว และ
(3) ข้อมูลเพจที่มีส่วนร่วม 3 แถว

รวมทั้งสิ้นมีข้อมูลในตารางโหนด จำนวน 442 แถว ข้อมูลในตารางเส้นความสัมพันธ์คือเส้นเชื่อมต่อระหว่างโหนดข้อมูลโพสต์ (1) และโหนดข้อมูลผู้ใช้งานหรือเพจที่มีส่วนร่วมกับโพสต์นั้นๆ (2, 3)

ยกตัวอย่างเช่นโพสต์ที่มีผู้ใช้มีส่วนร่วม 5 คนดังภาพด้านล่าง



รูป 13 ตัวอย่างโพสต์และการเก็บข้อมูลตาราง

ในตารางโหนดจะมีข้อมูลอยู่ 1 แถว ประกอบด้วยรายละเอียดของโพสต์นั้นคือ เลขกำกับโพสต์ (Label) URL ของโพสต์ ข้อความทั้งหมดของโพสต์ (Content) ความยาวข้อข้อความ จำนวนการมีส่วนร่วมของผู้ใช้ (Engagement) สื่อที่ใช้ (Media) และประเภทเนื้อหา (Category) ซึ่งแสดงในแถวที่ 9 ในตารางด้านล่างนี้

ตาราง 20 ตัวอย่างตารางโหนดในส่วนข้อมูลโพสต์

ID	Label	URL	Content	Length	Engagement	Media	Category
1	post 001	facebook.com//permalink.php?st...	คณะกรรมาธิ...	1008	5	Text	
2	post 002	facebook.com//117440689653208/...	ดุลยภาพ จา...	912	21	Text & Image	DemxInno
3	post 003	facebook.com//117440689653208/...	สรุปการชุม...	625	17	Text & Image	DemxInno
4	post 004	facebook.com//permalink.php?st...	ประกาศการเ...	1136	4	Text & Shared Post	
5	post 005	facebook.com//117440689653208/...	นับคะแนนเล...	41	5	Live	
6	post 006	facebook.com//117440689653208/...	นับคะแนนเล...	81	6	Live	
7	post 007	facebook.com//permalink.php?st...	การอบรมผู้...	1052	2	Text & Shared Post	
8	post 008	facebook.com//117440689653208/...	PVT หนทางส...	1106	1	Text & Image	
9	**post 009**	**facebook.com//117440689653208/...**	**ประกาศหาอา...**	**1255**	**5**	**Text & Image**	

อีกส่วนของตารางโหนดซึ่งเก็บข้อมูลผู้ใช้งานจะแสดงข้อมูลเลขกำกับผู้ใช้ URL ของผู้ใช้ ประเภทของการศึกษาและสถานที่ทำงาน เพศ ช่วงอายุ และคะแนนความโน้มเอียงของขั้วความคิดทางการเมือง โดยผู้ใช้ 5 ท่านที่มีส่วนร่วมกับโพสต์แถวที่ 9 มีข้อมูลบางส่วนในตารางดังนี้

ตาราง 21 ตัวอย่างตารางโหนดในส่วนข้อมูลผู้ใช้ที่มีส่วนร่วม

Id	Label	Title	Ed Pol Sci	Ed Law	Work-Type	Gender	Age	Bias
99	user 015	Nan ...	0	0	gov...	หญิง	Gen X	5
103	user 019	Pai ...	0	0	uni...	หญิง	Gen Y	2
122	user 038	...	0	0		ชาย	Gen X	0
133	user 049	Suth...	0	0	uni...	หญิง	Gen Y	2
134	user 050	Prat...	0	0		ชาย	Gen Y	0

ข้อมูลในตารางเส้นความสัมพันธ์มีจำนวนทั้งสิ้น 1078 แถว เป็นข้อมูลชุดเชื่อมต่อของโพสต์และผู้ใช้งาน จากตัวอย่างข้างต้นของโพสต์แถวที่ 9 ที่มีจำนวนผู้มีส่วนร่วม 5 ท่าน ในตารางเส้นความสัมพันธ์จะมีการจัดความสัมพันธ์ของต้นทาง (Source) และปลายทาง (Target) ดังนี้

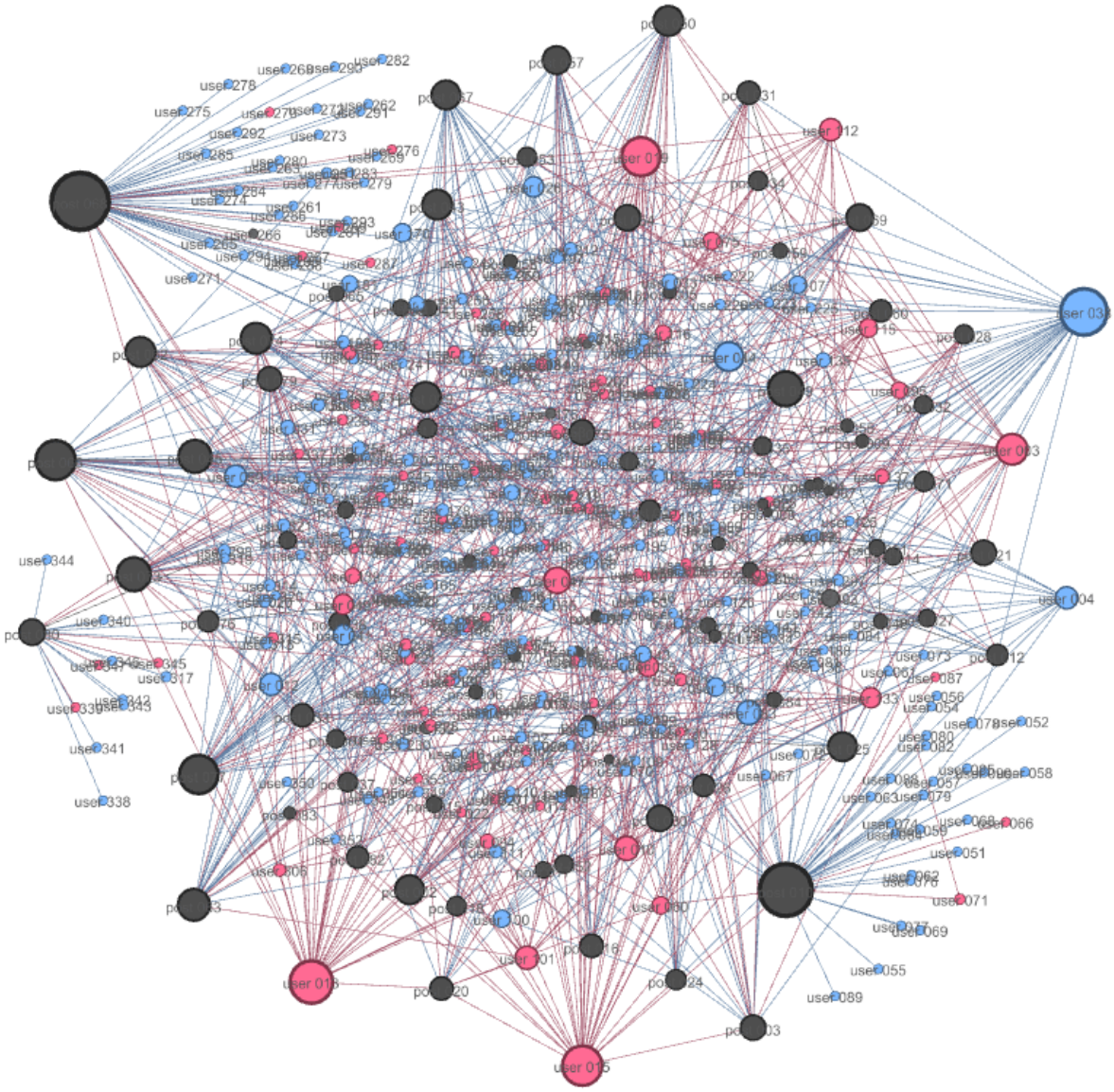
ตาราง 22 ตัวอย่างตารางเส้นความสัมพันธ์

Source	Target
user ...	post 008
user 015	post 009
user 019	post 009
user 038	post 009
user 049	post 009
user 050	post 009
user ...	post 010
user ...	post 010

ข้อมูลที่นำเข้าโปรแกรม Gephi 0.9.2 จะสามารถแสดงกราฟเครือข่ายสังคมได้ เห็นรูปทรงของเครือข่ายที่เชื่อมโยงกัน รวมถึงการกำหนดสีและขนาดให้กับคุณสมบัติต่างๆ ของโหนดและเส้นความสัมพันธ์ได้ เช่นการกำหนดสีให้กับผู้ใช้ตามคุณสมบัติเช่นสนานที่ทำงานหรือช่วงอายุ การกำหนดขนาดตามจำนวนการเชื่อมต่อระหว่างโหนด (degree) โดยในการออกแบบกราฟเครือข่ายสังคมในครั้งนี้เลือกใช้สีของโหนดแทนเพศที่ระบุได้ของผู้ใช้ เช่นสีชมพูแทนเพศหญิง สีฟ้าแทนเพศชาย และสีเทาแทนผู้ใช้ที่ระบุเพศไม่ได้และโหนดของโพสต์ที่ไม่มี

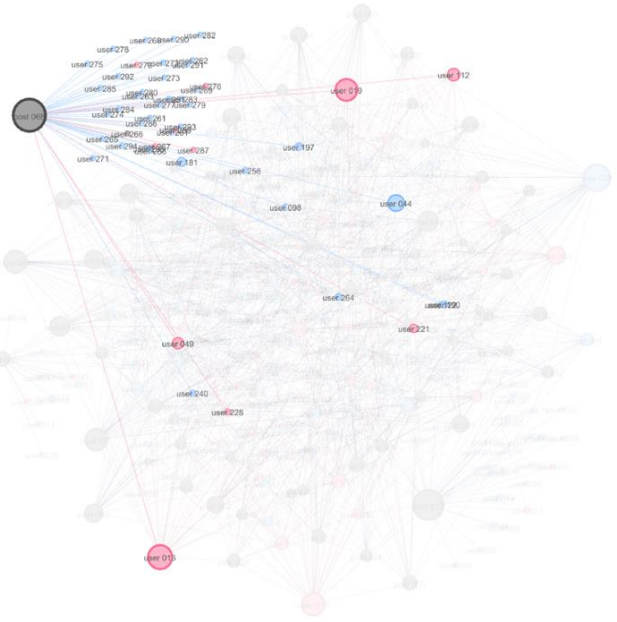
เพศเช่นกัน สำหรับขนาดของโหนดจะเป็นไปตามจำนวนเส้นที่เชื่อมต่อระหว่างโหนด (degree) ยกตัวอย่างเช่น โพสต์แถวที่ 9 ที่มีจำนวนผู้มีส่วนร่วม 5 ท่าน จะได้ขนาดเท่ากับ 5 จุดเป็นต้น

