สพร./PRD
พฤษภาคม 2567

มาตรฐานสำนักงานพัฒนารัฐบาลดิจิทัลอยู่ระหว่างการจัดทำ
ห้ามใช้หรือยึดร่างนี้เป็นมาตรฐาน

มาตรฐานสำนักงานพัฒนารัฐบาลดิจิทัลฉบับสมบูรณ์จะมีประกาศโดย
สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน)

ร่าง

มาตรฐานสำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน)
DGA Community Standard

ว่าด้วยแนวทางการจัดทำข้อมูลนิรนาม

GOVERNMENT DATA ANONYMIZATION GUIDELINE

สำหรับเสนอคณะกรรมการจัดทำร่างมาตรฐาน

สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน)
ชั้น 17 อาคารบางกอกไทยทาวเวอร์ 108 ถนนรางน้ำ แขวงถนนพญาไท เขตราชเทวี กรุงเทพฯ 10400
หมายเลขโทรศัพท์: 0 2612 6000 โทรสาร: 0 2612 6011 0 2612 6012

มาตรฐานสำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน)

DGA Community Standard

มสพร. x-2567

DGA x-2567

ว่าด้วย แนวทางการจัดทำข้อมูลนิรนาม

เวอร์ชัน 1.0

สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน)
สำนักนายกรัฐมนตรี

มาตรฐานสำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน)
ว่าด้วยแนวทางการจัดทำข้อมูลนิรนาม

มสพร. X-2567

สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน)
ชั้น 17 อาคารบางกอกไทยทาวเวอร์
108 ถนนรางน้ำ แขวงถนนพญาไท เขตราชเทวี กรุงเทพฯ 10400
หมายเลขโทรศัพท์: (+66) 0 2612 6000 โทรสาร: (+66) 0 2612 6011

ประกาศโดย
สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน)
สำนักนายกรัฐมนตรี
วันที่ ระบุวันที่ประกาศ

คณะกรรมการจัดทำร่างมาตรฐาน ข้อกำหนด และหลักเกณฑ์ ภายใต้พระราชบัญญัติการบริหารงานและการให้บริการภาครัฐผ่านระบบดิจิทัล พ.ศ. 2562

ที่ปรึกษา

นางไอรดา เหลืองวิไล	สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน)

ประธานกรรมการ

ผู้ช่วยศาสตราจารย์ณัฐวุฒิ หนูไพโรจน์	จุฬาลงกรณ์มหาวิทยาลัย

รองประธานกรรมการ

นายอาศิส อัญญะโพธิ์	สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน)

กรรมการ

นายมารุต บูรณรัช	ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ
นางสาวขนิษฐ์ ผาทอง	สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์ (องค์การมหาชน)
นายชลอ อินทพันธ์	สำนักบริหารการทะเบียน กรมการปกครอง
นางวณิสรา สุขวัฒน์	สำนักงานคณะกรรมการพัฒนาระบบราชการ
นายพัชโรดม ลิมปิษเฐียร	สำนักงานคณะกรรมการกฤษฎีกา
นายสันติ สิทธิเลิศพิศาล	สำนักงานมาตรฐานผลิตภัณฑ์อุตสาหกรรม
ดร.วีระ วีระกุล	สภาดิจิทัลเพื่อเศรษฐกิจและสังคมแห่งชาติ
รองศาสตราจารย์เกริก ภิรมย์โสภา	ประธานคณะทำงานเทคนิคด้านมาตรฐานกระบวนการและการดำเนินงานทางดิจิทัล
รองศาตราจารย์ธีรณี อจลากุล	ประธานคณะทำงานเทคนิคด้านมาตรฐานการบริหารจัดการข้อมูลภาครัฐ
ผู้ช่วยศาตราจารย์มารอง ผดุงสิทธิ์	ประธานคณะทำงานเทคนิคด้านมาตรฐานการเชื่อมโยงและแลกเปลี่ยนข้อมูลภาครัฐ

กรรมการและเลขานุการ

นางสาวอุรัชฎา เกตุพรหม	สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน)

คณะทำงานเทคนิคด้านมาตรฐานการบริหารจัดการข้อมูลภาครัฐ

ที่ปรึกษา

นางไอรดา เหลืองวิไล	สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน)
ผู้ช่วยศาสตราจารย์ณัฐวุฒิ หนูไพโรจน์	ประธานคณะกรรมการจัดทำร่างมาตรฐาน ข้อกำหนด และหลักเกณฑ์ ภายใต้พระราชบัญญัติการบริหารงานและการให้บริการภาครัฐผ่านระบบดิจิทัล พ.ศ. 2562
นางสาวฐิติรัตน์ ทิพย์สัมฤทธิ์กุล	มหาวิทยาลัยธรรมศาสตร์

ประธานคณะทำงาน

รองศาสตราจารย์ธีรณี อจลากุล	ผู้อำนวยการสถาบันข้อมูลขนาดใหญ่ (องค์การมหาชน)

รองประธานกรรมการ

ผู้ช่วยศาสตราจารย์โษฑศ์รัตต ธรรมบุษดี	มหาวิทยาลัยมหิดล

คณะทำงาน

นายพีระไทย พิศาลธรรมนนท์	กรมทรัพย์สินทางปัญญา
นายมารุต บูรณรัช	ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ
นางสาวปริสุทธิ์ จิตต์ภักดี	สถาบันข้อมูลขนาดใหญ่ (องค์การมหาชน)
นายธีระพงษ์ วงษ์สอาด	สำนักข่าวกรองแห่งชาติ
นายอภิสิทธิ์ สุขสาคร	สำนักงานคณะกรรมการกำกับหลักทรัพย์และตลาดหลักทรัพย์
นายอมรพันธุ์ นิติธีรานนท์	สำนักงานคณะกรรมการคุ้มครองข้อมูลส่วนบุคคล
นางสาวปศิญา เชื้อดี	สำนักงานคณะกรรมการข้อมูลข่าวสารของราชการ
นายสุมิทธิ์ เกศวพิทักษ์	สำนักงานคณะกรรมการพัฒนาระบบราชการ
นางกาญจนา ภู่มาลี	สำนักงานสถิติแห่งชาติ
นางสาวณัฐชยา ภาสสัทธา	สำนักงานสภาความมั่นคงแห่งชาติ
นายวันประชา เชาวลิตวงศ์	ธนาคารแห่งประเทศไทย
นายชรินทร์ ธีรฐิตยางกูร	สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน)
นางสาวมณฑา ชยาวิกรม	สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน)

คณะทำงานและเลขานุการ

นางสาวอุรัชฎา เกตุพรหม	สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน)

วิเคราะห์และจัดทำมาตรฐานรัฐบาลดิจิทัล

นางสาวสุภัทรา เรืองวานิช	สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน)
นางสาวศุภมาส พงษ์ภาคิน	สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน)
นายธนัชกฤศ เรืองฉวี	สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน)

แนวทางการจัดทำข้อมูลนิรนาม เวอร์ชัน 1.0 จัดทำขึ้นเพื่อเป็นตัวอย่างให้หน่วยงานภาครัฐนำไปใช้เป็นแนวทางในการจัดทำข้อมูลนิรนาม ซึ่งเป็นกระบวนการจัดทำให้ข้อมูลส่วนบุคคลหรือข้อมูลที่สามารถระบุตัวตนได้มาอยู่ในรูปแบบที่ไม่สามารถระบุตัวตนได้ เพื่อลดความเสี่ยงการระบุตัวตนของเจ้าของข้อมูล และเพื่อสร้างความมั่นใจให้แก่หน่วยงานภาครัฐในการสามารถนำข้อมูลไปใช้ประโยชน์ ทั้งยังเป็นการสร้างความเชื่อมั่นให้แก่ประชาชนถึงแนวทางการใช้ข้อมูลของภาครัฐ โดยแนวทางฉบับนี้ได้จัดทำตามแนวมาตรฐานและแนวปฏิบัติที่ดีของ

1. ประกาศคณะกรรมการพัฒนารัฐบาลดิจิทัล เรื่อง ธรรมาภิบาลข้อมูลภาครัฐ เวอร์ชัน 1.0
2. มรด. 6 : 2566 มาตรฐานรัฐบาลดิจิทัลว่าด้วยกรอบธรรมาภิบาลข้อมูลภาครัฐ ฉบับปรับปรุง: แนวปฏิบัติ
3. มสพร. 8-2565 มาตรฐานสำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน) ว่าด้วยหลักเกณฑ์การจัดระดับชั้นและการแบ่งปันข้อมูลภาครัฐ เวอร์ชัน 1.0
4. ศูนย์วิจัยกฎหมายและการพัฒนา คณะนิติศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย,แนวปฏิบัติเกี่ยวกับการคุ้มครองข้อมูลส่วนบุคคล 3.0 (TDPG 3.0)
5. มาตรฐาน NISTIR 8053: De-identification of Personal Information
6. The Personal Data Protection Commission ('PDPC'), Guide to Basic Data Anonymisation (31 March 2022)

และได้มีการจัดงานประชาพิจารณ์เพื่อเปิดรับฟังความคิดเห็นเป็นการทั่วไป และนำข้อมูล ข้อสังเกต ข้อคิดเห็นจากผู้ทรงคุณวุฒิและจากหน่วยงานที่เกี่ยวข้อง เพื่อให้ข้อเสนอแนะเกี่ยวกับมาตรฐานฉบับนี้มีความสมบูรณ์ครบถ้วน และสามารถนำไปปรับใช้ในทางปฏิบัติได้อย่างมีประสิทธิภาพ

