



มหาวิทยาลัยสุโขทัยธรรมมาธิราช
สาขาวิชานิเทศศาสตร์

เอกสารการสอน ฉบับเพิ่มเติม

ชุดวิชาการวิจัยการสื่อสารและ
การจัดการข้อมูลในยุคดิจิทัล

Communication Research and Data Management

15335

หน่วยที่ 1-4

คณะทำงานจัดทำเอกสารการสอนฉบับเพิ่มเติมชุดวิชาการวิจัยการสื่อสารและการจัดการข้อมูลในยุคดิจิทัล

อาจารย์ ดร.ชนาภา หนูนา

ประธานคณะทำงาน

รองศาสตราจารย์ไพบูรณ์ คะเชนทร์พรรค

บรรณาธิการ

รองศาสตราจารย์ ดร.สันหัต ทองรินทร์

คณะทำงาน

ผู้ช่วยศาสตราจารย์ ดร.ณัฐสุพงศ์ สุขโสภา

คณะทำงาน

พิมพ์ครั้งที่ 2 ภาคการศึกษาที่ 2/2566 จำนวนพิมพ์ 100 เล่ม

จัดพิมพ์โดย สำนักพิมพ์ มหาวิทยาลัยสุโขทัยธรรมาธิราช

สงวนลิขสิทธิ์

ลิขสิทธิ์ของมหาวิทยาลัยสุโขทัยธรรมาธิราช

คำนำ

เอกสารการสอนชุดวิชา ถือเป็นสื่อหลักในระบบการศึกษาทางไกลของมหาวิทยาลัยสุโขทัยธรรมมา-
ชิราช ซึ่งมีกระบวนการผลิตในรูปคณะกรรมการกลุ่มผลิตชุดวิชา ประกอบด้วย ผู้ทรงคุณวุฒิในสายวิชาการ
และวิชาชีพนั้นๆ ด้วยระบบและกระบวนการผลิตเอกสารการสอนที่มีมาตรฐานทำให้เอกสารการสอนชุดวิชา
เป็นสิ่งพิมพ์ที่มีคุณค่าทางวิชาการไม่เฉพาะต่อนักศึกษาของมหาวิทยาลัยสุโขทัยธรรมมาชิราชเท่านั้น
แต่รวมไปถึงนิสิต นักศึกษาของสถาบันการศึกษาอื่น นักวิชาการ ผู้ปฏิบัติงานในวิชาชีพ และผู้สนใจทั่วไป
ซึ่งจะได้ใช้ประโยชน์ในการศึกษาค้นคว้าความรู้เพื่อพัฒนาตนเองในด้านต่างๆ

อย่างไรก็ตาม ในยุคปัจจุบันที่เทคโนโลยีมีความก้าวหน้าไปอย่างรวดเร็ว ส่งผลให้กระบวนการ
สื่อสาร และองค์ความรู้ด้านนิเทศศาสตร์เปลี่ยนแปลงไปอย่างรวดเร็วเช่นกัน แม้ว่าสาขาวิชานิเทศศาสตร์
มหาวิทยาลัยสุโขทัยธรรมมาชิราช จะตระหนักถึงปัญหาและผลกระทบอันทำให้เกิดความล้าสมัยของเนื้อหาใน
บางเรื่อง โดยมีนโยบายการปรับปรุงเอกสารการสอนชุดวิชาอย่างต่อเนื่อง แต่ด้วยกระบวนการดำเนินงาน
อาจมีขั้นตอนทำให้ไม่สามารถดำเนินการผลิตเอกสารการสอนฉบับปรับปรุงได้อย่างรวดเร็ว ดังนั้น เพื่อให้
เอกสารการสอนชุดวิชามีความทันสมัย มีข้อมูลที่ถูกต้องและสอดคล้องกับการเปลี่ยนแปลงในบริบทปัจจุบัน
เพื่อเป็นประโยชน์ต่อนักศึกษามากที่สุด สาขาวิชานิเทศศาสตร์ จึงมีนโยบายให้มีการจัดทำเอกสารการสอน
ชุดวิชาฉบับเพิ่มเติม (Supplement) เป็นมาตรการเร่งด่วน

เอกสารการสอนฉบับเพิ่มเติมชุดวิชา 15335 การวิจัยการสื่อสารและการจัดการข้อมูลใน
ยุคดิจิทัลฉบับนี้ (ปรับมาจากชุดวิชาการวิจัยทางนิเทศศาสตร์) จึงมีการปรับเนื้อหาให้ทันสมัยมากขึ้น ด้วย
คณะทำงานเล็งเห็นว่าทางในปัจจุบัน ความก้าวหน้าของเทคโนโลยีไม่เพียงแต่ช่วยให้การทำงานด้านวิจัย
การสื่อสารมีความสะดวก รวดเร็ว และมีประสิทธิภาพมากขึ้นเท่านั้น หากแต่ยังทำให้ปริมาณข้อมูลที่
เกิดขึ้นมีมหาศาล โดยเฉพาะข้อมูลในระบบอินเทอร์เน็ตและการใช้สื่อสังคมออนไลน์ ซึ่งมีข้อมูลรูปภาพ
วิดีโอ เสียง หรือข้อความจากผู้ใช้ ระบบการจัดการข้อมูล และบิกดาต้า (Big data) จึงมีประโยชน์
อย่างมากในด้านต่างๆ เช่น ต่อการทำวิจัยและที่จะช่วยให้เกิดการบริหารจัดการข้อมูลให้ได้ประโยชน์สูงสุด
อย่างมีประสิทธิภาพ

ดังนั้น ในชุดวิชา 15335 การวิจัยการสื่อสารและการจัดการข้อมูลในยุคดิจิทัล จึงได้เพิ่มเนื้อหาที่
สำคัญเกี่ยวกับการจัดการข้อมูล และบิกดาต้า ทั้งสิ้น 4 ตอน เพื่อให้นักศึกษาได้เข้าใจและเรียนรู้ภาพรวม
ของบิกดาต้าว่าคืออะไร มีความสำคัญอย่างไร และสามารถนำไปประยุกต์ใช้ในงานด้านใดได้บ้าง นอกจากนี้
เนื้อหาทั้งหมดจะมีการนำไปใช้ในการออกข้อสอบเพื่อวัดผลทางความรู้ของนักศึกษาอีกด้วย

คณะทำงานหวังเป็นอย่างยิ่งว่า การอ่านเอกสารการสอนฉบับเพิ่มเติมนี้ จะทำให้นักศึกษามี
ความรู้ ความเข้าใจ และทักษะเกี่ยวกับกระบวนการวิจัย และหลักสำคัญของบิกดาต้ามากขึ้น ซึ่งสามารถ
นำไปประยุกต์ใช้ในวิชาชีพของนักศึกษาต่อไป

ประธานคณะทำงานจัดทำเอกสารการสอนฉบับเพิ่มเติมชุดวิชา
การวิจัยการสื่อสารและการจัดการข้อมูลในยุคดิจิทัล

สารบัญ

	หน้า
คำนำ.....	3
ตอนที่ 1 การวิจัยการสื่อสารและการจัดการข้อมูลในยุคดิจิทัล.....	5
เรื่องที่ 1.1 ความหมายและลักษณะสำคัญของบิกดาต้า.....	7
เรื่องที่ 1.2 ความสำคัญของบิกดาต้า.....	10
บรรณานุกรม.....	17
ตอนที่ 2 กระบวนการจัดการข้อมูลขนาดใหญ่.....	18
เรื่องที่ 2.1 ขั้นตอนการจัดการข้อมูลขนาดใหญ่.....	21
เรื่องที่ 2.2 แหล่งข้อมูลในการจัดการข้อมูลขนาดใหญ่.....	25
บรรณานุกรม.....	31
ตอนที่ 3 การวิเคราะห์ข้อมูลขนาดใหญ่.....	32
เรื่องที่ 3.1 ความหมายและความสำคัญของการวิเคราะห์ข้อมูลขนาดใหญ่.....	36
เรื่องที่ 3.2 ประเภทของการวิเคราะห์ข้อมูลขนาดใหญ่หรือบิกดาต้า.....	39
เรื่องที่ 3.3 แนวทางและขั้นตอนการวิเคราะห์ข้อมูลขนาดใหญ่.....	42
เรื่องที่ 3.4 เทคโนโลยีและเครื่องมือในการวิเคราะห์ข้อมูลขนาดใหญ่.....	45
เรื่องที่ 3.5 การนำข้อมูลขนาดใหญ่ไปใช้ประโยชน์.....	51
บรรณานุกรม.....	54
ตอนที่ 4 การรับฟังทางสังคม.....	56
เรื่องที่ 4.1 แนวคิดพื้นฐานและประโยชน์ของการรับฟังเสียงทางสังคม.....	58
เรื่องที่ 4.2 หลักการรับฟังเสียงทางสังคม.....	61
เรื่องที่ 4.3 ขั้นตอนการรับฟังเสียงทางสังคม.....	64
เรื่องที่ 4.4 เครื่องมือที่ใช้ในการรับฟังเสียงทางสังคม.....	66
บรรณานุกรม.....	71

ตอนที่ 1

การวิจัยการสื่อสารและการจัดการข้อมูลในยุคดิจิทัล

อาจารย์ ดร.ชานาภา หนูนา



ชื่อ
ผู้

อาจารย์ ดร.ชานาภา หนูนา
ศศ.บ. (สื่อสารมวลชน) (เกียรตินิยม) มหาวิทยาลัยเชียงใหม่
นศ.ม. (นิเทศศาสตร์) จุฬาลงกรณ์มหาวิทยาลัย
Ph.D. (Journalism & Mass Communications),
University of Kansas, USA

ตำแหน่ง

อาจารย์ประจำสาขาวิชานิเทศศาสตร์
มหาวิทยาลัยสุโขทัยธรรมมาธิราช

หน่วยที่เขียน

ตอนที่ 1

ตอนที่ 1

การวิจัยการสื่อสารและการจัดการข้อมูลในยุคดิจิทัล

โปรดอ่านหัวเรื่อง แนวคิด และวัตถุประสงค์ของตอนที่ 1 แล้วจึงศึกษารายละเอียดต่อไป

หัวเรื่อง

เรื่องที่ 1.1 ความหมายและลักษณะสำคัญของบิกดาต้า

เรื่องที่ 1.2 ความสำคัญของบิกดาต้า

แนวคิด

1. บิกดาต้า คือ ชุดข้อมูลขนาดใหญ่จำนวนมากและข้อมูลมีความหลากหลาย มีการจัดเก็บอย่างมีเป้าหมาย เพื่อวัตถุประสงค์บางประการ มีการจัดเรียงอย่างเป็นระบบ ต้องมีการถอดรหัสวิเคราะห์ข้อมูล และรู้วัตถุประสงค์ชัดเจนในการนำไปใช้ประโยชน์ มีลักษณะสำคัญ ได้แก่ ขนาดข้อมูลที่มีปริมาณมาก ความหลากหลายข้อมูล เช่น ตัวอักษรพิมพ์ ภาพ เสียงบรรยากาศ เป็นต้น ความเร็วข้อมูล การประมวลผลต้องอาศัยความต่อเนื่องและความรวดเร็ว ความแม่นยำข้อมูล ที่ต้องการความถูกต้องและความแม่นยำ และคุณค่าของข้อมูลที่มีประโยชน์และความสอดคล้องกับวัตถุประสงค์ทางธุรกิจ
2. บิกดาต้ามีความสำคัญต่อหลายภาคส่วน ได้แก่ ภาครัฐ ในแง่การกำหนดยุทธศาสตร์การพัฒนาประเทศ การกำหนดนโยบายของหน่วยงาน การสื่อสารกับประชาชน ในด้านต่างๆ ทั้งในสถานการณ์ปกติและภาวะวิกฤต และการปรับปรุงการบริการและตอบสนองตรงตามความต้องการของประชาชน ภาคเอกชน เช่น การพัฒนาผลิตภัณฑ์และการตอบสนองต่อความต้องการของลูกค้า การวิเคราะห์ความเสี่ยงในงานธนาคาร การพัฒนาระบบขนส่งสินค้าและโลจิสติกส์ การพัฒนาระบบการบริการตอบคำถามและให้ข้อมูลแก่ลูกค้า และภาคสื่อสารมวลชน ในด้านการผลิตและรายงานข่าว และด้านการผลิตและปรับปรุงรายการบันเทิงเพื่อเผยแพร่ผ่านสื่อ

วัตถุประสงค์

เมื่อศึกษาตอนที่ 1 จบแล้ว นักศึกษาสามารถ

1. อธิบายความหมายและลักษณะสำคัญของบิกดาต้าได้
2. อธิบายความสำคัญของบิกดาต้าได้

เรื่องที่ 1.1

ความหมายและลักษณะสำคัญของบิกดาต้า

นักศึกษาเคยสังเกตเห็นว่า เมื่อนักศึกษาค้นหาข้อมูลสินค้าหรือบริการบน Google เช่น เครื่องซักผ้า จากนั้นนักศึกษาจะเห็นโฆษณาเครื่องซักผ้าบนเฟซบุ๊ก (Facebook) หรือแอปพลิเคชันการชื้อขายต่างๆ เสมือนว่ากำลังมีใครคนใดคนหนึ่งติดตามและรู้ใจของนักศึกษาว่าต้องการอะไรเสมอ เหตุการณ์ดังกล่าวเกี่ยวข้องกับการรวบรวมข้อมูลของนักศึกษารวมไปถึงผู้ใช้อินเทอร์เน็ตอื่นๆ ทั่วโลก จนกลายเป็นปริมาณข้อมูลมหาศาลที่เรียกกันว่า บิกดาต้า (Big Data) (Big Data คืออะไร สำคัญอย่างไร, 2563)

บิกดาต้าเป็นแนวคิดที่มีความสำคัญอย่างมากในการทำงานวิจัยและการใช้ประโยชน์ต่อองค์กรทั้งรัฐและเอกชนในปัจจุบัน ถึงแม้ว่าแนวคิดเรื่องบิกดาต้าจะดูเป็นเรื่องใหม่สำหรับวงการวิจัย ธุรกิจและอุตสาหกรรม แต่แท้จริงแล้ว บิกดาต้าได้เริ่มต้นมาตั้งแต่คริสต์ทศวรรษ 1960 (ตั้งแต่ปี ค.ศ. 1960 ถึง 1969) และคริสต์ทศวรรษ 1970 (ปี ค.ศ. 1970 ถึง ค.ศ. 1979) ซึ่งมีการจัดตั้งศูนย์ข้อมูลแห่งแรก และพัฒนาระบบข้อมูลที่จัดเก็บรวบรวมข้อมูลขึ้น จนกระทั่งในปี 2005 หลังจากที่มีสื่อสังคมออนไลน์เกิดขึ้นและได้รับความนิยมอย่างล้นหลาม เช่น Facebook และ YouTube คนในแวดวงที่เกี่ยวข้อง จึงเริ่มตระหนักถึงปริมาณข้อมูลจำนวนมากที่ถูกสร้างขึ้นจากผู้ใช้ นอกจากนี้ในช่วงเวลาเดียวกันมีการพัฒนากรอบแนวคิดที่เรียกว่า Hadoop ซึ่งถูกออกแบบมาเพื่อรองรับการทำงานบนระบบคอมพิวเตอร์และการประมวลผลข้อมูล การพัฒนาดังกล่าวจึงถือว่าเป็นการเติบโตและการพัฒนาระบบการจัดการเก็บและประมวลผลข้อมูลมาจนถึงปัจจุบัน (What is Big Data?, n.d.)

ความหมายของบิกดาต้า

ความหมายของบิกดาต้า (Big Data) หรือข้อมูลขนาดใหญ่ถูกอธิบายไว้มากมายซึ่งใกล้เคียงกัน เช่น บิกดาต้า คือ ปริมาณของข้อมูลอันมหาศาล ซึ่งส่งผลต่อการบันทึกหรือจัดเก็บข้อมูลอย่างไม่เคยปรากฏมาก่อน โดยบิกดาต้ามีลักษณะสำคัญคือ ขนาดข้อมูลใหญ่ (volume) มีความหลากหลาย (variety) และการประมวลผลข้อมูลสูง (high-speed processing) ซึ่งเรียกรวมว่า 3Vs (Diebold, 2003; Laney, 2001).

นอกจากนี้ บิกดาต่ายังหมายถึงข้อมูลที่มีปริมาณจำนวนมาก มีความหลากหลาย และใช้เทคโนโลยีระดับสูงในการประมวลผลข้อมูล เพื่อช่วยในการวิเคราะห์ข้อมูลเชิงลึกและตัดสินใจในการดำเนินงานต่างๆ (Big Data, n.d.)

นอกจากนี้ บิกดาต้า ยังหมายถึง ชุดข้อมูลขนาดใหญ่จำนวนมากและข้อมูลมีความหลากหลาย มีการจัดเก็บ อย่างมีเป้าหมาย เพื่อวัตถุประสงค์บางประการ มีการจัดเรียงอย่างเป็นระบบ ต้องมีการถอดรหัสวิเคราะห์ข้อมูล และรู้วัตถุประสงค์ชัดเจนในการนำไปใช้ประโยชน์ (อังชิตา ลิมบีปัทมปาณี และคณะ, 2563)

ลักษณะสำคัญของบิกดาต้า

โดยทั่วไปแล้วบิกดาต้าต่างกับข้อมูลธรรมดา คือ ต้องมีระบบ (data systematic) มีการจัดการ (data management) การจัดเก็บ (data storage) และการนำไปใช้ให้เกิดประโยชน์ (data goal & usage) แต่ในส่วนของคุณลักษณะที่สำคัญของบิกดาต้า มีนักวิชาการและนักวิชาชีพกล่าวถึงคุณลักษณะของบิกดาต้าไว้หลากหลาย ในที่นี้ จะกำหนดลักษณะสำคัญของบิกดาต้าไว้ 5 ประการ หรือเรียกว่า 5Vs ดังนี้ (อัสมา กุลวานิชไชยพันธ์, 2561; Roy, 2020)

1. ขนาดของข้อมูล (volume) ในปัจจุบัน ข้อมูลที่ปรากฏตามอินเทอร์เน็ตหรือสื่อสังคมออนไลน์ มีปริมาณมาก ซึ่งทำให้ยากต่อการจัดเก็บ เพราะมีโปรแกรมทางคอมพิวเตอร์จำนวนมาก โดยเฉพาะสื่อสังคมออนไลน์ซึ่งมีหลากหลายและมีผู้ใช้งานจำนวนมากทั่วโลก เช่น Facebook มีการเก็บข้อมูลที่เกิดจากระบบและผู้ใช้งาน จำนวนกว่า 4 เพตะไบต์ (PB) ต่อวัน โดยมีการจัดเก็บ Hive : เป็นเครื่องมือใช้เตรียมข้อมูลที่เป็นลักษณะคลังข้อมูล (data warehouse) ซึ่งบรรจุข้อมูลประมาณ 300 เพตะไบต์ (Roy, 2020)

ข้อมูล เพตะไบต์เปรียบเทียบกับจำนวนภาพยนตร์ 250 เรื่อง หรือ ปริมาณน้ำ 1 ล้านล้านลิตร หรือจำนวนรูปถ่าย จำนวน 250,000 รูป ที่ถ่ายจากกล้องขนาด 12 ล้านพิกเซล (Petabytes-คืออะไร-และมีขนาดใหญ่แค่ไหน, ม.ป.ป.)

2. ความหลากหลายข้อมูล (variety) โดยบิกดาต้า ประกอบด้วยข้อมูลหลายรูปแบบ เช่น ตัวอักษรพิมพ์ ตัวอักษรเขียนด้วยลายมือ ภาพลายเส้น ภาพเอกซเรย์ ภาพถ่าย ภาพวาด ภาพเขียน ภาพกราฟิก ภาพเคลื่อนไหว คลิป วิดีโอ เสียงพูด เสียงดนตรี เสียงบรรยาย เสียงบรรยายภาค เสียงประกอบฉาก เสียงร้อง เสียงจากธรรมชาติ เสียงสร้างขึ้นจากเครื่องมือของมนุษย์ เป็นต้น รวมไปถึง ความหลากหลายของแหล่งข้อมูลที่มาจากทั้งในและนอกองค์กรอีกด้วย (สรุปรัด! “5V” องค์ประกอบและคุณลักษณะที่สำคัญของ Big Data, 2565)

3. ความเร็วข้อมูล (velocity) เนื่องจากบิกดาต้ามีปริมาณข้อมูลจำนวนมาก และหลากหลายรูปแบบ ดังนั้น การประมวลผลข้อมูล ต้องรวดเร็วพอที่จะทราบผลทันที ซึ่งต้องอาศัยความต่อเนื่องของกระบวนการประมวลผลและความรวดเร็ว ยกตัวอย่างเช่น ในการตรวจสอบนักเรียนจำนวนมาก เครื่องอ่านข้อมูลในคอมพิวเตอร์จะรันกระดาษคำตอบที่ละแผ่น ซึ่งอาจจะมีจำนวนเพียงหลักหมื่น และประมวลผลรวดเร็วพร้อมกันจึงจะเห็นโครงสร้างหรือภาพรวมทั้งหมด แต่บิกดาต้าจะต้องรันข้อมูลนับล้านข้อมูล ซึ่งเทคโนโลยีดั้งเดิมไม่สามารถทำได้ รวมทั้งหากมีข้อมูลบางจุดผิดพลาด ต้องจัดการแก้ไขอย่างรวดเร็ว (บุญลาภ ภูสุวรรณ, กมล ขาววิทย์ และอังชิตา ลิ้มปี๋ทมปาณี, 2563)

4. ความแม่นยำข้อมูล (veracity) เนื่องจากบิกดาต้ามีข้อมูลจำนวนมากและมีรูปแบบของข้อมูลหลากหลาย จึงต้องการความถูกต้องและความแม่นยำของข้อมูลที่รวบรวม ดังนั้น จึงต้องตรวจสอบ แก้ไข จัดรูปแบบ และคัดกรองข้อมูลให้อยู่ในสภาพที่พร้อมใช้งานที่สุด หรือที่เรียกว่า Data Cleaning (Data Cleansing คืออะไร สำคัญอย่างไรกับการจัดการข้อมูล, ม.ป.ป.) ผลออกมาจึงจะถูกต้องแม่นยำ ซึ่งเป็นขั้นตอนยากลำบากและใช้เวลานานมากที่สุด ยิ่งข้อมูลมีขนาดใหญ่มากเท่าไรยิ่งต้องใช้เวลานาน (อังชิตา ลิ้มปี๋ทมปาณี, และคณะ, 2563) นอกจากนี้ ข้อมูลจำนวนมากที่ขาดความจริงหรือครบถ้วนอาจส่งผลเสียได้ เช่น ในด้านการแพทย์ หากข้อมูลเกี่ยวกับยาที่ผู้ป่วยกำลังรับประทานไม่ครบถ้วน ชีวิตของผู้ป่วยอาจตกอยู่

ในอันตรายเป็นได้ เพราะฉะนั้น Big Data จึงควรมีคุณลักษณะของความถูกต้องอยู่ด้วย (สรุปชัด! “5V” องค์ประกอบและคุณลักษณะที่สำคัญของ Big Data, 2565)

5. คุณค่าของข้อมูล (value) ข้อมูลที่ถูกจัดเก็บไว้จำนวนมาก แต่ไม่ได้นำมาใช้ประโยชน์ก็จะไม่เกิดคุณค่า ดังนั้น ข้อมูลที่มีทั้งหมด แม้แต่ข้อมูลขยะ (junk data) ข้อมูลสมบูรณ์ (corrupted data) หรือข้อมูลที่จัดเก็บในรูปแบบที่หลากหลาย ก็สามารถทำให้เกิดคุณค่าได้ เมื่อเกิดกระบวนการต่างๆ เช่น การทำความสะอาดข้อมูล การเปลี่ยนแปลงรูปแบบข้อมูล หรือกระบวนการประมวลผล เพื่อนำไปใช้งานให้เกิดประโยชน์สอดคล้องกับวัตถุประสงค์ทางธุรกิจ นอกจากนี้ การเกิดคุณค่าของบิกดาต้าจำเป็นต้องมีการวิเคราะห์เชิงลึก และนำมาประมวลผล เพื่อนำไปใช้ประโยชน์และสร้างมูลค่ามหาศาลให้กับองค์กร (Big Data คืออะไร ลักษณะข้อมูลเป็นอย่างไร, ม.ป.ป.; Big data and analytics: definitions, value, trends and applications, n.d.; SME Thailand, 2565)

กิจกรรม 1.1

ลักษณะสำคัญของบิกดาต้ามีอะไรบ้าง

แนวตอบกิจกรรม 1.1

ลักษณะสำคัญของบิกดาต้า ได้แก่

1. ขนาดข้อมูล (volume) ที่มีปริมาณมาก
 2. ความหลากหลายข้อมูล (variety) เช่น ตัวอักษรพิมพ์ ภาพ เสียงบรรยายกาศ เป็นต้น
 3. ความเร็วข้อมูล (velocity) การประมวลผลต้องอาศัยความต่อเนื่องและความรวดเร็ว
 4. ความแม่นยำข้อมูล (veracity) ที่ต้องการความถูกต้องและความแม่นยำ
 5. คุณค่าของข้อมูล (value) ที่มีประโยชน์และความสอดคล้องกับวัตถุประสงค์ทางธุรกิจ
-

เรื่องที่ 1.2

ความสำคัญของบิกดาต้า

บิกดาต้ามีความสำคัญต่อหลายภาคส่วน ได้แก่ ต่อภาครัฐ ต่อภาคเอกชน และต่อภาคสื่อสารมวลชน ดังรายละเอียดต่อไปนี้

1. ความสำคัญของบิกดาต้าต่อภาครัฐ บิกดาต้าเป็นเครื่องมือทำให้การทำงานของภาครัฐง่ายขึ้น เพราะหน่วยงานของรัฐมีปริมาณข้อมูลเกี่ยวกับประชาชนจำนวนมาก บิกดาต้าจึงช่วยให้รัฐมีข้อมูลเพื่อนำมาจัดการ วิเคราะห์ ตั้งแต่การวางแผนกำหนดทิศทางประเทศ และทิศทางการทำงาน ทำให้มองเห็นปัญหาของประชาชนได้ชัดเจน สามารถเห็นข้อมูลภาพรวมทั้งระดับกระทรวง ทบวง กรม จังหวัด อำเภอ ตำบล หมู่บ้าน (อังธิดา ลิมปิปัทมปาณี และคณะ, 2563) นอกจากนี้ บิกดาต่ายังช่วยในการพัฒนาประสิทธิภาพในการสื่อสารกับประชาชน พัฒนาการบริการภาครัฐให้ตรงต่อความต้องการของประชาชน และใช้งบประมาณน้อยลง อีกทั้งช่วยแก้ไขปัญหา ซึ่งเจงนโยบาย หรือดำเนินการเกี่ยวกับประชาชนได้ เพื่อให้คุณภาพชีวิตของประชาชนดีขึ้น โดยความสำคัญของบิกดาต้าต่อภาครัฐสามารถอธิบายในรายประเด็นดังต่อไปนี้

1.1 การกำหนดยุทธศาสตร์การพัฒนาประเทศ บิกดาต้า ถูกนำมาใช้เพื่อเป็นข้อมูลเชิงประจักษ์ประกอบการตัดสินใจในการวางแผนและกำหนดยุทธศาสตร์การพัฒนาประเทศในระดับต่างๆ เพื่อให้ข้อมูลสอดคล้องกับเป้าหมายการพัฒนาภายใต้ยุทธศาสตร์ชาติและเชื่อมโยงแผนต่างๆ เช่น แผนแม่บทภายใต้ยุทธศาสตร์ชาติ แผนการปฏิรูปประเทศ แผนพัฒนาเศรษฐกิจและสังคมแห่งชาติ แผนความมั่นคงแห่งชาติ และแผนปฏิบัติการด้านต่างๆ ของส่วนของราชการ เป็นต้น

ยกตัวอย่างเช่น กระทรวงศึกษาธิการ สามารถใช้บิกดาต้าพิจารณาว่านักเรียนนักศึกษาจบในสาขาใดบ้างในแต่ละปี นักศึกษามีทักษะความรู้ความสามารถและความเชี่ยวชาญด้านใดบ้างที่ตอบโจทย์ตลาดแรงงานและการพัฒนาประเทศหรือไม่ ประเทศและตลาดแรงงานยังขาดแคลนบัณฑิตในสาขาวิชาใดบ้าง มีจำนวนเท่าไร ดังนั้น ข้อมูลของบิกดาต้าที่ถูกจัดเก็บอย่างเป็นระบบและประมวลผลแล้ว สามารถช่วยให้กระทรวงศึกษาธิการวางแผนการทำงานในอนาคต ออกแบบกลไกการบริหารและการจัดการศึกษาอย่างเป็นระบบมากขึ้น

อีกกรณีหนึ่งของกระทรวงการท่องเที่ยวและกีฬาหน้าที่ ได้พัฒนาระบบบิกดาต้า เพื่อการบริการด้านการท่องเที่ยวที่เกิดจากการมีส่วนร่วมของภาคส่วนต่างๆ ในการปรับปรุงข้อมูลดิจิทัล เช่น ข้อมูลทั่วไปของแหล่งท่องเที่ยว ผู้ประกอบการ การบอกเล่าหรือแชร์ประสบการณ์การท่องเที่ยวโดยนักท่องเที่ยว การบันทึกพฤติกรรมกรรมการท่องเที่ยวของนักท่องเที่ยว หรือกิจกรรมท่องเที่ยวโดยเจ้าของแหล่งท่องเที่ยว เจ้าของที่พัก เจ้าของร้านอาหาร หรือเจ้าของสินค้าในชุมชน รวมทั้งผู้ให้บริการด้านสุขภาพ เป็นต้น โดยเมื่อข้อมูลเหล่านี้ถูกจัดการและประมวลผลผ่านระบบบิกดาต้า ก็จะเชื่อมโยงกับระบบส่วนกลางและนำข้อมูลที่เป็นประโยชน์ต่อผู้มีส่วนได้เสียไปเผยแพร่บนเว็บไซต์ทางกระทรวงการท่องเที่ยวและกีฬา เช่น เป็นข้อมูล

ประกอบการวางแผนการเดินทาง การวางแผนกลยุทธ์ การสื่อสารการตลาดให้สอดคล้องกับกลุ่มเป้าหมาย เป็นต้น (กระทรวงท่องเที่ยวและกีฬา Big Data เพื่อการท่องเที่ยว, 2561)

1.2 การกำหนดนโยบายของหน่วยงาน บิ๊กดาต้าช่วยให้หน่วยงานของภาครัฐออกนโยบายและมาตรการที่เหมาะสมและสอดคล้องกับยุทธศาสตร์ชาติ รวมทั้งภารกิจของหน่วยงานนั้นๆ อีกทั้งยังช่วยติดตามวัดผลความสำเร็จของการดำเนินตามนโยบาย เพื่อให้เกิดการปรับปรุงและพัฒนาการกำหนดนโยบายให้เหมาะสมมากขึ้น (Data-driven Policy การดำเนินนโยบายธนาคารกลางบนพื้นฐานของข้อมูล, 2563) เช่น หน่วยงานกรมบัญชีกลาง มีนโยบายการจัดซื้อจัดจ้างยา ตัวอย่างพาราเซตามอลเป็นยาที่ใช้ในโรงพยาบาลทั่วประเทศ ยี่ห้อเดียวกันแต่การซื้อของแต่ละหน่วยงานในภาครัฐแตกต่างกัน การจัดซื้อจัดจ้างแต่ละโรงพยาบาลมีวิธีการอย่างไร การกำหนดราคา การประมูล กรมบัญชีกลาง กระทรวงการคลังสามารถมองเห็นข้อมูลทั้งประเทศและนำข้อมูลเหล่านั้นไปกำหนดนโยบายต่อไป

1.3 การสื่อสารกับประชาชน บิ๊กดาต้าช่วยให้ภาครัฐสื่อสารกับประชาชนในด้านต่างๆ ทั้งในสถานการณ์ปกติและภาวะวิกฤต โดยการบริหารจัดการข้อมูลได้อย่างมีประสิทธิภาพ นำเอาข้อมูลที่มีอยู่มาประยุกต์ให้เกิดประโยชน์ ยกตัวอย่างในสถานการณ์การแพร่ระบาดของโรคโควิด-19 ภาครัฐใช้บิ๊กดาต้าช่วยบริหารจัดการข้อมูลผ่านสื่อดิจิทัล เช่น เว็บไซต์ thaifightcovid.depa.or.th แอปพลิเคชัน COVID Tracker by MOPH แอปพลิเคชัน หมอชนะ เป็นต้น เพื่อรวบรวมข้อมูลเกี่ยวกับโควิด-19 เพื่อให้ทราบข้อมูลสถานการณ์โควิด-19 ทั้งในไทยและทั่วโลก การเดินทางของประชาชน ยอดจำนวนผู้ติดเชื้อไวรัสโควิด-19 สถานพยาบาล การติดตามผู้ติดเชื้อ หรือการประเมินตัวเองความเสี่ยงในการติดโรคของตนเอง ทั้งนี้ เพื่อนำไปสู่การจัดทำระบบสืบค้น วางมาตรการเฝ้าระวังความเดือดร้อนของผู้ได้รับผลกระทบจากสถานการณ์การแพร่ระบาดโควิด-19 และสื่อสารไปยังประชาชนให้เข้าใจและมีแนวทางปฏิบัติที่ถูกต้อง และเป็นแนวทางให้หน่วยงานภาครัฐสามารถสื่อสารแนวทางการแก้ไขปัญหา และช่วยเหลือประชาชนในด้านต่างๆ นอกจากนี้ ประชาชนเองสามารถทราบข้อมูลพิกัดสถานที่หากมีผู้ที่ได้รับเชื้อไวรัสโควิด-19 ทำให้สามารถหลีกเลี่ยงได้ (เมืองเพลง ปัญญาชาติ, ม.ป.ป.)