กราฟเครือข่ายสังคมที่กำหนดคุณสมบัติต่างๆ ดังที่กล่าวมาข้างต้นแสดงได้ดังนี้



รูป 14 กราฟเครือข่ายสังคมของโพสต์และผู้ใช้งานที่มีส่วนร่วม

โพสต์ที่ผู้ใช้มีส่วนร่วมมากเป็นอันดับหนึ่งจากชุดข้อมูลที่เก็บในการวิจัยครั้งนี้ได้แก่โพสต์ที่ 068 ซึ่งถูกจัดอยู่ในกลุ่มเนื้อหาอื่นๆ (Others) มีจำนวนการมีส่วนร่วม 50 ครั้ง ดังภาพด้านล่าง



รูป 15 กราฟเครือข่ายสังคมและภาพหน้าจอของโพสต์อันดับหนึ่ง

โดยโพสต์นี้มีผู้ใช้ที่มีส่วนร่วมจากหลายรุ่นดังนี้

ตาราง 23 ตารางรายละเอียดข้อมูลการมีส่วนร่วมของโพสต์อันดับหนึ่ง

Generation	Female	Male	N/A	Total
Gen X		2		2
Gen Y	10	28		38
Gen Z	1	6		7
N/A		2	1	3
			รวมการมีส่วนร่วม	**50**

โพสต์ที่ผู้ใช้มีส่วนร่วมมากเป็นอันดับสองจากชุดข้อมูลที่เก็บในการวิจัยครั้งนี้ได้แก่โพสต์ที่ 010 มีจำนวน ซึ่งถูกจัดอยู่ในกลุ่มโพสต์ที่อ้างอิงบทความของสำนักฯ (Democracy x Innovation) ได้รับการมีส่วนร่วม 45 ครั้ง ดังภาพด้านล่าง

รูป 16 กราฟเครือข่ายสังคมและภาพหน้าจอของโพสต์อันดับสอง

โดยโพสต์นี้มีรายละเอียดผู้ใช้ที่มีส่วนร่วมดังนี้

ตาราง 24 ตารางรายละเอียดข้อมูลการมีส่วนร่วมของโพสต์อันดับสอง

Generation	Female	Male	N/A	Total
Gen X	4	9		13
Gen Y	5	26		31
Gen Z				
N/A		1		1
			รวมการมีส่วนร่วม	**45**

โพสต์ที่ผู้ใช้มีส่วนร่วมมากเป็นอันดับที่สามจากชุดข้อมูลที่เก็บในการวิจัยครั้งนี้ได้แก่โพสต์ที่ 066 มีจำนวนจัดอยู่ในกลุ่มโพสต์แบ่งปันพอดแคสต์การมีส่วนร่วม 32 ครั้ง ดังภาพด้านล่าง

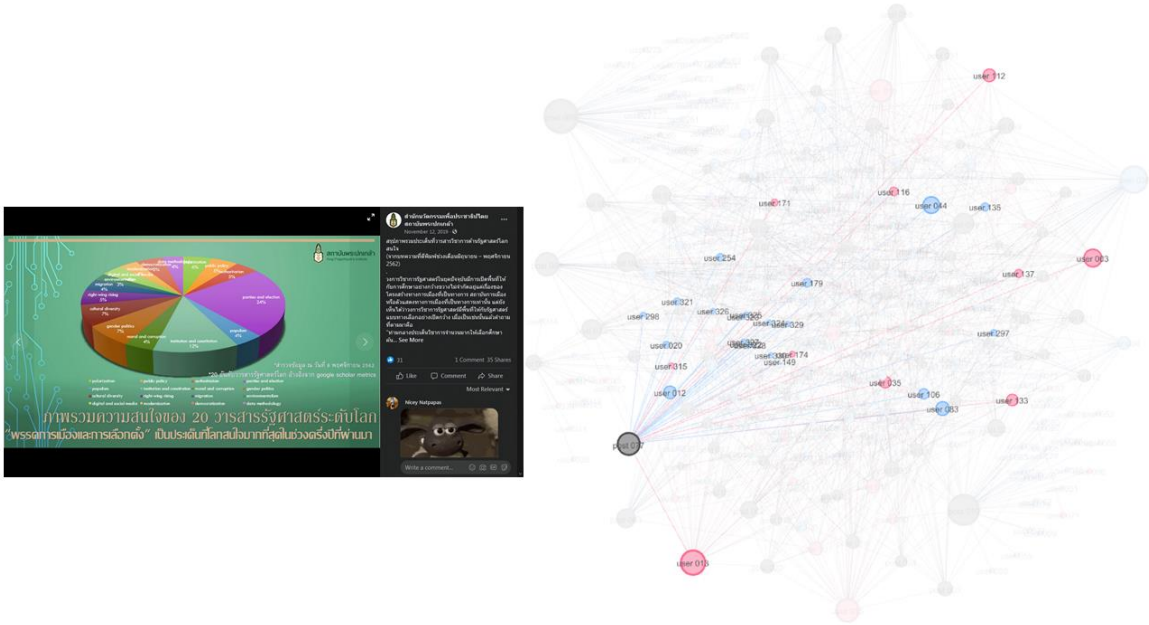
รูป 17 กราฟเครือข่ายสังคมและภาพหน้าจอของโพสต์อันดับสาม

โดยโพสต์นี้มีรายละเอียดผู้ใช้ที่มีส่วนร่วมดังนี้

ตาราง 25 ตารางรายละเอียดข้อมูลการมีส่วนร่วมของโพสต์อันดับสาม

Generation	Female	Male	N/A	Total
Gen X	1	1		2
Gen Y	6	15		21
Gen Z	1	4		5
N/A	2	1	1	4
			รวมการมีส่วนร่วม	**32**

โพสต์ที่ผู้ใช้มีส่วนร่วมมากเป็นอันดับที่สี่จากชุดข้อมูลที่เก็บในการวิจัยครั้งนี้ได้แก่โพสต์ที่ 077 ซึ่งถูกจัดอยู่ในกลุ่มเนื้อหาอื่นๆ (Others) มีจำนวนการมีส่วนร่วม 31 ครั้ง ดังภาพด้านล่าง



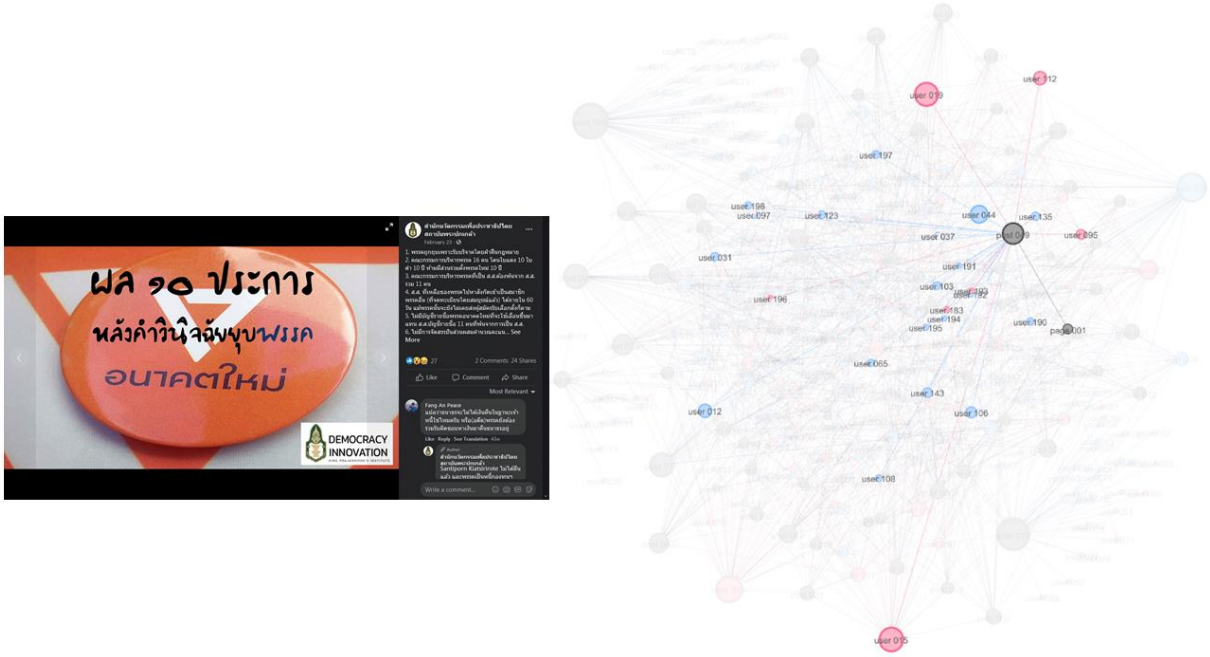
รูป 18 กราฟเครือข่ายสังคมและภาพหน้าจอของโพสต์อันดับสี่

โดยโพสต์นี้มีรายละเอียดผู้ใช้ที่มีส่วนร่วมดังนี้

ตาราง 26 ตารางรายละเอียดข้อมูลการมีส่วนร่วมของโพสต์อันดับสี่

Generation	Female	Male	N/A	Total
Gen X	3	3		6
Gen Y	7	15		22
Gen Z				
N/A		3		3
			รวมการมีส่วนร่วม	**31**

โพสต์ที่ผู้ใช้มีส่วนร่วมมากเป็นอันดับที่ห้าจากชุดข้อมูลที่เก็บในการวิจัยครั้งนี้ได้แก่โพสต์ที่ 049 ซึ่งถูกจัดอยู่ในกลุ่มเนื้อหาอื่นๆ (Others) เช่นกัน และมีจำนวนการมีส่วนร่วม 27 ครั้ง ดังภาพด้านล่าง