แนวทางการจัดทำข้อมูลนิรนาม เวอร์ชัน 1.0 ฉบับนี้จัดทำโดยฝ่ายมาตรฐานดิจิทัลภาครัฐ สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน) สำนักนายกรัฐมนตรี

สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน)
ชั้น 17 อาคารบางกอกไทยทาวเวอร์
108 ถนนรางน้ำ แขวงถนนพญาไท เขตราชเทวี กรุงเทพฯ 10400
หมายเลขโทรศัพท์: (+66) 0 2612 6000 โทรสาร: (+66) 0 2612 6011
E-mail: sd-g1_division@dga.or.th
Website: www.dga.or.th

คำนำ

ในยุคข้อมูลขนาดใหญ่ที่มีผลต่อการเชื่อมต่อทางดิจิทัลทั้งภาครัฐ ภาคเอกชนที่เพิ่มขึ้นอย่างรวดเร็ว ความจำเป็นในการปกป้องข้อมูลส่วนบุคคลและความเป็นส่วนตัวกลายเป็นสิ่งที่ไม่สามารถละเลยได้ กระบวนการจัดทำข้อมูลนิรนามจึงเป็นกระบวนการหนึ่งที่ช่วยให้สามารถใช้ข้อมูลสำหรับการวิเคราะห์ การวิจัย การพัฒนานวัตกรรม หรือเทคโนโลยีใหม่ โดยไม่เปิดเผยตัวตนของผู้ที่เกี่ยวข้องกับข้อมูลดังกล่าว กระบวนการนี้มีความสำคัญอย่างยิ่งในการรักษาความเชื่อมั่นและความน่าเชื่อถือในการจัดการข้อมูล ซึ่งเป็นปัจจัยสำคัญที่ส่งผลต่อความสำเร็จในการดำเนินงานขององค์กรและการรักษาความไว้วางใจจากผู้มีส่วนได้เสีย การจัดทำข้อมูลให้เป็นนิรนามนั้น ไม่เพียงแต่เป็นการลบหรือปรับเปลี่ยนข้อมูลที่สามารถระบุตัวตนของบุคคลได้ (อาทิเช่น ชื่อ, ที่อยู่, เลขประจำตัวประชาชน) แต่ยังรวมถึงการวิเคราะห์และปรับเปลี่ยนข้อมูลอื่น ที่อาจนำไปสู่การระบุตัวตนได้โดยอ้อม ซึ่งเป็นไปตามมาตรา 8 แห่งพระราชบัญญัติการบริหารงานและการให้บริการภาครัฐผ่านระบบดิจิทัล พ.ศ. 2562 (2) และประกาศคณะกรรมการพัฒนารัฐบาลดิจิทัลเรื่อง ธรรมาภิบาลข้อมูลภาครัฐ ข้อ 4 (5) เพื่อให้เกิดกระบวนการบริหารจัดการและคุ้มครองข้อมูลที่ครบถ้วน โดยกระบวนการนี้ต้องดำเนินการอย่างรอบคอบ เพื่อให้มั่นใจได้ว่าข้อมูลนิรนามที่ได้ ยังคงมีประโยชน์สำหรับวัตถุประสงค์ที่ต้องการนำใช้งาน ในขณะเดียวกันก็ไม่สามารถนำไปสู่การระบุตัวตนของบุคคลได้ ทั้งนี้การประมวลผลข้อมูลนิรนามมีด้วยกันหลายขั้นตอน ซึ่งรวมถึงการวิเคราะห์ความเสี่ยงในการระบุตัวตน การเลือกและปรับใช้เทคนิคในการทำให้ข้อมูลเป็นนิรนาม และการทดสอบความเป็นนิรนามของข้อมูลหลังจากผ่านกระบวนการ ทั้งหมดนี้ต้องดำเนินการภายใต้หลักการปกป้องข้อมูลและความเป็นส่วนตัวที่เข้มงวด เพื่อรับรองว่าข้อมูลที่ได้นั้นปลอดภัยและไม่เป็นอันตรายต่อบุคคลที่เกี่ยวข้อง

สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน) เล็งเห็นว่าการจัดทำข้อมูลนิรนามเป็นส่วนสำคัญในการสร้างความไว้วางใจและความน่าเชื่อถือในยุคสมัยที่ข้อมูลเป็นสิ่งจำเป็นต่อการดำเนินชีวิตและการทำงานของเราในชีวิตประจำวัน ตลอดถึงการขับเคลื่อนองค์กรให้ดำเนินไปยังเป้าหมายได้อย่างมีประสิทธิภาพ ดังนั้นการปฏิบัติตามหลักเกณฑ์และมาตรฐานที่เหมาะสมในการทำให้ข้อมูลเป็นข้อมูลนิรนามจะช่วยให้สามารถใช้ประโยชน์จากข้อมูลได้อย่างเต็มที่ โดยไม่ละเมิดสิทธิ์และความเป็นส่วนตัวของบุคคล ทั้งนี้กระบวนการในการจัดทำข้อมูลนิรนามอาจมีความซับซ้อน จึงจำเป็นต้องมีความรู้ ความเข้าใจ และดำเนินการอย่างระมัดระวังสูงสุด

สารบัญ

สารบัญรูป

สารบัญตาราง

มาตรฐานสำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน)
ว่าด้วยแนวทางการจัดทำข้อมูลนิรนาม

1. บทนำ

มาตรฐานสำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน) ว่าด้วยแนวทางการจัดทำข้อมูลนิรนาม จัดทำขึ้นเพื่อเป็นตัวอย่างให้หน่วยงานภาครัฐนำไปใช้เป็นแนวทางในการจัดทำข้อมูลนิรนาม ซึ่งเป็นกระบวนการจัดทำให้ข้อมูลส่วนบุคคลหรือข้อมูลที่สามารถระบุตัวตนได้มาอยู่ในรูปแบบที่ไม่สามารถระบุตัวตนเจ้าของข้อมูลได้ เพื่อการนำข้อมูลนำไปใช้ประโยชน์ต่อได้โดยไม่ขัดกับกฎหมาย เช่น การนำข้อมูลไปวิเคราะห์เพื่อขับเคลื่อนงานตามภารกิจองค์กร การแบ่งปันระหว่างหน่วยงานภาครัฐ เพื่อสร้างความมั่นใจให้แก่หน่วยงานภาครัฐในการสามารถนำข้อมูลไปใช้ประโยชน์ ทั้งยังเป็นการสร้างความเชื่อมั่นให้แก่ประชาชนถึงการใช้ข้อมูลของภาครัฐ ประกอบไปด้วย 5 บท ดังนี้

บทที่ 1 บทนำ กล่าวถึง ความเป็นมา วัตถุประสงค์ของการจัดทำข้อมูลนิรนาม ขอบข่าย คำนิยาม และ กฎหมายที่เกี่ยวข้อง เพื่อให้เห็นถึงความจำเป็นในการจัดทำข้อมูลนิรนาม **เหมาะสำหรับผู้บริหารระดับสูงของหน่วยงาน**

บทที่ 2 แนวคิดในการจัดทำข้อมูลนิรนาม กล่าวถึง ความสำคัญของการจัดทำข้อมูลนิรนาม แนวคิดและอธิบายถึงวิธีการจัดทำข้อมูลนิรนามโดยสังเขป เพื่อช่วยสร้างความเข้าใจต่อภาพรวมและพื้นฐานในการจัดข้อมูลนิรนาม **เหมาะสำหรับผู้บริหารระดับสูงของหน่วยงานและเจ้าหน้าที่ทั่วไป**

บทที่ 3 กระบวนการจัดทำข้อมูลนิรนาม กล่าวถึง การพิจารณาข้อมูลมีหลักการอย่างไร การขจัดข้อมูลระบุตัวตน หลักเกณฑ์การจัดทำข้อมูลนิรนาม โดยมีหลักเกณฑ์การพิจารณาและข้อเสนอแนะเพื่อให้เพื่อเป็นแนวทางให้แก่เจ้าของข้อมูลใช้ในการประกอบการพิจารณาเพื่อจัดทำข้อมูลนิรนามได้อย่างเหมาะสม รวมถึงการวิเคราะห์ความเสี่ยงและมาตรการจัดการความเสี่ยงที่อาจเกิดขึ้น เพื่อให้เห็นถึงกระบวนการจัดข้อมูลนิรนามที่เป็นขั้นตอน **เหมาะสำหรับผู้ปฏิบัติงานด้านข้อมูล**

บทที่ 4 ภาคผนวก (เครื่องมือ Open Source) กล่าวถึง เครื่องมือการจัดทำข้อมูลนิรนามเบื้องต้นโดยการใช้โปรแกรม Microsoft Excel รายการเครื่องมือการจัดทำข้อมูลนิรนามเชิงพาณิชย์ และกรณีศึกษาการจัดทำข้อมูลนิรนามของสถาบันข้อมูลขนาดใหญ่ (องค์การมหาชน) กรณีศึกษาการจัดทำข้อมูลนิรนามของธนาคารแห่งประเทศไทย เพื่อให้มีความเข้าใจต่อการจัดทำข้อมูลนิรนามและการใช้เครื่องมือต่างๆ **เหมาะสำหรับผู้ปฏิบัติงานด้านข้อมูลและเจ้าหน้าที่ทั่วไป**