1.4 การปรับปรุงการบริการและตอบสนองตรงตามความต้องการของประชาชน บิ๊กดาต้ามีส่วนช่วยปรับปรุงประสิทธิภาพพัฒนาระบบข้อมูลที่เกี่ยวข้องกับประชาชนซึ่งมีปริมาณมากให้สามารถรองรับการทำงานและบริการประชาชนผ่านระบบดิจิทัลได้มีประสิทธิภาพมากขึ้น และเชื่อมโยงข้อมูลระหว่างหน่วยงานต่างๆ ให้เข้าถึงง่ายขึ้น ด้วยการสนับสนุนงานด้านต่างๆ เช่น งานสารบรรณ งานแผน บัญชี เป็นต้น ทำให้บุคลากรมีข้อมูลที่ทันสมัยรองรับ สามารถดำเนินงานถูกต้อง และบริการประชาชนทำได้อย่างรวดเร็ว ลดขั้นตอนในการทำงาน ส่งผลให้การลดค่าใช้จ่าย ลดเรื่องร้องเรียนของหน่วยงานรัฐ และเพิ่มความโปร่งใสของการทำงานของภาครัฐ (Big Data หรือข้อมูลขนาดใหญ่มหาศาสตร์ที่มีประโยชน์กับทุกภาคส่วน, 2019)

2. ความสำคัญของบิ๊กดาต้าต่อภาคเอกชน ภาคเอกชนหรือภาคธุรกิจดำเนินการพัฒนา สร้างระบบจัดเก็บและใช้บิ๊กดาต้า ไว้รองรับการสร้างธุรกิจใหม่ๆ และตอบสนองความต้องการของผู้บริโภค ไม่ว่าจะเป็นการวิเคราะห์พฤติกรรมผู้บริโภค รู้ว่าผู้บริโภคชอบสินค้าอะไร ผู้บริโภคกำลังค้นหาสินค้าอะไร สินค้าใดขายดีติดอันดับ 1 ใน 10 บริเวณใดจะเป็นทำเลทองของธุรกิจ รวมทั้งสำรวจตลาดไปในตัวด้วยการใช้บิ๊กดาต้า ตัวอย่างของบิ๊กดาต้าต่อภาคเอกชน ได้แก่ (อังชิตา ลิมปีปัทมปาณี และคณะ, 2563)

2.1 การพัฒนาผลิตภัณฑ์และการตอบสนองต่อความต้องการของลูกค้า บริษัทต่างๆ เช่น Netflix และ Procter & Gamble ใช้บิกดาต้าเพื่อวิเคราะห์ความต้องการของผู้บริโภค บริษัทต่างๆ สร้างโมเดลในการทำนายผลิตภัณฑ์ใหม่ และบริการ โดยการจัดประเภทข้อมูลที่สำคัญเพื่อสร้างความสัมพันธ์ระหว่างผลิตภัณฑ์ในปัจจุบันและอดีต ว่ามีการนำเสนอและดำเนินการตลาดโฆษณาอย่างไรให้ประสบความสำเร็จ นอกจากนี้ Procter & Gamble ยังใช้ข้อมูลและการวิเคราะห์จากผลการวิจัยจากการสนทนากลุ่ม ข้อมูลบนสื่อสังคมออนไลน์ การทดสอบตลาดและการวางแผนผลิตและเปิดตัวผลิตภัณฑ์ (What is Big Data?, n.d.)

นอกจากนี้ ตัวอย่างหนึ่งที่น่าสนใจของการใช้บิกดาต้ามาการตอบสนองต่อความต้องการของลูกค้า คือ แอมะซอน.คอม (Amazon.com) เว็บไซต์ในลักษณะอีคอมเมิร์ซ ซึ่งขายสินค้าออนไลน์ได้ใช้เทคโนโลยีบิกดาต้าในการวิเคราะห์ข้อมูลขนาดใหญ่ของผู้ซื้อสินค้า มาพัฒนาระบบแนะนำสินค้าแบบอัตโนมัติ ทำให้ลูกค้ามีทางเลือกในการซื้อสินค้า และตัดสินใจซื้อในที่สุด ลองนึกถึงเมื่อลูกค้าเข้าเว็บไซต์ Amazon.com และทำการค้นหาสินค้าประเภทต่างๆบนเว็บไซต์ ระบบจะคัดเลือก และแนะนำประเภทสินค้าต่างๆ ที่ใกล้เคียงหรือน่าสนใจมาให้ ทั้งนี้ การประมวลผลของบิกดาต้า จะใช้ข้อมูลสถิติที่เก็บจากการรวบรวมข้อมูลของผู้ที่เข้ามาบนเว็บไซต์ถึงประวัติการค้นหา และการสั่งซื้อ รวมทั้งการร้องเรียนปัญหาต่างๆ ซึ่งทำให้บริษัท Amazon สามารถนำเสนอสินค้าบริการที่ตอบสนองต่อความต้องการของลูกค้าได้ (Apipoj Piasak, 2020)

The screenshot shows a web browser window with the Amazon.com URL. Below the browser, there is a sponsored advertisement for 'amazfit' with a price of \$119.99 and a coupon for \$10 off. The main content area is titled 'Inspired by your browsing history' and displays five book recommendations. Each recommendation includes a book cover, title, author, star rating, price, and shipping information. The books are: 'Media Literacy' by W. James Potter (144 reviews, \$85.00), 'Interviewing: Principles and Practices' by Charles Stewart (120 reviews, \$122.50), 'Persuasion: Theory and Research' by Daniel J. O'Keefe (54 reviews, \$88.10), 'Interviewing: A guide for journalists and writers' by Gail Sedorkin (21 reviews, \$20.18), and 'Communication in Everyday Life: A Survey of Communication' by Steve Duck (121 reviews, \$46.59).

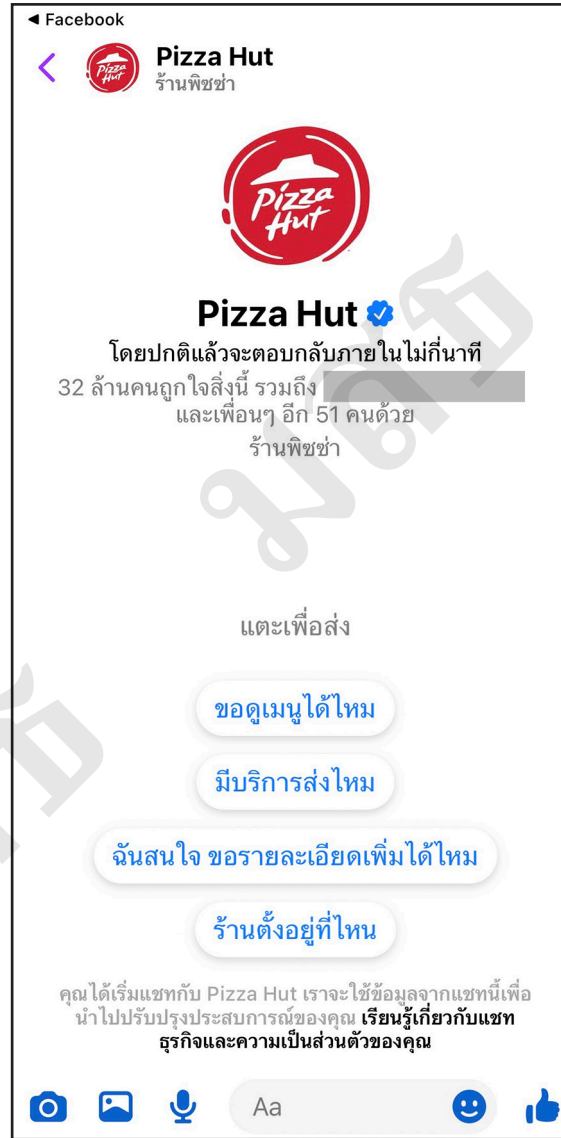
ภาพที่ 1.1 ตัวอย่างของการแนะนำสินค้าอัตโนมัติของเว็บไซต์ Amazon.com หลังจากที่คุณเขียนค้นหาหนังสือการสื่อสารมวลชน ระบบของเว็บไซต์จะแนะนำหนังสือด้านนิเทศศาสตร์อื่นๆ ที่ใกล้เคียง เช่น การรู้เท่าทันสื่อ การโน้มน้าวใจ เป็นต้น

2.2 การวิเคราะห์ความเสี่ยงในงานธนาคาร บิ๊กดาต้าถูกนำมาใช้ในการทำธุรกรรมการเงินและการไหลเข้าออกบัญชีของลูกค้า คัดกรองลูกค้าชั้นดีที่มีวินัยทางการเงินดีและสามารถกำหนดสินเชื่อปล่อยกู้ให้ลูกค้า พร้อมกำหนดดอกเบี้ยที่เหมาะสมของบุคคลนั้นๆ เพราะสามารถวิเคราะห์พฤติกรรมการใช้เงินของกลุ่มลูกค้าแต่ละกลุ่มได้ว่า มีการออมเงินอย่างไร หรือกำหนดประเภทบริการทางการเงินหรือการลงทุนเพื่อตอบโจทย์แต่ละกลุ่มลูกค้าได้ อย่างไรก็ดี การใช้บิ๊กดาต้าในภาคธุรกิจมีผลกระทบต่อลูกค้าผู้บริโภคหรือสังคมหลายประการ จำเป็นต้องพิจารณาให้รอบคอบในการนำมาใช้

2.3 การพัฒนาระบบขนส่งสินค้าและโลจิสติกส์ ภาคธุรกิจขนส่งและโลจิสติกส์ได้นำเทคโนโลยีบิ๊กดาต้ามาช่วยในการทำงานร่วมกับเทคโนโลยีอื่นๆ เช่น ระบุตำแหน่งบนพื้นโลก (GPS) ข้อมูลสัญญาณโทรศัพท์ (Mobile Data) หรือการบริการเทคโนโลยีด้านแผนที่ (Google Maps) เพื่อเก็บรวบรวมข้อมูล วิเคราะห์ข้อมูลพฤติกรรมกรรมการขนส่งสินค้าและการเดินทาง ยกตัวอย่างเช่น บริษัทขนส่งสินค้า A ได้บันทึกข้อมูลความเร็วในการขับขี่ การเบรกกะทันหัน ประวัติการเกิดอุบัติเหตุของคนขับ เพื่อมาวิเคราะห์และประเมินว่าพฤติกรรมกรรมการขับรถของคนขับได้มาตรฐานความปลอดภัยหรือไม่ นอกจากนี้ บิ๊กดาต้ายังสามารถช่วยให้บริษัทขนส่งวิเคราะห์ประวัติการบำรุงรักษา การซ่อมแซม และประสิทธิภาพของรถ เพื่อนำมาปรับปรุงพัฒนาให้การขนส่งสินค้ามีประสิทธิภาพ รวดเร็วและลดต้นทุนอื่นๆ รวมถึงการเกิดอุบัติเหตุบนท้องถนนได้ ด้านการประเมินความเสียหายและการบำรุงรักษา (น้ำเซ, 2565)

2.4 การพัฒนาระบบการบริการตอบคำถามและให้ข้อมูลแก่ลูกค้า บิ๊กดาต้ามีส่วนช่วยธุรกิจในการพัฒนาระบบการบริการของลูกค้าบนเว็บไซต์หรือแอปพลิเคชันต่างๆ ในการตอบคำถามหรือให้ข้อมูลแก่ลูกค้า โดยการประมวลผลข้อมูลจากประวัติการถามคำถาม หรือการใช้บริการ เช่น เทคโนโลยี Chatbot เป็นโปรแกรมคอมพิวเตอร์ที่สื่อสารกับลูกค้าผ่านเสียง หรือข้อความได้ เสมือนเป็นมนุษย์คนหนึ่งในการช่วยแก้ไขปัญหาหรือตอบคำถามของลูกค้า แต่คำถามของลูกค้าจะต้องมีความชัดเจน เพื่อไม่ให้เกิดความสับสนและได้รับคำตอบไม่ถูกต้อง (Chatbot สุดยอดผู้ช่วยอัจฉริยะ, ม.ป.ป.)

ในปัจจุบันการใช้บิ๊กดาต้ามาพัฒนาระบบ Chatbot ได้รับความนิยมมาก โดยเฉพาะผู้ประกอบการซื้อขายสินค้าออนไลน์ ซึ่งลูกค้าหรือผู้ที่สนใจสินค้ามักจะสอบถามข้อมูลสินค้าจำนวนมากก่อนการตัดสินใจซื้อผ่านการตอบแชท เช่น บนเพจ Facebook ลูกค้า แต่เพราะมีลูกค้าหรือผู้ที่สนใจสินค้าจำนวนมาก ผู้ประกอบไม่สามารถตอบคำถามได้อย่างรวดเร็ว Chatbot จึงช่วยแก้ปัญหาดังกล่าวได้อย่างรวดเร็ว



ภาพที่ 1.2 ตัวอย่างการนำบิกดาต้ามาใช้ในการพัฒนาระบบ Chatbot

3. ความสำคัญของบิกดาต้าต่องานสื่อสารมวลชน บิกดาต้ามีความสำคัญต่องานสื่อสารมวลชน ทั้งด้านการผลิตและรายงานข่าว และด้านการผลิตและปรับปรุงรายการบันเทิงเพื่อเผยแพร่ผ่านสื่อ ดังรายละเอียดดังนี้

3.1 ด้านการผลิตและรายงานข่าว รายการข่าวจำเป็นต้องอาศัยบิกดาต้าในการวิเคราะห์ ความสนใจของผู้ชม และพฤติกรรมการเปิดรับสาร เพื่อให้สามารถผลิตรายการข่าวที่ตรงตามความต้องการ ผู้รับสารให้มากที่สุด และสร้างความได้เปรียบทางการแข่งขันกับสถานีข่าวคู่แข่งถึงแม้ว่า BBC จะไม่ใช่ องค์กรแสวงหากำไรอย่างแท้จริง ยกตัวอย่างเช่น ในปี พ.ศ. 2556 องค์กรกระจายเสียงสาธารณะของ

สหราชอาณาจักร หรือ บีบีซี (BBC) ได้ริเริ่มโครงการการนำบิกดาต้ามาช่วยในการพัฒนาแพลตฟอร์มและกรอบสำหรับกรวิเคราะห์ข้อมูลและสร้างแบบจำลองทำนายพฤติกรรมผู้ชม จากการเก็บรวบรวมข้อมูลบนสื่อสังคมออนไลน์ รายงานข่าว และข้อมูลโปรแกรมรายการโทรทัศน์ โดยผลที่ได้คือการแนะนำเนื้อหาที่สามารถผลิตให้ตรงตามความต้องการของผู้ชมได้มากขึ้น อีกทั้ง BBC ยังสามารถนำเสนอรายการที่เหมาะสมกับพฤติกรรมกรรับชมรายการของผู้ชมกลุ่มต่างๆ ที่แตกต่างกันด้วย (Marr, 2016)

ในด้านกรรายงานข่าวนักข่าวหรือบรรณาธิการข่าวต้องนำข้อมูลที่รวบรวมมา หรือจากคลังข้อมูลบิกดาต้า มาวิเคราะห์ ถอดรหัส ประมวลผล เพื่อประกอบการรายงานข่าว ข้อมูลขนาดใหญ่นี้เป็นแหล่งข้อมูลที่สำคัญ จากนั้นนำข้อมูลที่จัดระบบมาวิเคราะห์และประมวลผล ทำให้ข้อมูลมีคุณค่า ได้ข้อมูลที่มีความน่าสนใจที่จะรายงานเป็นข่าว มองเห็นข้อมูลขนาดใหญ่ สามารถนำเสนอเป็นประเด็นข่าว ซึ่งมีหลายมิติในข้อมูลเดียวกัน ความสำคัญของบิกดาต้าที่วิเคราะห์และประมวลผลแล้ว มีคุณค่าข่าวทั้งความสำคัญ ความน่าสนใจ และความรวดเร็ว ต่อกรรายงานให้ประชาชนทราบ (อังชิตา ลิมปีปัทมปาณี และคณะ, 2563)

3.2 ด้านกรผลิตและปรับปรุงรายการบันเทิงเพื่อเผยแพร่ผ่านสื่อ ในการผลิตรายการบันเทิง ไม่ว่าจะเป็นรายการโทรทัศน์ ละคร หรือแม้แต่ภาพยนตร์ บิกดาต้าช่วยในการพัฒนาสร้างสรรค์และผลิตรายการต่างๆ โดยช่วยในการวิเคราะห์กระแสความนิยม ข้อมูล ความคิดเห็นต่างๆ เกี่ยวกับรายการนั้นๆ ผ่าน สื่อสังคมออนไลน์ เพื่อนำความคิดเห็นของผู้คนที่พูดถึงความสนใจ กระแส หรือทัศนคติต่อนักแสดงมาวิเคราะห์ เพื่อผลิตเนื้อหา และรูปแบบให้ตรงตามความต้องการของผู้รับสาร นอกจากนี้ บิกดาต้ายังช่วยในการวางแผนการสื่อสารการตลาด ประชาสัมพันธ์ให้ตรงตามลักษณะของกลุ่มเป้าหมายรับชมอีกด้วย เพราะรายการที่ผลิตจะมีการกำหนดกลุ่มเป้าหมายทั้งเพศ อายุและพฤติกรรมอื่น ที่เกี่ยวข้อง

เน็ตฟลิกซ์ (Netflix) เป็นตัวอย่างหนึ่งที่แสดงถึงความสำคัญของบิกดาต้าในการผลิตรายการจนประสบความสำเร็จ เมื่อมีการใช้บิกดาต้าวิเคราะห์ความต้องการของกลุ่มเป้าหมายก่อนการสร้างซีรีส์เรื่อง House of cards ซึ่งนำเสนอเนื้อหาเกี่ยวกับการเมืองอันร้อนแรงและการดำรงตำแหน่งประธานาธิบดีของประเทศสหรัฐอเมริกา ซึ่งมีต้นทุนกรสร้างสูงถึง 100 ล้านดอลลาร์สหรัฐอเมริกา จนได้รับความนิยมอย่างล้นหลาม

Netflix ใช้บิกดาต้าในการวิเคราะห์ข้อมูลว่าเนื้อหาซีรีส์ลักษณะใดที่กลุ่มผู้ชมชื่นชอบแต่ยังไม่ได้มีการผลิตหรือออกสู่ตลาด รวมทั้งข้อมูลทัศนคติต่อนักแสดงและผู้กำกับภาพยนตร์หรือรายการโทรทัศน์ ผลกรวิเคราะห์พบว่าซีรีส์/ภาพยนตร์ที่กลุ่มผู้ชม Netflix ชอบมาก คือ 1) ซีรีส์กรเมืองเรื่อง House of Cards ซึ่งเคยมีการผลิตในประเทศอังกฤษ และออกอากาศผ่านช่อง BBC ในปี ค.ศ. 1990 จำนวน 4 ตอน 2) ภาพยนตร์เรื่อง The Social Network โดยมี David Fincher เป็นผู้กำกับ และ 3) ผู้ที่ชื่นชอบ House of Cards ซึ่งผลิตในปี ค.ศ. 1990 มักจะชอบภาพยนตร์ที่ Kevin Spacey แสดงนำ หรือ David Fincher เป็นผู้กำกับ ซึ่งผลกรวิเคราะห์ดังกล่าวนำไปสู่กรตัดสินใจผลิตซีรีส์เรื่อง House of Cards ในฉบับประเทศสหรัฐอเมริกา โดยมีแสดงนำคือ Kevin Spacey และ David Fincher เป็นผู้กำกับ จนประสบความสำเร็จ (Nattapon Muangtum, 2021)

กล่าวโดยสรุป จะเห็นได้ว่าบิกดาต้าเป็นเทคโนโลยีที่มีบทบาทสำคัญต่อทุกภาคส่วน ทั้งภาครัฐ ภาคเอกชน และต่องานสื่อสารมวลชน ในการพัฒนาปรับปรุงประสิทธิภาพของการสื่อสารไปยังประชาชน ในรูปแบบต่างๆ อย่างไรก็ตาม การทำความเข้าใจเกี่ยวกับกระบวนการทำงานและการจัดการบิกดาต้า เป็นสิ่งสำคัญเช่นกัน เพื่อให้สามารถนำมาประยุกต์ใช้งานได้อย่างมีประสิทธิภาพตรงตามเป้าหมายหรือวัตถุประสงค์ที่องค์กรต่างๆ วางไว้ โดยนักศึกษาจะได้เรียนรู้การทำงานของบิกดาต้าในตอนที่ 2 กระบวนการจัดการข้อมูลขนาดใหญ่ (Big Data)

กิจกรรม 1.2

บิกดาต้ามีความสำคัญต่อภาครัฐในด้านใด และอย่างไรบ้าง

แนวตอบกิจกรรม 1.2

บิกดาต้ามีความสำคัญต่อหลายภาคส่วน ได้แก่

1. ภาครัฐ ในแง่การกำหนดยุทธศาสตร์การพัฒนาประเทศ การกำหนดนโยบายของหน่วยงาน การสื่อสารกับประชาชน ในด้านต่างๆ ทั้งในสถานการณ์ปกติและภาวะวิกฤต และการปรับปรุงการบริการ และตอบสนองตรงตามความต้องการของประชาชน
 2. ภาคเอกชน เช่น การพัฒนาผลิตภัณฑ์และการตอบสนองต่อความต้องการของลูกค้า การวิเคราะห์ความเสี่ยงในงานธนาคาร การพัฒนาระบบขนส่งสินค้าและโลจิสติกส์ การพัฒนาระบบการบริการ ตอบคำถามและให้ข้อมูลแก่ลูกค้า
 3. ภาคสื่อสารมวลชน ในด้านการผลิตและรายงานข่าว และด้านการผลิตและปรับปรุงรายการบันเทิงเพื่อเผยแพร่ผ่านสื่อ
-

บรรณานุกรม

- กระทรวงท่องเที่ยวฯ’ เดินหน้าพัฒนา Big Data เพื่อการท่องเที่ยว. (2561, 2 กุมภาพันธ์). *ไทยรัฐ*. สืบค้นจาก. <https://www.thairath.co.th/business/economics/1193920>
- หน้าเซ. (2565, 4 เมษายน). *ประโยชน์ของ BIG DATA กับงานด้านขนส่งและโลจิสติกส์*. Bus & Truck. สืบค้นจาก <https://www.busandtruckmedia.com/34182/>
- เมืองเพลง ปัญญาชาติ. (ม.ป.ป.). *บิ๊กดาต้า (Big data) กับเหตุการณ์วิกฤติไวรัสโคโรนา 19*. สืบค้นจาก <https://www.depa.or.th/th/article-view/big-data-covid19>
- สรุปชัด! “5V” องค์ประกอบและคุณลักษณะที่สำคัญของ Big Data. (2565, 4 เมษายน). สถาบันนวัตกรรมและชมรมอภิบาลข้อมูล. สืบค้นจาก <https://digi.data.go.th/blog/big-data-5v>
- อังธิดา ลิ้มปัทมปาณี, บุญลาภ ภูสุวรรณ และกมล ชาวาลวิทย์. (2563). *มหัศจรรย์ข้อมูล (Big Data)*. ใน *เอกสารการสอนชุดวิชาการรายงานข่าวและการบรรณาธิการ* (หน่วยที่ 3). นนทบุรี: มหาวิทยาลัยสุโขทัย-ศรีนครินทร์.
- Apipoj Piasak. (2020). *Amazon ใช้ Big Data อย่างไรจนประสบความสำเร็จกลายเป็นร้านค้าออนไลน์ที่มีมูลค่ามากที่สุดในโลก*. สืบค้นจาก <https://th.linkedin.com/pulse/amazon-%E0%B9%83%E0%B8%8A-big-data-apipoj-piasak>
- Big Data* หรือข้อมูลขนาดใหญ่มหาศาลที่มีประโยชน์กับทุกภาคส่วน. (2562). สืบค้นจาก. <https://www.securitysystems.in.th/2019/08/big-data/>
- Big Data* คืออะไร ลักษณะข้อมูลเป็นอย่างไร. (ม.ป.ป.). สืบค้นจาก <https://thewisdom.co/content/what-is-big-data/>
- Big data and analytics: definitions, value, trends and applications*. (n.d.). Retrieved from <https://www.i-scoop.eu/big-data-action-value-context/>
- Big Data*. (n.d.). Retrieved from <https://www.gartner.com/en/information-technology/glossary/big-data>
- Chatbot* สุดยอดผู้ช่วยอัจฉริยะ. (ม.ป.ป.). สืบค้นจาก <https://dgti.dga.or.th/chatbot-2/>
- Data Cleaning* หรือ *Data Cleansing* คืออะไร? (ม.ป.ป.). สืบค้นจาก <https://1stcraft.com/what-is-data-cleansing/>
- Data-driven Policy* การดำเนินนโยบายธนาคารกลางบนพื้นฐานของข้อมูล. (ม.ป.ป.). สืบค้นจาก https://www.bot.or.th/Thai/BOTMagazine/Pages/256306TheKnowledge_DataDrivenPolicy.aspx
- Diebold, F. (2003). *Big data dynamic factor models for macroeconomic measurement and forecasting in Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress of the Econometric Society*, (edited by Dewatripont, M., Hansen, L.P. and Turnovsky, S.).
- Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety. *META group research note*, 6(70).