รูป 19 กราฟเครือข่ายสังคมและภาพหน้าจอของโพสต์อันดับห้า

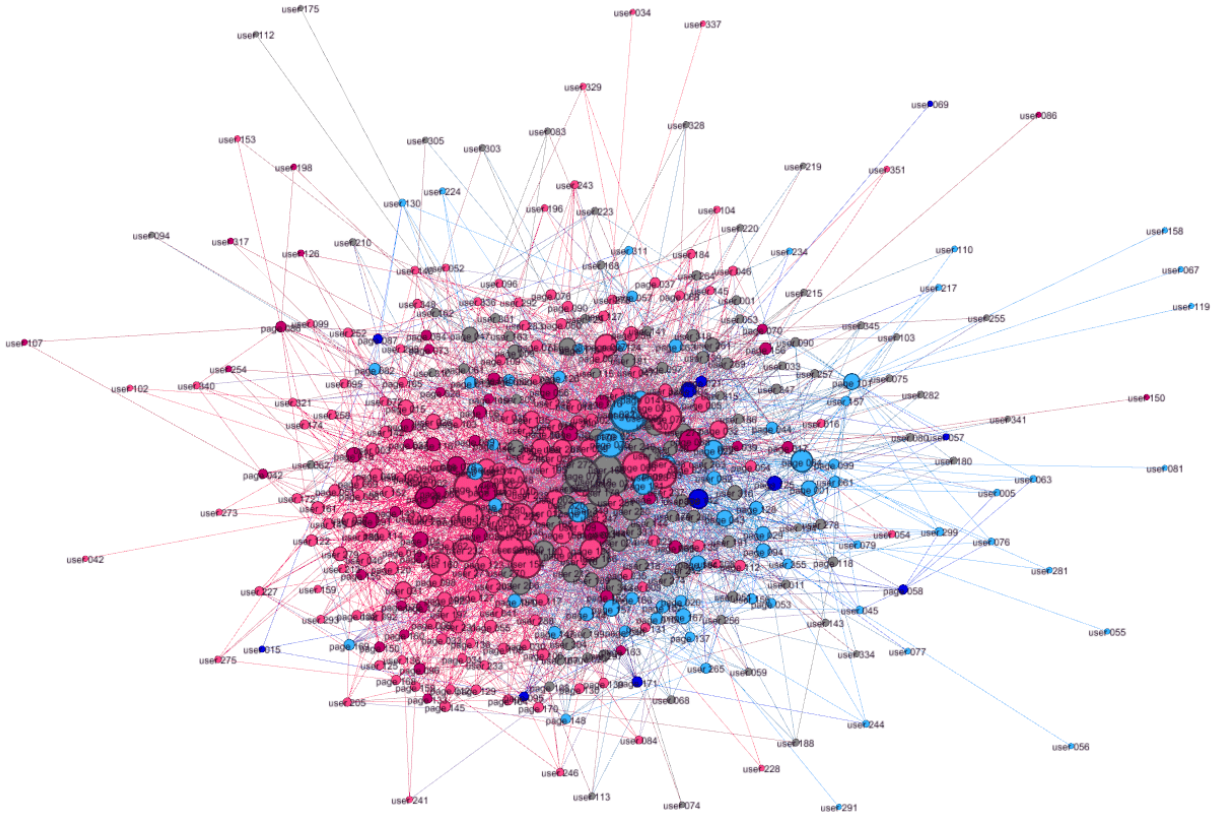
โดยโพสต์นี้มีรายละเอียดผู้ใช้ที่มีส่วนร่วมดังนี้

ตาราง 27 ตารางรายละเอียดข้อมูลการมีส่วนร่วมของโพสต์อันดับห้า

Generation	Female	Male	N/A	Total
Gen X	2	1		3
Gen Y	3	15		18
Gen Z	2	3		5
N/A			1	1
			รวมการมีส่วนร่วม	**27**

ในส่วนของคะแนนความโน้มเอียงของขั้วความคิดทางการเมืองนั้น ในภาพรวมซึ่งได้รายงานในตารางที่ 19 จะเห็นว่าผู้ใช้ส่วนใหญ่อยู่ในกลุ่มเอียงซ้าย (46%) และกลุ่มที่เป็นกลาง (38.40%) โดยมีค่าเฉลี่ยของผู้ใช้งานทั้งหมดอยู่ที่ 2.62 มัธยฐานอยู่ที่ 3.00 ซึ่งตีความได้ว่าโดยเฉลี่ยแล้วผู้ใช้งานของเพจสถาบันฯ มีความเป็นกลางทางการเมือง

ในภาพรวมของเครือข่ายสังคมของเพจที่ถูกกดไลก์และผู้ใช้งาน เมื่อลงรหัสสีให้เพจหรือผู้ใช้ตามความโน้มเอียงของขั้วความคิดทางการเมือง โดยให้สีแดงเข้มแสดงกลุ่มซ้ายจัด สีแดงชมพูแสดงกลุ่มเอียงซ้าย สีเทาแสดงกลุ่มเป็นกลาง สีฟ้าแสดงกลุ่มเอียงขวา และสีน้ำเงินเข้มแสดงกลุ่มขวาจัด ขนาดของโหนด (nodes) แทนจำนวนการเชื่อมต่อระหว่างโหนด หรือดีกรี (degree) จะเห็นว่ากลุ่มโหนดสีแดง (เอียงซ้ายและซ้ายจัด) มีจำนวนมากกว่าสีเทา (เป็นกลาง) และสีฟ้ากับน้ำเงิน (เอียงขวาและขวาจัด) แต่มีโหนดเพจเอียงขวาสองโหนดที่มีผู้ใช้งานกดไลก์มากติด 5 อันดับแรกคือ THE STANDARD และ The Momentum ดังจะแสดงให้เห็นในช่วงถัดไป



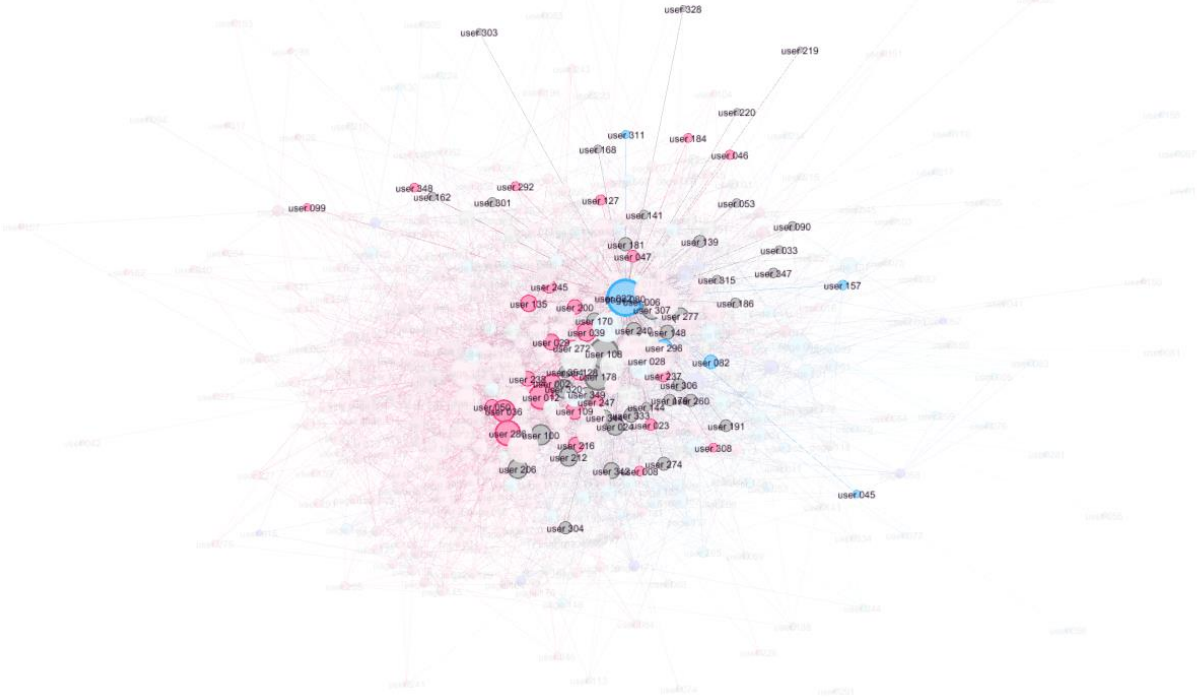
รูป 20 กราฟเครือข่ายสังคมของผู้ใช้และเพจที่กดไลก์

สำหรับเพจที่ติด 5 อันดับที่มีผู้ใช้กดไลก์มากที่สุดจากตารางที่ 18 คือ THE STANDARD (เลขโหนดที่ 080) The Reporters (เลขโหนดที่ 079) The MATTER (เลขโหนดที่ 074) The Momentum (เลขโหนดที่ 075) และ The101.world (เลขโหนดที่ 083)

โพสต์ที่ผู้ใช้กดไลก์มากเป็นอันดับแรกคือ THE STANDARD มีจำนวนไลก์ 75 ไลก์ โดยมีรายละเอียดความโน้มเอียงของขั้วความคิดทางการเมืองของผู้ใช้ดังนี้

ตาราง 28 จำนวนผู้กดไลก์เพจอันดับแรกแบ่งตามขั้วความคิดทางการเมือง

ขั้วความคิดทางการเมือง	จำนวนเพจ	ร้อยละ
ซ้ายจัด (1)	0	0.00%
เอียงซ้าย (2)	28	37.33%
กลาง (3)	42	56.00%
เอียงขวา (4)	5	6.67%
ขวาจัด (5)	0	0.00%
	75	**100.00%**