1.1. ความเป็นมา

ในโลกปัจจุบันที่ข้อมูลเป็นทรัพยากรสำคัญและมีบทบาทต่อการตัดสินใจในหลายด้าน การจัดการข้อมูลอย่างมีความรับผิดชอบและปลอดภัยกลายเป็นเรื่องที่ต้องให้ความสำคัญยิ่งขึ้น การจัดทำข้อมูลนิรนามเป็นหนึ่งในกระบวนการที่ช่วยให้สามารถใช้ข้อมูลสำหรับวัตถุประสงค์ทางการวิเคราะห์ การวิจัย การพัฒนานวัตกรรม หรือเทคโนโลยีใหม่ โดยไม่เปิดเผยข้อมูลส่วนบุคคลที่สามารถระบุตัวตนได้ แม้กระบวนการนี้จะฟังดูเรียบง่าย แต่ในการดำเนินการจริง การจัดทำข้อมูลให้เป็นนิรนามเป็นกระบวนการที่ต้องการความเข้าใจในหลายด้าน รวมถึงจริยธรรม กฎหมาย และการประเมินความเสี่ยง

สำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน) มุ่งเน้นให้หน่วยงานภาครัฐ ตระหนักถึงความสำคัญของการจัดทำข้อมูลนิรนามขั้นพื้นฐาน เพื่อให้เข้าใจถึงกระบวนการที่เกี่ยวข้อง และความท้าทายในการรักษาคุณภาพของข้อมูลให้คงที่ การทำให้ข้อมูลเป็นนิรนามอาจส่งผลให้ข้อมูลสูญเสียความหมายหรือความสามารถในการใช้งานสำหรับวัตถุประสงค์บางอย่าง การหาจุดสมดุลระหว่างการป้องกันความเป็นส่วนตัวกับการรักษาความสามารถในการใช้งานข้อมูลจึงเป็นสิ่งที่จำเป็น ที่อาจเกิดขึ้นระหว่างการทำให้ข้อมูลไม่สามารถระบุตัวตนได้ การพิจารณาทั้งเทคนิคและประเด็นด้านจริยธรรมในการจัดทำข้อมูลนิรนามจะถูกหยิบยกมาพูดถึง โดยเน้นย้ำถึงความจำเป็นในการหาจุดสมดุลระหว่างการใช้ประโยชน์จากข้อมูลและการรักษาความเป็นส่วนตัวของบุคคล การจัดทำข้อมูลนิรนามไม่เพียงแต่เกี่ยวข้องกับการลบข้อมูลที่สามารถระบุตัวตนได้โดยตรงเท่านั้น แต่ยังรวมถึงหลักคิดในการวิเคราะห์ว่าข้อมูลที่เหลืออยู่สามารถถูกใช้ในการระบุตัวตนได้หรือไม่ หากมีความเป็นไปได้จำเป็นต้องมีการปรับเปลี่ยนหรือเพิ่มเทคนิคการป้องกันเพื่อลดความเสี่ยงนี้ กระบวนการดังกล่าวต้องการความรู้ทางเทคนิคและความเข้าใจลึกซึ้งเกี่ยวกับทั้งข้อมูลที่จะถูกทำให้เป็นข้อมูลนิรนามและบริบทของการใช้งานข้อมูล เพื่อให้สามารถใช้ประโยชน์จากข้อมูลอย่างเต็มที่ โดยไม่ละเมิดสิทธิ์และความเป็นส่วนตัวของบุคคล

1.2. วัตถุประสงค์

การจัดทำข้อมูลให้เป็นนิรนามเป็นกระบวนการที่สำคัญในการป้องกันข้อมูลส่วนบุคคลและความเป็นส่วนตัว ในขณะเดียวกันก็ยังคงสามารถใช้ประโยชน์จากข้อมูลเหล่านั้นในการวิเคราะห์และการนำไปใช้เพื่อการตัดสินใจซึ่งวัตถุประสงค์หลักของกระบวนการนี้ครอบคลุมหลายด้าน รวมถึงการปกป้องข้อมูลส่วนบุคคล การรักษาความเป็นส่วนตัว และการสนับสนุนการใช้ข้อมูลอย่างมีจริยธรรมและถูกต้องตามกฎหมายที่กำหนด

1. การปกป้องข้อมูลส่วนบุคคล

การปกป้องข้อมูลส่วนบุคคลจากการถูกเข้าถึงหรือใช้งานโดยไม่ได้รับอนุญาต ซึ่งรวมถึงการป้องกันข้อมูลจากการระบุตัวตนได้โดยตรงหรืออ้อม การจัดทำข้อมูลนิรนามช่วยให้แน่ใจว่าข้อมูลที่เก็บรวบรวมและวิเคราะห์ไม่สามารถนำไปสู่การระบุตัวตนของบุคคลได้

2. การรักษาความเป็นส่วนตัว

การรักษาความเป็นส่วนตัวเป็นสิ่งสำคัญในยุคดิจิทัล การจัดทำข้อมูลนิรนามช่วยให้ยกระดับการใช้ข้อมูลในการวิเคราะห์ การวิจัย การพัฒนานวัตกรรมหรือเทคโนโลยีใหม่ และนำไปสู่การตัดสินใจได้ โดยไม่ละเมิดความเป็นส่วนตัวของบุคคลที่ข้อมูลนั้นอ้างอิงถึง

3. การใช้ข้อมูลอย่างมีจริยธรรม

การจัดทำข้อมูลนิรนามช่วยให้สามารถใช้ข้อมูลสำหรับวัตถุประสงค์ที่มีจริยธรรม เช่น การวิจัยทางการแพทย์ การศึกษา หรือ การพัฒนาบริการหรือผลิตภัณฑ์โดยไม่ละเมิดสิทธิ์หรือความเป็นส่วนตัวของบุคคล

4. การปฏิบัติตามกฎหมายและข้อกำหนด

การจัดทำข้อมูลนิรนามช่วยให้องค์กรสามารถปฏิบัติตามกฎหมายและข้อกำหนดที่เกี่ยวข้องกับการคุ้มครองข้อมูลและความเป็นส่วนตัว ไม่ว่าจะเป็น พระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ.2562 ในประเทศไทย , GDPR ในสหภาพยุโรป หรือ CCPA ในแคลิฟอร์เนีย เป็นต้น

5. การเพิ่มมูลค่าข้อมูล

การจัดทำข้อมูลนิรนามยังช่วยให้สามารถเพิ่มมูลค่าของข้อมูลโดยการอนุญาตให้แชร์และวิเคราะห์ข้อมูลได้โดยไม่เสี่ยงต่อการระบุตัวตนของบุคคลที่เกี่ยวข้อง ซึ่งเปิดโอกาสใหม่ ในการใช้ข้อมูลในทางที่เป็นประโยชน์

การจัดทำข้อมูลนิรนามจึงมีบทบาทสำคัญในการสร้างสมดุลระหว่างการใช้ประโยชน์จากข้อมูลและการปกป้องข้อมูลส่วนบุคคลและความเป็นส่วนตัว ด้วยการเน้นย้ำวัตถุประสงค์หลักเหล่านี้ การจัดทำข้อมูลนิรนามสามารถช่วยให้หน่วยงานทั้งภาครัฐ ภาคเอกชน และองค์กร ใช้ข้อมูลได้อย่างมีความรับผิดชอบและเป็นประโยชน์สูงสุด

1.3. ขอบข่าย

แนวทางการจัดทำข้อมูลนิรนาม เวอร์ชัน 1.0 จัดทำขึ้นเพื่อเป็นตัวอย่างให้หน่วยงานภาครัฐนำไปใช้เป็นแนวทางในการจัดทำข้อมูลนิรนาม ซึ่งเป็นกระบวนการจัดทำให้ข้อมูลส่วนบุคคลหรือข้อมูลที่สามารถระบุตัวตนได้มาอยู่ในรูปแบบที่ไม่สามารถระบุตัวตนเจ้าของข้อมูลได้ เพื่อลดความเสี่ยงการระบุตัวตนของเจ้าของข้อมูล และนำข้อมูลนำไปใช้ประโยชน์ต่อได้โดยไม่ขัดกับกฎหมาย เช่น การนำข้อมูลไปวิเคราะห์เพื่อขับเคลื่อนงานตามภารกิจองค์กร การแบ่งปันระหว่างหน่วยงานภาครัฐ เพื่อสร้างความมั่นใจให้แก่หน่วยงานภาครัฐในการสามารถนำข้อมูลไปใช้ประโยชน์ ทั้งยังเป็นการสร้างความเชื่อมั่นให้แก่ประชาชนถึงการใช้ข้อมูลของภาครัฐ โดยแนวทางฉบับนี้ได้จัดทำตามแนวมาตรฐานและแนวปฏิบัติที่ดีของ

1.3.1 ประกาศคณะกรรมการพัฒนารัฐบาลดิจิทัล เรื่อง ธรรมาภิบาลข้อมูลภาครัฐ เวอร์ชัน 1.0

1.3.2 มรด. 6 : 2566 มาตรฐานรัฐบาลดิจิทัลว่าด้วยกรอบธรรมาภิบาลข้อมูลภาครัฐ ฉบับปรับปรุง: แนวปฏิบัติ