- Marr, B. (2016). *Big data in practice: How 45 successful companies used big data analytics to deliver extraordinary results*. United Kingdom: John Wiley and Sons Ltd.
- Nattapon Muangtum. (2021). 3 บทเรียน Netflix กับการใช้ Big Data สร้างเป็นซีรีส์ดัง House of Cards. สืบค้นจาก <https://www.everydaymarketing.co/business-and-marketing-case-study/entertainment/netflix-big-data-analytics-house-of-cards/>
- Petabytes คืออะไร และมีขนาดใหญ่แค่ไหน. (ม.ป.ป.). สืบค้นจาก <https://www.quickserv.co.th/>
- Roy, A. S. (2020). *How does Facebook handle the 4+ petabyte of data generated per day? Cambridge Analytica - Facebook data scandal*. Retrieved from <https://medium.com/>
- SME Thailand. (2565, 21 เมษายน). แชร้เทคนิค จัดการข้อมูล Big Data อย่างไรให้เพอร์เฟ็คในยุคที่ AI ครองเมือง. *SME StartUp*. สืบค้นจาก <https://www.smethailandclub.com/startup-techstart-up/8003.html>
- What is Big Data?* (n.d.). Retrieved from <https://www.oracle.com/big-data/what-is-big-data/>

ตอนที่ 2

กระบวนการจัดการข้อมูลขนาดใหญ่

รองศาสตราจารย์ ดร.มนวิภา วงรุจิระ



ชื่อ
วุฒิ

รองศาสตราจารย์ ดร.มนวิภา วงรุจิระ
ว.บ. (การหนังสือพิมพ์) มหาวิทยาลัยธรรมศาสตร์
M.A. (Journalism) Ohio State University, USA
Ph.D. (Mass Communication) Florida State University, USA

ตำแหน่ง

รองศาสตราจารย์ประจำสาขาวิชานิติศาสตร์
มหาวิทยาลัยสุโขทัยธรรมราช

หน่วยที่เขียน

ตอนที่ 2

ตอนที่ 2

กระบวนการจัดการข้อมูลขนาดใหญ่

โปรดอ่านหัวเรื่อง แนวคิด และวัตถุประสงค์ของตอนที่ 2 แล้วจึงศึกษารายละเอียดต่อไป

หัวเรื่อง

เรื่องที่ 2.1 ขั้นตอนการจัดการข้อมูลขนาดใหญ่ (Big Data)

เรื่องที่ 2.2 แหล่งข้อมูลในการจัดการข้อมูลขนาดใหญ่

แนวคิด

1. การจัดการข้อมูลขนาดใหญ่มีขั้นตอนหลักๆ 3 ขั้นตอน คือ การรวบรวมข้อมูล การประมวลผลข้อมูล และการวิเคราะห์และนำเสนอข้อมูล ซึ่งในบางกรณีอาจใช้วิธีการตั้งเป้าหมายของการใช้ข้อมูลขนาดใหญ่มาเป็นจุดเริ่มต้นก่อนนำไปสู่ขั้นตอนการรวบรวมข้อมูล วิเคราะห์และนำเสนอก็ได้
2. ข้อมูลขนาดใหญ่แบ่งตามลักษณะของข้อมูลจะมี 3 ประเภท ได้แก่ ข้อมูลแบบมีโครงสร้าง ข้อมูลแบบไม่มีโครงสร้าง และข้อมูลแบบกึ่งโครงสร้างที่เป็นการผสมข้อมูลที่ทั้งมีโครงสร้างและไม่มีโครงสร้าง ส่วนแหล่งข้อมูลสามารถแบ่งได้ตามที่มาของข้อมูล ได้แก่ แหล่งข้อมูลของภาครัฐ ของภาคเอกชน และภาคประชาสังคม นอกจากนี้ยังสามารถแบ่งประเภทของแหล่งข้อมูลตามช่องทางได้อีก 2 ประเภท ได้แก่ แหล่งข้อมูลในระบบออนไลน์ และในระบบออฟไลน์

วัตถุประสงค์

เมื่อศึกษาตอนที่ 2 จบแล้ว นักศึกษาสามารถ

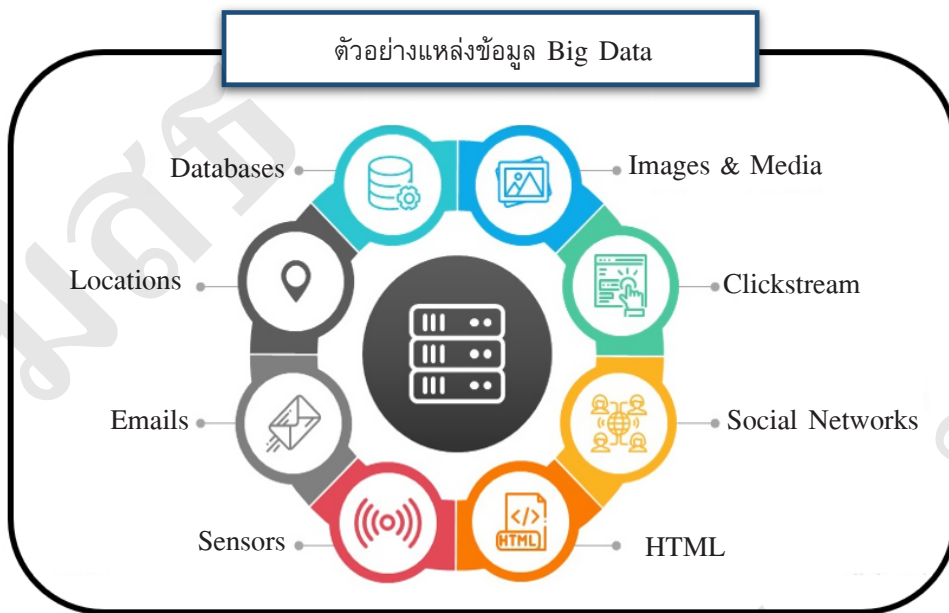
1. อธิบายขั้นตอนการจัดการข้อมูลขนาดใหญ่ได้
2. ระบุแหล่งที่มาของข้อมูลในการจัดการข้อมูลขนาดใหญ่ได้

เรื่องที่ 2.1

ขั้นตอนการจัดการข้อมูลขนาดใหญ่

ในการจัดการข้อมูลขนาดใหญ่มีขั้นตอนหลักๆ 3 ขั้นตอน กล่าวคือ

1. **การรวบรวมข้อมูล (Storage)** เป็นการรวบรวมข้อมูลจากแหล่งต่างๆ ทั้งข้อมูลที่มีคุณภาพ ข้อมูลที่คาดว่าจะมีประโยชน์ ซึ่งอาจจะครบถ้วนหรือไม่ครบถ้วน ข้อมูลหลากหลายประเภทอย่างข้อความ รูปภาพ วิดีโอ ไฟล์เสียงทั้งหลายจะส่งมาจัดเก็บที่ถังข้อมูล ตัวอย่างข้อมูลที่องค์กรรวบรวมไว้อยู่แล้ว เช่น ข้อมูลลูกค้าหรือสมาชิกที่มีทั้งรายละเอียดตัวบุคคลอย่าง เพศ อายุ รายได้ หรือพฤติกรรมกรซื้อสินค้า/บริการ หรือข้อมูลการเข้าใช้งานเว็บไซต์หรือสื่อสังคมออนไลน์ที่จะระบุพฤติกรรมกรใช้งานทั้งความถี่ ระยะเวลา เนื้อหาที่เข้าชม การแสดงความคิดเห็น การแบ่งปันเนื้อหา เป็นต้น



ภาพที่ 2.1 ตัวอย่างข้อมูลหลากหลายรูปแบบ

ที่มา: <https://www.bigdata-otp.com/activity.html>

2. **การประมวลผล (Processing)** เมื่อข้อมูลต่างๆ ได้นำมารวมกันไว้ในที่เดียวแล้ว ข้อมูลที่ได้มาอาจจะไม่ทั้งที่ครบถ้วนและไม่ครบถ้วน ทั้งที่เป็นรูปแบบที่หลากหลาย จะต้องมีการจัดระบบระเบียบ จะต้องนำไปจัดหมวดหมู่ข้อมูลที่มีความเกี่ยวข้องสัมพันธ์กัน ให้ผลคล้ายคลึงกัน หรือปรับข้อมูลให้มีรูปแบบที่สามารถนำไปประมวลผลได้ เช่น ข้อมูลที่เป็นข้อความในรูปแบบของไฟล์ PDF อาจไม่สามารถนำ

มาใช้วิเคราะห์ข้อความได้ก็ต้องแปลงข้อมูลเหล่านั้นให้เป็นข้อความที่สามารถใช้โปรแกรมคอมพิวเตอร์ในการวิเคราะห์ได้ เป็นต้น หลังจากจัดระบบข้อมูลแล้ว ก็จะนำข้อมูลเหล่านั้นมาเข้าสู่กระบวนการประมวลผล ซึ่งในขั้นตอนนี้จะมีการใช้เทคโนโลยีและโปรแกรมคอมพิวเตอร์ต่างๆ ในการประมวลผล แล้วนำมาเปลี่ยนเป็นรูปแบบข้อมูลเพื่อเอาเข้าระบบคลังข้อมูลที่ผ่านการประมวลผลแล้ว

3. การวิเคราะห์และนำเสนอ (Analysis) จากนั้นข้อมูลมากมายทั้งหมดที่ถูกจัดเรียงหรือประมวลผลในหลายมิติจะถูกนำมาวิเคราะห์หารูปแบบหรือ pattern ของข้อมูลที่ไม่เห็นด้วยตาเปล่า เช่น หารูปแบบความสัมพันธ์ที่ซ่อนอยู่ หาแนวโน้มการตลาดหรือพฤติกรรมของลูกค้าหรือผู้ใช้งานสื่อ เทรนด์ความชอบของลูกค้า และข้อมูลอื่นๆ ที่เป็นประโยชน์ต่อองค์กร และนำมาจัดทำเป็นรูปแบบการนำเสนอที่เข้าใจง่ายผ่านทางสถิติ กราฟ อินโฟกราฟิก หรือรูปภาพ

ในอีกมุมมองหนึ่งกระบวนการจัดการข้อมูลขนาดใหญ่ หรือ บิ๊กดาต้า (Big Data) อาจไม่ได้เริ่มที่การรวบรวมข้อมูลเป็นเบื้องต้น แต่เริ่มด้วยโจทย์หรือเป้าหมายที่ต้องการก่อนแล้วจึงนำไปสู่ขั้นตอนการรวบรวมข้อมูล ดังกรณีตัวอย่างการจัดการข้อมูลขนาดใหญ่เพื่อนำไปสู่การรายงานข่าวเชิงข้อมูล (อังกฤษ: ลิ้มปี๋ทมปาณี และคณะ, 2563) ประกอบด้วย 5 ขั้นตอน ต่อไปนี้

ขั้นตอนแรก: การวางแผนและออกแบบ (Design) เป็นการตั้งเป้าหมาย ตั้งโจทย์หรือสมมติฐานที่ต้องการจะใช้ข้อมูลเพื่อนำเสนอเรื่องอะไร การออกแบบ ‘โมเดล’ ของซีรีส์ข้อมูลประเด็นข่าว เป็นการกำหนดประเด็นข่าวหรือสารคดีเชิงข่าวที่ต้องการรายงานเป็นชุดเนื้อหาต่อเนื่องกัน เปรียบเทียบเท่ากับการสร้างโครงงานวิจัย เพื่อตอบโจทย์ หรือสมมติฐาน หรือวัตถุประสงค์ที่ต้องการ ซึ่งต้องมีโมเดลหรือแบบจำลองของงานวิจัย

ขั้นตอนที่สอง: การเก็บสะสมข้อมูล (Information collection) หรือการรวบรวมข้อมูลจากแหล่งต่างๆ โดยแหล่งที่มาข้อมูลเป็นจุดสำคัญที่ต้องเรียนรู้ ต้องเชื่อมโยงกับการค้นหาข้อมูล การเลือกใช้ประเภทข้อมูล แหล่งที่มาข้อมูลขนาดใหญ่จะเก็บไว้ที่ใด เรียกใช้อย่างไรให้ตรงเป้าหมายที่ต้องการ องค์กรจะมีการสะสมข้อมูลอย่างต่อเนื่องในช่วงระยะเวลาหนึ่ง บางข้อมูลจะมาจากการสะสมขององค์กร บางข้อมูลอาจจะมาจากแหล่งอื่นๆ การเก็บสะสมข้อมูลอาจใช้ระยะเวลานับเดือนหรือหลายปี อีกทั้งข้อมูลที่มีความกว้างหลายมิติ ซึ่งข้อมูลอาจอยู่ในแบบมีโครงสร้าง หรือข้อมูลแบบไม่มีโครงสร้าง หรือข้อมูลผสม ทั้งแบบมีโครงสร้างและไม่มีโครงสร้าง แต่อย่างไรก็ตามจะต้องเป็นข้อมูลที่มากพอจึงจะทำได้ หากข้อมูลน้อยเกินไปหรือเป็นข้อมูลที่ไม่สัมพันธ์เกี่ยวข้อง อาจไม่สามารถใช้เป็นบิ๊กดาต้าได้ จึงควรรู้จักแหล่งที่มาข้อมูลและเทคโนโลยีที่จะนำไปใช้ในบิ๊กดาต้า

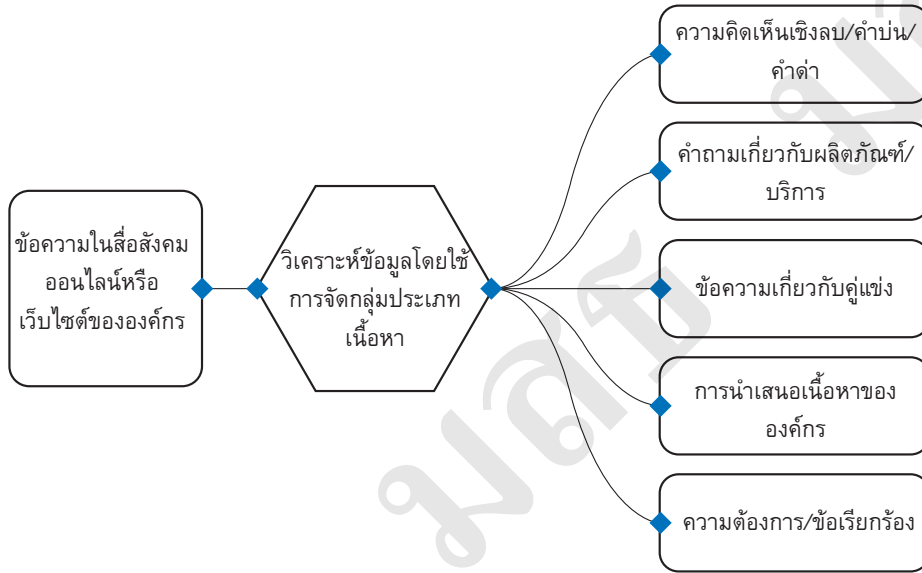
ขั้นตอนที่สาม: การทำความสะอาดข้อมูล (Data cleaning) เนื่องจากข้อมูลขนาดใหญ่ที่ได้อาจมีจำนวนมาก และข้อมูลที่น่ามาใช้อาจมีรูปแบบและลักษณะแตกต่างกัน รวมทั้งจะมีข้อมูลบางอย่างที่จะสามารถตอบโจทย์หรือสมมติฐานได้ ดังนั้น จึงต้องมีกระบวนการทำความสะอาดข้อมูลเพื่อให้ได้รูปแบบและลักษณะข้อมูลที่นำไปสู่การทำงานในขั้นตอนต่อไปได้ การทำความสะอาดข้อมูลมีวัตถุประสงค์เพื่อกำจัดข้อมูลที่ผิดพลาด (error) หรือบกพร่อง (missing) ทำให้ข้อมูลเกิดความแม่นยำ ทำนายผลได้ เครื่องมือที่ใช้เป็นโปรแกรมคอมพิวเตอร์ เช่น โปรแกรมเอกเซล โปรแกรมทางคณิตศาสตร์อื่นหรืออาจใช้อัลกอริทึม (Algorithm) (การใช้ชุดคำสั่งหรือโปรแกรมซอฟต์แวร์โดยกำหนดเงื่อนไขแบบที่ละ

ขั้นตอนที่จะทำให้คอมพิวเตอร์หรืออุปกรณ์อิเล็กทรอนิกส์นั้นทำสิ่งที่เรากำหนดให้ ซึ่งต้องใช้ชุดคำสั่งสำหรับการนั้นโดยเฉพาะ) ทำให้เห็นข้อมูลที่ผิดปกติ ข้อมูลผิดพลาด ข้อมูลบกพร่อง แล้วทำความสะอาดให้ข้อมูลนำมาใช้ได้ถูกต้อง การทำความสะอาดข้อมูลประกอบด้วย

- 1) แจกแจงข้อมูล ให้เป็นหัวข้อตรงกับข้อมูลที่จะใช้ประโยชน์
- 2) แก้ไขชุดข้อมูลที่ผิดหรือมีข้อบกพร่องให้มีความถูกต้อง และใช้งานได้
- 3) ทำข้อมูลให้เป็นมาตรฐานเดียวกัน
- 4) ลบข้อมูลที่ซ้ำซ้อนหรือไม่เกี่ยวข้องออกไป

ขั้นตอนที่สี่: การวิเคราะห์และประมวลผลข้อมูล (Data analysis) ต้องใช้เทคโนโลยีและเครื่องมือสนับสนุนมาช่วยในการวิเคราะห์และประมวลผลข้อมูลขนาดใหญ่ให้ออกมาอย่างรวดเร็วถูกต้อง ได้ผลลัพธ์ตรงตามโจทย์ การวิเคราะห์ข้อมูลขนาดใหญ่ต้องให้ความสำคัญกับความรวดเร็ว ความแม่นยำ ความถูกต้อง โดยอาศัยเครื่องมือวิเคราะห์ทางสถิติ ทางคณิตศาสตร์ เช่น ค่าสัมประสิทธิ์ (Correlations) ค่าสัมประสิทธิ์สหสัมพันธ์ (Correlations Matrix) เป็นต้น หรืออาจใช้เทคโนโลยีและเครื่องมือสมัยใหม่ เช่น ระบบฮาตูป ระบบคลาวด์ริทา ระบบฮอ์ตั้นเว็ริก เป็นต้น เพื่อมาช่วยสนับสนุนการทำงานรวดเร็ว ซึ่งข้อมูลแต่ละส่วนจะอธิบาย ภายใต้ ‘ซีรีส์ข้อมูลประเด็นข่าว’ ทีละประเด็น จากนั้น จึงสรุปข้อมูลรวม เพื่อตอบโจทย์หรือสมมติฐานที่ตั้งไว้ ในส่วนนี้นักข่าวหรือบรรณาธิการข่าวจะทำงานร่วมกับวิศวกรออกแบบระบบ และนักวิทยาศาสตร์ออกแบบข้อมูล

ขั้นตอนที่ห้า: การนำเสนอ (Data visualization/presentation) เป็นขั้นตอนที่ต้องนำเสนอข้อมูลเพื่อสื่อสารกับผู้อ่าน ผู้ชม ผู้ฟัง ด้วยวิธีการแตกต่างในแต่ละช่องทางสื่อหรือแพลตฟอร์มซึ่งมีคุณสมบัติต่างกัน การนำเสนอข้อมูลนี้สามารถนำเสนอในรูปแบบคอมพิวเตอร์กราฟิก รูปภาพ ลายเส้น แอนิเมชัน สื่อประสม คลิป การ์ตูน หรือการเคลื่อนไหวแบบสตอปโมชัน ฯลฯ เหล่านี้ เป็นการนำเสนอให้เกิดความเข้าใจใน ‘ภาพรวมหรือโครงสร้าง’ ของสิ่งที่ค้นพบตามประเด็นข่าว โดยต้องเป็นไปอย่างเหมาะสมตามช่องทางของสื่อแต่ละประเภท การนำเสนอข้อมูลขนาดใหญ่ที่มีการวิเคราะห์และประมวลผลแล้วให้เป็นที่เข้าใจได้ไม่ยากนั้นต้องอาศัยความคิดสร้างสรรค์และออกแบบ จากข้อมูลที่ซับซ้อน หลากหลาย จำนวนมหาศาล ต้องย่อยแล้วจัดเรียงใหม่ ใช้เทคโนโลยีคอมพิวเตอร์ต่างๆ และศิลปะการออกแบบมาสร้างสรรค์นำเสนออย่างน่าสนใจ



ภาพที่ 2.2 ตัวอย่างการวิเคราะห์ข้อความในสื่อออนไลน์โดยการจัดกลุ่มประเภทเนื้อหา

จากภาพที่ 2.2 เป็นตัวอย่างของการวิเคราะห์ข้อมูลจากข้อความที่ปรากฏในสื่อสังคมออนไลน์หรือในเว็บไซต์ขององค์กร ซึ่งโดยทั่วไปข้อความจำนวนมากเหล่านั้นหากใช้คนอ่านและวิเคราะห์จัดหมวดหมู่ก็จะใช้เวลานานมาก แต่เมื่อนำโปรแกรมคอมพิวเตอร์มาช่วยในการทำงานก็จะทำให้สามารถจัดระบบข้อความเหล่านั้นได้เป็นกลุ่มเนื้อหาต่างๆ เช่น คำร้องเรียนหรือข้อความเชิงลบของผู้ใช้เกี่ยวกับองค์กรหรือผลิตภัณฑ์ ข้อความที่มีการกล่าวถึงคู่แข่ง ความต้องการหรือข้อเรียกร้องของผู้ใช้หรือลูกค้า เป็นต้น การจัดกลุ่มประเภทเนื้อหาเหล่านี้จะช่วยให้องค์กรเห็นภาพรวมของความคิดเห็น ความต้องการของลูกค้าหรือผู้บริโภค และสามารถนำข้อมูลที่ได้ผ่านการวิเคราะห์จัดกลุ่มไว้แล้วไปใช้ประโยชน์ในการวางแผนงานหรือแก้ไขปัญหาต่างๆ ต่อไป รายละเอียดเกี่ยวกับการวิเคราะห์ข้อมูลจะได้อธิบายถึงในตอนต่อไป

กิจกรรม 2.1

ขั้นตอนการจัดการข้อมูลขนาดใหญ่ประกอบด้วยอะไรบ้าง

แนวตอบกิจกรรม 2.1

ขั้นตอนการจัดการข้อมูลขนาดใหญ่ประกอบด้วย 3 ขั้นตอน คือ

1. การรวบรวมข้อมูล
2. การประมวลผลข้อมูล
3. การวิเคราะห์และนำเสนอข้อมูล

เรื่องที่ 2.2

แหล่งข้อมูลในการจัดการข้อมูลขนาดใหญ่

การได้มาซึ่งข้อมูลบิกดาต้าจะต้องมีกระบวนการและขั้นตอนก่อนจะได้มา แหล่งที่มาของข้อมูลเป็นต้นทางสำคัญ โดย เบอร์นาร์ด มาร์ (Bernard Marr, 2015) ได้อธิบายว่าข้อมูลขนาดใหญ่สามารถแบ่งออกเป็น 3 ประเภทหลักๆ คือ

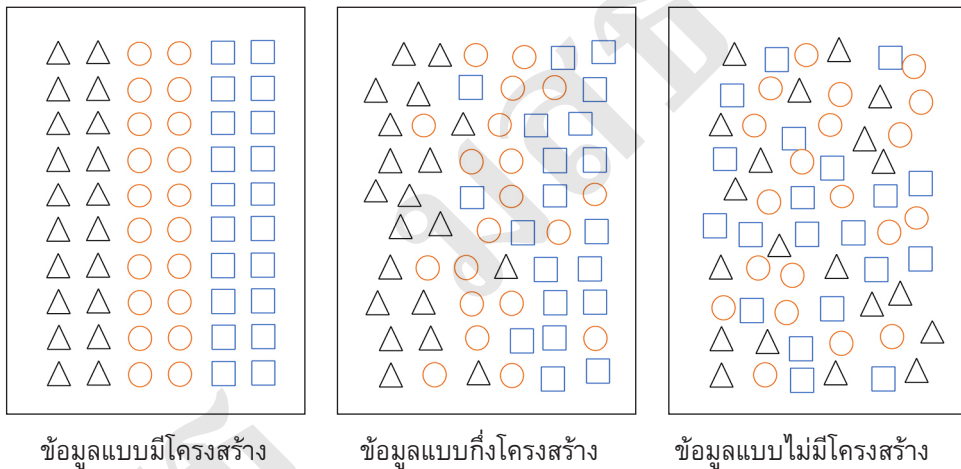
1. ข้อมูลแบบมีโครงสร้าง (Structured data) ชุดข้อมูลที่มีการจัดเรียงโครงสร้างอย่างเป็นระเบียบ มีความชัดเจน หรือระบุได้ด้วยตัวเลข พร้อมใช้งานได้ทันที เช่น จำนวนการซื้อขายกับลูกค้า เปอร์เซ็นต์ความเคลื่อนไหวภายในตลาดหุ้น เป็นต้น เป็นข้อมูลที่ง่ายต่อการค้นหาและจัดระบบ สามารถบรรจุ ‘สูตร’ หรือ ‘โมเดล’ ลงบนโปรแกรมเอกเซลในแนวแถวและแนวคอลัมน์ โปรแกรมจะวางข้อมูลให้ตรงคุณสมบัติ และหัวข้อของข้อมูล ข้อมูลจะจัดวางบนเอกเซล (Excel program) บนชีททำงาน (Spread sheet) ตามที่วางแผน ตามสูตรหรือโมเดล เช่น ข้อมูลการขายระดับภูมิภาค วางตามหัวข้อลูกค้าหรือวางตามหัวข้อผลิตภัณฑ์ ข้อมูลแบบมีโครงสร้างจะจัดกลุ่มเข้าด้วยกัน เพื่อให้ง่ายต่อการหาความสัมพันธ์ระหว่างสองหัวข้อนี้ เช่น ลูกค้ามีความพึงพอใจผลิตภัณฑ์หรือบริการอย่างไร ข้อมูลแบบมีโครงสร้างจะง่ายต่อการจัดเก็บ วิเคราะห์และค้นหาข้อมูล

2. ข้อมูลแบบไม่มีโครงสร้าง (Unstructured data) หมายถึง ชุดข้อมูลที่มีโครงสร้างไม่ชัดเจนหรือไม่สามารถระบุความแน่นอนของข้อมูลนั้นๆ ได้ ยังไม่สามารถประมวลผลเพื่อนำไปใช้ได้ทันที อย่างเช่น บทสนทนาโต้ตอบกับลูกค้าทาง Social Media ง่ายๆ คือ ข้อมูลอะไรก็ตามที่มีการรวบรวมไว้จำนวนมากมาก่อนซึ่งอาจจัดเก็บแบบไม่เป็นระบบ และไม่อยู่ในรูปลักษณะเดียวกัน นับเป็นข้อมูลแบบไม่มีโครงสร้าง อาจอยู่ในรูปแบบตัวอักษร ภาพ ตัวเลข ไฟล์เสียงและภาพเคลื่อนไหว เนื้อหาบนสื่อสังคมออนไลน์ คลิป สัญญาณภาพจากดาวเทียม ไฟล์นำเสนองาน ไฟล์พีดีเอฟ แบบสอบถามแบบปลายเปิด เว็บไซต์ ข้อมูลจากคอลเซ็นเตอร์และการบันทึกข้อมูลเสียง ข้อความอีเมล เป็นต้น เป็นข้อมูลที่ไม่สามารถบรรจุลงบนโปรแกรมเอกเซลในแถวและคอลัมน์ของดาต้าเบส เป็นข้อมูลไม่มีรูปแบบตายตัว

การไม่มีโครงสร้างของข้อมูลทำให้ยากต่อการค้นหา จัดการและวิเคราะห์นี่จึงเป็นเหตุว่าทำไมองค์กรหรือหน่วยงานต้องจัดการกับข้อมูล โดยเฉพาะข้อมูลไม่มีโครงสร้าง ใดๆ ก็ดี คุณค่าของข้อมูลแบบนี้คือสามารถเข้าถึงได้ง่ายและกระจัดกระจายอยู่ทั้งในและนอกองค์กร

3. ข้อมูลกึ่งมีโครงสร้าง (Semi-Structured Data) เป็นข้อมูลผสมเป็นกึ่งกลางของทั้งสองแบบ ระหว่างข้อมูลมีโครงสร้างกับข้อมูลแบบไม่มีโครงสร้าง ชนิดของข้อมูลผสมสามารถบอกได้ว่ามีลักษณะอย่างไรบ้าง แต่บางครั้งก็ไม่มีลักษณะโครงสร้างและยืดหยุ่น ดังนั้น ข้อมูลแบบนี้จึงมีทั้งที่จัดระบบ แยกชนิดและมีขนาดใหญ่ เพื่อให้การจัดระบบง่ายขึ้น แต่ยังคงความยืดหยุ่นในข้อมูล ตัวอย่างเช่น เนื้อหาในอีเมล มีทั้งข้อมูลแบบไม่มีโครงสร้าง แต่ก็มีโครงสร้าง เช่น ชื่ออีเมลแอดเดรสของผู้ส่งและผู้รับ เวลาส่ง เป็นต้น หรือตัวอย่างภาพดิจิทัล โดยตัวภาพเป็นข้อมูลที่ไม่มีโครงสร้าง แต่ถ้าภาพถ่ายด้วยสมาร์ทโฟน

จะปรากฏวันที่ เวลา สถานที่ถ่าย และยังระบุเจ้าของสมาร์ทโฟนได้ด้วย การเก็บภาพในคลังยังติดป้าย (Tag) แบ่งโครงสร้างประเภทภาพเป็นหัวข้อภาพบุคคล แหล่งท่องเที่ยว ทิวทัศน์ สัตว์เลี้ยง เป็นต้น ยังมีข้อมูลอีกหลายอย่างที่ระบุว่าเป็นข้อมูลแบบไม่มีโครงสร้าง แต่ที่จริงแล้วอาจเป็นข้อมูลผสมก็ได้ เช่น สเตตัสใน Social Media เป็นข้อมูลที่ไม่มีโครงสร้าง แต่ในกรณีที่มี Hashtag (#) เข้ามาช่วยในการจัดหมวดหมู่ จะทำให้ข้อมูลมีความเป็นระเบียบขึ้นมาเล็กน้อย



ภาพที่ 2.3 เปรียบความแตกต่างระหว่างข้อมูลแบบมีโครงสร้าง กึ่งโครงสร้าง และไม่มีโครงสร้าง

ที่มา: ปรับปรุงจาก <https://www.everydaymarketing.co/knowledge/data-101-how-different-structured-unstructured-and-semi-structured-data/>

ความแตกต่างของข้อมูลทั้ง 3 แบบ หากอธิบายเปรียบเทียบกับกรสัมภาษณ์สมัครงาน จะพบว่า มีคำถาม 3 ส่วน ที่ควรสัมภาษณ์ มีทั้งข้อมูลแบบมีโครงสร้าง ข้อมูลแบบไม่มีโครงสร้าง และข้อมูลผสม โดยข้อมูลแบบมีโครงสร้างคือ คำถามที่ฝ่ายบุคคลตั้งเป็นประเด็นคำถามเพื่อสัมภาษณ์ผู้สมัครทุกคน ส่วนข้อมูลแบบไม่มีโครงสร้าง คือ คำถาม คำสั่ง คำขอร้อง ให้ผู้สมัครแสดงศักยภาพ เพื่อทดสอบความสามารถ ส่วนข้อมูลผสม มีองค์ประกอบทั้งสองส่วน มีทั้งเชิงปริมาณที่กำหนดหัวข้อ แต่ก็ให้อิสระผู้ตอบจะตอบในแต่ละหัวข้อ จึงทำให้ส่วนหลังเป็นข้อมูลแบบไม่มีโครงสร้าง ดังนั้น ข้อมูลแบบมีโครงสร้างจะง่ายต่อการจัดการ ค้นหา ติดตาม ส่วนข้อมูลแบบไม่มีโครงสร้างจะมีความซับซ้อนและมีปริมาณมาก ทำให้ประสิทธิภาพการจัดการระบบลดลง หรือทำให้หาความสัมพันธ์ของข้อมูลได้ยาก

ทั้งนี้ การใช้แหล่งที่มาข้อมูลของ บิ๊กดาต้า ทั้งข้อมูลแบบมีโครงสร้างและแบบไม่มีโครงสร้าง ขึ้นอยู่กับความต้องการขององค์กรว่าจะนำข้อมูลขนาดใหญ่มาใช้เพื่อตอบโจทย์อะไรบ้าง แหล่งที่มาข้อมูลจะเชื่อมโยงกับเป้าหมายที่ต้องการไปใช้งานมากน้อยเพียงใด แหล่งที่มาของ บิ๊กดาต้า สามารถแบ่งได้หลากหลาย ดังนี้ (อังธิตา ลิ้มปีฬมปาณี และคณะ, 2563)

1. แหล่งที่มาข้อมูลแบ่งตามผู้ดูแลข้อมูล

1.1 แหล่งที่มาข้อมูลจากภาครัฐ ขณะนี้ข้อมูลจำนวนมากที่สุดและขนาดใหญ่ที่สุดที่เป็น บิ๊กดาต้า อยู่ที่หน่วยงานภาครัฐ ถ้าหากหน่วยงานใดมีบิ๊กดาต้าอยู่ แล้วอาจขอข้อมูลทั้งหมดได้ แต่หากขอเปิดเผย ไม่ได้ต้องอาศัยพระราชบัญญัติข้อมูลข่าวสารของทางราชการ พ.ศ. 2540 เพื่อขอเปิดเผยข้อมูลข่าวสารที่ต้องการจากกระทรวง กรม กองต่าง ๆ ซึ่งเมื่อเดือนพฤษภาคม 2561 รัฐบาลพลเอกประยุทธ์ จันทร์โอชา ได้ประกาศให้ความสำคัญเรื่องบิ๊กดาต้า โดยให้ทุกกระทรวง กรม กอง จัดทำฐานข้อมูลให้ทันสมัย รองรับกฎหมายใหม่และนโยบายไทยแลนด์ 4.0 แหล่งที่มาข้อมูลภาครัฐจึงจัดทำบิ๊กดาต้า เช่น กระทรวงการคลัง กรมบัญชีกลาง กำลังรวบรวมข้อมูลให้ประชาชนเข้าใช้ สามารถติดตามการเบิกจ่ายงบประมาณ เช่น หน่วยงานใดซื้ออะไรบ้าง ใครประมูล ใครเป็นคู่ค้ากับภาครัฐ ติดตามการใช้เงินงบประมาณของภาครัฐ ได้ง่ายขึ้น มองเห็นภาพรวมของประเทศ เป็นต้น

แหล่งที่มาข้อมูลจากภาครัฐอีกส่วนหนึ่งคือรัฐวิสาหกิจ ได้นำระบบมาตรฐาน COST มาใช้ในการทำงานของรัฐวิสาหกิจทั้ง 56 แห่ง โดยต้องเปิดเผยข้อมูลให้ประชาชนทราบ ทั้งข้อมูลเบื้องต้นและข้อมูลเชิงลึก ประชาชนสามารถตรวจสอบและรับรู้ข้อมูลการทำงานของหน่วยงานได้ ประกอบกับนโยบายไทยแลนด์ 4.0 ทำให้หน่วยงานรัฐวิสาหกิจจัดทำบิ๊กดาต้าขององค์กร ดังนี้

สำนักงานรัฐบาลอิเล็กทรอนิกส์ (องค์การมหาชน) (สรอ.) (Electronic Government Agency: EGA) เป็นองค์การมหาชน ตามพระราชกฤษฎีกาการจัดตั้งสำนักงานรัฐบาลอิเล็กทรอนิกส์ พ.ศ. 2554 เป็นหน่วยงานกลางรวบรวมข้อมูลของภาครัฐ มีหน้าที่ให้บริการและบริหารจัดการโครงสร้างพื้นฐานสารสนเทศเกี่ยวกับรัฐบาลอิเล็กทรอนิกส์ อาทิ การให้บริการระบบคลาวด์ภาครัฐ (Government Cloud) การให้บริการระบบเครือข่ายสื่อสารข้อมูลเชื่อมโยงหน่วยงานภาครัฐ (Government Information Network: GIN) เป็นต้น