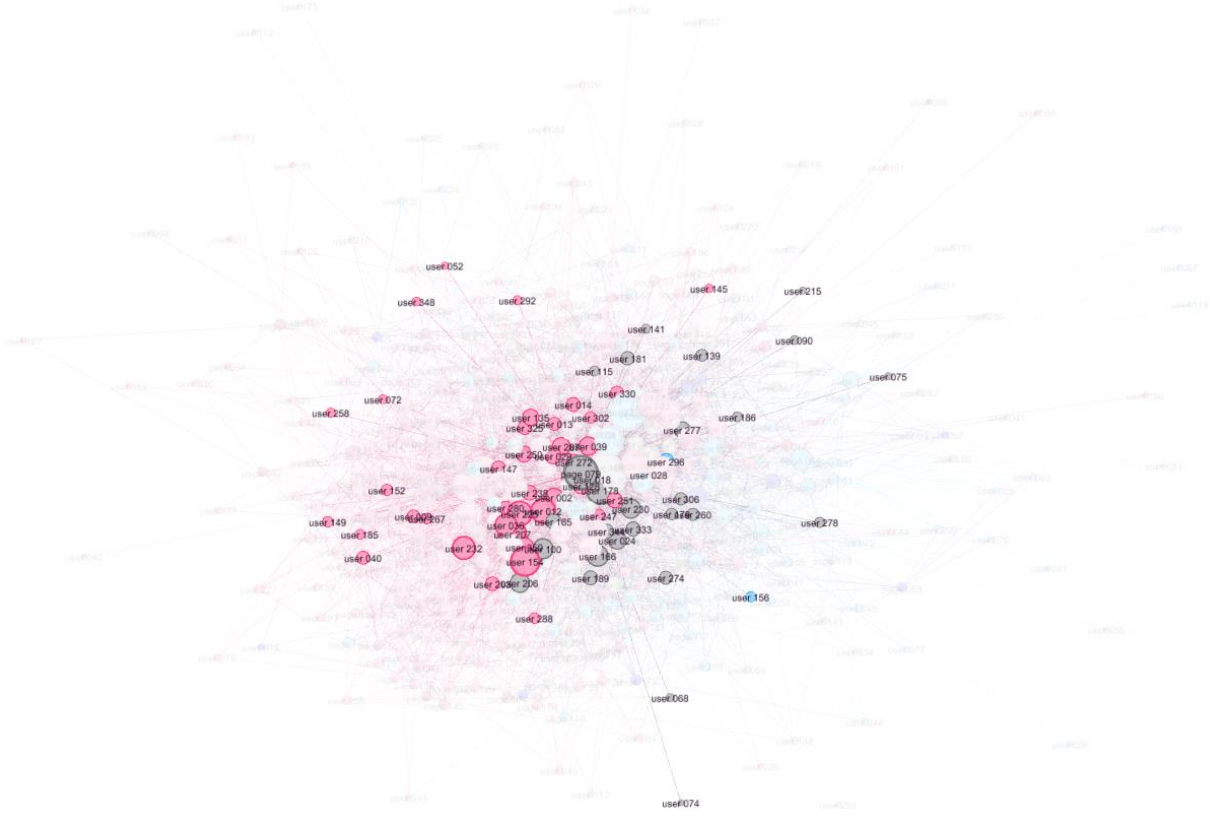


รูป 21 กราฟเครือข่ายสังคมของผู้ใช้และเพจที่กดไลก์เพจอันดับแรก

โพสต์ที่ผู้ใช้กดไลก์มากเป็นอันดับที่สอง คือ The Reporters มีจำนวนไลก์ 69 ไลก์ โดยมีรายละเอียดความโน้มเอียงของขั้วความคิดทางการเมืองของผู้ใช้ดังนี้

ตาราง 29 จำนวนผู้กดไลก์เพจอันดับที่สองแบ่งตามขั้วความคิดทางการเมือง

ขั้วความคิดทางการเมือง	จำนวนเพจ	ร้อยละ
ซ้ายจัด (1)	0	0.00%
เอียงซ้าย (2)	40	57.97%
กลาง (3)	27	39.13%
เอียงขวา (4)	2	2.90%
ขวาจัด (5)	0	0.00%
	69	**100.00%**

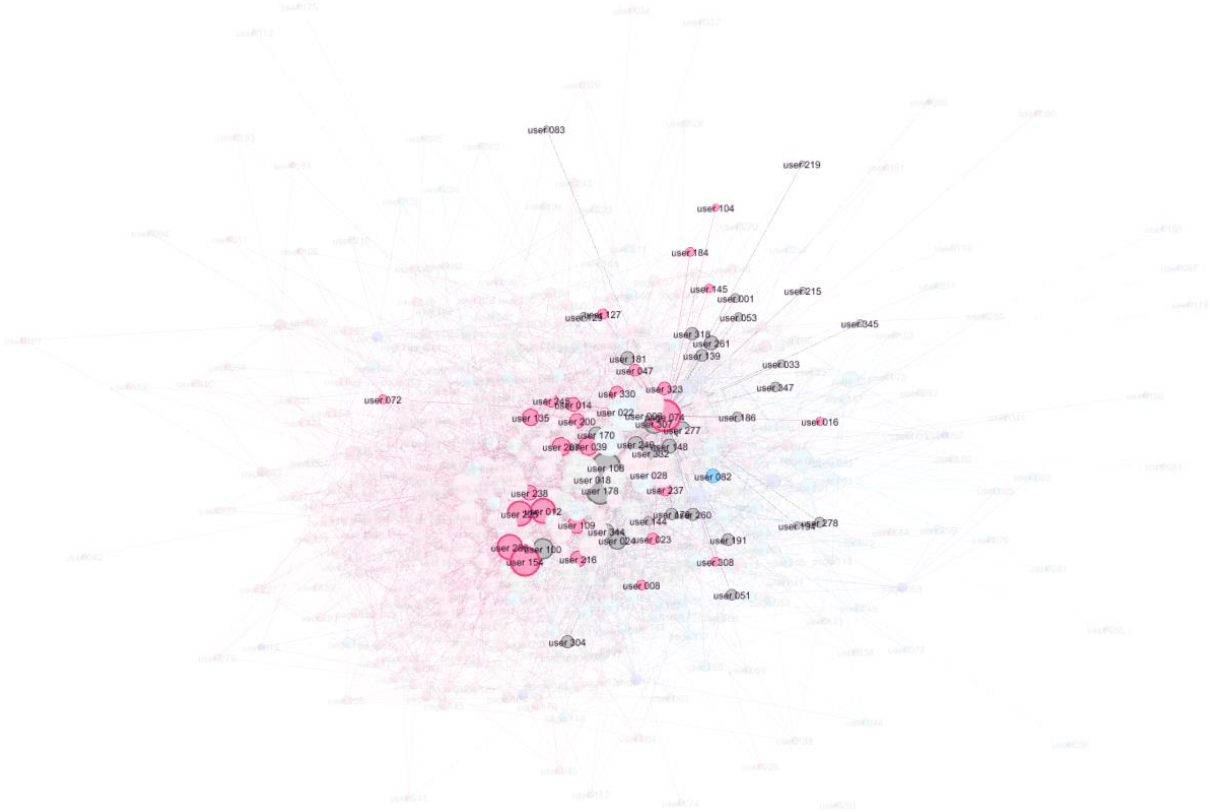


รูป 22 กราฟเครือข่ายสังคมของผู้ใช้และเพจที่กดไลก์เพจอันดับที่สอง

โพสต์ที่ผู้ใช้กดไลก์มากเป็นอันดับที่สาม คือ The MATTER มีจำนวนไลก์ 64 ไลก์ โดยมีรายละเอียดความโน้มเอียงของขั้วความคิดทางการเมืองของผู้ใช้ดังนี้

ตาราง 30 จำนวนผู้กดไลก์เพจอันดับที่สามแบ่งตามขั้วความคิดทางการเมือง

ขั้วความคิดทางการเมือง	จำนวนเพจ	ร้อยละ
ซ้ายจัด (1)	0	0.00%
เอียงซ้าย (2)	28	43.75%
กลาง (3)	35	54.69%
เอียงขวา (4)	1	1.56%
ขวาจัด (5)	0	0.00%
	64	**100.00%**

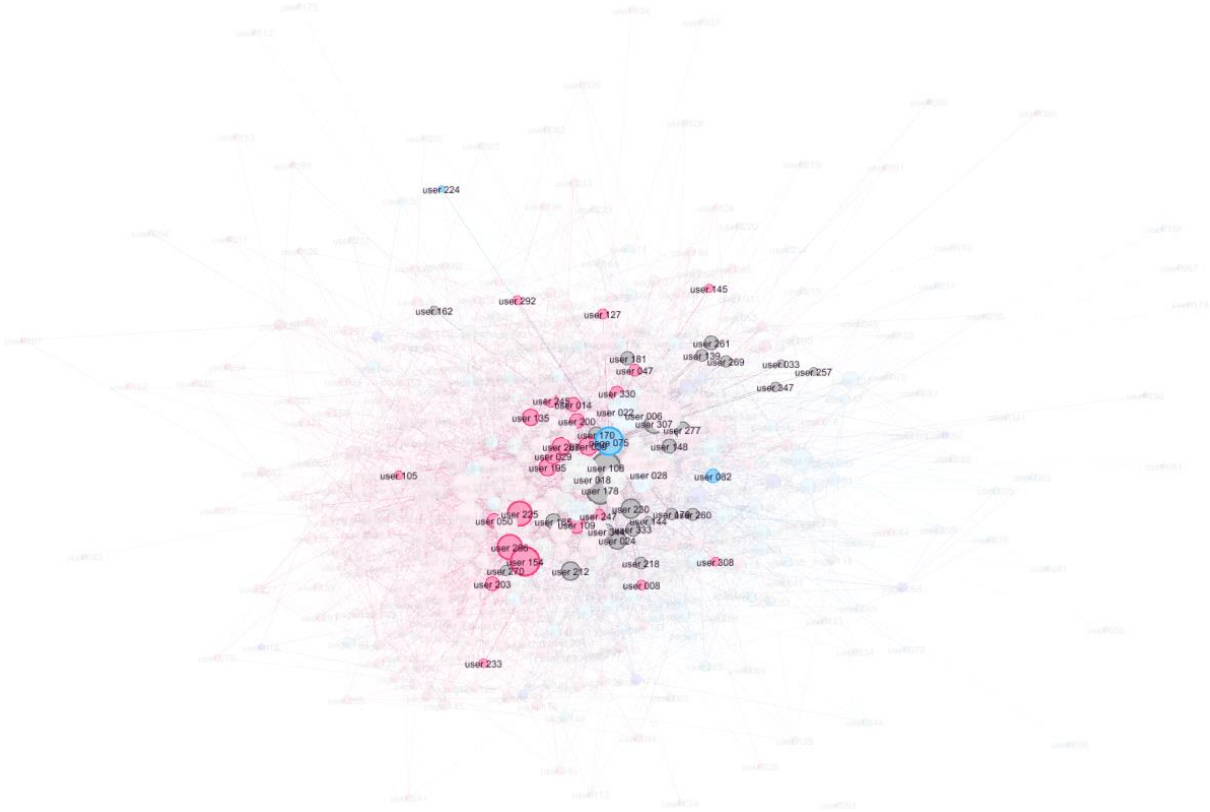


รูป 23 กราฟเครือข่ายสังคมของผู้ใช้และเพจที่กดไลก์เพจอันดับที่สาม

โพสต์ที่ผู้ใช้กดไลก์มากเป็นอันดับที่สี่ คือ The Momentum มีจำนวนไลก์ 55 ไลก์ โดยมีรายละเอียดความโน้มเอียงของขั้วความคิดทางการเมืองของผู้ใช้ดังนี้