1.3.3 มสพร. 8-2565 มาตรฐานสำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน) ว่าด้วยหลักเกณฑ์การจัดระดับชั้นและการแบ่งปันข้อมูลภาครัฐ เวอร์ชัน 1.0

1.3.4 ศูนย์วิจัยกฎหมายและการพัฒนา คณะนิติศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย,แนวปฏิบัติเกี่ยวกับการคุ้มครองข้อมูลส่วนบุคคล 3.0 (TDPG 3.0)

1.3.5 มาตรฐาน NISTIR 8053: De-identification of Personal Information

1.3.6 The Personal Data Protection Commission ('PDPC'), Guide to Basic Data Anonymisation (31 March 2022)

1.4. บทนิยาม

ข้อมูลส่วนบุคคล (Personal Data) หมายความว่า ข้อมูลเกี่ยวกับบุคคลซึ่งทำให้สามารถระบุตัวบุคคลนั้นได้ ไม่ว่าทางตรงหรือทางอ้อม แต่ไม่รวมถึงข้อมูลของผู้ถึงแก่กรรมโดยเฉพาะ

ผู้ควบคุมข้อมูลส่วนบุคคล (Data Controller) หมายความว่า บุคคลหรือนิติบุคคลซึ่งมีอำนาจหน้าที่ตัดสินใจเกี่ยวกับการเก็บรวบรวม ใช้ หรือเปิดเผยข้อมูลส่วนบุคคล

ผู้ประมวลผลข้อมูลส่วนบุคคล (Data Processor/Personal Data Processor) หมายความว่า บุคคลหรือนิติบุคคลซึ่งดำเนินการเกี่ยวกับ การเก็บรวบรวม ใช้ หรือเปิดเผยข้อมูลส่วนบุคคลตามคำสั่งหรือในนามของผู้ควบคุมข้อมูลส่วนบุคคล ทั้งนี้ บุคคลหรือนิติบุคคลซึ่งดำเนินการดังกล่าวไม่เป็นผู้ควบคุมข้อมูลส่วนบุคคล

ข้อมูลนิรนาม (Anonymous Data) หมายความว่า ข้อมูลที่ผ่านกระบวนการซึ่งทำให้ไม่สามารถระบุตัวตนหรือแสดงตัวตนได้ทั้งในปัจจุบันและอนาคต และไม่เป็นข้อมูลส่วนบุคคลตามกฎหมายว่าด้วยการคุ้มครองข้อมูลส่วนบุคคล

การจัดทำข้อมูลนิรนาม (Data Anonymization) หมายความว่า กระบวนการทำให้ข้อมูลส่วนบุคคลไม่สามารถระบุหรือเชื่อมโยงข้อมูลไปถึงตัวบุคคลได้ทั้งในปัจจุบันและอนาคต โดยใช้เทคนิคหลายเทคนิคร่วมกันจนมั่นใจว่าไม่สามารถระบุตัวตนได้ตามกฎหมายว่าด้วยการคุ้มครองข้อมูลส่วนบุคคล

การจัดทำข้อมูลแฝง (Data Pseudonymization) หมายความว่า กระบวนการเปลี่ยนแปลงข้อมูลส่วนบุคคลด้วยการใช้อักขระแฝงหรือวิธีการอื่นใด เช่น การเข้ารหัสข้อมูล (Encryption) การเข้าฟังก์ชันแฮช (Hashing) การเก็บข้อมูลแยกส่วนโดยเชื่อมผ่านโทเค็น (Tokenization) โดยยังสามารถเชื่อมโยงข้อมูลเพื่อระบุตัวตนได้เมื่อมีข้อมูลเพิ่มเติมประกอบและไม่ถือเป็นข้อมูลนิรนาม

ข้อมูลส่วนบุคคลรั่วไหล (Personal Data Breach) หมายความว่า การรั่วไหลหรือละเมิดมาตรการความมั่นคงปลอดภัยต่อข้อมูลส่วนบุคคลทำให้เกิดความเสียหาย สูญหาย เปลี่ยนแปลง เปิดเผยหรือเข้าถึงข้อมูลส่วนบุคคลที่ใช้งาน โดยไม่ได้รับอนุญาต

1.5. กฎหมายและแนวทางที่เกี่ยวข้อง

1.5.1. พระราชบัญญัติข้อมูลข่าวสารของทางราชการ พ.ศ. 2540 มาตรา 21- 25

1.5.2. ระเบียบว่าด้วยการรักษาความลับของทางราชการ พ.ศ. 2544 และที่แก้ไขเพิ่มเติม

1.5.3. พระราชบัญญัติการบริหารงานและการให้บริการภาครัฐผ่านระบบดิจิทัล พ.ศ. 2562 มาตรา 7 และมาตรา 8 ธรรมาภิบาลข้อมูลภาครัฐ

1.5.4. พระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562 มาตรา 37 (3) ประกอบ มาตรา 33 วรรค 5

1.5.5. ประกาศคณะกรรมการพัฒนารัฐบาลดิจิทัลเรื่อง ธรรมาภิบาลข้อมูลภาครัฐ ข้อ 4 ธรรมาภิบาลข้อมูลภาครัฐในระดับหน่วยงาน (5) การจำแนกหมวดหมู่ของข้อมูล เพื่อกำหนดนโยบายข้อมูลหรือกฎเกณฑ์เกี่ยวกับผู้มีสิทธิเข้าถึงและใช้ประโยชน์จากข้อมูล ภายในหน่วยงาน สำหรับให้ผู้ซึ่งมีหน้าที่เกี่ยวข้องปฏิบัติตามนโยบายหรือกฎเกณฑ์ได้อย่างถูกต้อง และสอดคล้องตามกฎหมายที่เกี่ยวข้อง อันจะนำไปสู่การบริหารจัดการข้อมูลภาครัฐอย่างเป็นระบบ

2. แนวคิดการจัดทำข้อมูลนิรนาม

2.1. ความสำคัญของการจัดทำข้อมูลนิรนาม

หลายหน่วยงานทั้งภาครัฐและเอกชนหันมาให้ความสนใจเรื่องข้อมูลส่วนบุคคลที่เพิ่มขึ้น เนื่องจากเป็นข้อมูลที่มีความสำคัญในระดับบุคคล สามารถนำไปใช้วิเคราะห์ประมวลผลเพื่อก่อให้เกิดประโยชน์ต่อหน่วยงานได้เป็นอย่างสูง ในทางกลับกันการนำข้อมูลส่วนบุคคลไปใช้ในทางที่มิชอบก็จะก่อให้เกิดความเสียหายต่อบุคคลที่เป็นเจ้าของข้อมูลได้ ปัจจุบันจึงมีพระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562 ลงประกาศในราชกิจจานุเบกษาเมื่อวันที่ 27 พฤษภาคม 2562 โดยมีผลบังคับใช้ในวันที่ 28 พฤษภาคม 2563 โดยในมาตรา 5 แห่งพระราชบัญญัตินี้ให้ใช้บังคับแก่การเก็บรวบรวม ใช้ หรือเปิดเผยข้อมูลส่วนบุคคลโดยผู้ควบคุมข้อมูลส่วนบุคคลหรือผู้ประมวลผลข้อมูลส่วนบุคคลซึ่งอยู่ในราชอาณาจักร ไม่ว่า การเก็บรวบรวม ใช้ หรือเปิดเผยนั้น ได้กระทำในหรือนอกราชอาณาจักรก็ตาม ซึ่งเป็นกฎหมายสำคัญที่ช่วยเรื่องการคุ้มครองข้อมูลส่วนบุคคล ป้องกันการละเมิดข้อมูลส่วนบุคคล รวมถึงการจัดเก็บและนำไปใช้โดยไม่ได้รับความยินยอมจากเจ้าของข้อมูลเสียก่อน โดยมีสำนักงานคณะกรรมการคุ้มครองข้อมูลส่วนบุคคล เป็นหน่วยงานที่มีวัตถุประสงค์เกี่ยวกับการคุ้มครองข้อมูลส่วนบุคคล รวมทั้งส่งเสริมและสนับสนุนให้เกิดการพัฒนาด้านการคุ้มครองข้อมูลส่วนบุคคลของประเทศ

ข้อมูลส่วนบุคคลถือเป็น 1 ใน 5 หมวดหมู่ของข้อมูล ตาม มรด. 6 : 2566 มาตรฐานรัฐบาลดิจิทัล ว่าด้วยกรอบธรรมาภิบาลข้อมูลภาครัฐ ฉบับปรับปรุง: แนวปฏิบัติ

รูปที่ 1: หมวดหมู่ของข้อมูล

สามารถศึกษาข้อมูลเพิ่มเติมได้โดยการแสกน QR Code
มรด. 6 : 2566 มาตรฐานรัฐบาลดิจิทัลว่าด้วยกรอบธรรมาภิบาลข้อมูลภาครัฐ ฉบับปรับปรุง: แนวปฏิบัติ