ปัจจุบันกำหนดให้สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน) (สพร.) (Digital Government Development Agency: DGA) ภายใต้การกำกับดูแลของนายกรัฐมนตรี สังกัดสำนักนายกรัฐมนตรี เป็นผู้กำกับดูแล เมื่อรัฐบาลมีคำสั่งให้ทุกหน่วยงานจัดทำบิ๊กดาต้าแล้ว สรอ. เป็นผู้รวบรวมเชื่อมโยงทุกหน่วยด้วยกัน (Central link) และให้เกิดความปลอดภัยในการเข้าถึงข้อมูลและการเชื่อมโยง

ขณะนี้ สำนักงานอิเล็กทรอนิกส์ภาครัฐ กระทรวงการคลัง กรมบัญชีกลาง ร่วมกันทำเว็บไซต์เพื่อตอบโจทย์ ‘ภาษีไปไหน’ ด้วยการใช้ บิ๊กดาต้าแสดงให้เห็นว่า งบประมาณแผ่นดินทั้งส่วนราชการและรัฐวิสาหกิจมีอะไรบ้างและใช้จ่ายอะไรบ้าง เบิกจ่ายงบเรื่องใดไปบ้าง ประมวลด้วยวิธีการอย่างไร

กรมพัฒนาธุรกิจการค้า กระทรวงพาณิชย์ ตัวอย่างบิ๊กดาต้าของกระทรวงพาณิชย์ที่มีมานาน แต่ภาครัฐต้องจ่ายเงินเพื่อซื้อข้อมูลของตัวเอง กรมพัฒนาธุรกิจการค้า (กรมส่งเสริมการค้าระหว่างประเทศ) นำข้อมูลกรมพัฒนาธุรกิจการค้าขึ้นบนเว็บไซต์ แต่เป็นเว็บไซต์ปิดให้บริการเฉพาะผู้เป็นสมาชิกประจำ ซึ่งส่วนใหญ่ทั้งภาครัฐและภาคเอกชนต้องใช้ข้อมูลดังกล่าวในเรื่องการส่งออกหรือเรื่องผู้ถือหุ้น แต่ผู้ริเริ่มพัฒนาและดูแลข้อมูลในเว็บไซด์นี้เป็นประเทศสิงคโปร์ ปัจจุบันธุรกิจออนไลน์ (Business Online: BOL) ทั้งข้อมูลด้านบัญชี ด้านภาษี ด้านการขาย นักข่าว ภาครัฐ ภาคเอกชน ต่างก็ต้องอาศัยเว็บไซด์นี้ค้นหาข้อมูลด้วยการซื้อข้อมูลในแต่ละครั้ง

1.2 แหล่งที่มาข้อมูลจากภาคเอกชน ภาคเอกชนจะจัดเก็บข้อมูลขององค์กรเพื่อใช้ในธุรกิจตัวเอง องค์กรเอกชนหรือองค์กรธุรกิจจัดเก็บแหล่งข้อมูลอย่างมีเป้าหมาย หากเปิดเผยได้ก็จะแถลงข่าวให้สาธารณชน บางองค์กรมีการจัดเก็บและประมวลผลข้อมูลเพื่อนำเสนอแก่สาธารณะโดยมีเป้าหมายเพื่อการตลาดและประชาสัมพันธ์ขององค์กรเช่นกัน อาทิ ศูนย์วิจัยธนาคารหลายแห่งที่มีการทำวิจัยการบริโภคทุกไตรมาส หรือสำรวจความสนใจผลิตภัณฑ์การเงินและการลงทุนทุกไตรมาสหรือบริษัทโซเชียลมีเดียที่ให้บริการวิเคราะห์ข้อมูลเกี่ยวกับประเด็นที่อยู่ในความสนใจของสังคม (trend) ที่เรียกว่า social listening

1.3 แหล่งที่มาข้อมูลจากภาคสังคมหรือภาคประชาชน แหล่งที่มาข้อมูลจากภาคสังคมหรือภาคประชาชนอาจยังไม่ค่อยพบ เนื่องจากการเก็บข้อมูลขนาดยักษ์ต้องใช้งบประมาณมหาศาล ยกเว้นภาคสังคมหรือภาคประชาชนที่มีองค์กรช่วยดำเนินการ อาจมีความทันสมัยในการวางระบบข้อมูล และข้อมูลเหล่านั้นย่อมเป็นแหล่งที่มาข้อมูลที่เป็นประโยชน์ต่อการใช้งาน เช่น ข้อมูลการบริจาคจากองค์กรสาธารณกุศล ข้อมูลการรณรงค์ต่างๆ ประเด็นใดได้รับความสนใจจากประชาชน ข้อมูลการค้นหาข้อมูลของประชาชนสูงสุดในรอบสัปดาห์ รอบไตรมาส ข้อมูลกิจกรรมตามเทศกาล อาทิ ลอยกระทงอิเล็กทรอนิกส์ ฐานเทศกาลคริสต์มาส

2. แหล่งที่มาข้อมูลแบ่งตามช่องทาง

2.1 แหล่งที่มาข้อมูลระบบออนไลน์ (Online) ได้แก่

1) เว็บไซต์ (Website) ใช้ดาวน์โหลดข้อมูลได้จากเว็บไซต์ต่างๆ ซึ่งส่วนใหญ่เป็นของภาครัฐ ส่วนภาคเอกชนไม่สามารถเข้าถึงได้ ยกเว้นภาคเอกชนจะเปิดเผยต่อสาธารณะ จึงจะสามารถนำมาใช้ได้

2) รายงานประจำปี (Annual Report) หน่วยงานต่างๆ โดยเฉพาะภาครัฐหรือภาคเอกชนที่เข้าตลาดหลักทรัพย์ ต้องจัดทำรายงานผลประกอบการประจำปี เพื่อแสดงผลงาน ผลผลิต งบประมาณ หรือผลประกอบการประจำปีเพื่อให้เป็นไปตามกฎหมายกำหนด

3) ข้อมูลจากการบริโภคสินค้าและบริการของผู้บริโภค (Customer Data) ข้อมูลของผู้บริโภคเป็นแหล่งข้อมูลที่สำคัญซึ่งจะมีทั้งข้อมูลที่มาจากรัฐหรือภาคเอกชน เช่น ข้อมูลผู้ใช้บริการรถไฟฟ้า ข้อมูลของผู้เปิดชมรายการโทรทัศน์ การเปิดรับฟังรายการวิทยุทางแอปพลิเคชัน ข้อมูลผู้ใช้บัตรเครดิตในแต่ละปี ข้อมูลผู้ใช้บริการธนาคาร ข้อมูลผู้ลงทุนในตลาดหลักทรัพย์ ข้อมูลผู้ใช้เงินอิเล็กทรอนิกส์ ข้อมูลผู้เข้าถึงเพจสินค้าหรือบริการ ผู้ใช้บริการขนส่งสินค้าและระบบไปรษณีย์ ผู้ซื้อสินค้าทางออนไลน์ ข้อมูลจากคิวอาร์โค้ด (QR code) เป็นต้น

4) ข้อมูลจากออนไลน์ (Crowd Source in Social Media) ส่วนใหญ่มักเป็นข้อมูลที่ใช้สื่อสังคมออนไลน์ ได้แก่ เฟซบุ๊ก ไลน์ ทวิตเตอร์ ยูทูป อินสตาแกรม กูเกิล บางครั้งการใช้ข้อมูลเหล่านี้ อาจต้องระดมความคิดเห็น สำรวจความเห็น จากการตั้งประเด็น หรือตั้งโจทย์ แล้วจึงนำข้อมูลคำตอบจากผู้ใช้สื่อสังคมออนไลน์มาประมวลผลอีกครั้ง

5) ระบบคลาวด์ (Cloud) แหล่งข้อมูลออนไลน์ระบบคลาวด์เป็นแหล่งข้อมูลต่อเนื่องจากการให้บริการเก็บข้อมูลขนาดใหญ่ มีหลายองค์กรที่ให้บริการเหล่านี้ เช่น กูเกิล แอปเปิ้ล เฟซบุ๊ก ไลน์ อินสตาแกรม เป็นต้น หรือผู้ให้บริการเทคโนโลยีโทรคมนาคม เช่น ดีแทค เอไอเอส ทู เป็นต้น

อย่างไรก็ดี ข้อมูลจากแอ่งข้อมูล (Data Lake) โดยตรง ซึ่งเป็นข้อมูลเชิงลึก (Insight Data) ถูกเก็บไว้หลายประเภท ผู้ดูแลแอปพลิเคชันหรือผู้ดูแลระบบคลาวด์นั้นเป็นผู้นำข้อมูลไปใช้ จึงยังไม่สามารถเข้าถึงโดยตรงยกเว้นต้องซื้อข้อมูลจากผู้ให้บริการ

2.2 แหล่งข้อมูลระบบออฟไลน์ (Data Collection Offline) ปัญหาของประเทศไทยคือการเก็บข้อมูลในรูปแบบที่สามารถนำไปใช้ในการวิเคราะห์ข้อมูลขนาดใหญ่ยังไม่มีระบบและไม่มีประสิทธิภาพมากพอ รูปแบบการเก็บข้อมูลที่เป็นดิจิทัลมีไม่มากนัก ทำให้มีแหล่งข้อมูลต่างๆ ที่ต้องอาศัยเวลาและกำลังคนไปเก็บรวบรวมมาที่เรียกว่าเป็นแหล่งข้อมูลระบบออฟไลน์ ซึ่งมีหลายประเภท ดังนี้

1) งานวิจัย ส่วนใหญ่รายงานผลวิจัยจะรวบรวมข้อมูลระดับหนึ่งแล้วและอาจมีข้อมูลต่อยอดจากพื้นฐานเดิมมากขึ้น เป็นแหล่งข้อมูลที่น่าเชื่อถือระดับหนึ่ง องค์กรสามารถใช้แหล่งข้อมูลนี้มาจัดเรียงข้อมูลขึ้นใหม่ เพื่อวิเคราะห์และประมวลผลเป็นบิกดาต้าต่อไป

2) องค์กรที่มีข้อมูลพื้นฐาน เช่น ภาครัฐ ทั้งกระทรวง กรม กอง เป็นต้น กระทรวงเกษตรมีข้อมูลเกี่ยวกับเกษตรกรทั่วประเทศ กระทรวงพาณิชย์มีข้อมูลเกี่ยวกับการจดทะเบียนบริษัท การยกเลิกกิจการการส่งออก-นำเข้าสินค้า ตลาดหลักทรัพย์มีข้อมูลเกี่ยวกับการซื้อขายหลักทรัพย์และบริษัทจดทะเบียนต่างๆ กระทรวงศึกษาธิการมีข้อมูลเกี่ยวกับนักเรียน ครู โรงเรียน โรงพยาบาลภาครัฐและภาคเอกชนมีข้อมูลเกี่ยวกับโรคระบาด อุบัติเหตุ เป็นต้น นอกจากนี้ องค์กรที่มีข้อมูลพื้นฐาน เช่น ภาคสังคม ทั้งองค์กรสาธารณกุศล องค์กรช่วยเหลือกิจการสาธารณะ ศูนย์วิจัยเพื่อสังคม เป็นต้น ก็เป็นแหล่งข้อมูลระบบปิดที่สำคัญได้

3) ลงพื้นที่จริง การดำเนินการเก็บข้อมูลด้วยตัวเองและรวบรวมอย่างเป็นระบบทุกครั้งที่ลงพื้นที่ สามารถนำมาเก็บเป็นฐานข้อมูลขนาดใหญ่ เมื่อจำนวนข้อมูลมากพอ หรือระยะเวลาานพอ จะสังเกตเห็นประเด็นที่น่าสนใจ การลงพื้นที่จริงๆ ยุคใหม่นี้ สามารถลงพื้นที่ทั้งในระบบออฟไลน์ และลงพื้นที่จริงในระบบออนไลน์

4) การสำรวจความคิดเห็น แหล่งข้อมูลอีกแหล่งหนึ่งคือ ผลจากการสำรวจความคิดเห็นของประชาชน เพราะความคิดของประชาชนต่อประเด็นทางสังคมสามารถเห็นแนวโน้มของกระแสสังคมในเรื่องใดเรื่องหนึ่ง และนำไปสู่การวางแผนการทำงานต่อไปได้ เทคโนโลยีสมัยใหม่สามารถทำให้แหล่งข้อมูลระบบออฟไลน์กลายเป็นแหล่งข้อมูลระบบออนไลน์ และประมวลผลเพื่อนำมาใช้ได้อย่างรวดเร็ว

5) ฐานข้อมูลสะสมข่าว กรณีเป็นแหล่งข้อมูลสำหรับสำนักข่าว เพราะข่าวเดิมที่เคยเสนอไปแล้ว เมื่อเก็บสะสมได้ระยะหนึ่งย่อมทำให้เกิดเป็นฐานข้อมูล แต่ยังไม่ใช้บิกดาต้าหากสำนักข่าวนำข่าวเดิมที่สะสมมาจัดระบบใหม่ ทำให้เห็นความสัมพันธ์ระหว่างข้อมูลบางอย่างหรือนำไปต่อยอดเป็นประเด็นข่าวใหม่ ฐานข้อมูลข่าวสะสมก็สามารถเป็นแหล่งข่าวระบบออฟไลน์ในบิกดาต้าได้เช่นกัน

อย่างไรก็ดี แหล่งที่มาข้อมูลที่เป็นแอ่งข้อมูลของบิกดาต้าส่วนใหญ่มักเชื่อมต่อ (Connect) เชื่อมโยงข้อมูลเข้าด้วยกันเป็นลักษณะแหล่งข้อมูลออนไลน์ เนื่องจากอุปกรณ์ดิจิทัลต่างๆ ออกแบบให้เก็บข้อมูล ทั้งแบบรายบุคคลและแบบโครงสร้างทั้งหมด การจัดระบบให้ข้อมูลและออกแบบเพื่อการใช้งาน เป็นเรื่องที่ต้องทำงานร่วมกับผู้เชี่ยวชาญ ได้แก่ วิศวกรผู้ออกแบบระบบ (Data Engineer) นักวิทยาศาสตร์

ออกแบบการวิเคราะห์ข้อมูล (Data Scientist) เพื่อให้การใช้งานบิกดาต้าตอบโจทย์ตอบสนองมาตรฐานได้ถูกต้องคุ้มค่ามากที่สุด

กิจกรรม 2.2

1. ข้อมูลขนาดใหญ่มีกี่ประเภท อะไรบ้าง
 2. แหล่งข้อมูลที่แบ่งตามช่องทางมีกี่ประเภท อะไรบ้าง
-

แนวตอบกิจกรรม 2.2

1. ข้อมูลขนาดใหญ่มี 3 ประเภท ได้แก่ ข้อมูลแบบมีโครงสร้าง ข้อมูลแบบไม่มีโครงสร้าง และข้อมูลแบบผสมคือทั้งมีโครงสร้างและไม่มีโครงสร้าง
 2. แหล่งข้อมูลที่แบ่งตามช่องทางมี 2 ประเภท ได้แก่ แหล่งข้อมูลในระบบออนไลน์ และในระบบออฟไลน์
-

บรรณานุกรม

- ระดับของการวิเคราะห์ข้อมูล (*Data Analytics*). (2563). สืบค้นจาก <https://datacubeth.ai/level-of-data-analytics/>
- อังธิตา ลิ้มปัทมปาณี, บุญลาภ ภูสุวรรณ และกมล ชวาลวิทย์. (2563). มหิมามหาข้อมูล (Big Data). ใน *เอกสารการสอนชุดวิชาการรายงานข่าวและการบรรณาธิกร* (หน่วยที่ 3). นนทบุรี: มหาวิทยาลัยสุโขทัย-ธรรมมาธิราช
- Big Data Analytics* คืออะไร?. (2563). สืบค้นจาก <https://www.ko.in.th/big-data-analytics-คืออะไร/>
- Gartner Says Advanced Analytics Is a Top Business Priority*. (2014). Retrieved from <https://www.gartner.com/en/newsroom/press-releases/2014-10-21-gartner-says-advanced-analytics-is-a-top-business-priority>
- Marr, Bernard. (2019). *What the Differences between Structured, Semi-Structured and Unstructured Data*. Retrieved from <https://www.bernardmarr.com/default.asp?contentID=1966>
- Nattapon Muangtum. (2021). *Data 101 มา รู้จักความต่าง Structured, Unstructured และ Semi-structured data*. สืบค้นจาก <https://www.everydaymarketing.co/knowledge/data-101-how-different-structured-unstructured-and-semi-structured-data/>

សំរេង

សំរេង

សំរេង

សំរេង

សំរេង

សំរេង

សំរេង

សំរេង

ตอนที่ 3

การวิเคราะห์ข้อมูลขนาดใหญ่

ผู้ช่วยศาสตราจารย์ ดร.ณัฐสุพงศ์ สุขโสภา



ชื่อ
วุฒิ

ผู้ช่วยศาสตราจารย์ ดร.ณัฐสุพงศ์ สุขโสภา
ว.บ. ว.ม. (สื่อสารมวลชน) มหาวิทยาลัยธรรมศาสตร์

ตำแหน่ง

นศ.ด. (นิเทศศาสตร์) จุฬาลงกรณ์มหาวิทยาลัย
ผู้ช่วยศาสตราจารย์ประจำสาขาวิชานิเทศศาสตร์
มหาวิทยาลัยสุโขทัยธรรมมาธิราช

หน่วยที่ปรับปรุง

ตอนที่ 3

ตอนที่ 3

การวิเคราะห์ข้อมูลขนาดใหญ่

โปรดอ่านหัวเรื่อง แนวคิด และวัตถุประสงค์ของตอนที่ 3 แล้วจึงศึกษารายละเอียดต่อไป

หัวเรื่อง

- เรื่องที่ 3.1 ความหมายและความสำคัญของการวิเคราะห์ข้อมูลขนาดใหญ่
- เรื่องที่ 3.2 ประเภทของการวิเคราะห์ข้อมูลขนาดใหญ่
- เรื่องที่ 3.3 แนวทางและขั้นตอนการวิเคราะห์ข้อมูลขนาดใหญ่
- เรื่องที่ 3.4 เทคโนโลยีและเครื่องมือในการวิเคราะห์ข้อมูลขนาดใหญ่
- เรื่องที่ 3.5 การนำข้อมูลขนาดใหญ่ไปใช้ประโยชน์

แนวคิด

1. การวิเคราะห์ข้อมูลขนาดใหญ่หรือการวิเคราะห์ข้อมูลบิกดาต้า เป็นการนำข้อมูลที่มีอยู่เป็นจำนวนมากมาทำการวิเคราะห์ เพื่อนำมาหารูปแบบของข้อมูลที่ไม่สามารถมองเห็นได้ด้วยตาเปล่า ผลจากการวิเคราะห์บิกดาต้า ทำให้มีข้อมูลที่เป็นข้อเท็จจริงที่น่าเชื่อถือซึ่งผ่านการวิเคราะห์อย่างเป็นระบบ มีความสำคัญและจำเป็นต่อการนำไปใช้งานทั้งในภาครัฐและเอกชนในหลายด้าน
2. ประเภทของการวิเคราะห์ข้อมูลขนาดใหญ่มี 4 ประเภท ได้แก่ การวิเคราะห์เชิงพรรณนาการวิเคราะห์เชิงวินิจฉัย การวิเคราะห์เชิงคาดการณ์ล่วงหน้า และการวิเคราะห์เชิงกำหนดแนวทาง
3. แนวทางและขั้นตอนการวิเคราะห์ข้อมูลขนาดใหญ่หรือบิกดาต้าประกอบด้วยขั้นตอนสำคัญ 5 ขั้นตอน คือ การกำหนดเป้าหมาย/วัตถุประสงค์ในการใช้ข้อมูล การทำความเข้าใจข้อมูล การเก็บข้อมูล การสะสมข้อมูล และการทำความสะอาดข้อมูล การใช้เทคโนโลยีเพื่อวิเคราะห์ข้อมูล และการประมวลผล และการนำเสนอข้อมูล
4. เทคโนโลยีและเครื่องมือที่นำมาใช้วิเคราะห์ข้อมูลขนาดใหญ่แบ่งออกได้เป็นเครื่องมือขั้นพื้นฐาน และเครื่องมือขั้นสูง ซึ่งเทคโนโลยีและเครื่องมือที่นำมาใช้กับบิกดาต้ามีหลายอยู่ระบบ ในที่นี้ จะยกตัวอย่างโปรแกรมที่เปิดให้ใช้งานฟรี ได้แก่ ระบบฮาดูป ระบบคลาวด์รีจา ระบบฮอว์ตันเวิร์ก ระบบแมพพาร์ ระบบแอมะซอน ระบบไมโครซอฟท์ และระบบอัลติสเกล
5. การนำข้อมูลขนาดใหญ่ไปใช้ประโยชน์สามารถนำไปใช้ได้หลายด้าน ได้แก่ การศึกษา การเกษตร อุตสาหกรรม คมนาคมและการจัดการโลจิสติกส์ สุขภาพและการแพทย์ และการตลาดและการขาย

วัตถุประสงค์

เมื่อศึกษาตอนที่ 3 จบแล้ว นักศึกษาสามารถ

1. อธิบายความหมายและความสำคัญของการวิเคราะห์ข้อมูลขนาดใหญ่ได้
2. อธิบายประเภทของการวิเคราะห์ข้อมูลขนาดใหญ่ได้
3. ระบุถึงแนวทางและขั้นตอนการวิเคราะห์ข้อมูลขนาดใหญ่ได้
4. อธิบายการนำเทคโนโลยีและเครื่องมือไปใช้ในการวิเคราะห์ข้อมูลขนาดใหญ่ได้
5. อธิบายการนำข้อมูลขนาดใหญ่ไปใช้ประโยชน์ในภาครัฐและเอกชนได้

เรื่องที่ 3.1

ความหมายและความสำคัญของการวิเคราะห์ข้อมูลขนาดใหญ่

จากเนื้อหาในตอนที่ผ่านมา นักศึกษาได้มีความรู้เกี่ยวกับความหมาย กระบวนการจัดการที่เกี่ยวข้องกับข้อมูลขนาดใหญ่หรือบิ๊กดาต้า (Big Data) ไปพอสมควรแล้ว สำหรับเนื้อหาในตอนนี้จะมุ่งเน้นในขั้นตอนสำคัญอีกขั้นตอนหนึ่งในกระบวนการจัดการบิ๊กดาต้า ซึ่งก็คือ ขั้นตอนการวิเคราะห์ข้อมูล

บิ๊กดาต้า (Big Data) หมายถึง ชุดข้อมูลขนาดใหญ่จำนวนมาก โดยที่ข้อมูลมีความหลากหลาย ทั้งแบบมีโครงสร้าง แบบไม่มีโครงสร้าง แบบผสม รวมถึงแหล่งที่มาหรือช่องทางที่มาของข้อมูลก็มีได้มากมาย อาทิ แหล่งข้อมูลออนไลน์/ออฟไลน์ อินเทอร์เน็ต เว็บไซต์ อุปกรณ์ดิจิทัลต่างๆ อินเทอร์เน็ตในทุกสิ่ง (Internet of Things: IoT) เป็นต้น ซึ่งต้องมีเป้าหมายในการจัดเก็บ มีการจัดเก็บและเรียบเรียงข้อมูลที่ได้มาอย่างเป็นระบบ ต้องมีการถอดรหัสวิเคราะห์ข้อมูล และรู้วัตถุประสงค์ที่ชัดเจนในการนำไปใช้ประโยชน์

1. ความหมายของการวิเคราะห์ข้อมูลขนาดใหญ่

การวิเคราะห์ข้อมูล (Data Analytic) ขนาดใหญ่ คือ การนำข้อมูลที่มีอยู่เป็นจำนวนมากมามหาศาลมาทำการวิเคราะห์ ซึ่งการวิเคราะห์ข้อมูลบิ๊กดาต้า (Big Data) อาศัยหลักการพื้นฐานบางอย่างเพื่อพัฒนาเป็นเทคนิคในการดึงข้อมูลสำคัญออกจากชุดฐานข้อมูลขนาดใหญ่ที่เต็มไปด้วยความซับซ้อนของข้อมูลที่หลากหลาย เพื่อนำมาหารูปแบบ (pattern) ของข้อมูลที่ไม่สามารถมองเห็นได้ด้วยตาเปล่า ไม่ว่าจะเป็นการหารูปแบบความสัมพันธ์ที่ซ่อนอยู่ การหาแนวโน้มของสังคม การหาแนวโน้มหรือกระแสความขึ้นความชอบของลูกค้า ผู้บริโภค และผู้รับบริการต่างๆ และข้อมูลอื่นๆ ที่เป็นประโยชน์กับองค์กรหน่วยงานทั้งภาครัฐและภาคเอกชน ไม่ว่าจะเป็นบริษัท ผู้ประกอบการทางธุรกิจ เจ้าของสินค้าและบริการต่างๆ ซึ่งผลจากการวิเคราะห์บิ๊กดาต้า ทำให้มีข้อมูลที่เป็นข้อเท็จจริงที่น่าเชื่อถือซึ่งผ่านการวิเคราะห์อย่างเป็นระบบ เพื่อวัตถุประสงค์บางประการ เช่น เพื่อใช้ประกอบการตัดสินใจ เพื่อนำไปใช้ประโยชน์ของภาครัฐและเอกชนในหลายๆ ด้าน โดยระดับของการวิเคราะห์ก็เป็นได้หลากหลายขึ้นอยู่กับวัตถุประสงค์การนำไปใช้งานหรือการนำไปใช้ประโยชน์

ด้วยเหตุนี้เอง ในการวิเคราะห์บิ๊กดาต้า จำเป็นต้องมีการนำเทคโนโลยีและเครื่องมือสนับสนุนมาช่วยในการวิเคราะห์และประมวลผลข้อมูลขนาดใหญ่เพื่อให้ออกมาอย่างรวดเร็ว ถูกต้อง และแม่นยำ ตรงตามโจทย์/วัตถุประสงค์ที่ได้กำหนดไว้

ทั้งนี้การวิเคราะห์ข้อมูลขั้นพื้นฐานอาจไม่จำเป็นต้องใช้บุคลากรผู้ที่มีความเชี่ยวชาญโดยตรงในการจัดการ โดยปกติจะสามารถใช้ซอฟต์แวร์แบบชุดเดียวได้ หากแต่ในกรณีของการวิเคราะห์ข้อมูลขนาดใหญ่จำเป็นต้องมีผู้เชี่ยวชาญในการวิเคราะห์ข้อมูลโดยใช้ซอฟต์แวร์โดยเฉพาะ

บิกดาต้าเป็นประเด็นที่ต้องให้ความสำคัญ เพราะเป็นแหล่งที่มาของการทำความเข้าใจข้อมูลในเชิงลึก ในอันที่จะตอบสนองการทำงานและการแก้ไขปัญหาต่างๆ ได้มากขึ้น ซึ่งเป็นประโยชน์อย่างยิ่งในการตัดสินใจเชิงบริหารและการวางนโยบายต่างๆ และการดำเนินธุรกิจขององค์กร ดังนั้นการใช้ประโยชน์จากบิกดาต้าจำเป็นอย่างยิ่งที่จะต้องเข้าใจว่าการวิเคราะห์ประเภทใดที่จำเป็นต้องนำมาใช้เพื่อการใด เพื่อให้ได้ประโยชน์สูงสุดจากข้อมูลที่มีอยู่ รวมถึงการเปลี่ยนข้อมูลขนาดใหญ่ที่ไม่ได้นำมาใช้ประโยชน์ให้เป็นการใช้ได้เปรียบเชิงกลยุทธ์

2. ความสำคัญของการวิเคราะห์ข้อมูลขนาดใหญ่

การวิเคราะห์ข้อมูลขนาดใหญ่หรือการวิเคราะห์ข้อมูลบิกดาต้า (Big Data) มีความสำคัญและจำเป็นต่อการทำงานทั้งในภาครัฐและเอกชน ดังรายละเอียดต่อไปนี้

2.1 ภาครัฐกับบิกดาต้า ข้อมูลบิกดาต้าสามารถนำมาใช้กำหนดทิศทางการทำงานของภาครัฐ ทำให้มองเห็นแนวโน้มสังคมชัดเจน จึงนำมาใช้วางยุทธศาสตร์ ตอบโจทย์หรือดำเนินการเกี่ยวกับประชาชน เป็นเครื่องมือทำให้การทำงานของภาครัฐง่ายขึ้น ตั้งแต่การวางแผนกำหนดทิศทางประเทศ ซึ่งเป็นเรื่องใหญ่ สามารถเห็นข้อมูลภาพรวม (perspective) ตั้งแต่ระดับกระทรวง ทบวง กรม จังหวัด อำเภอ ตำบล ไปจนถึงหมู่บ้าน ทั้งเด็ก ผู้ใหญ่ ผู้สูงอายุ ผู้คนในวัยทำงาน ผู้ป่วยติดเตียง ผู้ป่วยเรื้อรังทั้งหมด ไม่ว่าจะอยู่ในภาคเหนือ ภาคตะวันออกเฉียงเหนือ ภาคกลาง ภาคตะวันตก หรือภาคใต้ ระบุได้ทุกภูมิภาคทุกจังหวัด รัฐมีข้อมูลแล้วนำมาจัดการทำให้คุณภาพชีวิตของประชาชนดีขึ้น จะใช้เพื่อพัฒนาอุตสาหกรรม หรือพัฒนาเศรษฐกิจของประเทศให้ตรงจุด ตรงความต้องการ อีกทั้งยังจัดสรรงบประมาณ เพื่อให้เกิดการพัฒนา ปรับปรุงหรือแก้ไขได้อย่างมีประสิทธิภาพ

ยิ่งกว่านั้น เมื่อการพัฒนาของรัฐบรรลุถึงเป้าหมายที่ตอบโจทย์ชัดเจน ย่อมทำให้การรับเรื่องร้องเรียน ลดลง ไม่ต้องรอรัฐบาลมาแก้ไขปัญหา เช่น สาธารณูปโภคขั้นพื้นฐานต้องสร้างที่ใดบ้าง ระยะเวลาใดต้องซ่อมแซม มีการปรับปรุงหรือดำเนินการไปแล้วถึงขั้นตอนใด เกิดประสิทธิภาพและประสิทธิผล ลดค่าใช้จ่าย เพิ่มรายได้ ลดการคอร์รัปชัน สังคมรับรู้การทำงานของรัฐและรัฐบาล เกิดความโปร่งใส เห็นขั้นตอนดำเนินงานชัดเจน ตรวจสอบได้ตลอดกระบวนการ (อังชิตา ลิมปีทมปาณี และคณะ, 2563)

บิกดาต้าจึงมีความสำคัญอย่างยิ่งยวดต่อการทำงานในภาครัฐและรัฐบาล ทำให้มองเห็นข้อมูลทั้งหมดทั้งประเทศ ขึ้นอยู่กับว่าจะตัดสินใจเลือกอย่างไร เมื่อมีข้อมูลจำนวนมากพร้อมใช้งานวางอยู่ตรงหน้า ซึ่งภาครัฐต้องรู้ว่าจะใช้บิกดาต้าไปเพื่ออะไรและจะใช้อย่างไร