ตาราง 31 จำนวนผู้กดไลก์เพจอันดับที่สี่แบ่งตามขั้วความคิดทางการเมือง

ขั้วความคิดทางการเมือง	จำนวนเพจ	ร้อยละ
ซ้ายจัด (1)	0	0.00%
เอียงซ้าย (2)	26	47.27%
กลาง (3)	27	49.09%
เอียงขวา (4)	2	3.64%
ขวาจัด (5)	0	0.00%
	55	**100.00%**

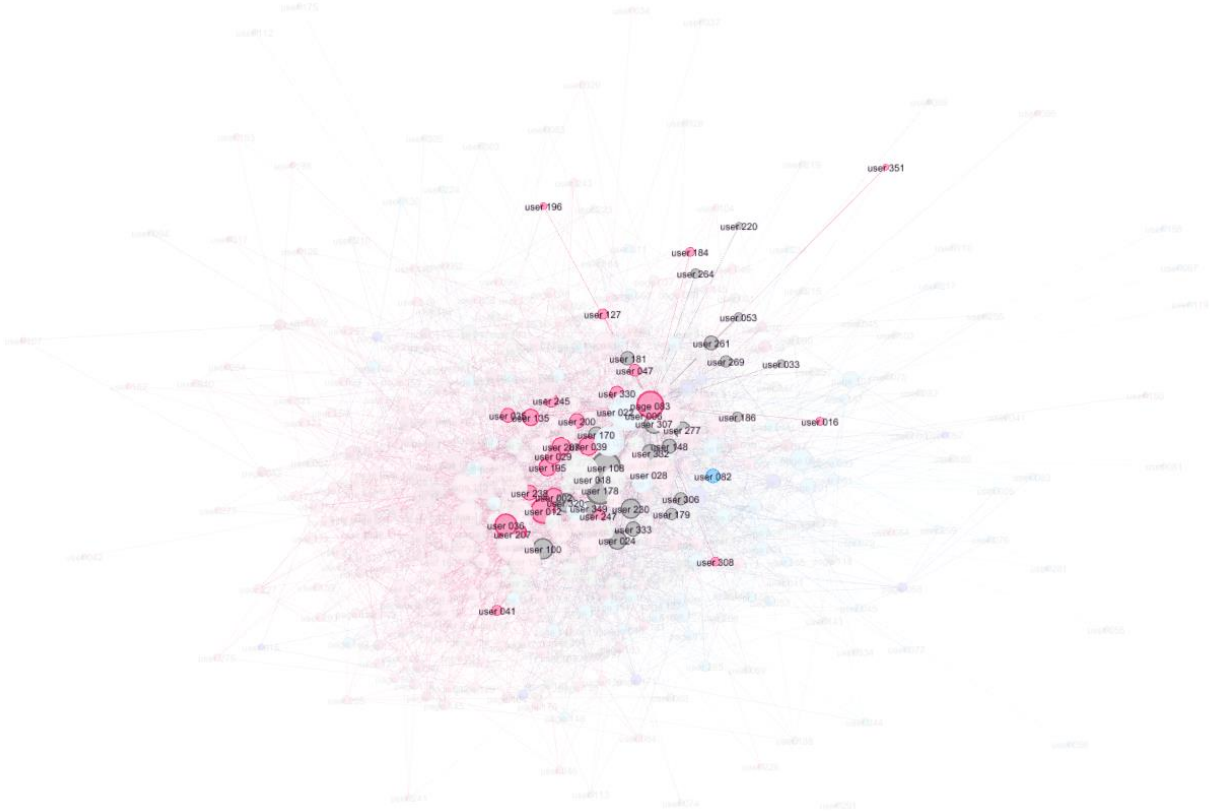


รูป 24 กราฟเครือข่ายสังคมของผู้ใช้และเพจที่กดไลก์เพจอันดับที่สี่

โดยโพสต์ที่ผู้ใช้กดไลก์มากเป็นอันดับที่ห้า คือ The101.world มีจำนวนไลก์ 51 ไลก์ โดยมีรายละเอียดความโน้มเอียงของขั้วความคิดทางการเมืองของผู้ใช้ดังนี้

ตาราง 32 จำนวนผู้กดไลก์เพจอันดับที่ห้าแบ่งตามขั้วความคิดทางการเมือง

ขั้วความคิดทางการเมือง	จำนวนเพจ	ร้อยละ
ซ้ายจัด (1)	0	0.00%
เอียงซ้าย (2)	25	49.02%
กลาง (3)	25	49.02%
เอียงขวา (4)	1	1.96%
ขวาจัด (5)	0	0.00%
	51	**100.00%**



รูป 25 กราฟเครือข่ายสังคมของผู้ใช้และเพจที่กดไลก์เพจอันดับที่ห้า

เมื่อพิจารณาถึงการเลือกรับสื่อหรือติดตามบุคคลสาธารณะแบบไหน ทางทีมงานได้แบ่งกลุ่มผู้ใช้ทั้ง 5 กลุ่ม และนำมาวิเคราะห์ร่วมกับเพจสื่อหรือบุคคลสาธารณะทั้ง 5 กลุ่ม โดยมีบัญชีผู้ใช้ทั้งสิ้น 250 คน และผู้ใช้เหล่านี้มีข้อมูลเพจที่กดไลก์ไว้ทั้งหมดที่เลือกมา 3,537 เพจ ทำมาแสดงในรูปแบบตารางการแจกแจงแบบไขว้ (CrossTabs หรือ PivotTable) ทั้งในรูปแบบจำนวนผู้ใช้และร้อยละได้ดังนี้

ตาราง 33 ตารางแจกแจงแบบไขว้ของจำนวนผู้ใช้และเพจแบ่งตามขั้วความคิดทางการเมือง

ขั้วความคิดทางการเมือง	ซ้ายจัด	เอียงซ้าย	กลาง	เอียงขวา	ขวาจัด	รวม
ซ้ายจัด (7 คน)	12	2				**14**
เอียงซ้าย (115 คน)	582	907	140	297	28	**1954**
กลาง (96 คน)	198	515	110	485	78	**1386**
เอียงขวา (29 คน)	4	20	14	111	26	**175**
ขวาจัด (3 คน)				3	5	**8**
รวม	**796**	**1444**	**264**	**896**	**137**	**3537**

ตาราง 34 ตารางแจกแจงแบบไขว้ของร้อยละของผู้ใช้และเพจแบ่งตามขั้วความคิดทางการเมือง

ขั้วความคิดทางการเมือง	ซ้ายจัด	เอียงซ้าย	กลาง	เอียงขวา	ขวาจัด
ซ้ายจัด (7 คน)	86%	14%	0%	0%	0%
เอียงซ้าย (115 คน)	30%	46%	7%	15%	1%
กลาง (96 คน)	14%	37%	8%	35%	6%
เอียงขวา (29 คน)	2%	11%	8%	63%	15%
ขวาจัด (3 คน)	0%	0%	0%	38%	63%

จะเห็นได้จากตาราง 34 ด้านบนว่ากลุ่มคนในแต่ละขั้ว รับข่าวสารจากขั้วตรงข้ามบ้าง ที่น่าสนใจคือกลุ่มเอียงซ้าย (115 คน) ที่รับข้อมูลจากฝั่งขวา (เอียงขวาและขวาจัด) เป็นจำนวน 16% และกลุ่มเอียงขวาที่รับข้อมูลจากฝั่งซ้าย (เอียงซ้ายและซ้ายจัด) เป็นจำนวน 13%

ทางทีมงานได้จัดทำตารางการแจกแจงแบบไขว้แบบมากกว่า 2 ตัวแปร คือในแกนนอนวางตัวแปรสองชั้น ได้แก่เพศและขั้วความคิดทางการเมืองของผู้ใช้ ส่วนในแกนตั้งใช้ตัวแปรเพจสื่อหรือบุคคลสาธารณะเพื่อแสดงแบบแผนที่อาจซ่อนอยู่ในข้อมูลโดยใช้ตัวแปรได้ ทั้งในรูปแบบจำนวนผู้ใช้และร้อยละได้ดังนี้

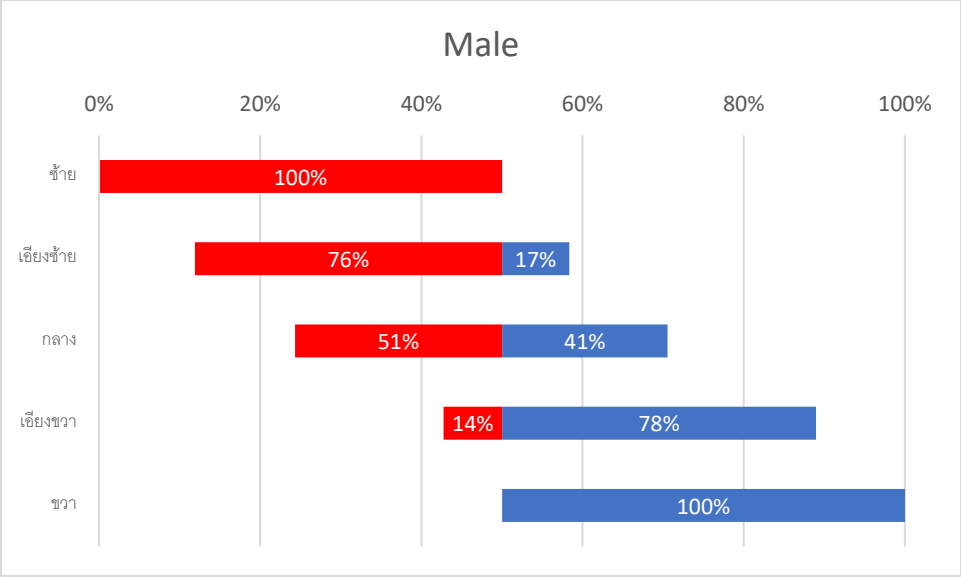
ตาราง 35 ตารางแจกแจงแบบไขว้ของจำนวนผู้ใช้และเพจแบ่งตามเพศและขั้วความคิดทางการเมือง