ซึ่งเป็นการบริหารจัดการข้อมูล โดยการจัดหมวดหมู่ข้อมูลเป็นส่วนหนึ่งของการกำกับดูแลข้อมูลให้ข้อมูลมีคุณภาพ และเชื่อมโยงข้อมูลด้วยกันอย่างมั่นคงปลอดภัย ภายใต้นโยบายและแนวปฏิบัติด้านการจัดทำธรรมาภิบาลข้อมูลของหน่วยงานภาครัฐนั้น เป็นการยกระดับให้บริการของหน่วยงานภาครัฐด้านดิจิทัลให้มีประสิทธิภาพ นอกจากนี้เพื่อให้ข้อมูลนำไปสู่การเปิดเผยและแบ่งปันข้อมูล จึงจำเป็นต้องมีการจัดระดับชั้นข้อมูลซึ่งเป็นการบริหารจัดการข้อมูลภายในหน่วยงานก่อนการเปิดเผย โดยพิจารณาการจัดระดับชั้นข้อมูลภาครัฐที่มีความอ่อนไหวให้สอดคล้องตามเกณฑ์การแบ่งระดับชั้นข้อมูลภาครัฐ ตามที่ สพร. ได้ประกาศเป็นข้อเสนอแนะให้กับหน่วยงานภาครัฐสามารถนำไปปฏิบัติใช้ตาม มสพร. 8-2565 มาตรฐานสำนักงานพัฒนารัฐบาลดิจิทัลว่าด้วยหลักเกณฑ์การจัดระดับชั้นและการแบ่งปันข้อมูลภาครัฐ

Data Class. Level / Data Category	เปิดเผย (Open)	เผยแพร่ภายในองค์กร (Private)	ลับ (Confidential / sensitive)	ลับมาก (Secret / Medium Sensitive)	ลับที่สุด (Top secret / Highly Sensitive)
ข้อมูลสาธารณะ	พ.ร.บ. ข้อมูลข่าวสารของทางราชการ 2540 (มาตรา 7, 9) และ แนวทางการเปิดเผยข้อมูลภาครัฐ				
ข้อมูลใช้ภายใน		ISO 27001: 2013			
ข้อมูลส่วนบุคคล		พ.ร.บ. คุ้มครองข้อมูลส่วนบุคคล 2562 (มาตรา 24 – มาตรา 27) พ.ร.บ. ข้อมูลข่าวสารของทางราชการ 2540 (มาตรา 9 และมาตรา 15 ที่เปิดเผยได้)			
ข้อมูลข่าวสารลับ			• ระเบียบว่าด้วยการรักษาความลับของทางราชการ 2544 • ระเบียบสำนักนายกรัฐมนตรีว่าด้วยการรักษาความปลอดภัยแห่งชาติ พ.ศ. 2552		พ.ร.บ. ข้อมูลข่าวสารของทางราชการ 2540 (มาตรา 14 – 15 อาจมีคำสั่งมิให้เปิดเผย)
ข้อมูลความมั่นคง			• นโยบายและแผนระดับชาติ ว่าด้วยความมั่นคงแห่งชาติ		

รูปที่ 2: การจัดหมวดหมู่และระดับชั้นข้อมูลภาครัฐ

สามารถศึกษาข้อมูลเพิ่มเติมได้โดยการแสกน QR Code
มสพร. 8-2565 มาตรฐานสำนักงานพัฒนารัฐบาลดิจิทัลว่าด้วยหลักเกณฑ์การจัดระดับชั้นและแบ่งปันข้อมูลภาครัฐ

อย่างไรก็ดี ถึงแม้ว่ามีกฎหมายที่บัญญัติขึ้น เพื่อคุ้มครองข้อมูลส่วนบุคคล หรือ มีการส่งเสริมให้หน่วยงานภาครัฐมีการจัดทำธรรมาภิบาลข้อมูลภาครัฐ เพื่อกำกับดูแลข้อมูลเป็นอย่างดีแล้วนั้น ปัจจุบันยังพบว่ายังมีปัญหาการหลุดรั่วของข้อมูลส่วนบุคคลที่เกิดขึ้นตลอดหลายปีที่ผ่านมา ซึ่ง PDPA Thailand ได้มีการรวบรวมเหตุการณ์สำคัญ ที่เกิดขึ้นมาตั้งแต่ปี พ.ศ. 2561 จนถึง พ.ศ. 2566 เผยแพร่ผ่านทางเว็บไซต์

รูปที่ 3: ตัวอย่างข้อมูลรั่วไหล (PDPA Thailand, 2023)

เห็นได้ชัดว่าไม่ว่าหน่วยงานภาครัฐหรือหน่วยงานเอกชนยังคงมีความเสี่ยงในการถูกโจมตี ก่อให้เกิดการรั่วไหลของข้อมูลได้ ถึงแม้จะมีมาตรการในการปกป้องคุ้มครองข้อมูล หรือมีระบบป้องกันที่มีประสิทธิภาพ แต่ยังเกิดปัญหาการรั่วไหลของข้อมูลได้ซึ่งอาจเกิดจากแฮกเกอร์หรือกลุ่มผู้ไม่หวังดีที่ต้องการนำข้อมูลส่วนบุคคลไปใช้ในเชิงพาณิชย์ด้วยการโจรกรรมข้อมูลโดยอาศัยกระบวนการทำงานใดกระบวนการหนึ่ง

ที่หละหลวมในการดูแลข้อมูลเป็นช่องว่างให้ ทำให้ข้อมูลส่วนบุคคลยังคงเป็นเป้าหมายสำคัญที่อาจสร้างมูลค่าให้กับผู้ไม่หวังดี แต่กลับสร้างผลกระทบเป็นวงกว้างในระดับบุคคลได้ เช่น ข้อมูลทางการแพทย์ ข้อมูลบัตรเครดิตหรือข้อมูลทางการเงิน เป็นต้น ดังนั้นการจัดทำข้อมูลนิรนามจึงเป็นเครื่องมือสำคัญที่ช่วยในการปกป้องคุ้มครองข้อมูลส่วนบุคคลและลดความเสี่ยงที่อาจก่อให้เกิดความเสียหายที่อาจเกิดขึ้นกับตัวบุคคลได้ ในขณะเดียวกันเป็นการสร้างความมั่นใจแก่หน่วยงานภาครัฐ ให้สามารถนำข้อมูลมาใช้สำหรับวิเคราะห์ วิจัย และพัฒนานวัตกรรม หรือเทคโนโลยีใหม่ โดยไม่เปิดเผยตัวตนของบุคคลผู้ที่เกี่ยวข้องกับข้อมูล

รูปที่ 4: การปกป้องคุ้มครองข้อมูลส่วนบุคคลด้วยการทำให้เป็นข้อมูลนิรนาม

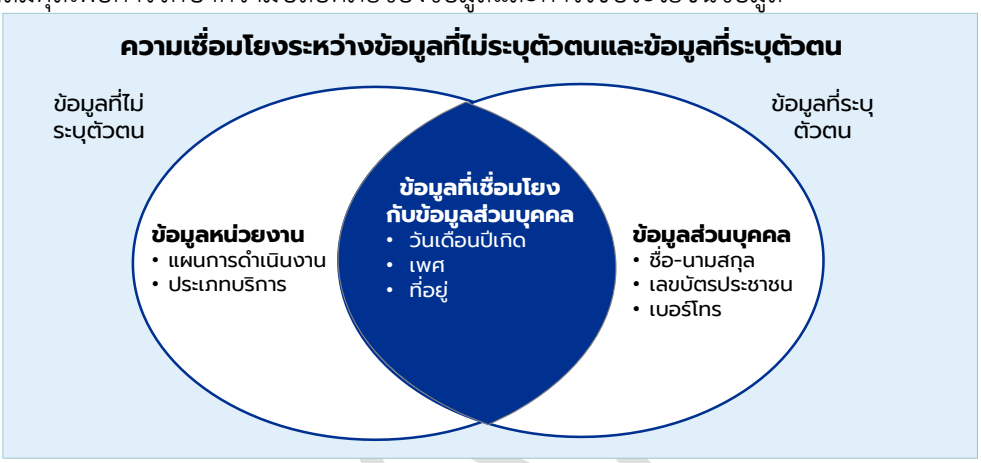
2.2. แนวคิดการจัดทำข้อมูลนิรนาม

การจัดทำข้อมูลนิรนามเป็นการสร้างความมั่นใจให้แก่หน่วยงานภาครัฐในการนำข้อมูลไปใช้ในการวิเคราะห์หรือเพื่อประโยชน์ ได้อย่างถูกต้องตามกฎหมายและมีธรรมาภิบาล และยังสร้างความเชื่อมั่นให้แก่ประชาชนต่อแนวทางการใช้ข้อมูลของภาครัฐ ในการคุ้มครองข้อมูลให้มีความปลอดภัยและรักษาความเป็นส่วนตัว โดยข้อมูลที่มีความอ่อนไหวที่ต้องใช้ความระมัดระวังเป็นพิเศษ คือข้อมูลที่ผ่านการจัดระดับชั้นข้อมูลตามมาตรฐานสำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน) ว่าด้วยหลักเกณฑ์การจัดระดับชั้นและการแบ่งปันข้อมูลภาครัฐ (มสพร. 8-2565) โดยข้อมูลอ่อนไหวที่กล่าวมาจะครอบคลุมถึง 1) ข้อมูลหมวดหมู่ข้อมูลส่วนบุคคลที่มีระดับชั้นข้อมูลระดับเผยแพร่ภายในองค์กร (Private) - ลับที่สุด (Top Secret) 2) ข้อมูลใช้ภายในหน่วยงานข้อมูลส่วนบุคคลที่มีระดับชั้นข้อมูลระดับเผยแพร่ภายในองค์กร (Private) - ลับที่สุด (Top Secret) เช่น ข้อมูลทางการเงิน และ 3) ข้อมูลในหมวดหมู่ข้อมูลความลับทางราชการและข้อมูลความมั่นคง ซึ่งข้อมูลอ่อนไหวที่กล่าวมาอาจมีข้อมูลที่ระบุตัวตนได้รวมอยู่ด้วย เช่น ในบางกรณีข้อมูลส่วนบุคคลและข้อมูลหน่วยงานสามารถเป็นข้อมูลที่ทับซ้อนกันได้ เช่น ข้อมูลเพศ ข้อมูลที่อยู่ ซึ่งเป็นข้อมูลส่วนบุคคลที่สามารถเชื่อมโยงไปยังตัวบุคคลได้ และเป็นข้อมูลที่หน่วยงานนำไปใช้ประโยชน์เพื่อดำเนินการตามภารกิจของหน่วยงาน ซึ่งในการนำใช้ข้อมูลที่อยู่ในพื้นที่ทับซ้อนไปใช้ประโยชน์อาจมีความเสี่ยงที่จะนำไปสู่การระบุตัวตนได้ สามารถสรุปได้ 3 รูปแบบ (Article 29 Data Protection Working Party (European Commission), 2014) ดังนี้