2.2 ภาคเอกชนกับบิกดาต้า การวิเคราะห์ข้อมูลปริมาณมหาศาลแสดงให้เห็นประโยชน์ต่างๆ ก็คือ ความรวดเร็วและประสิทธิภาพ ขณะที่หลายปีก่อนหน้านั้น เอกชนหลายธุรกิจต้องรวบรวมข้อมูลดำเนินการวิเคราะห์ และค้นหาข้อมูลที่สามารถนำมาใช้ประกอบการตัดสินใจในอนาคตได้ ส่วนในทุกวันนี้ ธุรกิจสามารถระบุข้อมูลเชิงลึกสำหรับการกระทำที่ต้องการการตัดสินใจอย่างทันทั่วทั้งที่ ความสามารถในการทำงานได้เร็วขึ้นและยังคงไว้ซึ่งความคล่องตัว ช่วยให้องค์กรธุรกิจสร้างความได้เปรียบด้านการแข่งขันอย่างที่ไม่เคยมีมาก่อน

ภาคเอกชนหรือภาคธุรกิจดำเนินการพัฒนา สร้างระบบ จัดเก็บและใช้บิกดาต้า ไว้รองรับการสร้างธุรกิจใหม่ๆ เข้าใจและตอบสนองความต้องการของผู้บริโภค ไม่ว่าจะเป็นการวิเคราะห์พฤติกรรมผู้บริโภค รู้ว่าผู้บริโภคชอบสินค้าอะไร ผู้บริโภคกำลังค้นหาสินค้าอะไร หรือวิเคราะห์ความต้องการของตลาดในอนาคต เช่น บริเวณใดจะเป็นทำเลทองของธุรกิจ รวมทั้งเป็นการต่อยอดวางแผนเพื่อโอกาสทางธุรกิจได้อย่างมีประสิทธิภาพ

การเก็บสะสมข้อมูลภาคธุรกิจของแต่ละองค์กรหรือบริษัทธุรกิจส่วนใหญ่เป็นการเก็บโดยใช้เทคโนโลยีสารสนเทศที่พัฒนาขึ้นเอง หรืออาจใช้สื่อสังคมออนไลน์อย่างเฟซบุ๊ก ไลน์ อินสตาแกรม ยูทูป แล้วนำข้อมูลมาใช้ในการบริหารจัดการและคาดการณ์แนวโน้ม คาดการณ์ความเสี่ยงที่จะเกิดขึ้นกับองค์กร/บริษัท

ทั้งนี้ พฤติกรรมการใช้งานสื่อสังคมออนไลน์ เช่น เฟซบุ๊ก ไลน์ ทวิตเตอร์ อินสตาแกรม การใช้เสิร์ชเอนจินกูเกิล (Google) เป็นต้น การใช้งานต่างๆ เหล่านี้จะถูกบันทึกไว้ ดังที่เรียกกันว่ารอยเท้าบนโลกดิจิทัล (Digital Footprint) ซึ่งสามารถติดตามร่องรอยการใช้งานในออนไลน์ของผู้บริโภคได้หมดและสามารถนำข้อมูลมาวิเคราะห์ กำหนดวิธี เวลา ช่องทางการเข้าถึงผู้บริโภคอย่างใกล้ชิด ทำให้การตลาดยุคใหม่ตอบโจทย์ผู้บริโภค และคาดการณ์แนวโน้มพฤติกรรมผู้บริโภคได้แม่นยำยิ่งขึ้น

ดังนั้น บิกดาต้าจึงมีความสำคัญและจำเป็นต่อการทำงานทั้งในภาครัฐและเอกชน/ธุรกิจต่างๆ ซึ่งการนำข้อมูลขนาดใหญ่หรือบิกดาต้าไปใช้ประโยชน์นั้นจะกล่าวถึงโดยละเอียดอีกครั้งในเรื่องที่ 3.5

กิจกรรม 3.1

การวิเคราะห์ข้อมูลขนาดใหญ่หรือบิกดาต้ามีความหมายอย่างไร

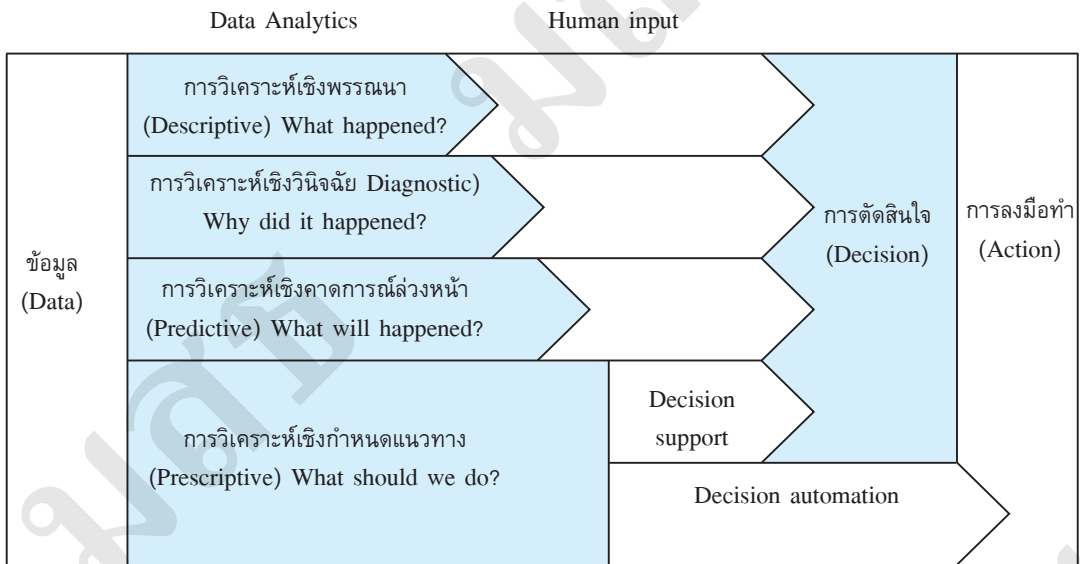
แนวตอบกิจกรรม 3.1

การวิเคราะห์ข้อมูลขนาดใหญ่หรือการวิเคราะห์ข้อมูลบิกดาต้า เป็นการนำข้อมูลที่มีอยู่เป็นจำนวนมากมาทำการวิเคราะห์ โดยอาศัยหลักการพื้นฐานบางอย่างในการดึงข้อมูลสำคัญออกจากชุดฐานข้อมูลขนาดใหญ่ที่เต็มไปด้วยความซับซ้อนของข้อมูลที่หลากหลาย เพื่อนำมาหารูปแบบของข้อมูลที่ไม่สามารถมองเห็นได้ด้วยตาเปล่า ผลจากการวิเคราะห์บิกดาต้า ทำให้มีข้อมูลที่เป็นข้อเท็จจริงที่น่าเชื่อถือซึ่งผ่านการวิเคราะห์อย่างเป็นระบบ เพื่อวัตถุประสงค์บางประการ เช่น เพื่อใช้ประกอบการตัดสินใจ เพื่อนำไปใช้ประโยชน์ของภาครัฐและเอกชนในหลายๆ ด้าน โดยระดับของการวิเคราะห์ก็เป็นได้หลากหลายขึ้นอยู่กับวัตถุประสงค์การนำไปใช้งานหรือการนำไปใช้ประโยชน์

เรื่องที่ 3.2

ประเภทการวิเคราะห์ข้อมูลขนาดใหญ่

การดำเนินการวิเคราะห์ข้อมูลมี 4 ประเภทใหญ่ๆ (Gartner, 2014) ได้แก่ การวิเคราะห์เชิงพรรณนา การวิเคราะห์เชิงวินิจฉัย การวิเคราะห์เชิงคาดการณ์ล่วงหน้า และการวิเคราะห์เชิงกำหนดแนวทาง ดังรายละเอียดต่อไปนี้



ภาพที่ 3.1 ลักษณะของการวิเคราะห์ข้อมูลขนาดใหญ่ 4 ประเภท

ที่มา: Gartner. (2014).

1. การวิเคราะห์เชิงพรรณนา (Descriptive analytics) เป็นการวิเคราะห์ข้อมูลขั้นพื้นฐานที่สุด และเป็นแกนหลักสำหรับการวิเคราะห์ในขั้นอื่นๆ ต่อไป การวิเคราะห์เชิงพรรณนานั้นใช้เพื่อทำความเข้าใจภาพรวมของกระบวนการหรือสถานการณ์ที่เกิดขึ้นขององค์กร จากมุมมองที่หลากหลายในระยะสั้น เพื่อตอบคำถามว่า เกิดอะไรขึ้น? (what happened?) สิ่งที่เกิดขึ้นนั้นเหมาะสมหรือดีสำหรับองค์กรภายในช่วงเวลานั้นๆ หรือไม่ เนื่องจากการวิเคราะห์เชิงพรรณนานั้นเป็นการวิเคราะห์ขั้นพื้นฐาน ดังนั้นจึงมีการใช้ประเภทนี้ในองค์กรธุรกิจและภาครัฐ เช่น การตลาด การค้าขายสินค้าออนไลน์ ไปจนถึงการธนาคาร และการดูแลสุขภาพ เป็นต้น

การวิเคราะห์ลักษณะนี้ช่วยในการตัดสินใจได้ค่อนข้างน้อยจึงจำเป็นต้องมีผู้เชี่ยวชาญ (หรือผู้บริหาร) เข้ามาเกี่ยวข้องในการตัดสินใจดังแสดงในส่วนที่เป็น Human Input (ในภาพที่ 3.1) ตัวอย่าง เช่น หากฝ่ายผลิตต้องการจะกำหนดจำนวนการผลิตสินค้าชนิดนี้จะต้องผลิตเท่าไร การวิเคราะห์แบบนี้จะให้ข้อมูลมาว่าที่ผ่านมามีการขายได้เท่าไร ผู้บริหารจะต้องมาคิดเพื่อตัดสินใจอีกครั้ง

สำหรับเครื่องมือวิเคราะห์เชิงพรรณนาที่เป็นที่คุ้นเคย คือ Google Analytics ที่ใช้อธิบายพฤติกรรมผู้ใช้งานออนไลน์ เช่น ในช่วงเวลาหนึ่งเมื่อมีการทำกิจกรรมทางการตลาดหรือทำแคมเปญใดๆ บนเว็บไซต์ แล้วการวิเคราะห์เชิงพรรณนาของ Google จะแสดงสถิติการโต้ตอบกับเนื้อหาเหล่านั้น ได้แก่ ใครเข้ามาบ้าง (Who? ซึ่งก็คือข้อมูล ID ผู้ใช้) สถานการณ์การเข้าใช้ (แหล่งที่มา - เข้าโดยตรงหรือโดยอ้อมจากแหล่งอื่น เช่น เฟซบุ๊ก ทวิตเตอร์) เข้ามาเมื่อใด (วันที่) ใช้เวลากับเว็บนานเพียงใด และใช้เวลาในส่วนเนื้อหาใดมาก

2. การวิเคราะห์เชิงวินิจฉัย (Diagnostic analytics) วัตถุประสงค์ของการวิเคราะห์การวินิจฉัย คือ การทำความเข้าใจว่า ทำไมบางสิ่งเกิดขึ้น (why did it happened?) สิ่งที่ทำให้เกิดเหตุการณ์นี้ การวิเคราะห์เชิงวินิจฉัยเป็นการตรวจสอบที่มุ่งศึกษาเหตุและผลกระทบและเพื่อปรับแผนงานที่เหมาะสมกับสถานการณ์ นอกไปจากนั้นยังหมายถึง

- การตรวจจับความผิดปกติ ความผิดปกติคือสิ่งใดก็ตามที่ทำให้เกิดคำถามเกี่ยวกับการวิเคราะห์ ไม่ว่าจะอะไรก็ตามที่ไม่เข้ากับบรรทัดฐาน อาจเป็นกิจกรรมที่ไม่คาดคิด เช่น อัตราการสมัครสมาชิกของหน้าสื่อออนไลน์ขององค์กรลดลงอย่างฉับพลัน

- การตรวจสอบความผิดปกติ ในการทำบางสิ่งของลูกค้าหรือผู้บริโภคหรือสถานการณ์ที่เกิดขึ้นองค์กรต้องเข้าใจว่ามันเกิดขึ้นได้อย่างไร กระบวนการนี้รวมถึงการระบุแหล่งที่มาของความผิดปกติ และการค้นหารูปแบบความผิดปกติในแหล่งข้อมูลต่างๆ

- การกำหนดความสัมพันธ์เชิงสาเหตุ หลังจากเหตุการณ์ที่ทำให้เกิดความผิดปกติ การวิเคราะห์แบบนี้จะระบุที่มาและจะนำไปสู่การที่ต้องเชื่อมต่อจุดต่างๆ เพื่อให้เห็นภาพรวมของสิ่งผิดปกติเหล่านั้น

ตัวอย่างเช่น ยอดขายของสินค้าในรอบเดือนที่ผ่านมาในแต่ละสาขา ซึ่งอาจจะทำให้เราเห็นฤดูกาล (seasonal) ความนิยมของสินค้าในแต่ละช่วงเวลาได้ จึงทำให้การเข้ามามีส่วนร่วมของผู้บริหารในการตัดสินใจน้อยลงเมื่อเทียบกับ Descriptive Analytics (พื้นที่ human input จะมีขนาดสั้นลง)

3. การวิเคราะห์เชิงคาดการณ์ล่วงหน้า (Predictive analytics) การวิเคราะห์เชิงคาดการณ์ได้รับการออกแบบมาเพื่อคาดการณ์สิ่งที่อนาคตอาจเกิดขึ้นได้ในระดับหนึ่ง (what will happened?) แสดงถึงผลลัพธ์ที่เป็นไปได้อย่างหลากหลาย โดยการคาดการณ์สถานการณ์ในลักษณะต่างๆ ในเชิงธุรกิจ จัดเป็นการใช้ข้อมูลขนาดใหญ่เพื่อดำเนินการเชิงรุก ดังนั้น การวิเคราะห์เชิงคาดการณ์ล่วงหน้าช่วยให้ผู้บริหารคาดการณ์แนวทางที่เป็นไปได้เพื่อนำไปสู่การตัดสินใจทางธุรกิจในมิติต่างๆ การวิเคราะห์เชิงทำนายเช่นนี้ต้องใช้ข้อมูลขนาดใหญ่ที่มีจากแหล่งข้อมูลที่เกี่ยวข้องทั้งหมด ซึ่งอาจเป็นแหล่งข้อมูลเดียวหรือผสมผสานระหว่างระบบทรัพยากรบุคคล ระบบการผลิต ระบบคลังสินค้า ระบบบริการลูกค้า ซึ่งนอกจากจะใช้ข้อมูลปัจจุบันจำนวนมากแล้วยังสามารถใช้ข้อมูลในอดีตมาคาดการณ์สิ่งที่เกิดขึ้นในอนาคต

ต่อไป เช่น ใช้ข้อมูลยอดขายแต่ละสาขารายเดือนมาสร้างแบบจำลอง (model) เพื่อคาดการณ์ว่าในอีก 1 เดือนถัดไปจะต้องผลิตสินค้าไปวางจำหน่ายในสาขาใดเป็นจำนวนเท่าไรบ้าง เนื่องจากวิธีการนี้มีการคาดการณ์มาให้จากระบบแล้ว คนจึงมีส่วนเกี่ยวข้องน้อยลงเมื่อเทียบกับ 2 วิธีการก่อนหน้านี้ (พื้นที่ Human input จะมีขนาดสั้นลง)

4. การวิเคราะห์เชิงกำหนดแนวทาง (Prescriptive analytics) เป็นขั้นตอนที่ช่วยคาดการณ์ว่าจะมีอะไรเกิดขึ้นต่อไป แล้วแนะนำแนวทางที่เหมาะสมให้อีกด้วย (what should we do?) ถือว่าเป็นขั้นสูงสุดของการวิเคราะห์ข้อมูล สำหรับวิธีการนี้จะเห็นว่าไม่ต้องมีคนเข้ามาเกี่ยวข้อง (ไม่มีพื้นที่ Human input) จากตัวอย่างที่ผ่านมาก็อาจจะอธิบายได้ว่าระบบจะคาดการณ์จำนวนการผลิตสินค้าได้แม่นยำมากขึ้น และเมื่อเกิดเหตุการณ์สินค้าขาดสต็อกก็จะแนะนำว่าควรที่จะผลิตเพิ่มเติมเท่าไรให้ได้ การวิเคราะห์ในระดับนี้มักนิยมใช้สำหรับการจัดการบางเรื่อง เช่น ปรับขั้นตอนหรือกระบวนการในองค์กรเพื่อพัฒนากระบวนการปฏิบัติงานให้เหมาะสม การจัดการวางแผนและปรับแต่งแคมเปญ การจัดการสินค้าคงคลังสินค้าและความสัมพันธ์กับลูกค้า เป็นต้น

ดังที่กล่าวมาแล้วทั้งหมดจะเห็นได้ว่าการวิเคราะห์ข้อมูลในระดับที่สูงขึ้นจะมีประโยชน์มากขึ้น แต่ทว่าก็มีความยากในการวิเคราะห์อยู่เช่นกัน ดังนั้นในการทำงานของบริษัทหรือองค์กรขนาดเล็กก็อาจใช้การวิเคราะห์ในระดับเชิงพรรณนาก็น่าจะเพียงพอแล้ว

กิจกรรม 3.2

ลักษณะของการวิเคราะห์ข้อมูลขนาดใหญ่หรือบิ๊กดาต้ามีกี่ประเภท อะไรบ้าง

แนวตอบกิจกรรม 3.2

ลักษณะการวิเคราะห์ข้อมูลขนาดใหญ่หรือบิ๊กดาต้ามี 4 ประเภท ได้แก่

1. การวิเคราะห์เชิงพรรณนา (Descriptive analysis)
2. การวิเคราะห์เชิงวินิจฉัย (Diagnostic analysis)
3. การวิเคราะห์เชิงคาดการณ์ (Predictive analysis)
4. การวิเคราะห์เชิงกำหนดแนวทาง (Prescriptive analysis)

เรื่องที่ 3.3

แนวทางและขั้นตอนการวิเคราะห์ข้อมูลขนาดใหญ่

การได้มาของข้อมูลขนาดใหญ่หรือบิ๊กดาต้า จะต้องมีกระบวนการและขั้นตอนก่อนที่จะได้มา กระบวนการวิเคราะห์บิ๊กดาต้า อาจมองได้ว่าเป็นเรื่องใหม่ ต้องใช้เทคโนโลยี การจัดการ คน ที่มามองมาช่วยสนับสนุน อีกทั้งมีความจำเป็นที่จะต้องเตรียมความพร้อมของบุคลากรด้วย โดยอาจทำการฝึกอบรมความรู้และทักษะเพิ่มเติมหรืออาจจำเป็นต้องจ้างบริษัทภายนอกที่มีความเชี่ยวชาญโดยเฉพาะ (outsourc) มาช่วยงาน

นอกจากนั้นพึงระมัดระวังการเก็บข้อมูลส่วนบุคคล ตามพระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562 (Personal Data Protection Act หรือ PDPA) ที่เป็นกฎหมายที่ออกมาคุ้มครองสิทธิเกี่ยวกับข้อมูลส่วนบุคคล ซึ่งทำให้สามารถระบุตัวตนบุคคลนั้นได้ ไม่ว่าจะเป็นทางตรงหรือทางอ้อม เพื่อที่จะการควบคุมไม่ให้องค์กร/หน่วยงานทั้งภาครัฐและธุรกิจต่าง ๆ นำข้อมูลไปใช้โดยไม่ได้รับความยินยอมจากตัวเจ้าของข้อมูล เช่น ในกรณีการที่เราได้รับโฆษณาบนสื่อสังคมออนไลน์ โดยมีการนำข้อมูลการใช้งานของเราต่างๆ ไปใช้โดยที่เราไม่รู้ตัว หรือไม่ได้ยินยอมให้องค์กรเก็บข้อมูล เป็นต้น

สำหรับขั้นตอนการวิเคราะห์ข้อมูลขนาดใหญ่หรือบิ๊กดาต้าประกอบด้วยขั้นตอนสำคัญตามลำดับดังนี้

- 1) การกำหนดเป้าหมาย/วัตถุประสงค์ในการใช้ข้อมูล
- 2) การทำความเข้าใจข้อมูล
- 3) การเก็บข้อมูล การสะสมข้อมูล และการทำความสะอาดข้อมูล
- 4) การใช้เทคโนโลยีเพื่อวิเคราะห์ข้อมูล
- 5) การประมวลผล และการนำเสนอข้อมูล

ทั้งนี้ ในแต่ละขั้นตอนมีรายละเอียดดังต่อไปนี้

กระบวนการจัดการข้อมูลขนาดใหญ่ หรือบิ๊กดาต้า (Big Data) อาจไม่ได้เริ่มที่การรวบรวมข้อมูล เป็นเบื้องต้น แต่เริ่มด้วยโจทย์หรือเป้าหมายที่ต้องการก่อนแล้วจึงนำไปสู่ขั้นตอนการรวบรวมข้อมูล

1. การกำหนดเป้าหมาย/วัตถุประสงค์ในการใช้ข้อมูล ก่อนที่จะมีการดำเนินการวางแผนจัดเก็บข้อมูลใดๆ หรือการตัดสินใจเกี่ยวกับเทคโนโลยี/เครื่องมือ/โปรแกรมที่จะนำมาใช้ในการจัดการข้อมูล องค์กร หน่วยงาน บริษัท จำเป็นต้องมีการตรวจสอบก่อนว่าจะนำข้อมูลที่ได้มานั้นไปใช้เพื่ออะไร ซึ่งก็คือการกำหนดเป้าหมายหรือวัตถุประสงค์ในการใช้บิ๊กดาต้าล่วงหน้านั่นเอง

2. การทำความเข้าใจข้อมูล ในขั้นตอนนี้เป็นการทำความเข้าใจอย่างลึกซึ้งว่า องค์กร/บริษัทมีความต้องการจะนำข้อมูลไปทำอะไร และจะใช้วิธีการใดในการวิเคราะห์ข้อมูลในอันที่จะตอบวัตถุประสงค์หรือเป้าหมายที่ตั้งไว้ในขั้นตอนแรก ทั้งนี้มีประเด็นที่จะต้องตระหนักถึงก็คือ 1) ข้อมูลที่ต้องการนั้นเป็นข้อมูลอะไรบ้าง 2) ต้องการข้อมูลจากแหล่งใดและเป็นข้อมูลประเภทใด เพื่อที่จะได้ทราบว่าจะนำข้อมูลที่

ต้องการเหล่านั้นมาจากแหล่งใดได้บ้าง รวมถึงเทคโนโลยี ซอฟต์แวร์ใดที่จะนำมาใช้ในการจัดการข้อมูล และจำเป็นต้องต่อเชื่อมกับเครื่องมือใดได้บ้าง และ 3) ต้องการวิธีใดในการได้มาซึ่งคำตอบ ต้องการวิเคราะห์แบบใด คำถามอย่างไร รวมถึงรูปแบบคำตอบที่ต้องการ เช่น กราฟ หน้าต่างรายงานผล (report dashboard) เป็นต้น

3. การเก็บข้อมูล การสะสมข้อมูลที่เกี่ยวข้อง และการทำความสะอาดข้อมูล การเก็บสะสมข้อมูล อาจใช้ระยะเวลาสั้นนับเดือนหรือหลายปี ซึ่งข้อมูลที่ต้องการนั้นอาจมีหลายมิติ เช่น อาจอยู่ในแบบมีโครงสร้าง หรือข้อมูลแบบไม่มีโครงสร้าง หรือข้อมูลแบบผสมผสานทั้งแบบมีโครงสร้างและไม่มีโครงสร้าง อย่างไรก็ตาม หากต้องการให้รวดเร็ว เพื่อตอบวัตถุประสงค์/โจทย์และการแก้ไขปัญหา ควรเก็บข้อมูลแบบมีโครงสร้างที่ชัดเจน มีแผนและระบบรองรับ ทำให้ประหยัดเวลาและทรัพยากร

ข้อมูลที่มาจากแหล่งข้อมูลต่างๆ จะต้องมีการจัดระบบโดยจัดทำเป็นแอ่งข้อมูลหรือโกดังข้อมูล (data lake/data warehouse) เพื่อนำข้อมูลขนาดใหญ่มาวิเคราะห์และประมวลผล ข้อมูลของบิกดาต้าส่วนใหญ่มักเชื่อมต่อ (connect) เชื่อมโยงข้อมูลเข้าด้วยกันเป็นลักษณะแหล่งข้อมูลออนไลน์ เนื่องจากอุปกรณ์ดิจิทัลต่างๆ ออกแบบให้เก็บข้อมูล ทั้งแบบรายบุคคลและแบบโครงสร้างทั้งหมด การจัดระบบให้ข้อมูลและออกแบบเพื่อการใช้งาน เป็นเรื่องที่ต้องทำงานร่วมกับผู้เชี่ยวชาญ ได้แก่ วิศวกรผู้ออกแบบระบบ (Data Engineer) นักวิทยาศาสตร์ออกแบบการวิเคราะห์ข้อมูล (Data Scientist) เพื่อให้การใช้งานบิกดาต้าตอบโจทย์ตอบสนองมาตรฐานได้ถูกต้องคุ้มค่ามากที่สุด

เมื่อได้เก็บข้อมูลมากเพียงพอหรือได้ข้อมูลผ่านการจัดระบบแล้ว จำเป็นต้องมีการทำความสะอาดข้อมูลซึ่งเป็นกระบวนการตรวจสอบ สะสาง แก้ไข หรือจัดรูปแบบข้อมูลให้อยู่ในสภาพที่พร้อมใช้งานที่สุด มีความเป็นมาตรฐานเดียวกัน รวมไปถึงคัดกรองข้อมูลที่ไม่ถูกต้องหรือไม่จำเป็นออกไปจากชุดข้อมูลที่จะใช้วิเคราะห์หรือประมวลผล เพื่อให้ชุดข้อมูลที่จะใช้มีความสมบูรณ์ มีคุณภาพ พร้อมนำไปวิเคราะห์และใช้ประโยชน์ตามเป้าหมายหรือวัตถุประสงค์ที่กำหนดไว้แน่นอน

ในการทำความสะอาดข้อมูลนั้นจำเป็นต้องทำทุกขั้นตอน ตั้งแต่เก็บข้อมูล รวบรวมข้อมูล จัดมาตรฐานข้อมูล วิเคราะห์ข้อมูล แปลผลข้อมูลตามสมมติฐานหรือโจทย์ ซึ่งการทำความสะอาดข้อมูลนั้นอาจต้องทำหลายรอบ เพื่อกำจัดข้อมูลที่ผิดพลาด (error) หรือบกพร่อง (miss) ทำให้ข้อมูลเกิดความแม่นยำ มีความสมบูรณ์ ทำนายผลได้ถูกต้องที่สุด เครื่องมือที่ใช้เป็นโปรแกรมคอมพิวเตอร์ เช่น โปรแกรมเอกเซล (excel) โปรแกรมทางคณิตศาสตร์อื่น หรืออาจใช้อัลกอริทึม (algorithm) ทำให้เห็นข้อมูลที่ผิดปกติ ข้อมูลผิดพลาด ข้อมูลบกพร่อง แล้วทำความสะอาดให้ข้อมูลนำมาใช้ได้ถูกต้อง

4. การใช้เทคโนโลยีเพื่อวิเคราะห์ข้อมูล ภายหลังจากการเก็บและรวบรวมข้อมูลอย่างต่อเนื่องในช่วงเวลาหนึ่งจนได้ข้อมูลขนาดใหญ่แล้ว และได้ผ่านกระบวนการทำความสะอาดข้อมูลจนได้ข้อมูลที่สมบูรณ์เพียงพอแล้ว ก็จะก้าวเข้าสู่ขั้นตอนการวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลเป็นส่วนสำคัญ ที่จะได้ข้อมูลขนาดใหญ่มาหาศาล ต้องอาศัยความรวดเร็ว ความแม่นยำ ความถูกต้อง และการออกแบบที่ทำให้เห็น ‘โครงสร้าง’ หรือ ‘ภาพรวม’ ทั้งหมด โดยอาศัยเครื่องมือวิเคราะห์ทางสถิติ ทางคณิตศาสตร์ เช่น ค่าสัมประสิทธิ์ (Correlations) ค่าสัมประสิทธิ์สหสัมพันธ์ (Correlations Matrix) เป็นต้น หรืออาจใช้เทคโนโลยีและเครื่องมือสมัยใหม่ เช่น ระบบฮาดูป ระบบคลาวด์

ระบบฮาร์ดแวร์ เป็นต้น (โดยจะกล่าวโดยละเอียดอีกครั้งในเรื่อง 3.4 เทคโนโลยีและเครื่องมือวิเคราะห์ข้อมูลต่อไป) เพื่อมาช่วยสนับสนุนการทำงานรวดเร็ว ในส่วนนี้อาจจะต้องทำงานร่วมกับวิศวกรออกแบบระบบ และนักวิทยาศาสตร์ออกแบบข้อมูล

ปัจจุบันมีเทคโนโลยีเครื่องมือที่ใช้ในการวิเคราะห์ข้อมูลอย่างมากมาย มีทั้งต้องเสียค่าใช้จ่ายและไม่เสียค่าใช้จ่าย ซึ่งก็มีช่วงเวลาที่ทดลองใช้ก่อนได้เพื่อเป็นการทดสอบความสะดวกในการใช้งาน

อย่างไรก็ตามประเด็นสำคัญก็คือ การที่จะเลือกใช้เทคโนโลยีและเครื่องมือต่างๆ มาใช้งานนั้นจะต้องแน่ใจว่าสามารถที่จะนำมาใช้จัดการกับข้อมูลได้อย่างคุ้มค่าและมีประสิทธิภาพต่อการวิเคราะห์เพื่อตอบโจทย์/คำถามหรือเพื่อบรรลุวัตถุประสงค์และเป้าหมายตามที่ต้องการ

5. การประมวลผล และนำเสนอข้อมูล ภายหลังจากได้ผ่านขั้นตอนการวิเคราะห์ข้อมูลขนาดใหญ่แล้ว ก็จะเป็นการประมวลผลให้ได้ข้อสรุป เช่น ได้มองเห็นข้อมูลสองส่วนมีความสัมพันธ์กันหรือข้อมูลหลายส่วนมีความสัมพันธ์ ข้อมูลที่ได้มีความหมายสำคัญในการตอบประเด็น ตอบสมมติฐาน หรือตอบวัตถุประสงค์ที่ได้ตั้งไว้ในขั้นตอนแรก ซึ่งสุดท้ายนำข้อมูลเหล่านั้นมาจัดทำเป็นรูปแบบการนำเสนอที่เข้าใจง่ายผ่านทางสถิติ กราฟ อินโฟกราฟิก รูปภาพ ลายเส้น แอนิเมชัน สื่อประสม ฯลฯ เพื่อให้ผู้บริหารได้รับทราบข้อมูลในท้ายที่สุด

กิจกรรม 3.3

จงระบุถึงขั้นตอนการวิเคราะห์ข้อมูลขนาดใหญ่

แนวตอบกิจกรรม 3.3

ขั้นตอนการวิเคราะห์ข้อมูลขนาดใหญ่หรือบิ๊กดาต้าประกอบด้วยขั้นตอนสำคัญ ดังนี้

1. การกำหนดเป้าหมาย/วัตถุประสงค์ในการใช้ข้อมูล
2. การทำความเข้าใจข้อมูล
3. การเก็บข้อมูล การสะสมข้อมูล และการทำความสะอาดข้อมูล
4. การใช้เทคโนโลยีเพื่อวิเคราะห์ข้อมูล
5. การประมวลผล และการนำเสนอข้อมูล