ขั้วความคิดทางการเมือง	**ซ้ายจัด**	**เอียงซ้าย**	**กลาง**	**เอียงขวา**	**ขวาจัด**	**รวม**
ชาย (190 คน)	**615**	**1132**	**201**	**720**	**119**	**2787**
ซ้ายจัด (7 คน)	12	2				**14**
เอียงซ้าย (82 คน)	436	697	105	226	22	**1486**
กลาง (75 คน)	163	414	84	390	72	**1123**
เอียงขวา (24 คน)	4	19	12	102	22	**159**
ขวาจัด (2 คน)				2	3	**5**
หญิง (60 คน)	**181**	**312**	**63**	**176**	**18**	**750**
ซ้ายจัด (0 คน)						
เอียงซ้าย (33 คน)	146	210	35	71	6	**468**
กลาง (21 คน)	35	101	26	95	6	**263**
เอียงขวา (5 คน)		1	2	9	4	**16**
ขวาจัด (1 คน)				1	2	**3**
รวม (250 คน)	**796**	**1444**	**264**	**896**	**137**	**3537**

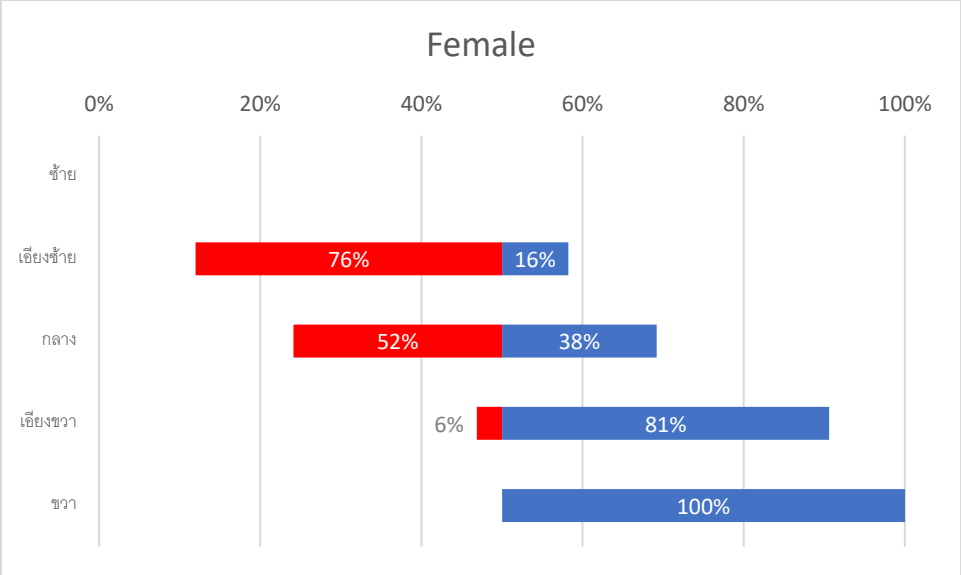
ตาราง 36 ตารางแจกแจงแบบไขว้ของร้อยละของผู้ใช้และเพจแบ่งตามเพศและขั้วความคิดทางการเมือง

ขั้วความคิดทางการเมือง	ซ้ายจัด	เอียงซ้าย	กลาง	เอียงขวา	ขวาจัด
ชาย	**22%**	**41%**	**7%**	**26%**	**4%**
ซ้ายจัด	86%	14%	0%	0%	**0%**
เอียงซ้าย	29%	47%	7%	15%	**1%**
กลาง	15%	37%	7%	35%	**6%**
เอียงขวา	3%	12%	8%	64%	**14%**
ขวาจัด	0%	0%	0%	40%	**60%**
หญิง	**24%**	**42%**	**8%**	**23%**	**2%**
ซ้ายจัด	0%	0%	0%	0%	**0%**
เอียงซ้าย	31%	45%	7%	15%	**1%**
กลาง	13%	38%	10%	36%	**2%**
เอียงขวา	0%	6%	13%	56%	**25%**
ขวาจัด	0%	0%	0%	33%	**67%**
Total	**23%**	**41%**	**7%**	**25%**	**4%**

จากตาราง x ซึ่งแบ่งข้อมูลเป็น 2 กลุ่มตามเพศชายและหญิงจะเห็นว่ามีค่าใกล้เคียงกันในทุกกลุ่มขั้วความคิดทางการเมือง เพียงแต่กลุ่มเพศหญิงไม่ปรากฏผู้ใช้ที่มีขั้วความคิดทางการเมืองแบบขวาจัด และเพื่อให้เห็นภาพชัดยิ่งขึ้น ทีมงานได้รวมข้อมูลของแต่ละฝั่ง (รวมซ้ายจัดกับเอียงซ้าย และรวมขวาจัดกับเอียงขวา) เพื่อทำกราฟแสดงสัดส่วนในรูปแบบปิรามิดประชากร (Population Pyramid) ดังนี้



รูป 26 ปิรามิดประชากรเพศชายแบ่งตามขั้วความคิดทางการเมืองซ้ายและขวา



รูป 27 ปิรามิดประชากรเพศหญิงแบ่งตามขั้วความคิดทางการเมืองซ้ายและขวา

ปิรามิดประชากรของผู้ใช้เพศชายและหญิงแสดงให้เห็นว่ากลุ่มผู้ใช้เพศหญิงที่มีขั้วความคิดทางการเมืองแบบ “เอียงขวา” เลือกกดไลก์เพจสื่อหรือบุคคลสาธารณะทางฝั่งซ้ายเพียง 6% ซึ่งน้อยกว่ากึ่งหนึ่งของกลุ่มผู้ใช้เพศชายที่มีขั้วความคิดทางการเมืองแบบ “เอียงขวา” ซึ่งเลือกกดไลก์เพจกลุ่มตรงข้ามถึง 14%

ในการตอบคำถามเกี่ยวกับขั้วความคิดทางการเมืองระหว่างรุ่น ทางทีมงานได้จัดทำตารางการแจกแจงแบบไขว้แบบมากกว่า 2 ตัวแปร คือในแกนนอนวางตัวแปรสองชั้นได้แก่เพศและช่วงอายุ ส่วนในแกนตั้งใช้ตัวแปรเพจสื่อหรือบุคคลสาธารณะเพื่อแสดงแบบแผนที่อาจซ่อนอยู่ในข้อมูลโดยใช้ตัวแปรได้ ทั้งในรูปแบบจำนวนผู้ใช้และร้อยละได้ดังนี้

ตาราง 37 ตารางแจกแจงแบบไขว้ของจำนวนผู้ใช้และเพจแบ่งตามเพศและช่วงอายุ

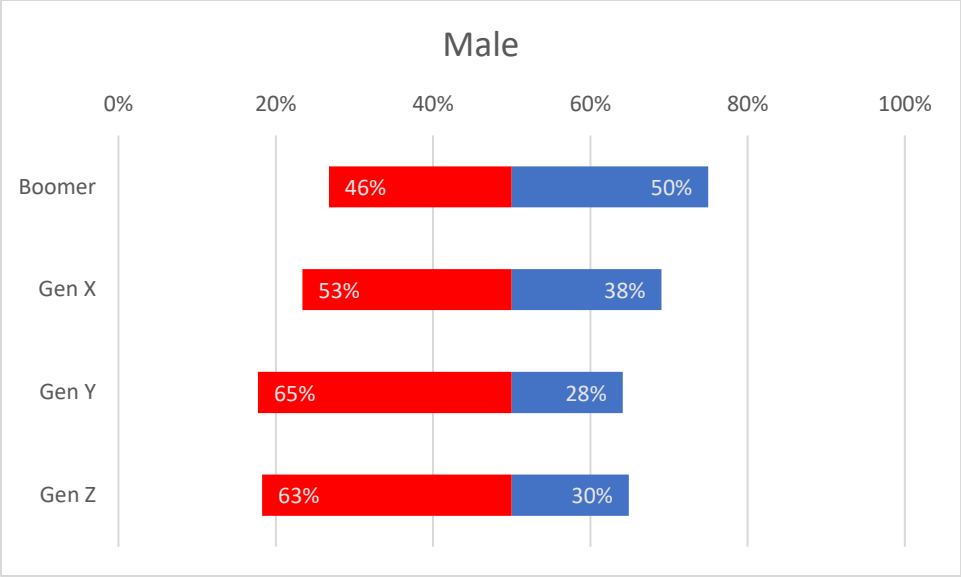
ขั้วความคิดทางการเมือง	ซ้ายจัด	เอียงซ้าย	กลาง	เอียงขวา	ขวาจัด	รวม
ชาย (190 คน)	**615**	**1132**	**201**	**720**	**119**	**2787**
Boomer (7 คน)	18	35	4	50	7	**114**
Gen X (29 คน)	52	107	26	102	12	**299**
Gen Y (129 คน)	453	819	143	472	85	**1972**
Gen Z (17 คน)	70	128	21	80	13	**312**
N/A (8 คน)	22	43	7	16	2	**90**
หญิง (60 คน)	**181**	**312**	**63**	**176**	**18**	**750**
Boomer (2 คน)	8	19	6	18	1	**52**
Gen X (14 คน)	16	39	7	41	5	**108**
Gen Y (30 คน)	98	167	36	66	5	**372**
Gen Z (11 คน)	52	65	8	41	6	**172**
N/A(3 คน)	7	22	6	10	1	**46**
รวม (250 คน)	**796**	**1444**	**264**	**896**	**137**	**3537**