- การแปลกแยกออกจากกลุ่ม (Singling out) คือ การที่ข้อมูลสามารถระบุตัวตนได้ เนื่องจากข้อมูลมีลักษณะแปลกแยกจากกลุ่มมากเป็นพิเศษ ส่งผลให้เชื่อมโยงไปยังข้อมูลส่วนบุคคลได้ เช่น ชื่อ-นามสกุล เลขบัตรประชาชน
- ความสามารถเชื่อมโยง (Linkability) คือ ข้อมูลบ่งชี้ทางอ้อมที่สามารถเชื่อมโยงไปยังข้อมูลส่วนบุคคล หรือข้อมูลที่บ่งชี้ทางตรง และจะสามารถระบุตัวตนได้หากมีการนำข้อมูลไปเชื่อมโยงกับข้อมูลอื่นประกอบกัน เช่น อายุ เพศ กรุ๊ปเลือด วันเดือนปีเกิด

● การอนุมาน (Inference) คือ การที่ตัวตนถูกระบุได้เนื่องจากสามารถคาดเดาค่าจริงของข้อมูลส่วนที่ถูกอำพรางได้ โดยอาศัยการตีความจากข้อมูลอื่นประกอบ เช่น การพิจารณาข้อมูลระหว่างตำแหน่งงาน เพศ และอายุงานก็จะสามารถเชื่อมโยงไปยังตัวบุคคลได้

ความเสี่ยงในการระบุตัวตนอาจส่งผลไปต่อการใช้ประโยชน์ข้อมูล โดยข้อมูลที่สามารถระบุตัวตนได้ก็จะมีความเสี่ยงในการระบุตัวตนสูง ส่งผลให้ระดับการนำไปใช้ประโยชน์ค่อนข้างต่ำ เพราะต้องการคุ้มครองข้อมูลให้ปลอดภัย ในขณะที่ข้อมูลที่ไม่สามารถเชื่อมโยงหรือระบุตัวตนได้จะสะท้อนว่า ข้อมูลนั้นมีระดับความเสี่ยงในการระบุตัวตนต่ำ ส่งผลให้ข้อมูลสามารถนำไปใช้ประโยชน์ได้มากขึ้น หน่วยงานจึงต้องมีการพิจารณาสร้างความสมดุลเพื่อการรักษาความปลอดภัยของข้อมูลและการใช้ประโยชน์ข้อมูล



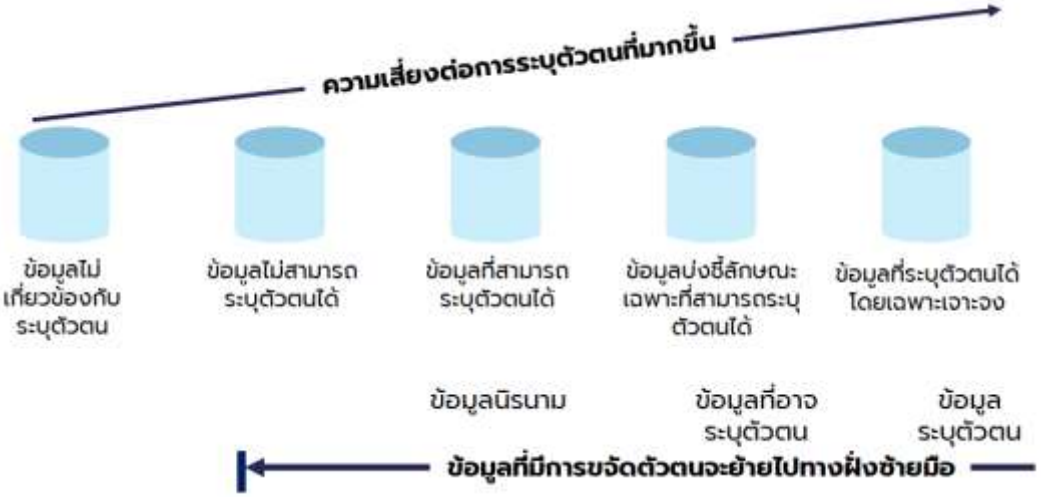
รูปที่ 5: ความเชื่อมโยงระหว่างข้อมูลที่ไม่ระบุตัวตนกับข้อมูลที่ระบุตัวตน

ดังนั้น การจัดทำข้อมูลนิรนามจะเป็นเครื่องมือในการขจัดตัวตนหรือลดความเสี่ยงในการระบุตัวตนได้ โดยมีหลักการพิจารณาข้อมูลเพื่อจัดทำข้อมูลนิรนามจะเป็นการพิจารณาระหว่าง (1) คุณค่าจากการใช้ประโยชน์ของข้อมูล (Value) (2) การรักษาความลับของเจ้าของข้อมูล (Confidentiality) ซึ่งสอดคล้องตาม CIA ซึ่งน้ำหนักของการรักษาความลับของเจ้าของข้อมูลนั้น (Confidentiality) จะต้องไม่มากเกินไปกว่าคุณค่าจากการใช้ประโยชน์ของข้อมูล (Value) ก็ย่อมถือว่ามีการจัดทำข้อมูลนิรนามในระดับที่เหมาะสมเพื่อลดความเสี่ยงต่อการระบุตัวตนที่มากขึ้นให้สามารถนำข้อมูลมาใช้ประโยชน์ต่อได้

ในรูปจะเห็นว่า ข้อมูลด้านซ้ายสุดจะเป็นข้อมูลที่ไม่เกี่ยวข้องกับการระบุตัวตน เช่น แผนการดำเนินงาน ซึ่งเป็นข้อมูลที่ไม่มีความเสี่ยงต่อการระบุตัวตน แต่ในทางกลับกันในส่วนของข้อมูลด้านขวาสุดจะเป็นข้อมูลที่ระบุตัวตนได้โดยเฉพาะเจาะจง เช่น ชื่อ-นามสกุล ซึ่งสามารถเชื่อมโยงไปยังตัวตนบุคคลได้โดยตรง ดังนั้น ข้อมูลที่ผ่านการขจัดตัวตนก็จะกลายเป็นข้อมูลนิรนาม เพื่อป้องกันไม่ให้สามารถระบุตัวตนได้ โดยข้อมูลยังคงคุณค่าจากการใช้ประโยชน์ของข้อมูล (Value) ที่ต้องการไว้ได้ ซึ่งจะทำให้หน่วยงานสามารถนำข้อมูลมาใช้ประโยชน์ต่อได้ (Garfinkel, October 2015)

การนำข้อมูลไปประมวลผลการดำเนินการของหน่วยงาน ซึ่งรวมไปถึงการแบ่งปันและเปิดเผยข้อมูลต่อหน่วยงานอื่น จะต้องมีการจัดทำข้อมูลนิรนาม โดยแบ่งความเข้มข้นในการจัดทำข้อมูลนิรนามคือ การแฝงข้อมูล (Pseudonymization) เป็นวิธีการในการแทนที่สิ่งที่ระบุตัวตนของเจ้าของข้อมูลโดยตรงและสามารถถอดรหัส/แปลงข้อมูลให้ย้อนกลับไปเป็นข้อมูลที่ระบุตัวตนได้ ซึ่งยังคงถือว่าเป็นข้อมูลส่วนบุคคล การขจัดตัวตน (De-identification) คือการลบข้อมูลในส่วนที่จะเชื่อมโยงไปข้อมูลที่ระบุตัวตนได้ การจัดทำข้อมูล

นิรนาม (Anonymization) คือการขจัดตัวตนทำให้ไม่สามารถระบุตัวตนได้ ซึ่งรวมถึงการแฝงข้อมูล เพื่อลดความเสี่ยงในการระบุตัวตน โดยหน่วยงานสามารถพิจารณาได้ตามที่จะกล่าวต่อไปในบทที่ 3



รูปที่ 6: ความเสี่ยงต่อการระบุตัวตน

ทั้งนี้ แม้ว่าข้อมูลที่จัดทำเป็นข้อมูลนิรนามจะไม่ถือว่าเป็นข้อมูลส่วนบุคคลหรือข้อมูลที่ระบุตัวตนได้ทางอ้อม และอยู่นอกขอบเขตการคุ้มครองตามกฎหมายว่าด้วยการคุ้มครองข้อมูลส่วนบุคคล แต่เนื่องจากเป็นการใช้ข้อมูลที่มีลักษณะอ่อนไหว หน่วยงานจึงควรมีการใช้งานอย่างระมัดระวัง (Garfinkel, October 2015)