เรื่องที่ 3.4

เทคโนโลยีและเครื่องมือในการวิเคราะห์ข้อมูลขนาดใหญ่

ดังที่กล่าวมาบ้างแล้วว่า ลักษณะสำคัญของบิกดาต้าที่มีปริมาณข้อมูลมากมายมหาศาล อีกทั้งยังมีความหลากหลายของข้อมูล ดังนั้นจึงมีความจำเป็นที่จะต้องใช้เทคโนโลยีและเครื่องมือสนับสนุนมาช่วยในการวิเคราะห์และประมวลผลข้อมูลขนาดใหญ่ให้ออกมาอย่างรวดเร็วถูกต้อง ได้ผลลัพธ์ตรงตามโจทย์ หรือวัตถุประสงค์/เป้าหมายที่ได้กำหนดไว้ล่วงหน้า

สำหรับเทคโนโลยีและเครื่องมือดำเนินการกับบิกดาตานั้นเป็นเทคโนโลยีและเครื่องมือต่างๆ ที่สามารถติดตามหรือเก็บข้อมูลได้จากหลายๆ แหล่งที่มา ซึ่งเทคโนโลยีเหล่านี้มีทั้งใช้กับองค์กรหรือธุรกิจที่มีขนาดใหญ่ไปจนถึงองค์กรหรือธุรกิจที่มีขนาดปานกลางและขนาดเล็ก

การวิเคราะห์ข้อมูลเป็นส่วนสำคัญที่จะได้บิกดาต้าต้องอาศัยความรวดเร็ว ความแม่นยำ ความถูกต้อง และการออกแบบที่ทำให้เห็น “โครงสร้าง” หรือ “ภาพรวม” ทั้งหมด โดยอาศัยเครื่องมือวิเคราะห์ทางสถิติ ทางคณิตศาสตร์ เช่น ค่าสัมประสิทธิ์ (Correlations) ค่าสัมประสิทธิ์สหสัมพันธ์ (Correlations Matrix) เป็นต้น

ทั้งนี้ เทคโนโลยีและเครื่องมือที่นำมาใช้วิเคราะห์บิกดาต้าอาจแบ่งเป็น 2 ชั้นใหญ่ๆ (อังชิตา ลิมปัทมปาณี และคณะ, 2563) คือ เครื่องมือชั้นพื้นฐาน และเครื่องมือชั้นสูง ดังนี้

1. **เครื่องมือชั้นพื้นฐานของบิกดาต้า (Basic Level)** คือ การใช้ซอฟต์แวร์ของวินโดวส์ (Microsoft Window) ในโปรแกรมเอกเซล (Excel) เป็นเครื่องมือที่สามารถนำมาใช้ได้ เช่น

- 1) การเรียงลำดับมากไปหาน้อย หรือเรียงลำดับจากน้อยไปหามาก
- 2) การจัดหมวดหมู่ เพื่อใช้ค้นหาจัดแบ่งเป็นกลุ่ม
- 3) การคำนวณสูตรทางคณิตศาสตร์ ฟังก์ชัน เพื่ออธิบายโจทย์ที่ตั้งไว้ตั้งแต่ตอนแรก
- 4) การแสดงผลกราฟิก เป็นเส้น เป็นกราฟ เป็นแท่งกราฟ เป็นวงกลมแบ่งสัดส่วน
- 5) การเสนอข้อมูลตัวเลขเป็นตาราง ให้ความเหมือนหรือความต่าง การเปรียบเทียบข้อมูล

ให้เห็น

2. **เครื่องมือชั้นสูงของบิกดาต้า (Advance Level)** ได้แก่ โปรแกรมทาโบลาร์ (TABLEAU) ซึ่งเป็นซอฟต์แวร์เฉพาะ ต้องซื้อจากประเทศสหรัฐอเมริกา ซึ่งในการทำงานด้วยซอฟต์แวร์ชั้นสูงต้องทำงานเป็นที่ระหว่างนักบริหาร วิศวกรออกแบบระบบ และนักวิทยาศาสตร์ออกแบบข้อมูล



ภาพที่ 3.2 โปรแกรมทาโบลว (TABLEAU)

ที่มา: <https://getintopc.com/software/office-tools/tableau-desktop-pro-2019-free-download/>

สำหรับการใช้โปรแกรมขั้นสูง หลายองค์กรจำเป็นต้องซื้อซอฟต์แวร์เฉพาะ เช่น โปรแกรมทาโบลว ดังกล่าว จะคำนวณให้และจัดสัดส่วนออกมาแนะนำเสนอเป็นภาพตามแพลตฟอร์มที่เราเลือก มีกราฟิกต่างๆ ให้เลือกตามประเด็นที่ต้องการจะวิเคราะห์จากข้อมูลขนาดใหญ่ของบิกดาต้า เมื่อจัดสัดส่วนแล้วจะแสดงเป็นภาพให้เห็น จับประเด็นจากภาพกราฟิกดังกล่าวได้ง่ายขึ้น

อย่างไรก็ดี เทคโนโลยีและเครื่องมือที่นำมาใช้กับบิกดาต้ามีหลายอยู่ระบบ ในที่นี้จะยกตัวอย่าง โปรแกรมที่เปิดให้ใช้งานฟรี ได้แก่ 1) ระบบฮาดูป 2) ระบบคลาวด์รีธา 3) ระบบฮอว์ตันเวิร์ก 4) ระบบแมพอาร์ 5) ระบบแอมะซอน 6) ระบบไมโครซอฟท์ และ 7) ระบบอัลติสเทล ให้เข้าใจพอสังเขป ดังนี้

1. ระบบฮาดูป (Hadoop) เป็นชุดของโปรแกรมเปิดให้ใช้งานฟรี (Open Source Programs) รวมถึงโปรแกรมกระบวนการทำงานทั้งหมด เป็นโปรแกรมที่ได้รับความนิยมและทุกคนสามารถเข้าถึงได้ โปรแกรมฮาดูปมี 4 โมดูล (Bernard Marr.2019) โมดูลแต่ละส่วนมีหน้าที่เฉพาะเพื่อวิเคราะห์ชุดข้อมูล บิกดาต้า ประกอบด้วย

1.1 โมดูลระบบแจกแจงข้อมูล (Distributed file-system) มี 2 ส่วนสำคัญคือ การเรียกข้อมูลและการนำข้อมูลไปใช้ ต้องง่ายต่อการจัดการ การเชื่อมโยงระหว่างข้อมูลขนาดใหญ่ที่ต่อกับอุปกรณ์คอมพิวเตอร์กับการทำความสะอาดข้อมูล ซึ่งเป็นเครื่องมือพื้นฐานที่จะตรวจสอบข้อมูลให้ถูกต้อง ในระบบข้อมูล (file system) ของคอมพิวเตอร์จะใช้คอมพิวเตอร์เรียงเก็บข้อมูลขนาดยักษ์ ทำให้ง่ายต่อการค้นหาและนำไปใช้ ระบบฮาดูปใช้โปรแกรมไฟล์ซึ่งเป็นระบบของตัวเอง “เหนือ” ระบบโปรแกรมไฟล์ของเครื่องคอมพิวเตอร์ที่เราใช้งานอยู่ จึงสามารถใช้งานจากเครื่องคอมพิวเตอร์ไหนก็ได้ที่มีระบบปฏิบัติการ OS สนับสนุน

1.2 โมดูลเลือกข้อมูลตามโมเดล (Map Reduce) แมพรีดิวส์เป็นชื่อของระบบปฏิบัติการ ทำหน้าที่อ่านข้อมูลจากแหล่งเก็บข้อมูล แล้วเลือกข้อมูลตาม “สูตร” หรือ “โมเดล” ที่ใส่เข้าไป เพื่อวิเคราะห์ข้อมูลเรียกว่าแมพ (Map) จากนั้นใช้การคำนวณทางคณิตศาสตร์ประมวลผลเลือกข้อมูล เช่น ต้องการนับจำนวนผู้ชายอายุ 30 ปีขึ้นไปในฐานข้อมูลลูกค้า ซึ่งเรียกวิธีการนี้ว่า รีดิวส์ (reduce)

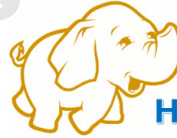
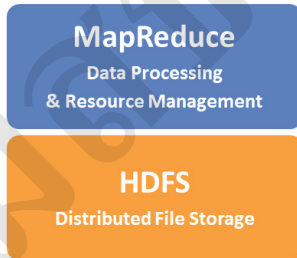
1.3 โมดูลฮาดูปคอมมอน (Hadoop Common) เป็นอีกส่วนหนึ่งของกระบวนการคอมพิวเตอรืที่ต้องอาศัยโปรแกรมจาวา (Java) ซึ่งเป็นโปรแกรมของคอมพิวเตอรืของผู้ใช้ เช่น โปรแกรมวินโดวส์ โปรแกรมยูนิคส์ เป็นต้น เพื่ออ่านข้อมูล ภายใต้ระบบการทำงานของระบบฮาดูป

1.4 โมดูลยานจัดเก็บและรันข้อมูล (YARN) เป็นส่วนสุดท้ายที่จัดการแหล่งข้อมูลขนาดยักษ์ ในการจัดเก็บและรันข้อมูล (Running) เพื่อวิเคราะห์ข้อมูลและประมวลผลในที่สุด

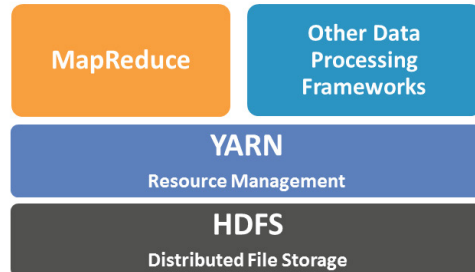
การทำงานทั้ง 4 โมดูลนี้ อยู่ภายใต้กรอบการทำงานของระบบฮาดูป ซึ่งใช้งานล่าสุดเมื่อไม่กี่ปีมานี้ ระบบฮาดูป เป็นโปรแกรมพัฒนาขึ้นจากวิศวกรที่คิดไปข้างหน้าและตระหนักกว่าจะเป็นซอฟต์แวร์ที่เป็นประโยชน์ต่อทุกคน ทั้งช่วยเก็บข้อมูลและวิเคราะห์ข้อมูลของบิกดาต้า มีความยืดหยุ่น องค์กรสามารถใช้ปรับเปลี่ยนข้อมูลและจัดการข้อมูลขนาดใหญ่ได้ในราคาถูก ความต้องการใช้ระบบนี้คือ ระบบการจัดเก็บข้อมูล และการประมวลผลโดยอาศัยฮาร์ดแวร์ศูนย์รวมข้อมูล



Hadoop v1.0



Hadoop v2.0



ภาพที่ 3.3 โปรแกรมระบบฮาดูป

ที่มา: <https://www.clusterkit.co.th/blogs/big-data-with-hadoop/>

ซอฟต์แวร์ออนไลน์ยักษ์ใหญ่ที่เรารู้จักและใช้กันทุกวันนี้ต่างให้บริการฟรี จึงมีผู้ใช้งานจำนวนมากมหาศาล เช่น แอมะซอน (Amazon) กูเกิล (Google) ซึ่งพัฒนาขึ้นโดยวิศวกรผู้เชี่ยวชาญและปรับปรุงให้ทันสมัยตลอด ซอฟต์แวร์ออนไลน์นี้เป็นการทำงานร่วมมือระหว่างผู้ใช้ที่เป็นอาสาสมัครกับภาคธุรกิจ ส่วนระบบฮาดูป (Hadoop) ที่พัฒนาโดย อาปาเช่ ซอฟต์แวร์ ฟาวเดชัน ต้องอาศัยผู้เชี่ยวชาญด้านไอที นักวิทยาศาสตร์ข้อมูลมาช่วย เนื่องจากเป็นเทคนิคเฉพาะ หลายธุรกิจที่มาใช้แล้วส่วนใหญ่จะพัฒนาขึ้นเป็น ระบบคลาวด์ริว (Cloudera) โปรแกรมนี้ติดตั้งง่าย อีกทั้งสามารถรันข้อมูลบนระบบฮาดูปรวดเร็ว

อย่างไรก็ตาม ยังมีโปรแกรมอื่นที่พัฒนาใช้ในบิกดาต้า เช่น โปรแกรมสปาร์ค (Spark) โปรแกรมคอฟคา (Kafka) เพื่อใช้วิเคราะห์ข้อมูลขนาดใหญ่ นอกจากเทคโนโลยีและเครื่องมือข้างต้นที่กล่าวมา ยังมีระบบอื่นที่นำมาใช้กับบิกดาต้า (Marr, 2019) ได้แก่

2. ระบบคลาวด์ทรีรา (Cloudera) หรือระบบคลาวด์ เป็นระบบแรกๆที่เริ่มใช้ร่วมกับระบบฮาดูป เวลาใช้งานต้องทำเป็นชุดการทำงานและต่อเนื่อง ระบบคลาวด์ มีชื่อเต็มว่าคลาวด์ทรีรา ซีดีเอช (Cloudera CDH) ซึ่งเป็นโปรแกรมที่เปิดให้บริการฟรี เป็นระบบปฏิบัติการซึ่งเป็นที่รู้จักอย่างรวดเร็ว ใช้โปรแกรมช่วยคือ SQL (Structured Query Language) เพื่อระบบฮาดูป ผู้ใช้บริการระบบคลาวด์จะได้รับประโยชน์จากระบบนี้ คือ พื้นที่ใช้ส่วนรวม ความปลอดภัย และพื้นที่ใช้ร่วมกับแอปพลิเคชันอื่นที่จะมาเชื่อมภายหลัง ระบบสนับสนุนทั้งหมดดูแลโดยบริษัทคลาวด์ทรีรา เอนเตอร์ไพรส์ ซับสคริปชัน เซอร์วิส (Cloudera Enterprise Subscription Service)



ภาพที่ 3.4 โปรแกรมระบบคลาวด์ทรีรา

ที่มา: <https://www.techtalkthai.com/easy-way-on-getting-started-with-hadoop-and-big-data-analytics-within-2-hours-with-cloudera/>

3. ระบบฮอว์ตันเวิร์ก (Hortonworks) เป็นระบบปฏิบัติการที่ให้บริการแบบเปิดฟรี หลายคนรู้จักระบบฮอว์ตันเวิร์กด้วยการค้นหาจากการใช้โค้ด (Code) และเผยแพร่ในแหล่งรวมโปรแกรมที่เปิดให้ใช้ฟรี บางบริษัทเห็นแนวโน้มและช่องทางการตลาดจากระบบฮอว์ตันเวิร์ก องค์กรที่ดี ไม่นานมานี้ ผู้พัฒนาระบบฮอว์ตันเวิร์กหยุดพัฒนาโปรแกรมไปช่วงหนึ่ง แต่แอมะซอนกับไอบีเอ็ม ก็เสนอระบบฮอว์ตันเวิร์กในแพลตฟอร์มของตัวเองเพียงคู่ไปกับระบบฮาดูป ทั้งนี้ ระบบฮอว์ตันเวิร์กยังเป็นแพลตฟอร์มที่สร้างสรรค์แบบระบบข้อมูลเปิด ซึ่งเป็นระบบที่ใช้งานง่าย มีมาตรฐาน ประหยัดค่าใช้จ่าย ในระยะยาวจึงคาดว่าระบบฮอว์ตันเวิร์กจะเป็นที่นิยมยิ่งขึ้น

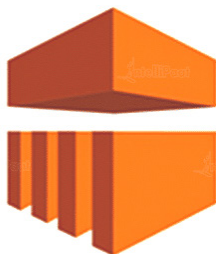


ภาพที่ 3.5 โปรแกรมระบบฮอว์ตันเวิร์กและระบบแมพอาร์

ที่มา: <https://telegra.ph/A-propos-des-grandes-distributions-Hadoop-Cloudera-Hortonworks-MapR-et-consorts-02-12>

4. ระบบแมพอาร์ (MapR) คล้ายกับระบบคลาวด์ที่รู้จักกับฮอว์ตันเวิร์ก ระบบแมพอาร์เป็นแพลตฟอร์มที่เน้นผู้ให้บริการมากกว่าการจัดการบริการ เช่นเดียวกับบริษัทแอมะซอนหรือบริษัทไมโครซอฟท์ แมพอาร์ทำงานแบบบูรณาการร่วมกับระบบฐานข้อมูล (Database) แมพอาร์-ดีบี (MapR-DB) ระบุว่าสามารถประมวลผลได้เร็วมากประมาณ 4-7 ครั้ง เร็วกว่าระบบฮาดูปในการรันข้อมูล 1 ครั้ง แมพอาร์มีระบบใช้พลังงานอย่างมีประสิทธิภาพและประมวลผลรวดเร็วมาก จึงเป็นตัวเลือกที่ดีที่สุดสำหรับบิ๊กดาต้า

5. ระบบแอมะซอน อีลาสติก แมพ รีดิวส์ (Amazon Elastic Map Reduce) แอมะซอนผู้ให้บริการระบบคลาวด์ ภายใต้การทำงานของระบบฮาดูปแพลตฟอร์ม ผ่านเว็บไซต์แอมะซอนเว็บเซอร์วิส แอมะซอนนำเสนอระบบที่ใช้แล้วคุ้มค่าคือ ระบบคลาวด์ ใช้สำหรับเก็บข้อมูลและประมวลผลข้อมูล สามารถนำข้อมูลส่งเข้าหรือเรียกข้อมูลออกมาใช้ตามความต้องการ เว็บไซต์แอมะซอนประกาศเมื่อเร็วๆ นี้ว่า ลูกค้าสามารถใช้โปรแกรม อาปาเช่ ฟลิ้งค์ (Apache Flink stream) วิเคราะห์และประมวลผลได้รวดเร็วแบบทันที (Real-time) บนแพลตฟอร์มเท่ากับเครื่องมืออื่นๆ เช่น โปรแกรมคาฟคา (Kafka) และโปรแกรมเพรสโต (Presto) นอกจากนี้ยังเชื่อมต่อกับแอมะซอนในระบบคลาวด์เซอร์วิส ด้วยโปรแกรมอีซีทู (EC2) เพื่อการประมวลผล (Cloud Processing) และโปรแกรมแอมะซอน เอสที (Amazon S3) กับโปรแกรมไดนาโม ดีบี (Dynamo DB) เพื่อการเก็บข้อมูล (Storage) รวมทั้งโปรแกรมเอ็ดบิลิวเอส ไอโอที (AWS IoT) เพื่อเก็บสะสมข้อมูลจากอุปกรณ์เชื่อมต่อคอมพิวเตอร์ (Internet of Things device) เป็นชุดการทำงานที่น่าสนใจและมีประสิทธิภาพมาก



Amazon EMR



ภาพที่ 3.6 โปรแกรมระบบแอมะซอน อีลาสติก แมพ รีดิวส์

ที่มา: <https://intellipaat.com/blog/amazon-elastic-mapreduce/>

6. ระบบไมโครซอฟท์ (Microsoft) ไมโครซอฟท์ อาซัวร์ เอชดี (Microsoft Azure HD) แพลตฟอร์มสำหรับวิเคราะห์ข้อมูลเชิงลึก (Insight) เป็นบริการระบบคลาวด์ เพื่อจัดการกับระบบคอมพิวเตอร์ที่ติดตั้งโปรแกรมเปิดให้ใช้ฟรีอย่างระบบฮาดูป ระบบฮอว์ตันเวิร์ก ระบบคลาวด์รีรา และระบบแมพอาร์ เป็นต้น ไมโครซอฟท์ อาซัวร์ เอชดี จะทำงานภายใต้แหล่งข้อมูลอาซัวร์ (Azure Data Lake) บนระบบคลาวด์ที่เก็บข้อมูลและวิเคราะห์ข้อมูล สามารถทำงานอย่างมีประสิทธิภาพเท่ากับระบบฮาดูป โปรแกรมแบบเอชดี อินไซต์ (HD Insight) มีหลายโปรแกรม เช่น สปาร์ค (Spark) ไฮฟ์วี่ (Hive) คาฟคา (Kafka) และสตอร์ม (Storm) ทั้งหมดเป็นระบบคลาวด์ที่ให้บริการดีที่สุด ระบบเหล่านี้มีการรักษาความปลอดภัยภายใต้การทำงานบนแพลตฟอร์มนั้นๆ



ภาพที่ 3.7 โปรแกรมระบบไมโครซอฟท์ อาซัวร์ เอชดี

ที่มา: <https://www.quickserv.co.th/knowledge-base/solutions/Microsoft-Azure>

7. ระบบอัลติสเกล (Altiscale) เป็นอีกบริษัทหนึ่งให้บริการระบบคลาวด์ เพื่อเก็บข้อมูลและการจัดการข้อมูลในระบบฮาดูป อัลติสเกล ดาต้า คลาวด์ (Altiscale Data Cloud) ให้บริการระบบปฏิบัติการอัตโนมัติ รวมถึงความปลอดภัยข้อมูล การจัดสเกลข้อมูล การค้นหาข้อมูล ควบคู่กับการทำงานหลักระบบฮาดูป ระบบข้อมูลให้บริการในระบบคลาวด์มีหลายระบบ เช่น สปาร์ค (Spark) ไฮฟ์วี่ (Hive) พิก (Pig) เป็นต้น อย่างไรก็ตาม อัลติสเกลต่างจากระบบคลาวด์ดังกล่าวคือ ต้องทำงานร่วมกับระบบฮาดูป ซึ่งเป็นโปรแกรมเปิด มากกว่าจะทำงานบนแพลตฟอร์มเดียว เหมือนระบบฮอว์ตันเวิร์ก หรือระบบแมพอาร์ ซึ่งเป็นโปรแกรมที่ต้องซื้อ



ภาพที่ 3.8 โปรแกรมระบบอัลติสเกล

ที่มา: http://insidebigdata.com/wp-content/uploads/2014/06/altiscale_logo.jpg

อย่างไรก็ดี การเลือกใช้เทคโนโลยีและเครื่องมือเพื่อทำกับบิกดาต้า ที่เป็นระบบประหยัดงบประมาณและครอบคลุมการทำงานนั้น ยังมีโปรแกรมหรือระบบอื่นๆ ที่พัฒนาอย่างต่อเนื่อง อย่างไรก็ตาม การเลือกใช้เทคโนโลยีใดจะต้องอาศัยการวางแผนและการมองเห็นประโยชน์จากการใช้ข้อมูลขนาดใหญ่

สำหรับการพัฒนาองค์กรในด้านต่าง ๆ ซึ่งต้องคำนึงถึงความคุ้มค่าของการใช้ข้อมูลมากกว่าพิจารณาด้านความก้าวหน้าทันสมัยของเทคโนโลยีเพียงอย่างเดียว

กิจกรรม 3.4

โปรแกรมระบบฮาดูปประกอบด้วย 4 โมดูลย่อยก็โมดูล แต่ละโมดูลมีหน้าที่อย่างไรบ้าง

แนวตอบกิจกรรม 3.4

โปรแกรมฮาดูปมี 4 โมดูลย่อย ซึ่งโมดูลแต่ละส่วนมีหน้าที่เฉพาะเพื่อวิเคราะห์ชุดข้อมูลบิกดาต้า ประกอบด้วย

1. โมดูลระบบแจกแจงข้อมูล
 2. โมดูลเลือกข้อมูลตามโมเดล
 3. โมดูลฮาดูปคอมมอน
 4. โมดูลยานจัดเก็บและรันข้อมูล
-

เรื่องที่ 3.5

การนำข้อมูลขนาดใหญ่ไปใช้ประโยชน์

ระบบบิกดาต้าจะช่วยให้สามารถวิเคราะห์ข้อมูลที่มีความซับซ้อน มีจำนวนมาก และทันเหตุการณ์ (real-time) ตัวอย่างของการประยุกต์ใช้งานข้อมูลขนาดใหญ่ในหลายด้าน ในที่นี้จะเน้นเฉพาะภาครัฐและเอกชนเป็นสำคัญ เช่น

1. การศึกษา กระทรวงศึกษาธิการสามารถใช้บิกดาต้าพิจารณาว่านักเรียนนักศึกษาจบในสาขาใดบ้างในแต่ละปี ตอบโจทย์การพัฒนาประเทศภาคเศรษฐกิจหรือตลาดแรงงานหรือไม่ ยังขาดแคลนบัณฑิตในสาขาวิชาใดบ้าง เท่าไร ควรวางแผนการทำงานในอนาคตล่วงหน้าอย่างไร เพื่อให้ตอบโจทย์ยุทธศาสตร์ชาติ

2. การเกษตร ในภาคการเกษตรมีโครงการที่ชื่อว่า การเกษตรแม่นยำ (Precision Farming) หมายถึง การเกษตรที่อาศัยการทำงานร่วมกันของระบบเซนเซอร์ที่สามารถวัดความชื้น ปริมาณแสงแดด อุณหภูมิ ระบบ ฐานข้อมูลพีช และระบบให้น้ำ ปรับปริมาณแสง และระบบปรับอุณหภูมิ ให้ทำงานสอดคล้องกันเพื่อสร้างสภาวะแวดล้อมที่เหมาะสมต่อการเจริญเติบโตของพืชมากที่สุด และแม่นยำที่สุด

ระบบดังกล่าวสามารถช่วยให้เกษตรกรใช้ทรัพยากรเท่าที่จำเป็น ประหยัด และสามารถประมาณการช่วงเวลาเก็บเกี่ยวและปริมาณพืชผลที่จะได้อีกด้วย

3. อุตสาหกรรม ในโรงงานและอุตสาหกรรมสามารถใช้เทคโนโลยีในการเชื่อมต่อข้อมูลจากฝ่ายต่างๆ เพื่อให้ได้มาซึ่งบิกดาต้าในการจัดการสินค้าคงคลัง ทรัพยากร/วัตถุดิบต่างๆ การเงิน การบัญชี การขนส่ง และอื่นๆ ในการช่วยให้โรงงาน/อุตสาหกรรมบริหารงานได้อย่างมีประสิทธิภาพและเป็นระบบ อาทิ ช่วยตรวจสอบความผิดปกติของเครื่องจักรได้ ช่วยใช้คาดการณ์เวลาที่จำเป็นต้องเปลี่ยนอะไหล่ของอุปกรณ์เมื่อถึงเวลาเสียได้ ซึ่งจะช่วยลดค่าใช้จ่ายในการเปลี่ยนอะไหล่ใหม่โดยไม่จำเป็น นอกจากนี้ ยังมี การเชื่อมต่อข้อมูลระหว่างร้านสะดวกซื้อ ระบบโลจิสติกส์ และโรงงาน จะช่วยให้สามารถบริหารการผลิต และกระจายสินค้าให้ได้ประสิทธิภาพการผลิตมากขึ้นและลดต้นทุนที่ไม่จำเป็น

4. คมนาคมและการจัดการโลจิสติกส์ บิกดาต้าจะเข้ามามีส่วนช่วยในการพัฒนาระบบคมนาคม และการจัดการโลจิสติกส์โดยช่วยสนับสนุนให้มีการเชื่อมต่อข้อมูลระหว่างยานพาหนะด้วยกัน หรือระหว่างยานพาหนะและระบบควบคุมการจราจรอื่น เช่น ระบบสัญญาณ การจราจร ระบบข้อมูลสภาพจราจรหรือการนำเอาระบบดังกล่าวมาใช้กับระบบขนส่งมวลชนที่จะช่วยให้การบริการมีความปลอดภัย สะดวก และตรงเวลามากยิ่งขึ้น นอกจากนี้ การนำระบบดังกล่าวไปใช้ในการขนส่งสินค้า จะทำให้สามารถทราบตำแหน่งยานพาหนะ ทราบสถานการณ์รับ-ส่ง สินค้า อันส่งผลให้การจัดการสินค้าคงคลังมีประสิทธิภาพมากยิ่งขึ้น

5. สุขภาพและการแพทย์ การใช้บิกดาต้าในด้านสุขภาพและการแพทย์มักใช้เพื่อเก็บข้อมูลจำนวนมากของผู้ป่วย ทั้งนี้หากมีการนำเทคโนโลยี IoT มาประยุกต์ใช้ด้วยโดยใช้อุปกรณ์ IoT ที่เก็บข้อมูลสุขภาพและสัญญาณทางร่างกาย (bio signals) เช่น สัญญาณชีพจร ความดันโลหิต คุณภาพการนอน การเคลื่อนไหว การหายใจ ผ่านการใช้อุปกรณ์สวมใส่ (wearable devices) เพื่อรวบรวมและประมวลผลออกมาเป็นข้อมูลสุขภาพ และอาการเจ็บป่วย ซึ่งสามารถเก็บข้อมูลการเจ็บป่วยที่มีประโยชน์ต่อการวินิจฉัย ก่อนที่คนไข้มาถึงการดูแลของแพทย์ การคาดการณ์และการวินิจฉัยการเจ็บป่วยล่วงหน้า (predictive diagnostic) ลดภาระในการตรวจซักประวัติ การแจ้งเตือนการเจ็บป่วยทันที และระบบติดตามการแพร่กระจายของโรค ซึ่งข้อมูลและค่าสถิติการเจ็บป่วยและสุขภาพของกลุ่มประชาชนโดยรวมจะเป็นประโยชน์ต่อการวางแผนทางสาธารณสุข สุขภาพ และการแพทย์

6. การตลาดและการขาย ในแง่ของการตลาดโดยเฉพาะการตลาดดิจิทัล (Digital Marketing) ซึ่งเป็นการตลาดที่ดำเนินการผ่านเครื่องมือดิจิทัล/แพลตฟอร์มดิจิทัลต่างๆ เช่น การทำการตลาดออนไลน์ (Search Engine Marketing: SEM), การทำการตลาดบนสังคมออนไลน์/ช่องทางโซเชียลมีเดีย (Social Media Marketing), การติดตามรวบรวมข้อมูล ความคิดเห็นของผู้บริโภค (Social Monitoring) ในอันที่จะทำความเข้าใจความต้องการของผู้บริโภค/ลูกค้าอย่างลึกซึ้งรอบด้าน ข้อมูลต่างๆ เกี่ยวกับลักษณะของผู้บริโภค เช่น ความต้องการ ทศนคติ วิถีชีวิต ลักษณะทางประชากรศาสตร์ ขั้นตอนการตัดสินใจซื้อ การใช้สินค้า/ผลิตภัณฑ์ และบริการ เป็นต้น เหล่านี้ ล้วนแต่เป็นข้อมูลที่มีความสำคัญและจำเป็นในการวิเคราะห์ผู้บริโภค เพื่อที่จะนำผลการวิเคราะห์ดังกล่าวไปประยุกต์ใช้ในการตลาดและการขายได้อย่างเหมาะสม สำหรับข้อมูลต่างๆ ของลูกค้าและผู้บริโภคนั้นเปรียบประดุจขุมทรัพย์ที่สร้างความได้เปรียบให้กับองค์กรบริษัท หรือแบรนด์นั้นๆ แต่ทว่าหัวใจสำคัญไม่ได้อยู่ที่การมีคลังข้อมูลมหาศาล หากแต่อยู่ที่ความสามารถ

ในการเก็บรวบรวมข้อมูลต่างๆ และนำข้อมูลที่มีมาทำการวิเคราะห์ และตีความอย่างแม่นยำได้มากขึ้นเพียงใด (เช่น รู้ว่าข้อมูลส่วนไหนที่สำคัญหรือไม่สำคัญ ข้อมูลใดจะส่งผลลัพธ์ได้ตามที่ต้องการ) และเพื่อประเมินผลถึงสิ่งที่ได้ทำมาแล้ว รวมถึงพยากรณ์ถึงสิ่งที่จะเกิดขึ้นในอนาคต ด้วยเหตุนี้เองปัจจุบันเกิดเทคโนโลยีที่ใช้ตอบโจทย์หรือแก้ปัญหาด้านข้อมูลผู้บริโภคอยู่มากมาย เช่น เทคโนโลยีการวิเคราะห์ข้อมูล (Data Analytics) เทคโนโลยีการแปลงข้อมูลเป็นภาพ (Data Visualization) เทคโนโลยีการจัดระเบียบข้อมูล (Data Management Platform) เทคโนโลยีการจัดเก็บ จัดกลุ่ม และรวบรวมข้อมูลของลูกค้า (Customer Data Platform) เป็นต้น ในอันที่จะส่งมอบผลิตภัณฑ์ สินค้าและบริการที่สามารถตอบสนองความต้องการให้กับผู้บริโภคได้ในที่สุด