ตาราง 38 ตารางแจกแจงแบบไขว้ของร้อยละของผู้ใช้และเพจแบ่งตามเพศและช่วงอายุ

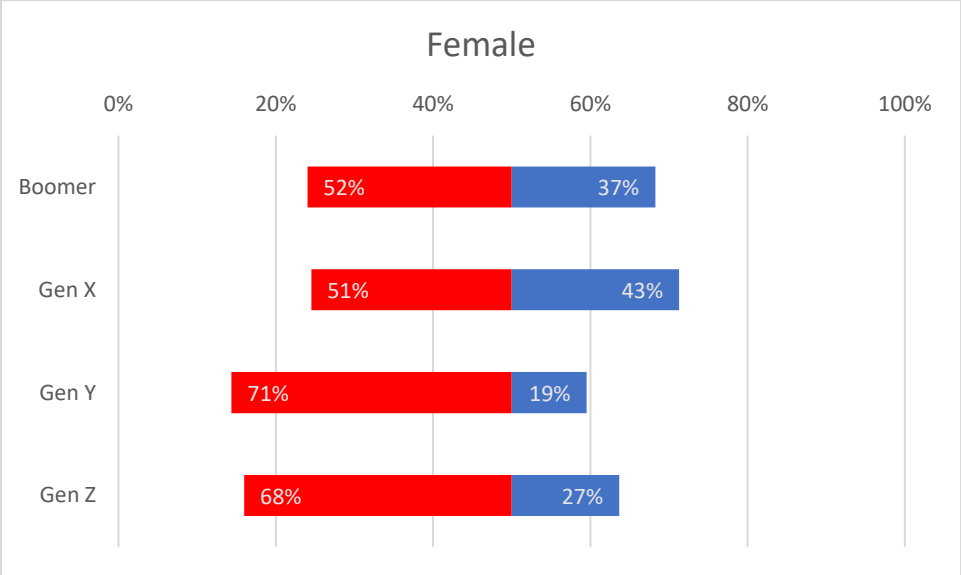
ขั้วความคิดทางการเมือง	ซ้ายจัด	เอียงซ้าย	กลาง	เอียงขวา	ขวาจัด
ชาย (190 คน)	**22%**	**41%**	**7%**	**26%**	**4%**
Boomer (7 คน)	16%	31%	4%	44%	6%
Gen X (29 คน)	17%	36%	9%	34%	4%
Gen Y (129 คน)	23%	42%	7%	24%	4%
Gen Z (17 คน)	22%	41%	7%	26%	4%
N/A (8 คน)	24%	48%	8%	18%	2%
หญิง (60 คน)	**24%**	**42%**	**8%**	**23%**	**2%**
Boomer (2 คน)	15%	37%	12%	35%	2%
Gen X (14 คน)	15%	36%	6%	38%	5%
Gen Y (30 คน)	26%	45%	10%	18%	1%
Gen Z (11 คน)	30%	38%	5%	24%	3%
N/A(3 คน)	15%	48%	13%	22%	2%
รวม (250 คน)	**23%**	**41%**	**7%**	**25%**	**4%**

จากตาราง x ซึ่งแบ่งข้อมูลเป็น 2 กลุ่มตามเพศชายและหญิง ในส่วนของเพศชายจะเห็นว่ากลุ่ม Boomer มีสัดส่วนเอียงทางขั้วฝั่งขวามากถึงร้อยละ 50 แต่กลุ่ม Gen X, Gen Y และ Gen Z มีสัดส่วนความสนใจเพจฝั่งขวาเพียงร้อยละ 39, 28 และ 30 ตามลำดับ ในส่วนของเพศหญิงจะเห็นว่ากลุ่ม Boomer กลุ่ม Gen X กลุ่ม Gen Y และกลุ่ม Gen Z มีสัดส่วนเอียงทางขั้วฝั่งขวาร้อยละ 37 ร้อยละ 43 ร้อยละ 19 และร้อยละ 27 ตามลำดับ

ในภาพรวมแสดงให้เห็นว่ายิ่งเป็นคนรุ่นใหม่ยิ่งมีแนวโน้มจะเอียงไปทางขั้วฝั่งขวาน้อยลงทั้งในกลุ่มเพศหญิงและเพศชาย ทั้งนี้เพื่อให้เห็นภาพชัดยิ่งขึ้น ทีมงานได้รวมข้อมูลของแต่ละฝั่ง (รวมซ้ายจัดกับเอียงซ้าย และรวมขวาจัดกับเอียงขวา) เพื่อทำกราฟแสดงสัดส่วนในรูปแบบปิรามิดประชากร (Population Pyramid) ดังนี้



รูป 28 ปิรามิดประชากรเพศชายแบ่งตามรุ่นและขั้วความคิดทางการเมืองซ้ายและขวา



รูป 29 ปิรามิดประชากรเพศหญิงแบ่งตามรุ่นและขั้วความคิดทางการเมืองซ้ายและขวา

ปิรามิดประชากรของเพศชายและหญิงในแต่ละช่วงอายุแสดงให้เห็นชัดเจนว่ากลุ่มคนรุ่นใหม่ ทั้ง Gen Y และ Gen Z เลือกกดไลก์เพจสื่อและบุคคลสาธารณะทางฝั่งซ้ายเป็นส่วนใหญ่ คือ 60% ขึ้นไป กลุ่มช่วงอายุที่มีความแตกต่างในการเลือกกดไลก์คือ Boomers ซึ่งในฝั่งเพศชายเลือกกดไลก์เพจสื่อและบุคคลสาธารณะทั้งฝั่งซ้าย

และขวาในจำนวนใกล้เคียงกัน คือเพจฝั่งซ้าย 46% และเพจฝั่งขวา 50% ในขณะที่กลุ่ม Boomers เพศหญิงเลือกกดไลก์เพจสื่อและบุคคลสาธารณะทั้งฝั่งซ้าย 52% แต่เพจฝั่งขวาเพียง 37%

บทที่ 4 สรุปผลการวิจัย

โครงการ “วิเคราะห์ความแตกต่างของความคิดและทัศนคติทางการเมืองระหว่างประชาชนต่างช่วงวัย” มีวัตถุประสงค์เพื่อวิเคราะห์ความแตกต่างของความคิดและทัศนคติทางการเมืองระหว่างประชาชนต่างช่วงวัยโดยกระบวนการทางวิทยาศาสตร์ข้อมูล และนำเสนอผลข้อมูลด้วยภาพข้อมูล (Data Visualization)

ในขั้นต้นได้มีการกำหนดกลุ่มประชาชนไว้ 3 กลุ่มหลัก คือกลุ่มผู้สูงวัย (60 ปีขึ้นไป) กลุ่มวัยกลางคน (30-59 ปี) กลุ่มเยาวชน (18-29 ปี) แต่ในภายหลังเมื่อได้ชี้แจงและปรึกษากับทางทีมงานของผู้ว่าจ้าง จึงเปลี่ยนเป็นการแบ่งตามรุ่น (generation) ได้แก่กลุ่ม Baby Boomer (เกิดปี พ.ศ.2489-2507) เจนเนอเรชัน X (เกิดปี พ.ศ. 2508-2522) เจนเนอเรชัน Y (เกิดปี พ.ศ.2523-2540) และเจนเนอเรชัน Z (เกิดหลังปี พ.ศ.2540) โดยได้ทำการเก็บข้อมูลแฟนเพจบริการเครือข่ายสังคม (social networking service) ที่นำเสนอข้อมูลความรู้เกี่ยวกับประชาธิปไตยของสถาบันพระปกเกล้าโดยกระบวนการทางวิทยาศาสตร์ข้อมูลด้วยการเก็บข้อมูลผ่านหน้าเว็บหรือ web scraping นำข้อมูลมาทำความสะอาด เตรียมข้อมูลเพื่อการวิเคราะห์โดยใช้ผู้ช่วยวิจัยประเมินข้อมูลที่เก็บมาได้บางส่วนเพื่อให้ข้อมูลมีความสมบูรณ์ที่สุด

การนำเสนอในบทนี้แบ่งออกเป็น 3 หัวข้อ ได้แก่การสรุปขั้นตอนการวิจัยและผลการวิเคราะห์ ข้อจำกัดในการเก็บข้อมูลและการนำเสนอผลการวิเคราะห์ และข้อเสนอแนะในอนาคต

4.1 สรุปขั้นตอนการวิจัยและผลการวิเคราะห์

หลังจากทำการวิเคราะห์ข้อมูลด้วยสถิติเชิงพรรณา และการใช้ภาพในรูปแบบต่างๆ เช่นทรีแมป (treemap) ตัวหนังสือภาพ (pictogram) และกราฟเครือข่ายสังคม (social network graph) ในภาพรวมนั้นพบว่าข้อมูลเกี่ยวกับการโพสต์ข้อความบนหน้าแฟนเพจของสถาบันฯ นั้น พบว่าไม่มีตัวแปรไหนที่มีส่วนสัมพันธ์กันอย่างมีนัยสำคัญดังแสดงไว้ในตารางที่ 10 ทั้งนี้ แสดงให้เห็นว่าปริมาณข้อความ (จำนวนตัวอักษร) ของโพสต์ไม่มีส่วนเกี่ยวข้องกับการมีส่วนร่วมแต่อย่างใด นอกจากนี้ประเภทของสื่อ (media) หรือประเภทของเนื้อหา (content category) ก็ไม่มีส่วนสัมพันธ์กับจำนวนการมีส่วนร่วมแต่อย่างใดเช่นกัน

ในประเด็นของกลุ่มประชากรที่มีความสนใจในเนื้อหาของสถาบันฯ ส่วนใหญ่นั้นอยู่ในกลุ่มคนรุ่นใหม่ (younger generation) หรือ Generation Y แบ่งเป็นเพศชายร้อยละ 49.86 และเพศหญิงร้อยละ 41.65 รวมทั้งหมดเป็นร้อยละ 64.51 ของข้อมูลทั้งหมด กลุ่มคนที่มีความสนใจรองลงมาคือกลุ่มวัยทำงานหรือ Generation X โดยแบ่งเป็นเพศชายร้อยละ 10.99 และเพศหญิงร้อยละ 6.48 รวมทั้งหมดเป็นร้อยละ 17.47 ของข้อมูลทั้งหมดดังแสดงไว้ในตารางที่ 13