รูปที่ 7: กรอบแนวคิดในการจัดทำข้อมูลนิรนาม

การจัดทำข้อมูลนิรนามที่กล่าวมาข้างต้น เป็นการทำให้ข้อมูลส่วนบุคคลไม่สามารถระบุตัวตนของบุคคลได้ เพื่อวัตถุประสงค์คุ้มครองข้อมูลส่วนบุคคล โดยการนำข้อมูลที่ใช้ระบุตัวตนของบุคคลได้ในชุดข้อมูลมาเข้ารหัส ลบ หรือทำลาย สรุปขั้นตอนดังนี้ (สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ, 2023)

- ขั้นตอนที่ 1: การพิจารณาว่าข้อมูลระบุตัวตนได้หรือไม่ : การพิจารณาประเภทข้อมูล เช่น ข้อมูลตัวบ่งชี้ทางตรง ข้อมูลตัวบ่งชี้ทางอ้อม ข้อมูลคุณลักษณะเฉพาะ (บทที่ 3.1)
- ขั้นตอนที่ 2: การเลือกใช้เทคนิคข้อมูลนิรนามกับข้อมูลบ่งชี้ทางตรง (บทที่ 2.3 และ 3.2)
- ขั้นตอนที่ 3: การเลือกใช้เทคนิคข้อมูลนิรนามกับข้อมูลบ่งชี้ทางอ้อม (บทที่ 2.3 และ 3.2)

- ขั้นตอนที่ 4: การประเมินความเสี่ยง : การพิจารณาสถานการณ์ของข้อมูล และการวิเคราะห์ความเสี่ยง (บทที่ 3.3)
- ขั้นตอนที่ 5: การบริหารความเสี่ยง : การกำหนดมาตรการจัดการความเสี่ยง (บทที่ 3.3) และ*การทบทวนกระบวนการ* (กลับไปขั้นตอนที่ 1 หากยังมีความเสี่ยง)

รูปที่ 8: ขั้นตอนการจัดทำข้อมูลนิรนาม

ตัวอย่างหน่วยงานที่มีการจัดทำข้อมูลนิรนาม

เมื่อหน่วยงานมีการจัดทำข้อมูลนิรนามแล้วก็จะสามารถเปิดเผย หรือแบ่งปันกับหน่วยงานภายนอกได้ โดยการเข้ารหัสในการรับส่งข้อมูลเพื่อคุ้มครองความปลอดภัยของข้อมูล ซึ่งข้อมูลที่ผ่านการทำเป็นข้อมูลนิรนามอาจส่งผลให้ระดับความเสี่ยงของข้อมูลมีการเปลี่ยนแปลง ซึ่งควรมีการจัดระดับชั้นข้อมูลใหม่ (Declassification) ตัวอย่างเช่น หากข้อมูลหมวดหมู่ใช้ภายในผ่านการจัดทำข้อมูลนิรนามจะได้ชุดข้อมูลใหม่ที่อาจมีการเปลี่ยนจากระดับชั้นลับเป็นเผยแพร่ภายในองค์กร เอกสารต้นฉบับยังคงอยู่ในระดับชั้นเดิม เป็นต้น

รูปที่ 9: ตัวอย่างการจัดทำข้อมูลนิรนาม

2.3. วิธีการจัดทำข้อมูลนิรนาม

เทคนิคที่ใช้ในการประมวลผลข้อมูลเพื่อปกป้องความเป็นส่วนตัวของผู้ใช้และข้อมูลที่ละเอียดอ่อนหลากหลายเทคนิคด้วยกัน ไม่ว่าจะเป็นการลบหรือการเข้ารหัสข้อมูลที่สามารถระบุตัวบุคคลได้ในชุดข้อมูล เป้าหมายคือเพื่อให้มั่นใจถึงความเป็นส่วนตัวของข้อมูล การลบข้อมูลระบุตัวตนจะช่วยลดความเสี่ยงของการรั่วไหลของข้อมูลเมื่อข้อมูลถูกย้ายข้ามขอบเขต หรือหากถูกผู้ไม่หวังดี โจรกรรมข้อมูลไป นอกจากนี้ยังรักษาโครงสร้างของข้อมูล ทำให้สามารถวิเคราะห์ข้อมูลภายหลังการลบ ข้อมูลระบุตัวตนได้ ซึ่งจากการศึกษาค้นคว้า จึงทำการรวบรวมเทคนิคที่ใช้หลากหลายเทคนิคมาให้ผู้อ่านได้ทราบและสามารถนำไปปรับใช้กับข้อมูลของตนได้อย่างเหมาะสม โดยเทคนิคที่ใช้มีดังต่อไปนี้ (สถาบันข้อมูลขนาดใหญ่, 2021) (Satori Cyber Ltd., 2021)

รูปที่ 10: วิธีการจัดทำข้อมูลนิรนาม

ตารางที่ 1: วิธีการจัดทำข้อมูลนิรนาม

เทคนิคการลบข้อมูลระบุตัวตน	คำอธิบาย
1.Attribute Suppression **การลบคุณลักษณะเฉพาะ** (การลบข้อมูลรายคอลัมน์)	การลบข้อมูลรายคอลัมน์ : การลบคุณลักษณะข้อมูล (เช่น คอลัมน์ข้อมูลที่เป็นข้อมูลเฉพาะ หรือ ข้อมูลทางตรง วิธีดังกล่าวจะใช้ต่อเมื่อ ในกรณีที่ไม่จำเป็นต้องใช้ข้อมูลดังกล่าวในชุดข้อมูล และ ไม่ต้องการเปิดเผยตัวตนค่าข้อมูลเฉพาะนั้นอีกต่อไป)
2.Record Suppression **การลบข้อมูลรายบันทึก** (การลบข้อมูลรายแถว)	การลบข้อมูลรายบันทึก เช่น แถวข้อมูล โดยเฉพาะอย่างยิ่งเมื่อข้อมูลดังกล่าวอาจมีค่าข้อมูลที่ไม่สามารถทำให้เป็นข้อมูลที่ไม่สามารถระบุตัวตนได้ต่อไป.
3.Character Masking **การปิดทับลักษณะข้อมูล**	การปิดบังข้อมูลเกี่ยวข้องกับการอนุญาตให้เข้าถึงข้อมูลที่เป็นความลับในรูปแบบที่ถูกแก้ไข สามารถทำได้โดยการแก้ไขข้อมูลในขณะที่เข้าถึง (การปิดบังข้อมูลแบบเปลี่ยนแปลงตลอดเวลา), หรือโดยการสร้างฐานข้อมูลเสมือนที่มีข้อมูลที่ถูกทำให้เป็นนิรนาม (การปิดบังข้อมูลแบบคงที่) การทำข้อมูลให้เป็นนิรนามสามารถทำได้ผ่านเทคนิคหลายเทคนิค รวมถึงการเข้ารหัส, การสับเปลี่ยนคำหรือตัวอักษร, หรือการแทนที่ด้วยคำจากพจนานุกรม ตัวอย่าง ของการปิดทับข้อมูล: - การแทนที่รายละเอียดและชื่อที่ระบุตัวบุคคลด้วยสัญลักษณ์และอักขระอื่น - การย้ายรายละเอียดไปจุดอื่นหรือการสุ่มข้อมูลที่ละเอียดอ่อน เช่น ชื่อหรือหมายเลขบัญชี - การแทนที่บางส่วนด้วยส่วนอื่นจากชุดข้อมูลเดียวกัน