ดังนั้นจะเห็นว่าหากประเทศไทยต้องการจะพัฒนาประเทศให้มีประสิทธิภาพมากขึ้น การนำบิ๊กดาต้ามาใช้จะสามารถช่วยได้ในหลายภาคส่วน ไม่ว่าจะเป็นด้านการศึกษา เกษตรกรรม อุตสาหกรรม คมนาคมและขนส่ง การตลาดและการขาย รวมไปถึงทางด้านสุขภาพและการแพทย์อีกด้วย

กิจกรรม 3.5

จงอธิบายการนำข้อมูลขนาดใหญ่ไปใช้ประโยชน์ในด้านสุขภาพและการแพทย์

แนวตอบกิจกรรม 3.5

ในด้านสุขภาพและการแพทย์มักใช้เก็บข้อมูลจำนวนมากของผู้ป่วย ทั้งที่เป็นข้อมูลสุขภาพและสัญญาณทางร่างกาย เช่น สัญญาณชีพจร ความดันโลหิต เป็นต้น เพื่อรวบรวมและประมวลผลออกมาเป็นข้อมูลสุขภาพ และอาการเจ็บป่วย ซึ่งสามารถเก็บข้อมูลการเจ็บป่วยที่มีประโยชน์ต่อการวินิจฉัยก่อนที่คนไข้มาถึงการดูแลของแพทย์ เป็นการคาดการณ์และการวินิจฉัยการเจ็บป่วยล่วงหน้า ลดภาระในการตรวจซักประวัติ ซึ่งข้อมูลและค่าสถิติการเจ็บป่วยและสุขภาพของกลุ่มประชาชนโดยรวมจะเป็นประโยชน์ต่อการวางแผนทางสาธารณสุข สุขภาพ และการแพทย์

บรรณานุกรม

- ระดับของการวิเคราะห์ข้อมูล (*Data Analytics*). (2563). สืบค้นจาก <https://datacubeth.ai/level-of-data-analytics/>
- อังธิตา ลิ้มปีปัทมปาณี, บุญลาภ ภูสุวรรณ และกมล ชวาลวิทย์. (2563). มหิมามหาข้อมูล (Big Data). ใน *เอกสารการสอนชุดวิชาการรายงานข่าวและการบรรณาธิกร* (หน่วยที่ 3). นนทบุรี: มหาวิทยาลัยสุโขทัยธรรมาธิราช
- Big Data Analytics คืออะไร?* (2563). สืบค้นจาก <https://www.ko.in.th/big-data-analytics-คืออะไร/>
- Gartner Says Advanced Analytics Is a Top Business Priority*. (2014). Retrieved from <https://www.gartner.com/en/newsroom/press-releases/2014-10-21-gartner-says-advanced-analytics-is-a-top-business-priority>
- Marr, Bernard. (2019). *What the Differences between Structured, Semi-Structured and Unstructured Data*. Retrieved from <https://www.bernardmarr.com/default.asp?contentID=1966>

ตอนที่ 4

การรับฟังทางสังคม

รองศาสตราจารย์ ดร.สันหัตถ์ ทองรินทร์



ชื่อ
วุฒิ

รองศาสตราจารย์ ดร.สันหัตถ์ ทองรินทร์
ค.บ. (เกียรตินิยม) ค.ม., ค.ด. (เทคโนโลยีและสื่อสารการศึกษา)
จุฬาลงกรณ์มหาวิทยาลัย

ตำแหน่ง

ว.ม. (สื่อสารมวลชน) มหาวิทยาลัยธรรมศาสตร์
รองศาสตราจารย์ประจำสาขาวิชานิติศาสตร์
มหาวิทยาลัยสุโขทัยธรรมราช

หน่วยที่เขียน

ตอนที่ 4

ตอนที่ 4

การรับฟังทางสังคม

โปรดอ่านหัวเรื่อง แนวคิด และวัตถุประสงค์ของตอนที่ 4 แล้วจึงศึกษารายละเอียดต่อไป

หัวเรื่อง

- เรื่องที่ 4.1 แนวคิดพื้นฐานและประโยชน์ของการรับฟังเสียงทางสังคม
- เรื่องที่ 4.2 หลักการรับฟังเสียงทางสังคม
- เรื่องที่ 4.3 ขั้นตอนการรับฟังเสียงทางสังคม
- เรื่องที่ 4.4 เครื่องมือที่ใช้ในการรับฟังเสียงทางสังคม

แนวคิด

1. การรับฟังเสียงทางสังคมเป็นอีกแนวทางหนึ่งของการวิจัยที่ต้องการเจาะลึกข้อมูลที่มีการกล่าวถึงองค์กร สินค้าหรือบริการที่ปรากฏในสื่อสังคมออนไลน์ ซึ่งข้อมูลที่สามารถนำไปใช้ประโยชน์ที่หลากหลาย อาทิ การวางแผนการสื่อสาร การจัดการภาวะวิกฤต การวิจัยและพัฒนาผลิตภัณฑ์
2. การรับฟังเสียงทางสังคมมีหลักการที่สำคัญนับตั้งแต่การกำหนดวัตถุประสงค์ การกำหนดข้อมูลเชิงลึก การบูรณาการกับแหล่งข้อมูลอื่น การได้ข้อมูลที่มีความถูกต้อง การนำไปสร้างกลยุทธ์ การกำหนดเครื่องมือ และการกำหนดผู้รับผิดชอบ
3. การรับฟังเสียงทางสังคมมีขั้นตอน ได้แก่ การกำหนดวัตถุประสงค์ การกำหนดคำสำคัญ การกำหนดขอบเขต การกำหนดเป้าหมาย การตรวจสอบแหล่งที่มาของข้อมูล และการนำข้อมูลไปใช้งาน
4. ตัวอย่างเครื่องมือที่ใช้ในการวิเคราะห์การรับฟังเสียงทางสังคมของไทย ได้แก่ โปรแกรม ZOCIAL EYE และ โปรแกรม Zanroo ซึ่งทั้งสองโปรแกรมจะมีการให้บริการหลักที่เหมือนกันคือ การใช้คำสำคัญในการตรวจจับข้อมูลแต่ก็มีรายละเอียดของฟีเจอร์การใช้งานที่แตกต่างกันบ้าง

วัตถุประสงค์

เมื่อศึกษาตอนที่ 4 จบแล้ว นักศึกษาสามารถ

1. ระบุประโยชน์ของการรับฟังเสียงทางสังคมได้
2. ระบุหลักการรับฟังเสียงทางสังคมได้
3. ระบุขั้นตอนการรับฟังเสียงทางสังคมได้
4. ยกตัวอย่างการทำงานของเครื่องมือที่ใช้ในการรับฟังเสียงทางสังคมได้

ความนำ

ในการวิจัยทางนิเทศศาสตร์หรือการสื่อสาร หากพิจารณาตามองค์ประกอบของการสื่อสารซึ่งได้แก่ ผู้ส่งสาร สาร ช่องทางการสื่อสาร และผู้รับสารนั้น ในส่วนของผู้รับสารหรือกลุ่มเป้าหมายของการสื่อสาร นับได้ว่ามีความสำคัญเป็นอย่างมาก เนื่องจากจะเป็นการวัดหรือประเมินว่าสารที่ผู้ส่งสารไปยังกลุ่มเป้าหมายนั้นมีการบรรลุวัตถุประสงค์ในการสื่อสารมากน้อยเพียงใด

การวัดหรือประเมินเกี่ยวกับผู้รับสาร สามารถใช้วิธีการวิจัยทั้งเชิงปริมาณและคุณภาพ ซึ่งจะให้ผลของการวิจัยในการตอบหรืออธิบายได้แตกต่างกัน อย่างไรก็ตาม การวิจัยหรือประเมินเกี่ยวกับกลุ่มเป้าหมายนั้นองค์กรหรือผู้ที่ทำการประเมินมักต้องการคำตอบในเชิงลึกเพื่อให้สามารถอธิบายปรากฏการณ์ได้อย่างลึกซึ้งและสามารถนำมาใช้ในการแก้ไขปัญหา การวางแผนการสื่อสารให้กลุ่มเป้าหมายได้อย่างมีประสิทธิภาพและตรงประเด็น สอดคล้องกับความต้องการของกลุ่มเป้าหมายได้มากที่สุด

ถึงแม้ว่าโดยหลักการวิจัยแล้ว จะมีความเชื่อเกี่ยวกับการวิจัยเชิงคุณภาพว่าสามารถให้คำตอบในเชิงลึกมากกว่าการวิจัยเชิงปริมาณ แต่ก็อาจยังให้ข้อมูลในเชิงลึกได้เพียงระดับหนึ่ง แต่เมื่อมีการนำเทคโนโลยีการสื่อสารโดยเฉพาะสื่อสังคมออนไลน์เข้ามาใช้ในชีวิตประจำวันของกลุ่มเป้าหมายมากขึ้น จึงเป็นเหตุให้องค์กร สินค้าหรือบริการหันมาสนใจว่าในสื่อสังคมออนไลน์ต่างๆ นั้น กลุ่มเป้าหมายได้มีการสนทนากล่าวถึง องค์กร สินค้าหรือบริการอย่างไรบ้างที่องค์กรอาจยังไม่ทราบ เพื่อที่จะได้นำข้อมูลต่างๆ มาใช้ในการปรับปรุงการดำเนินงานขององค์กร สินค้าหรือบริการให้ตรงกลุ่มเป้าหมายมากที่สุด

อาจกล่าวได้ว่า การรับฟังเสียงจากกลุ่มเป้าหมายที่มีการกล่าวถึงองค์กร สินค้าหรือบริการ เป็นวิธีการวิจัย ค้นคว้า วิธีการหนึ่งทางการสื่อสารที่ใช้ในการแสวงหาคำตอบให้กับปรากฏการณ์ที่เกิดขึ้น

เรื่องที่ 4.1

แนวคิดพื้นฐานและประโยชน์ของการรับฟังเสียงทางสังคม

1. แนวคิดพื้นฐานของการรับฟังเสียงทางสังคม

ปัจจุบันสื่อสังคมออนไลน์ ได้มีการนำมาใช้งานอย่างแพร่หลายจนแทบจะกลายเป็นส่วนหนึ่งในชีวิตประจำวันของคนส่วนใหญ่ โดยเฉพาะในการสื่อสารแลกเปลี่ยนข้อมูลระหว่างกัน การแบ่งปันความคิดเห็น ความรู้สึก และเชื่อมโยงกันและขยายวงกว้างระหว่างผู้ใช้งานมากขึ้น

นอกจากนั้น สื่อสังคมออนไลน์ยังเข้ามามีบทบาทในการสื่อสารจากองค์กรไปยังกลุ่มเป้าหมาย ซึ่งบางองค์กรได้มีการนำสื่อสังคมออนไลน์มาใช้เป็นช่องทางที่เป็นทางการขององค์กรทั้งภาครัฐและภาคเอกชน โดยเฉพาะในการดำเนินการทางธุรกิจ ซึ่งการนำสื่อสังคมออนไลน์มาใช้นั้นถือได้ว่าเป็นการพลิกโฉมการทำธุรกิจต่างๆ ไม่ว่าจะเป็นการนำมาใช้ในการสื่อสารการตลาด การสร้างความสัมพันธ์กับลูกค้า โดยเฉพาะการรับฟังเสียงจากลูกค้าในมิติต่างๆ ไม่ว่าจะเป็นการแบ่งปันประสบการณ์ ความคิดเห็น และการมีปฏิสัมพันธ์กับตราสินค้าหรือแบรนด์ ผลิตภัณฑ์ และบริการต่างๆ ซึ่งนอกจากจะถือว่าเป็นความท้าทายแล้ว ยังถือว่าเป็นโอกาสในการสร้างการเปลี่ยนแปลงใหม่ที่องค์กรต้องมีการเตรียมความพร้อมที่จะนำมาใช้งานจากที่ไม่เคยมีมาก่อน ซึ่งเป็นการสร้างความแตกต่างจากองค์กรอื่น รวมถึงเป็นการเตรียมการให้เป็นช่องทางการสื่อสารใหม่ที่ทรงพลังในสร้างความสัมพันธ์และประสบการณ์ใหม่ให้กับลูกค้าเพื่อให้บรรลุเป้าหมายของการรณรงค์ทางการตลาดหรือการสื่อสารขององค์กรต่างๆ

อย่างไรก็ตาม สิ่งที่สำคัญคือองค์กร ควรระลึกอยู่เสมอว่า องค์กรต้องการสื่อสารให้กลุ่มเป้าหมายเกิดความเข้าใจว่าองค์กรทำการสื่อสารข้อมูลอะไรออกไปแล้วกลุ่มเป้าหมายมีการพูดถึงองค์กร แบรนด์ สินค้าหรือบริการขององค์กรอย่างไร ซึ่งการพูดถึงของกลุ่มเป้าหมายในประเด็นต่างๆ ดังกล่าวแล้ว ยังถือว่าเป็นข้อมูลเชิงลึกที่มีต่อองค์กร ซึ่ง “การรับฟังเสียง” จากกลุ่มเป้าหมายดังกล่าวอาจไม่มีการวิจัยได้สามารถเข้าถึงข้อมูลเชิงลึกมาก่อน

ทั้งนี้ จุดมุ่งหมายสูงสุดขององค์กรต่างๆ โดยเฉพาะองค์กรด้านธุรกิจนั้น คือการมีชื่อเสียงได้รับการยอมรับจากกลุ่มเป้าหมาย หรือลูกค้า ทั้งที่เป็นกลุ่มเป้าหมายที่มีอยู่ในปัจจุบันหรือผู้ที่คาดว่าจะมีศักยภาพในการเป็นกลุ่มเป้าหมายใหม่ ดังนั้น การใช้สื่อสังคมออนไลน์จะทำให้เกิดการสื่อสารในการสร้างความผูกพัน หรือสร้างประสบการณ์ร่วมกับองค์กร ตราสินค้า สินค้า หรือบริการ และเมื่อกลุ่มเป้าหมายมีส่วนร่วมก็จะมีการขยายข้อมูลที่เป็นเชิงบวกจากประสบการณ์จริงออกไปผ่านช่องทางต่างๆ โดยเฉพาะการใช้สื่อสังคมออนไลน์ที่มีศักยภาพในการแพร่กระจายที่รวดเร็วและเป็นวงกว้าง และที่สำคัญคือ องค์กรได้ทำการเผยแพร่ข้อมูลโดยไม่ต้องลงทุนในการซื้อสื่อโฆษณาหรือประชาสัมพันธ์

การรับฟังเสียงทางสังคม (Social listening) คืออะไร

นับตั้งแต่มีการใช้สื่อสังคมออนไลน์อย่างแพร่หลาย ทำให้องค์กรต่างๆ ต่างมีความสนใจว่ากลุ่มเป้าหมายจะมีการกล่าวถึงองค์กรอย่างไรบ้าง โดยเฉพาะการกล่าวถึงประเด็นในเชิงลบ ซึ่งองค์กรจะทำการ

ติดตามจากสื่อสังคมออนไลน์ว่ากลุ่มเป้าหมายมีการกล่าวถึง องค์กร แรนด์ ลินค้ำ หรือ บริการอย่างไรบ้าง นอกจากนั้น สื่อสังคมออนไลน์ยังได้รับการยอมรับจากองค์กรต่างๆ โดยเฉพาะองค์กรด้านธุรกิจ ว่ามีความสะดวกและได้ข้อมูลรวดเร็ว รวมถึงข้อมูลจากสื่อสังคมออนไลน์นั้น “ค่อนข้าง” ที่จะสามารถยอมรับได้ที่องค์กรต่างๆหันมาให้ความสนใจ ซึ่งมีการคาดการณ์ว่าเกือบครึ่งหนึ่งขององค์กรทางธุรกิจมีการนำสื่อสังคมออนไลน์มาใช้เป็น “เครื่องมือ” ในการ “รับฟังเสียง” จากลูกค้า โดยการลงทุนจัดหาเครื่องมือรวบรวมข้อมูลเกี่ยวกับจำนวนของลูกค้า การตรวจสอบการสนทนา ข้อเสนอ หรือ การเรียกร้องของลูกค้าเกี่ยวกับองค์กร

แบรนด์ ลินค้ำและบริการ

ทั้งนี้ในแต่ละวัน จะมีการสนทนาของลูกค้าเป็นจำนวนมากในสื่อสังคมออนไลน์ ไม่ว่าจะเป็นเว็บไซต์ บล็อก ไลน์ ทวิตเตอร์ อินสตาแกรม หรือสื่ออื่นๆ ทั่วโลก ที่มีการกล่าวถึง องค์กร แรนด์ ลินค้ำและบริการ โดยข้อมูลดังกล่าวจะถูกรวบรวม โดยบริษัทที่มีความเชี่ยวชาญ โดยการรวบรวม และจัดเก็บข้อมูลทั่วไป ในเชิงปริมาณ หัวข้อสำคัญหรือความรู้สึกทั่วไปของเนื้อหาที่กลุ่มเป้าหมายสร้างขึ้น โดยที่ยังไม่มีการจัดระบบหรือจัดโครงสร้างของข้อมูล และหากองค์กรต่างๆ มีความสนใจว่ามีการสนทนาเกี่ยวกับองค์กรอย่างไรบ้าง บริษัทที่ทำหน้าที่ในการเก็บรวบรวมข้อมูลก็จะใช้หลักการของ algorithms ซึ่งเป็นแนวคิดของการใช้คำสั่ง หรือเงื่อนไขให้คอมพิวเตอร์ทำในสิ่งที่กำหนดให้ ซึ่งในการรับฟังเสียงทางสังคม จะเป็นการจัดระบบและตรวจจับข้อมูลด้วยคำสำคัญ (keyword) หรือวิธีการอื่นๆ แล้วแสดงผลข้อมูลตามที่ต้องการ เช่น ลูกค้าของบริษัทมีการกล่าวถึง ลินค้ำ หรือบริการอย่างไร เป็นต้น

2. ประโยชน์ของการรับฟังเสียงทางสังคม

เมื่อองค์กรได้ข้อมูลจากการรับฟังเสียงจากสื่อสังคมออนไลน์แล้ว องค์กรจะนำข้อมูลที่ได้มาแล้วในหลายลักษณะ ได้แก่

2.1 การนำไปใช้ในการแก้ไขปัญหาที่กลุ่มเป้าหมายหรือลูกค้ามีต่อองค์กร ในลักษณะของการปกป้องชื่อเสียง รวมถึงการจัดการชื่อเสียงขององค์กร

2.2 การนำไปใช้ในการวางแผนการสื่อสารให้ตรงกับความต้องการของกลุ่มเป้าหมาย รวมถึงการขยายหรือการสร้างกลุ่มเป้าหมายใหม่ขององค์กร

2.3 การจัดการภาวะวิกฤตขององค์กร โดยการรับฟังเสียงทางสังคมจะช่วยให้การตรวจสอบว่าเริ่มมีปัญหาอะไรบ้างที่เกี่ยวข้อง ที่มีแนวโน้มที่อาจเป็นอันตรายต่อชื่อเสียงขององค์กร แรนด์ ลินค้ำหรือบริการ ซึ่งฝ่ายประชาสัมพันธ์ หรือสื่อสารองค์กรจะนำข้อมูลจากการรับฟังเสียงทางสังคมมากำหนดขนาดและความสำคัญของปัญหา เพื่อที่จะได้ทำการสื่อสาร การตอบสนองหรือการให้คำแนะนำ ช่วยเหลือแก่กลุ่มเป้าหมายหรือลูกค้าขององค์กร ก่อนที่ปัญหาจะมีความรุนแรง และขยายวงกว้างออกไปจนเกินที่จะจัดการแก้ไขได้

2.4 การบริการลูกค้า เนื่องจากปัจจุบันองค์กรต่างๆ ได้มีการนำสื่อสังคมออนไลน์มาให้บริการกลุ่มเป้าหมายหรือลูกค้าเพิ่มมากขึ้น โดยใช้เป็นทั้งช่องทางเสริม บางองค์กรใช้เป็นช่องทางหลักที่เป็นทางเลือกเพิ่มขึ้นจากช่องทางเดิม เพื่อให้เหมาะสมกับพฤติกรรมการใช้งานของกลุ่มเป้าหมาย ซึ่งองค์กร

ไม่เพียงแต่จะใช้สื่อสังคมออนไลน์ในการตรวจสอบขอการสนับสนุน หากแต่ยังนำมาใช้ในการตรวจสอบความต้องการเพื่อสร้างความผูกพันกับกลุ่มเป้าหมายในเชิงรุก ซึ่งถือว่าเป็นสิ่งสำคัญในการที่จะส่งเสริมความเข้าใจให้กับกลุ่มเป้าหมายหรือลูกค้าได้ด้วยเช่นกัน

2.5 การสร้างโปรแกรมทางการตลาด โดยฝ่ายการตลาดจะนำข้อมูลจากการรับฟังทางสังคมหรือข้อเสนอแนะจากกลุ่มเป้าหมายหรือลูกค้า ซึ่งเป็นข้อมูลเชิงลึกนำไปใช้เป็นประโยชน์ในการวางแผนที่จะจัดทำข้อมูล เพื่อที่จะเป็นข้อเสนอ โปรแกรมทางสินค้าหรือบริการ และทำการสื่อสารไปยังกลุ่มเป้าหมายหรือลูกค้าตามความเหมาะสมของแต่ละกลุ่มเป้าหมายที่มีลักษณะเฉพาะกลุ่ม (segments) ที่มีความต้องการที่แตกต่างกัน

2.6 การให้ข้อมูลเกี่ยวกับสินค้าหรือบริการโดยองค์กรจะเสนอข้อมูลที่มีมูลค่า (valuable information) ของสินค้าหรือบริการใหม่ๆ แก่กลุ่มเป้าหมายหรือลูกค้า และรับฟังเสียงเพื่อสนับสนุนการพัฒนา ทดสอบและการปรับปรุงให้มีความเหมาะสมตามความต้องการ ซึ่งเป็นการประหยัดค่าใช้จ่ายในการทำการวิจัยและพัฒนาสินค้าหรือบริการ

2.7 การวิจัยทางการตลาด การใช้สื่อสังคมออนไลน์ในการตรวจสอบและการรับฟังเสียงทางสังคมทำให้สามารถช่วยฝ่ายวิจัยการตลาดขององค์กรสามารถระบุแนวโน้มของอุตสาหกรรม ตามความต้องการทั่วไป (customer wants) และความต้องการที่จำเป็น (customer needs) ของกลุ่มเป้าหมายของลูกค้าได้ชัดเจนมากยิ่งขึ้น รวมถึงยังสามารถติดตามสถานการณ์การแข่งขันสินค้าหรือบริการซึ่งจะช่วยให้องค์กรมีความเข้าใจในการกำหนดตำแหน่งทางการตลาดได้ รวมไปถึงการใช้ในการประเมินกลยุทธ์และการทำความเข้าใจว่าจะทำอย่างไรจึงจะทำให้สินค้าหรือบริการนั้นมีความโดดเด่นกลุ่มเป้าหมายหรือลูกค้า

2.8 การวิเคราะห์ผู้ทรงอิทธิพลหรืออินฟลูเอนเซอร์ (Influencer analysis) การรับฟังทางสังคมนอกจากองค์กรจะใช้ช่องทางสื่อสังคมออนไลน์แล้ว องค์กรยังสามารถพิจารณาว่ามีผู้ทรงอิทธิพลใดที่มีการทำงานในขอบข่ายเดียวกัน สอดคล้องหรือใกล้เคียงกันกับพันธกิจ สินค้าหรือบริการขององค์กร โดยการจัดทำบัญชีรายชื่อผู้ทรงอิทธิพล (influencer lists) ที่มีศักยภาพและความสามารถในการเข้าถึงกลุ่มเป้าหมายหรือลูกค้าได้ ทั้งนี้ผู้ทรงอิทธิพลจะมีส่วนในการรับฟังความคิดเห็นทางสังคม และสร้างการมีส่วนร่วมในการสร้างความเข้าใจ ความเชื่อมโยง ความผูกพันให้กับองค์กรได้

จากข้างต้นจะเห็นได้ว่า การรับฟังเสียงทางสังคมนั้นมีประโยชน์ให้กับองค์กร สินค้าหรือบริการ รวมไปถึงการเกิดประโยชน์ในการนำไปใช้งานของแต่ละแผนก ฝ่าย หรือหน่วยงานที่เกี่ยวข้อง และสามารถนำข้อมูลไปใช้ในการสื่อสารทั้งในเชิงรุก การแก้ปัญหา การวางแผนได้อย่างหลากหลายและมีประสิทธิภาพตามวัตถุประสงค์

กิจกรรม 4.1

การรับฟังเสียงทางสังคมมีประโยชน์อย่างไร

แนวตอบกิจกรรม 4.1

การรับฟังเสียงทางสังคมมีประโยชน์ คือ การนำไปใช้ในการแก้ไขปัญหา การวางแผนการสื่อสาร การจัดการภาวะวิกฤตขององค์กร การบริการลูกค้า การสร้างโปรแกรมทางการตลาด การให้ข้อมูลเกี่ยวกับสินค้าหรือบริการ การวิจัยทางการตลาด การวิเคราะห์ผู้ทรงอิทธิพล

เรื่องที่ 4.2

หลักการรับฟังเสียงทางสังคม

การรับฟังเสียงทางสังคมมีวิธีการ ดังต่อไปนี้

1. **การกำหนดวัตถุประสงค์** โดยองค์กร ต้องพิจารณาว่าจะใช้วิธีการรับฟังเสียงทางสังคมเพื่อจุดมุ่งหมายใด ต้องการทราบว่าสังคมกำลังสนทนาหรือกล่าวถึงองค์กร สินค้า หรือบริการอย่างไร ทั้งนี้เพื่อให้การกำหนดค่าสำคัญมีความตรงตามความต้องการมากที่สุด เช่น ต้องการนำไปใช้ในการวางแผนการสื่อสารการตลาด การประชาสัมพันธ์ การสร้างภาพลักษณ์ การแก้ไขภาวะวิกฤตขององค์กร เป็นต้น

2. **การกำหนดข้อมูลเชิงลึกที่สามารถนำไปปฏิบัติได้** โดยองค์กรจะต้องคิดว่าต้องการหาข้อมูลอย่างไรเพื่อให้ตรงความต้องการ มีความเหมาะสมและมั่นใจได้ว่าผลที่ได้สามารถนำไปใช้ประโยชน์ได้จริง เช่น การวิเคราะห์ข้อมูลที่เป็นความรู้สึกอาจไม่ได้มีความสำคัญ เช่น กลุ่มเป้าหมายหรือลูกค้าบางคนอาจมีความรู้สึกเพียงว่าชอบหรือไม่ชอบสินค้าหรือบริการ แต่การวิเคราะห์ข้อมูลในเชิงลึกว่าทำไม หรือมีเหตุผลอะไรที่ทำให้กลุ่มเป้าหมาย หรือลูกค้ามีความรู้สึกเช่นนั้น หรือมีความรู้สึกเชื่อมั่นเพิ่มขึ้น หรือลดลงอย่างไร อะไรคือสาเหตุ มีเหตุผลอย่างไร อะไรมีอิทธิพลต่อความคิดเช่นนั้น จึงจะทำให้ได้ประโยชน์ในการนำไปใช้งานมากกว่า

3. **การบูรณาการกับแหล่งข้อมูลอื่น** หากองค์กรไม่มีการบูรณาการข้อมูลจากสื่อขององค์กรที่มีอยู่ให้เข้ากับแหล่งข้อมูลอื่นๆ ในสื่อสังคม ก็ถือว่าการใช้สื่อสังคมไม่ศักยภาพ จึงมักจะส่งผลให้ไม่ประสบความสำเร็จในการตรวจสอบว่ามี การสนทนา หรือกล่าวถึงองค์กร สินค้า หรือบริการอย่างไรบ้าง ดังนั้นจึงควรมีการบูรณาการสื่อสังคมให้เข้ากับสื่อขององค์กรที่มีอยู่เดิม ทั้งนี้ในองค์กรต่างๆ จึงหันมาใช้กลยุทธ์สร้างทีมงานสำหรับด้านสื่อสังคมขึ้นมารองรับในการเป็นศูนย์กลางในการรวบรวมข้อมูลที่เกี่ยวข้องกับองค์กร

4. **การได้ข้อมูลที่มีความถูกต้องและตรงตามความต้องการ** ดังที่กล่าวมาแล้วว่า การรับฟังทางสังคมสามารถนำมาใช้ประโยชน์ให้กับหน่วยงานต่างๆ ขององค์กรที่มีความแตกต่างและหลากหลายหน้าที่ ความรับผิดชอบ ไม่ว่าจะเป็นส่วนของฝ่ายการบริการลูกค้า ฝ่ายการตลาด ฝ่ายทรัพยากรบุคคล ฝ่าย

พัฒนาผลิตภัณฑ์ เป็นต้น ซึ่งเมื่อมีการรวบรวมและวิเคราะห์ข้อมูลนั้นหมายความว่าข้อมูลนั้นๆ จำเป็นอย่างยิ่งที่ต้องมีความถูกต้องและเป็นข้อมูลในเชิงลึก มีความสัมพันธ์กับการนำไปใช้งานและที่สำคัญคือผู้ที่มีอำนาจในการตัดสินใจขององค์กรสามารถที่จะนำข้อมูลนั้นมาใช้ได้อย่างทันที่และเหมาะสมกับบริบท หรือสถานการณ์

5. การนำผลการรับฟังเสียงทางสังคมไปใช้ในการสร้างกลยุทธ์ โดยทั่วไปแล้วข้อมูลต่างๆ ที่มีอยู่ในโลกอินเทอร์เน็ต นั้นมีจำนวนอยู่มากมายมหาศาล ดังนั้นจึงเป็นสิ่งที่ท้าทายสำหรับองค์กรในการที่จะทำการบริหารจัดการ หรือจัดระเบียบของข้อมูลที่เกี่ยวข้องให้สามารถนำมาใช้งานได้ง่าย โดยองค์กรต้องมีการวางแผนเกี่ยวกับการนำข้อมูลมาใช้งาน ซึ่งหากองค์กรใช้วิธีการรับฟังทางสังคมในลักษณะเฉพาะกิจหรือใช้บางครั้งบางคราว ก็อาจจะทำให้องค์กรขาดข้อมูลที่เป็นคุณค่าที่แท้จริง (real value) เกี่ยวกับองค์กรที่ปรากฏในสังคม แต่ถ้าหากองค์กรใช้วิธีการรับฟังทางสังคมอย่างสม่ำเสมอจะทำให้องค์กรสามารถใช้ข้อมูลเชิงลึกที่มีความหมายที่แท้จริงในการนำไปกำหนดกลยุทธ์ที่เหมาะสมทั้งต่อตัวกลุ่มเป้าหมายและองค์กรเอง

อย่างไรก็ตาม การกำหนดกลยุทธ์เกี่ยวกับการรับฟังเสียงทางสังคมมักจะมีพื้นฐานที่เกี่ยวข้องกับการกำหนดเป้าหมายและวัตถุประสงค์ของการรับฟังเสียงทางสังคม และผลที่ได้สามารถรองรับสนับสนุนการใช้งานได้หลากหลาย และที่สำคัญคือต้องมีการวางแผนเกี่ยวกับการใช้เครื่องมือเพื่อให้ได้ผลที่มีความแม่นยำตามที่ต้องการ