ข้อมูลที่เป็นไปในทางเดียวกันนี้คือกลุ่มคนที่สนใจเนื้อหาของสถาบันฯ เท่าที่ระบุข้อมูลส่วนตัวได้ ส่วนหนึ่งมีภูมิหลังในสาขาวิชาที่เกี่ยวข้องเช่นด้านรัฐศาสตร์หรือนิติศาสตร์ รวมทั้งหมด 65 บัญชีผู้ใช้จาก 355 ราย คิดเป็นร้อยละ 18.3 หรือจากบัญชีผู้ใช้ทั้งสิ้น 238 รายที่แสดงข้อมูลการศึกษา คิดเป็นร้อยละ 27.3

นอกจากนี้กลุ่มคนที่สนใจเป็นส่วนใหญ่ประกอบอาชีพอยู่ในส่วนราชการ ซึ่งจากข้อมูลที่ระบุได้ทั้งหมด 147 บัญชีผู้ใช้นั้น มีจำนวนคนในกลุ่มนี้ถึง 65 ราย คิดเป็นร้อยละ 44.2

ในส่วนข้อมูลความโน้มเอียงของขั้วความคิดทางการเมืองพบว่าผู้ใช้ส่วนใหญ่อยู่ในกลุ่มเอียงซ้าย (46%) และกลุ่มที่เป็นกลาง (38.40%) โดยมีค่าเฉลี่ยของผู้ใช้งานทั้งหมดอยู่ที่ 2.62 มัธยฐานอยู่ที่ 3.00 ซึ่งตีความได้ว่าโดยเฉลี่ยแล้วผู้ใช้งานของเพจสถาบันฯ มีความเป็นกลางทางการเมือง

สำหรับเพจที่ติด 5 อันดับที่มีผู้ใช้กดไลก์มากที่สุด มีเพจที่เอียงซ้ายอยู่ 2 เพจคือ The MATTER และ The101.world เพจเอียงขวา 2 เพจคือ THE STANDARD และ The Momentum อีกเพจหนึ่งเป็นเพจที่เป็นกลางคือ The Reporters เมื่อวิเคราะว่ากลุ่มคนที่ถูกจัดอยู่แต่ละขั้วการเมือง เลือกกดไลก์เพจฝั่งตรงข้ามหรือไม่ พบว่ากลุ่มเอียงซ้าย (115 คน) ที่รับข้อมูลจากฝั่งขวา (เอียงขวาและขวาจัด) เป็นจำนวน 16% และกลุ่มเอียงขวาที่รับข้อมูลจากฝั่งซ้าย (เอียงซ้ายและซ้ายจัด) เป็นจำนวน 13% แสดงให้เห็นว่ากลุ่มคนจากขั้วการเมืองใด เลือกรับข้อมูลจากฝังตรงข้ามน้อยมาก

เมื่อวิเคราะห์ข้อมูลแบ่งตามเพศผู้ใช้งานพบว่าเพศชายและหญิงจะเห็นว่ามีค่าใกล้เคียงกันในทุกกลุ่มขั้วความคิดทางการเมืองเพียงแต่กลุ่มเพศหญิงไม่ปรากฏผู้ใช้ที่มีขั้วความคิดทางการเมืองแบบขวาจัด แต่เมื่อแสดงภาพในลักษณะปิรามิดประชากรกลับพบว่ากลุ่มผู้ใช้เพศหญิงที่มีขั้วความคิดทางการเมืองแบบ “เอียงขวา” เลือกกดไลก์เพจสื่อหรือบุคคลสาธารณะทางฝั่งซ้ายเพียง 6% ซึ่งน้อยกว่ากึ่งหนึ่งของกลุ่มผู้ใช้เพศชายที่มีขั้วความคิดทางการเมืองแบบ “เอียงขวา” ซึ่งเลือกกดไลก์เพจกลุ่มตรงข้ามถึง 14%

ในส่วนของการวิเคราะห์ช่วงอายุ (Generation) แบ่งตามเพศพบว่าเพศชายกลุ่ม Boomer มีสัดส่วนเอียงทางขั้วฝั่งขวามากถึงร้อยละ 50 แต่ในกลุ่ม Gen X ลงมาถึง Gen Z กลับมีสัดส่วนเอียงทางขั้วฝั่งซ้ายมากกว่าทั้งสิ้นเป็นจำนวนร้อยละ 53, 65 และ 63 ตามลำดับ ต่างกับกลุ่มเพศหญิงที่มีสัดส่วนเอียงทางขั้วฝั่งซ้ายในทุกช่วงอายุและมีมากที่สุดในกลุ่ม Gen Y คือร้อยละ 71 เมื่อแสดงข้อมูลในลักษณะปิรามิดประชากรจะเห็นได้ว่าแทบทุกกลุ่มช่วงอายุของเพศชายและเพศหญิงมีแนวโน้มติดตามข่าวสารในเพจสื่อและบุคคลสาธารณะในฝั่งซ้าย มีสัดส่วนมากน้อยต่างกันตามช่วงอายุ โดยอายุยิ่งน้อยก็ยิ่งมีสัดส่วนเอียงซ้ายมากขึ้น

4.2 ข้อจำกัดในการเก็บข้อมูลและการนำเสนอผลการวิเคราะห์

4.2.1 ความเปลี่ยนแปลงของซอร์สโค้ดหลังบ้านของเฟซบุ๊กในช่วงปลายเดือนกันยายน 2563

การใช้เทคนิค web scraping สามารถเข้าถึงข้อมูลรายชื่อผู้ใช้ (profiles) ที่เปิดเป็นสาธารณะได้เท่านั้น ทำให้ไม่สามารถเข้าถึงข้อมูลเชิงลึกได้ในหลายกรณี เช่นข้อมูลการศึกษาทั้งในอดีตและปัจจุบันสามารถระบุได้รวม 65 บัญชีผู้ใช้จากทั้งสิ้น 355 ราย คิดเป็นร้อยละ 18.3 หรือข้อมูลสถานที่ทำงานทั้งในอดีตและปัจจุบันสามารถระบุได้รวม 147 บัญชีผู้ใช้จากทั้งสิ้น 355 ราย คิดเป็นร้อยละ 41.4 ความเปลี่ยนแปลงของซอร์สโค้ดหลังบ้านของแพลตฟอร์มเฟซบุ๊กเองทำให้การเก็บข้อมูลจากหน้าเว็บของเฟซบุ๊กมีขั้นตอนที่ซับซ้อนและไม่สามารถรับประกันได้ว่าจะเก็บข้อมูลได้ครบถ้วน ซึ่งในการเก็บข้อมูลในโครงการวิจัยนี้สามารถเก็บข้อมูลหน้าเพจเฟซบุ๊กของสถาบันพระปกเกล้าได้ตั้งแต่เริ่มก่อตั้งเพจเมื่อเดือนกันยายน 2562 ถึงช่วงเดือนตุลาคม 2563

4.2.2 ข้อกำหนดของพระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562 (Personal Data Protection Act 2012 หรือ PDPA) และจริยธรรมการวิจัยในมนุษย์

เนื่องจากการวิจัยครั้งนี้มีการเก็บข้อมูลส่วนตัวของบัญชีผู้ใช้เป็นจำนวน 355 ราย ถึงแม้ข้อมูลเหล่านี้จะเป็นข้อมูลที่ผู้ใช้เปิดเป็นสาธารณะบนแพลตฟอร์มเฟซบุ๊ก แต่ยังเข้าข่ายการวิจัยในมนุษย์ ทางทีมงานวิจัยจึงทำการลบข้อมูลที่ระบุตัวตนผู้ใช้งานในการทำรายงานฉบับนี้ ข้อมูลรายละเอียดของการวิจัยทั้งหมดถูกเก็บอยู่ใน ที่ปลอดภัยและจะทำการแบ่งบันข้อมูลให้เฉพาะบัญชีอีเมลของเจ้าหน้าที่สถาบันฯ ที่มีส่วนเกี่ยวข้องเท่านั้น

4.3 ข้อเสนอแนะในอนาคต

4.3.1 การเลือกแพลตฟอร์มในการเก็บข้อมูลที่หลากหลาย

เนื่องจากเฟซบุ๊กถือเป็นแพลตฟอร์มที่มีความเป็นส่วนตัวสูงกว่าบริการเครือข่ายสังคมอื่น แม้จะเป็นแพลตฟอร์มที่มีการใช้อย่างแพร่หลาย การเก็บข้อมูลเท่าที่สามารถเก็บได้จากหน้าเพจ จึงควรวิเคราะควบคู่ไปกับการเก็บข้อมูลจากแหล่งอื่น (source triangulation) บริการเครือข่ายสังคมอื่นเช่นทวิตเตอร์ (Twitter) หรืออินสตาแกรม (Instagram) และใช้เทคนิคด้านการตลาดในการเชื่อมโยงข้อมูลดังรายละเอียดในข้อเสนอแนะข้อต่อไป

4.3.2 การใช้เทคนิคด้านการตลาดสนับสนุน

ปัญหาในการเชื่อมโยงข้อมูลส่วนตัวในหลายแพลตฟอร์มบริการเครือข่ายสังคมนั้น คือผู้วิจัยไม่สามารถระบุตัวตนของผู้ใช้งานได้ เช่นผู้ใช้งาน ก. อาจจะใช้ชื่อและนามสกุลจริงบนแพลตฟอร์มเฟซบุ๊ก แต่เลือกใช้ชื่อบัญชีเล่นหรือชื่อที่ตั้งขึ้นเองบนแพลตฟอร์มทวิตเตอร์ เพื่อเพิ่มโอกาสในการเชื่อมโยงข้อมูลหลายแพลตฟอร์ม ทาง

สถาบันฯ อาจใช้เทคนิคด้านการตลาดผ่านหน้าเพจโดยให้รางวัลกับผู้ใช้ที่ตอบแบบสอบถามความคิดเห็นต่อเนื้อหาที่ทางสถาบันเผยแพร่ และขอคำยินยอมเปิดเผยข้อมูลเพื่อทำการเชื่อมโยงไปยังแพลตฟอร์มอื่นด้วย

บรรณานุกรม

Randolph, J. J. (2005). *Free-marginal multirater kappa: An alternative to Fleiss´ fixed-marginal multirater kappa*. Paper presented at the Joensuu University Learning and Instruction Symposium 2005, Joensuu, Finland, October 14-15th, 2005.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications* (Vol. 8). Cambridge university press.