6. การกำหนดเครื่องมือที่ใช้และเนื้อหาจากการรับฟังเสียง นอกจากการใช้สื่อสังคมออนไลน์ในการรับฟังเสียงทางสังคมแล้ว สิ่งที่สำคัญอีกประการหนึ่งในการรับฟังเสียงทางสังคม คือเนื้อหาที่องค์กรต้องการทราบ ซึ่งองค์กร สินค้าหรือบริการจำเป็นต้องมีการพิจารณาว่าเนื้อหาใดจึงจะเป็นเนื้อหาที่มีคุณค่ามากที่สุด ดังนั้น จึงต้องมั่นใจว่าเนื้อหาที่ได้รับฟังเสียงจากสื่อสังคมออนไลน์นั้นมีความเกี่ยวข้องกับการนำไปใช้งาน

อย่างไรก็ตาม องค์กรควรต้องคำนึงถึงแหล่งที่มาของข้อมูลด้วย เช่น ข้อมูลที่ได้ควรถูกหาจากเครือข่ายสังคมใด ซึ่งต้องรู้ว่าเครือข่ายสื่อสังคมออนไลน์ใดที่กลุ่มเป้าหมาย หรือลูกค้าขององค์กรใช้งานอยู่ หรือไม่ใช้งาน มีการพูดหรือสนทนาเกี่ยวกับองค์กรอย่างไร มีบทสนทนาใดบ้างที่มีความสำคัญต่อองค์กร สินค้าหรือบริการ กลุ่มเป้าหมายหรือลูกค้าที่กล่าวถึงองค์กรเป็นเพศใด อายุเท่าใด เป็นต้น ทั้งนี้ถ้าหากองค์กรไม่มีการกำหนดขอบข่ายเนื้อหาและมีการค้นหาการสนทนาอย่างกว้างๆ ทั่วไป ผลลัพธ์ที่ได้ก็อาจไม่มีประโยชน์ที่จะนำมาใช้งานได้ ดังนั้นองค์กรจึงควรใช้เวลาในการพิจารณาว่าต้องการประเด็นเนื้อหาที่กำลังทำการรวบรวมนั้นมีความตรงตามวัตถุประสงค์จริงๆ

ส่วนเครื่องมือที่จะนำมาใช้นั้น องค์กรจะต้องตัดสินใจว่าเครื่องมือใดจะมีความเหมาะสมมากที่สุด ทั้งนี้อาจพิจารณาจากประสิทธิภาพของเครื่องมือที่มีอยู่ว่ามีศักยภาพเพียงใด รวมถึงการพิจารณาถึงต้นทุนในการนำมาใช้ด้วย ซึ่งก็จะมีทั้งแบบที่ให้ใช้งานโดยไม่คิดมูลค่า และแบบที่คิดราคาค่าบริการการใช้งาน

7. การกำหนดผู้รับผิดชอบในการฟังเสียงทางสังคม ในการรับฟังทางสังคมนั้น ประเด็นเกี่ยวกับความรับผิดชอบนั้นถือว่าเป็นเรื่องที่มีความสำคัญเป็นอย่างมาก โดยเฉพาะการที่จะกำหนดให้มีความชัดเจนว่าใครควรเป็นผู้รับผิดชอบในการวิเคราะห์ข้อมูลและใครควรเป็นผู้ดูแลรายละเอียดข้อมูล นอกจากนั้นในส่วน

ของเจ้าของเครื่องมือในการใช้วิเคราะห์การรับฟังเสียงทางสังคมก็จำเป็นที่จะต้องมีความรับผิดชอบในการนำเครื่องมือที่เหมาะสมและดีที่สุดมาใช้ในการวิจัยและวิเคราะห์ข้อมูล เช่น การออกแบบการเขียนโปรแกรมการวิเคราะห์ข้อมูล การรายงานผลในเชิงลึก รวมถึงการมีประสบการณ์ในการวิเคราะห์เนื้อหา (text analysis) เป็นต้น นอกจากนี้ผู้ที่ทำการวิเคราะห์ข้อมูลจากการรับฟังเสียงทางสังคมยังจำเป็นต้องมีความรู้ความเข้าใจในเชิงเทคนิควิธีการตรวจจับข้อมูล (data capture methods) มีจุดหรือตำแหน่งใดในสื่อสังคมออนไลน์ที่จะสามารถหาและเชื่อมโยงข้อมูลต่อได้ ดังนั้นผู้ที่ทำการวิเคราะห์ข้อมูลจึงต้องมีความเชี่ยวชาญและมีทักษะในการทำกรวิเคราะห์ข้อมูลที่อยู่บนออนไลน์และยังต้องมีความรู้ความเข้าใจเกี่ยวกับพฤติกรรมทางสังคมของกลุ่มเป้าหมายหรือลูกค้า ด้วยเพื่อให้สามารถทำการวิเคราะห์เนื้อหาได้เป็นอย่างดี

กิจกรรม 4.2

หลักการรับฟังเสียงทางสังคมมีอะไรบ้าง

แนวตอบกิจกรรม 4.2

การรับฟังเสียงทางสังคมมีหลักการ คือ การกำหนดวัตถุประสงค์ การกำหนดข้อมูลเชิงลึก การบูรณาการกับแหล่งข้อมูลอื่น การได้ข้อมูลที่มีความถูกต้องและตรงตามความต้องการ การนำผลการรับฟังเสียงทางสังคมไปใช้ในการสร้างกลยุทธ์ การกำหนดเครื่องมือที่ใช้และเนื้อหาจากการรับฟังเสียง และการกำหนดผู้รับผิดชอบในการฟังเสียงทางสังคม

เรื่องที่ 4.3

ขั้นตอนการรับฟังเสียงทางสังคม

เมื่อองค์กร สินค้าหรือบริการ ต้องการนำการรับฟังเสียงทางสังคมมาใช้ และมีการตัดสินใจเลือกที่จะใช้เครื่องมือใด ซึ่งอาจเป็นเครื่องมือที่ต้องเสียค่าใช้จ่ายหรือไม่มีค่าใช้จ่าย ในการนำมาจัดการวิเคราะห์เนื้อหาเพื่อให้ได้ข้อมูลเชิงลึกตามวัตถุประสงค์ที่ต้องการ โดยมีขั้นตอนของการรับฟังเสียงทางสังคมดังนี้

1. การกำหนดวัตถุประสงค์ (objectives) โดยองค์กร สินค้าหรือบริการต้องการรับฟังเสียงทางสังคม และนำผลมาใช้ในด้านใด เช่น การแก้ไขปัญหาที่มีการกล่าวถึงองค์กร การวางแผนการสื่อสาร ตลอดจนการพัฒนา ปรับปรุง เกี่ยวกับองค์กร สินค้าหรือบริการ เป็นต้น

2. การกำหนดคำสำคัญและประเด็น (keywords and topics) ที่ต้องการตรวจสอบหรือต้องการรับฟังจากสื่อสังคมออนไลน์ โดยการรับฟังเสียงทางสังคมจะเป็นการใช้หลักการของ “การค้นหาตามคำสำคัญ” (keyword-based search) ซึ่งอาจเป็นคำหรือประเด็นทั้งในทางบวกและทางลบ ซึ่งข้อมูลที่ได้จากทั้ง 2 ด้าน จะสามารถนำมาใช้ประโยชน์ร่วมกันได้

นอกจากนั้น การกำหนดคำสำคัญ หรือประเด็นที่ต้องการทราบนั้น อาจมีการใช้คำสำคัญมากกว่าหนึ่งคำ เช่น การใช้คำที่มีความหมายเดียวกันหรือใกล้เคียงกัน ตลอดจนอาจมีการผสมคำ เพื่อให้ได้ข้อมูลที่มากขึ้น โดยเครื่องมือที่ใช้จะดึงข้อมูลที่ตรง หรือสัมพันธ์กับคำสำคัญ หรือประเด็นที่ใช้ในการค้นหาแล้วแสดงผลออกมา

การค้นหาจากสื่อสังคมออนไลน์ที่องค์กรใช้งานหรือที่เป็นทางการขององค์กรจะให้ผลการค้นหาที่ตรงตามความต้องการมากกว่า แต่ถ้าเป็นผลที่ได้มาจากเครื่องมือการสืบค้นอื่นอาจได้ข้อมูลในเชิงกว้างมากกว่า และอาจจะได้ข้อมูลที่ไม่ครบถ้วน

3. การกำหนดขอบเขต (parameter) อื่นเพิ่มเติม เช่น ลักษณะของกลุ่มเป้าหมาย โดยมักจะกำหนดตามข้อมูลที่เป็นลักษณะทางบุคลิกภาพที่เป็นตัวแทนของกลุ่มเป้าหมาย (personas) ได้แก่

- ข้อมูลเกี่ยวกับลักษณะทางประชากร (demographics) ได้แก่ เพศ อายุ ภูมิภาค ซึ่งข้อมูลในลักษณะทางประชากรนี้อาจจะแสดงผลได้ไม่สมบูรณ์นัก ทั้งนี้อาจเนื่องจากผู้ใช้สื่อสังคมออนไลน์อาจระบุตัวตนของตัวเองหรือไม่ก็ได้ หรืออาจใช้การระบุตัวตนที่เป็นการสร้างตัวตนเทียมขึ้นมา ซึ่งผู้ที่ทำหน้าที่ในการวิเคราะห์ข้อมูลจะต้องอาศัยความชำนาญในการตีความจากข้อความ (text) ที่มีการกล่าวหรือสนทนากันในประเด็นต่าง ๆ เช่น ผู้ที่ทำการสนทนาอาจไม่ได้ระบุเพศ แต่มีการระบุข้อความในประโยค การสนทนา ด้วยคำแสดงเพศ อาทิ ผม ดิฉัน ครับ ค่ะ ก็จะสามารถอนุมานได้ว่าเป็นเพศหญิงหรือชาย แต่ก็อาจไม่ใช่ผลที่แท้จริงทั้งหมดก็ได้

- ความสนใจที่กลุ่มเป้าหมายกล่าวถึง (interests of people mentioning)

- ความรู้สึกที่กลุ่มเป้าหมายกล่าวถึง (sentiment of mentions) ได้แก่ การกล่าวถึงในเชิงบวก เป็นกลาง หรือเชิงลบ

ในประเด็นของความสนใจ และความรู้สึกที่กลุ่มเป้าหมายกล่าวถึง องค์กร สินค้าหรือบริการนั้น เครื่องมือในการรับฟังความคิดเห็นทางสังคมมักจะไม่สามารถทำการตรวจจับและวิเคราะห์ข้อมูลได้โดยตรง แต่อาจใช้วิธีหลายวิธีการร่วมกัน เช่น การแปลความหมายจากประสบการณ์ของผู้วิเคราะห์ การดูบริบท หรือสถานการณ์ประกอบกัน เป็นต้น

ทั้งนี้โดยธรรมชาติมนุษย์มักจะใช้การแสดงออกซึ่งความรู้สึกที่แตกต่างกัน บางคนแสดงออก มาอย่างตรงไปตรงมา เช่น ชอบ ประทับใจ ไม่ชอบ เฉยๆ เป็นต้น หรือบางครั้งอาจมีการใช้ภาษาแสดง ที่ไม่ได้แสดงความรู้สึกออกมาตรงๆ ก็ได้ ได้แก่ จึ้ง หมายถึง ดี เลิศ เฉียบ หมายถึง ถูกใจ โดนใจ บ๊ยะ/ บั้ง หมายถึง การดูดี สุดยอด, ยืนหนึ่ง หมายถึง ความโดดเด่น ตัวเต็ง ที่หนึ่ง

- ปัจจัยด้านผู้ทรงอิทธิพล (influencer factors) ได้แก่ จำนวนของผู้ที่ติดตามในสื่อสังคม ออนไลน์ของผู้ทรงอิทธิพลที่กล่าวถึง หรือมีปฏิสัมพันธ์ต่อประเด็นที่เกี่ยวกับองค์กร สินค้าหรือบริการ

ทั้งนี้ การกำหนดขอบเขตต่างๆ ดังข้างต้นจะทำให้สามารถจำแนกกลุ่มเป้าหมายได้ชัดเจน และทำให้ได้ทราบว่าการกลุ่มเป้าหมายดังกล่าวมีความสนใจในประเด็นใดและช่องทางการสื่อสารที่กลุ่มเป้าหมาย นิยมใช้งาน

อย่างไรก็ตาม การกำหนดขอบเขตเพิ่มเติมเพื่อให้ได้ข้อมูลที่มีความลึกมากขึ้นนี้ อาจจำเป็นต้อง ใช้เครื่องมือที่มีการออกแบบและเขียนโปรแกรมขึ้นมาโดยเฉพาะ แต่ถ้าหากเป็นเครื่องมือในการสืบค้น ข้อมูล (search engine) โดยทั่วไป ได้แก่ google ก็อาจไม่สามารถหาข้อมูลในเชิงลึกได้เท่าที่ต้องการ

4. การกำหนดเป้าหมาย (goals) ของการค้นหาการรับฟังเสียงทางสังคม โดยเป้าหมายจะ แสดงออกมาในรูปแบบของความต้องการให้เกิดในการรายงานผล เช่น การแสดงผลที่เป็นความถี่ของ ข้อมูลแบบรายวัน รายสัปดาห์ รายเดือน หรือรายไตรมาส เป็นต้น ซึ่งการแสดงความถี่จะทำให้เห็นระดับ ความสำคัญหรือความเข้มข้นของข้อมูลที่จะนำไปใช้งาน

5. การตรวจสอบแหล่งที่มาของข้อมูล (sources) อย่างระมัดระวัง รอบคอบ ถึงแม้ว่าการได้ใช้ งานเทคโนโลยีที่มีความก้าวหน้าและทันสมัย แต่การเข้าใจว่าเทคโนโลยีนั้นๆ มีที่มาที่ไปอย่างไรก็ถือว่าเป็น สิ่งสำคัญเช่นเดียวกัน เช่น ความน่าเชื่อถือของแหล่งข้อมูล เป็นต้น

อย่างไรก็ตาม สิ่งที่สำคัญของการรับฟังเสียงทางสังคม คือ การระลึกอยู่เสมอว่าข้อมูลที่ได้อาจยังไม่มีความครอบคลุมตามที่ต้องการ และรวมถึงข้อมูลที่ได้อาจยังไม่ได้ครอบคลุมในทุกช่องทางที่อาจยังเข้าไม่ถึง ดังนั้น การรับฟังเสียงทางสังคมจึงควรทำการวิเคราะห์ข้อมูลจากที่ได้รับผลมาจากหลายๆ แหล่ง ร่วมกันเพื่อให้ข้อมูลเกิดความสมบูรณ์รอบด้านมากที่สุด

6. การนำข้อมูลไปใช้งาน (usage) โดยหลังจากที่ผู้วิเคราะห์ได้ผลของการรับฟังเสียงทางสังคม แล้วอาจมีการนำผลนั้นรายงานส่งไปยังฝ่ายที่เกี่ยวข้องให้นำไปใช้ในการแก้ไขปัญหาเกี่ยวกับการสื่อสาร ด้านต่างๆ ที่องค์กร สินค้าหรือบริการกำลังประสบอยู่ได้อย่างทันท่วงที ก่อนที่จะมีความปานปลาย หรือ อาจนำไปใช้งานในการสื่อสารเชิงรุก เช่น การวางแผนการสื่อสารขององค์กร สินค้าหรือบริการไปยังกลุ่ม เป้าหมาย เป็นต้น

กิจกรรม 4.3

การรับฟังเสียงทางสังคมมีขั้นตอนอย่างไรบ้าง

แนวตอบกิจกรรม 4.3

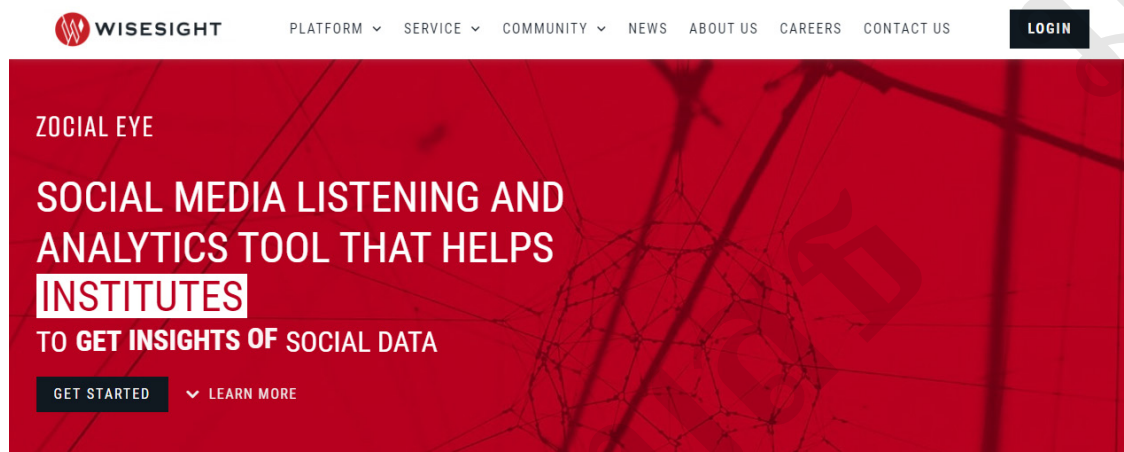
การรับฟังเสียงทางสังคมมีขั้นตอน ได้แก่ การกำหนดวัตถุประสงค์ การกำหนดคำสำคัญและประเด็น การกำหนดขอบเขต การกำหนดเป้าหมาย การตรวจสอบแหล่งที่มาของข้อมูล และการนำข้อมูลไปใช้งาน

เรื่องที่ 4.4

เครื่องมือที่ใช้ในการรับฟังเสียงทางสังคม

สำหรับในเอกสารนี้จะขอแนะนำตัวอย่างของเครื่องมือในการรับฟังเสียงทางสังคมที่ชื่อ “ZOCIAL EYE” ของบริษัทไวส์ไซต์ (WISESIGHT) ประเทศไทย (<https://wisesight.com>) โดยเครื่องมือ ZOCIAL EYE จะใช้ในการรวบรวมข้อมูลบนสื่อสังคมออนไลน์ เพื่อนำข้อมูลสำหรับช่วยในการวางแผนงานด้านการทำการตลาดออนไลน์ (Digital marketing)

1. ZOCIAL EYE



ภาพที่ 4.1 ZOCIAL EYE

ที่มา: <https://wisesight.com/zocialeye/#learn>

ZOCIAL EYE เป็นเครื่องมือวิเคราะห์ข้อมูลโซเชียลมีเดียจากบริษัท WISESIGHT ที่ออกแบบมาให้มีฟีเจอร์ที่สามารถดึงข้อมูลที่สำคัญได้แบบเรียลไทม์ พร้อมทั้งดูข้อมูลได้ในหลากหลายรูปแบบ โปรแกรมจะให้ผู้ใช้งานป้อนคำสำคัญ (Keyword) ซึ่งซึ่งการตั้งค่าคำสำคัญสามารถปรับแต่งได้อย่างละเอียด โดย

Main Keyword เป็นการใส่คำที่เฉพาะเจาะจงเกี่ยวกับเรื่องที่สนใจ ทั้งคำเขียนถูก เขียนผิด ภาษาไทย หรือภาษาอังกฤษ

Sub Keyword เป็นการใส่คำที่เกี่ยวข้องกับ Main Keyword เช่น ต้องการทราบคำที่พูดถึง เพลง นักร้อง ผลงานจากทุกคำที่ถูกจับใน Main Keyword

Exclude Keyword เป็นการใส่คำที่ไม่เกี่ยวข้อง เพื่อให้คำดังกล่าวไม่เข้ามาอยู่ในระบบ

ภาพที่ 4.2 ลักษณะหน้าจอสำหรับการป้อน Main Keyword/ Sub Keyword/ Exclude Keyword

ที่มา: <https://wisesight.com/news/blog/what-is-social-listening/#>

สำหรับการแสดงผลข้อมูลก็จะมีฟีเจอร์ต่าง ๆ ที่ออกแบบมาให้ดูข้อมูลแบบเจาะจงในแต่ละประเภทได้ละเอียดขึ้น แสดงผลภาพรวมของข้อมูลจากแพลตฟอร์มต่าง ๆ และเลือกที่จะดูแต่ละคำสำคัญ (Keyword) เป็นรายแพลตฟอร์ม รายหมวดหมู่ หรือดูแบบ Sentiment รวมถึง เลือกลงเป็นรายวัน (Day) รายสัปดาห์ (Week) และช่วงวันจันทร์-ศุกร์ (WeekDay) และรายเดือน (Month) หรือจะใช้เปรียบเทียบช่วงเวลาปัจจุบันกับอดีตที่ผ่านมาก็ได้ โดยข้อมูลเหล่านี้จะประมวลผลอยู่ภายใต้แคมเปญที่นักการตลาดตั้งค่าคำสำคัญเอาไว้ เช่น

Channel เป็นการสรุปผลของจำนวนข้อความ (Message) ที่เกี่ยวข้องกับแพลตฟอร์มต่างๆ โดยแบ่งเป็นเปอร์เซ็นต์ให้เห็นอย่างชัดเจน

Top Messages by engagement เป็นการสรุปผลข้อความไหนในโลกโซเชียลที่ได้เอ็นเกจเมนต์เยอะที่สุด ซึ่งจะใช้เพื่อดูว่า คอนเทนต์ที่ได้รับความสนใจเยอะว่าเป็นคอนเทนต์ลักษณะไหน และมาจากช่องทางใด

Statistics เป็นภาพรวมจากตัวเลขสถิติที่สรุปออกมาว่า มีจำนวนข้อความเท่าไร เอ็นเกจเมนต์ทั้งหมดเท่าไร มีค่าเฉลี่ยของข้อความและเอ็นเกจเมนต์ต่อวันเท่าไร มีการบอกจำนวนแอดเคานต์ทั้งหมด และจำนวนคำสำคัญ (Keyword) ในแต่ละแบบ

Sentiments ผลสรุปความรู้สึกของคนบนโซเชียลต่อเรื่องราวต่างๆ ว่ามีความรู้สึกเชิงบวก เชิงลบ หรือกลาง และมีจำนวนเท่าไรหรือเป็นเปอร์เซ็นต์ที่เท่าไร

Top Users จะบอกถึงผู้ที่ได้รับเอ็นเกจเมนต์มากที่สุดในแต่ละแพลตฟอร์ม ซึ่งจะช่วยในการหาอินฟลูเอนเซอร์ที่มีอิทธิพล

Demographics เป็นการแสดงผลลัพธ์ข้อมูลประชากร ซึ่งประกอบด้วยช่วงอายุและเพศ เพื่อให้ธุรกิจได้ทราบข้อมูลเชิงลึกของกลุ่มคนที่สนใจในเรื่องดังกล่าว

Main Keyword และ Sub Keyword เป็นจำนวนข้อความของคำสำคัญที่ใช้ในการเก็บข้อมูลในแคมเปญ ซึ่งจะแสดงผลออกมาเป็น เปอร์เซ็นต์และจำนวนตัวเลขที่ชัดเจน

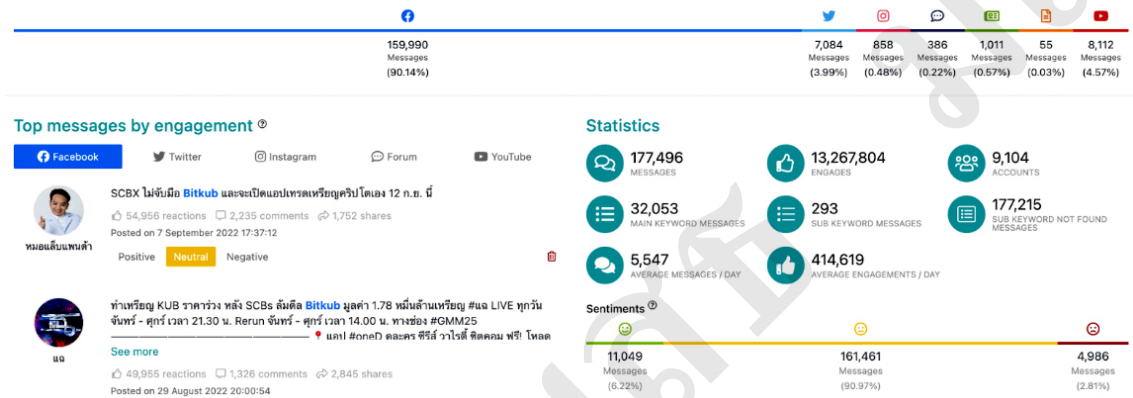
Account Labels เป็นการแสดงผลเกี่ยวกับการจัดประเภท Influencer ตามเกณฑ์มาตรฐานของ Wiselight โดยจะระบุว่า ในแคมเปญที่สร้างขึ้นมีธุรกิจใด (Categories) ที่กล่าวถึงเพียงใด เป็น Influencer หรือบัญชีประเภทใด (Type) มีจำนวนผู้ติดตามเท่าใด (Audience Size) และเคยได้รับรางวัล THAILAND SOCIAL AWARD หรือไม่ รวมถึงเป็นบัญชีผู้ใช้งานที่มีประสิทธิภาพการทำงาน (Performance) ดีมากน้อยเพียงใด

Word clouds และ Hashtag clouds เป็นการสรุปผลข้อมูลที่ทำให้เห็นว่า ในแคมเปญนี้คำสำคัญคำใดถูกพูดถึงมากที่สุด

Image Label Powered by KIRIN เป็นการสรุปผลลัพธ์ด้านรูปภาพว่า รูปภาพที่ใช้ในแคมเปญเป็นรูปภาพประเภทใดบ้าง

Heatmap เป็นการสรุปจังหวัดที่คนบนโซเชียลพิมพ์คำสำคัญนั้นๆ เช่น คนมักพูดถึงเกี่ยวกับคริปโตในกรุงเทพฯ เป็นต้น

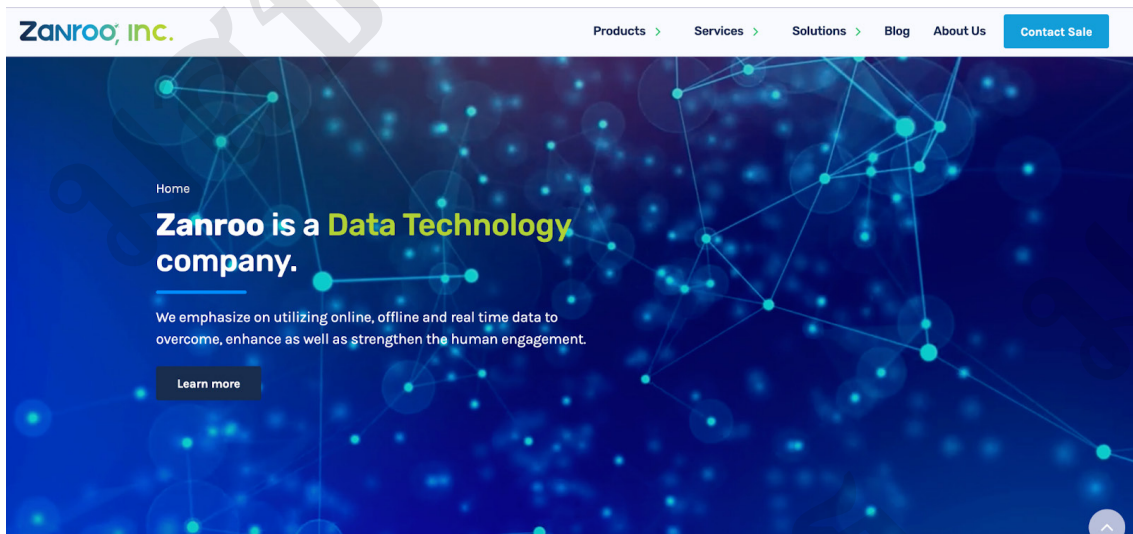
Channel



ภาพที่ 4.3 การแสดงผลการค้นหา

ที่มา: <https://wisesight.com/news/blog/what-is-social-listening/#>

2. Zanroo



ภาพที่ 4.4 เครื่องมือการวิเคราะห์การรับฟังเสียงทางสังคมของ Zanroo

ที่มา: https://contentshifu.com/blog/review-social-listening-tools#9_Zanroo

Zanroo (แซนรู) เป็นเครื่องมือ Social Listening & Analytics ของไทยมีการใช้งานที่เข้าใจง่าย สามารถรวบรวมข้อมูลมาได้หลากหลาย อีกทั้งยังมี NPL ซอฟต์แวร์รองรับหลายภาษา

จุดเด่นของ Zanroo คือ การใช้งานแบบเรียลไทม์ (Real-time Listening & Monitoring) ในการตรวจจับกระแสโซเชียลได้อย่างรวดเร็วไม่ต้องประมวลผลนาน รวมถึงการมอนิเตอร์ความเคลื่อนไหวที่อาจส่งผลกระทบต่อแบรนด์เพื่อการจัดการความเสี่ยง (Crisis management) ได้อย่างตรงจุด

NPL for multiple languages ที่สามารถรองรับภาษาที่หลากหลาย เช่น ไทย อังกฤษ จีน ญี่ปุ่น อินโดนีเซีย ฟิลิปปินส์ เวียดนาม พม่า และอื่นๆ อีกมากมาย

Multi-level Keywords structure และ Auto-Tagging functions ที่ทำให้สามารถเจาะลึกข้อมูลด้วยคำสำคัญ (key word) ที่ยืดหยุ่นไปถึงระดับบุคคล และนำมาจัด Category ด้วย Tag เพื่อการวิเคราะห์ที่ชัดเจนมากขึ้น

กิจกรรม 4.4

จงยกตัวอย่างของโปรแกรมเครื่องมือการวิเคราะห์การรับฟังเสียงทางสังคมที่มีในประเทศไทย

แนวตอบกิจกรรม 4.4

ตัวอย่างของเครื่องมือวิเคราะห์ข้อมูลการรับฟังทางสังคมในประเทศไทย ได้แก่

1. ZOICAL EYE ของบริษัท WISESIGHT โดยเป็นเครื่องมือที่ออกแบบมาให้มีพีเจอร์ที่สามารถดึงข้อมูลที่สำคัญได้แบบเรียลไทม์ พร้อมทั้งดูข้อมูลได้ในหลากหลายรูปแบบ โดยโปรแกรมจะให้ผู้ใช้งานป้อนคำสำคัญ (Keyword) ซึ่งการตั้งคำสำคัญสามารถปรับแต่งได้อย่างละเอียดตามวัตถุประสงค์ของผู้ใช้งาน

2. ZANROO เป็นเครื่องมือที่การออกแบบใช้งานแบบเรียลไทม์ ในการตรวจจับกระแสโซเชียลได้อย่างรวดเร็วไม่ต้องประมวลผลนาน รวมถึงการมอนิเตอร์ความเคลื่อนไหวที่อาจส่งผลกระทบต่อแบรนด์เพื่อการจัดการความเสี่ยง สามารถรองรับภาษาที่หลากหลาย รวมถึง Multi-level Keywords structure และ Auto-Tagging functions ที่ทำให้สามารถเจาะลึกข้อมูลด้วยคำสำคัญ (key word) ที่ยืดหยุ่นไปถึงระดับบุคคล และนำมาจัด Category ด้วย Tag เพื่อการวิเคราะห์ที่ชัดเจนมากขึ้น

บรรณานุกรม

- Alex Gonçalves. (2017). *Social Media Analytics Strategy: Using Data to Optimize Business Performance*. Retrieved from <https://doi.org/10.1007/978-1-4842-3102-9>
- MYcustomer.com. *Social Listening Guide: Advice and best practice for successfully monitoring Social media*.
https://contentshifu.com/blog/review-social-listening-tools#9_Zanroo
<https://wisesight.com>

សំរេង

សំរេង

សំរេង

សំរេង

សំរេង

សំរេង

សំរេង

សំរេង