เทคนิคการลบข้อมูลระบุตัวตน	คำอธิบาย
	- การลบหรือ " ลบล้าง" ค่าที่ละเอียดอ่อนภายในบันทึกข้อมูล - การเข้ารหัสข้อมูลเพื่อให้ผู้ใช้ที่ไม่ได้รับอนุญาตไม่สามารถเข้าถึงได้โดยไม่ต้องใช้คีย์ถอดรหัส ซึ่งการปกป้องตัวอักษร: การแทนที่บางตัวอักษรของค่าข้อมูลด้วยสัญลักษณ์ที่เป็นคงที่ (เช่น * หรือ x) เป็นต้น. **ตัวอย่างเช่น** การปกป้องรหัสไปรษณีย์โดยการเปลี่ยนจาก "10400" เป็น "10xxx".
4.Pseudonymization การแฝงข้อมูล	การใช้นามแฝงเป็นวิธีการหนึ่งของการลบข้อมูลระบุตัวตน โดยจะแทนที่ตัวระบุส่วนตัวด้วยนามแฝงหรือตัวระบุที่เป็นเท็จ เช่น ชื่อ “นาย ธนัชกฤศ” อาจเปลี่ยนเป็น “นาย สมชาย” เพื่อช่วยให้มั่นใจได้ถึงการรักษาความลับของข้อมูลและความแม่นยำทางสถิติ (ค่าที่สร้างขึ้นมาใหม่นี้ควรจะเป็นเอกลักษณ์และไม่ควรมีความสัมพันธ์กับค่าเดิม เพื่อไม่ให้สามารถหาค่าเดิมจากนามแฝงได้) โดยเทคนิคพื้นฐานในการแฝงข้อมูล เช่น - การเข้ารหัสข้อมูล (Encryption) - การเข้าฟังก์ชันแฮช (Hashing) - การเก็บข้อมูลแยกส่วนโดยเชื่อมผ่านโทเค็น (Tokenization)
5.Generalization การทำให้ข้อมูลเป็นสามัญ	การทำให้ข้อมูลเป็นสามัญ จำเป็นต้องยกเว้นข้อมูลบางอย่างเพื่อให้สามารถระบุตัวตนได้น้อยลง ข้อมูลสามารถเปลี่ยนแปลงเป็นช่วงของค่าที่มีขอบเขตตรรกะ ตัวอย่างเช่น อาจละเว้นเลขที่บ้านตามที่อยู่ที่ระบุ หรือแทนที่ด้วยช่วงภายใน 200 เลขที่บ้านของมูลค่าเดิม แนวคิดคือการลบตัวบ่งชี้บางอย่างออกโดยไม่กระทบต่อความถูกต้องของข้อมูล หรือ การลดความละเอียดของข้อมูล เช่น โดยการแปลงอายุของบุคคลเป็นช่วงอายุ. **ตัวอย่าง** การทำให้ข้อมูลอายุบุคคลจาก "26 ปี" เป็น "25-29 ปี"
6.Swapping/Shuffling/ Permutation การสลับข้อมูล	การสลับข้อมูล หรือที่เรียกว่าการสับเปลี่ยนข้อมูล จะจัดเรียงค่าข้อมูลเป็นชุดข้อมูลใหม่เพื่อไม่ให้ตรงกับข้อมูลเริ่มต้น การสลับคอลัมน์ ที่แสดงค่าที่จดจำได้ รวมถึงวันเกิด ตำแหน่ง เงินเดือน ซึ่งอาจมีอิทธิพลอย่างมากต่อการไม่ระบุตัวตน
7.Data Perturbation การรบกวนข้อมูล	การรบกวนข้อมูล เปลี่ยนชุดข้อมูลเริ่มต้นเล็กน้อยโดยใช้วิธีการปัดเศษและสัญญาณรบกวนแบบสุ่ม ค่าที่ใช้จำเป็นต้องสัมพันธ์กับการรบกวนที่ใช้ สิ่งสำคัญคือต้องเลือกฐานที่ใช้ในการแก้ไขค่าเดิมอย่างระมัดระวัง หากฐานเล็กเกินไป ข้อมูลจะไม่ถูกทำให้เป็นนิรนามอย่างเพียงพอ และหากฐานใหญ่เกินไป ข้อมูลอาจไม่สามารถระบุหรือใช้งานได้ การปรับค่าในข้อมูลโดยการเพิ่ม "สัญญาณรบกวน" ในข้อมูลต้นฉบับ (เช่น +/- ค่าสุ่มในข้อมูล) ระดับของการบิดเบือนควรสัมพันธ์กับช่วงค่าของข้อมูล

เทคนิคการลบข้อมูลระบุตัวตน	คำอธิบาย
	ตัวอย่างเช่น การบิดเบือนข้อมูลเงินเดือนของบุคคลจาก "256,654 บาท" เป็น " 300,000 บาท" โดยปัดข้อมูลขึ้นไปถึง 50,000 บาท
8.Synthetic Data การสังเคราะห์ข้อมูล	ข้อมูลสังเคราะห์เป็นข้อมูลที่สร้างขึ้นตามอัลกอริทึมโดยไม่มีการเชื่อมต่อกับกรณีจริงใดใด ข้อมูลนี้ใช้เพื่อสร้างชุดข้อมูลปลอม แทนที่จะใช้หรือแก้ไขชุดข้อมูลดั้งเดิม และลดทอนการปกป้องและความเป็นส่วนตัว วิธีข้อมูลนี้ใช้ระบบทางคณิตศาสตร์ตามรูปแบบหรือคุณลักษณะในชุดข้อมูลดั้งเดิม การถดถอยเชิงเส้น ส่วนเบี่ยงเบนมาตรฐาน ค่ามัธยฐาน และวิธีการทางสถิติอื่นอาจถูกนำมาใช้เพื่อสร้างผลลัพธ์สังเคราะห์ ข้อมูลสังเคราะห์จึงมีความเหมาะสำหรับการพัฒนา/ทดสอบแอปพลิเคชัน แต่ไม่เหมาะสำหรับการฝึกโมเดล AI.
9.Data Aggregation การรวมข้อมูล	เป็นกระบวนการที่รวมข้อมูลจากหลายแหล่งหรือจากหลายระเบียนเข้าด้วยกันเพื่อสร้างสรุปหรือข้อมูลที่มีมูลค่าเพิ่ม การรวมข้อมูลอาจช่วยในการลดความละเอียดของข้อมูลส่วนบุคคลที่อาจถูกระบุได้ โดยเปลี่ยนจากข้อมูลระดับประเภทเป็นข้อมูลรวม เช่น จากข้อมูลเกี่ยวกับการใช้จ่ายของบุคคลแต่ละคนเป็นข้อมูลเกี่ยวกับการใช้จ่ายเฉลี่ยของกลุ่มบุคคล วิธีนี้สามารถช่วยในการป้องกันการระบุตัวตนของบุคคลจากชุดข้อมูลโดยการทำให้ข้อมูลนั้นเป็นส่วนรวมมากขึ้นและลดรายละเอียดของข้อมูลที่อาจนำไปสู่การระบุตัวตนได้. **ตัวอย่างเช่น** ข้อมูลดิบสามารถถูกรวมกันในช่วงเวลาที่กำหนดเพื่อให้ได้สถิติเช่น ค่าเฉลี่ย ค่าต่ำสุด ค่าสูงสุด ผลรวม และจำนวนการนับ

3. กระบวนการจัดทำข้อมูลนิรนาม

ในการจัดทำข้อมูล ผู้ควบคุมข้อมูลส่วนบุคคลและผู้ประมวลผลข้อมูลส่วนบุคคล มีหน้าที่ตามกฎหมายในการจัดให้มีมาตรการรักษาความมั่นคงปลอดภัยที่เหมาะสม เพื่อป้องกันการสูญหาย เข้าถึง ใช้ เปลี่ยนแปลง แก้ไข หรือเปิดเผยข้อมูลส่วนบุคคลโดยปราศจากอำนาจหรือโดยมิชอบ ในส่วนนี้มีความมุ่งหมายที่จะแสดงให้เห็นถึงกรอบความคิดในการพิจารณาเลือกใช้วิธี ที่เหมาะสมในการจัดทำข้อมูลนิรนาม โดยประเมินจากปัจจัยที่เกี่ยวข้องทั้งที่เกี่ยวข้องกับตัวข้อมูลเอง และที่เกี่ยวข้องกับสิ่งแวดล้อมของข้อมูล เพื่อให้ผู้ควบคุมข้อมูล และผู้ประมวลผลข้อมูล สามารถปฏิบัติตามหลักการตามบทบัญญัติของมาตรา 37 พระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562

รูปที่ 11: ข้อมูลนิรนาม

ทั้งนี้ หลักการสำคัญสำหรับการจัดทำข้อมูลนิรนามคือ การทำให้ไม่อาจระบุคุณลักษณะของตัวเจ้าของข้อมูลได้จากข้อมูลดังกล่าว (non-attributable) เพราะในบางกรณีเจ้าของข้อมูลอาจถูกระบุคุณลักษณะได้ โดยที่ไม่จำเป็นต้องมีการระบุตัวตนอย่างชัดเจน การลดความเสี่ยงดังกล่าวด้วยวิธีการและมาตรการที่ถูกต้องเหมาะสม ย่อมสามารถคุ้มครองผู้ควบคุม และผู้ประมวลผลข้อมูลจากความรับผิดที่อาจเกิดขึ้นได้ในกรณีที่มี และยังเป็นการใช้ประโยชน์จากข้อมูลที่มี อาทิ จากข้อมูลที่อยู่ใน ระดับชั้น “ลับ” “ลับมาก” ซึ่งยากต่อการเข้าถึง และเสี่ยงต่อการนำมาประมวลผล เมื่อข้อมูลถูกจัดทำให้เป็นข้อมูลนิรนามแล้ว ผู้ควบคุมข้อมูลสามารถพิจารณาจัดระดับชั้นข้อมูลดังกล่าว เป็นข้อมูลที่อยู่ในระดับชั้น “เผยแพร่ภายในองค์กร” นำไปใช้ประโยชน์ นำมาวิเคราะห์ สร้างมูลค่ากับข้อมูลที่มีอีกด้วย แต่อย่างไรก็ดีถึงแม้ว่า **ข้อมูลจะถูกลดความน่าจะเป็นในการระบุตัวตนของเจ้าของข้อมูลแล้วนั้น ผู้ใช้ข้อมูลนิรนาม ยังคงจำเป็นต้องใช้ข้อมูลอย่างระมัดระวังเสมอ**

แนวทางที่ขอนำเสนอในเล่มข้อเสนอแนะนี้ เพียงเพื่อให้ผู้อ่านสามารถทำความเข้าใจถึงกระบวนการจัดทำข้อมูลนิรนามได้โดยง่าย และ เห็นถึงกรณีศึกษาจากหน่วยงานที่มีการจัดทำข้อมูลนิรนาม สพร. จึงอ้างอิงในแนวทางจาก Guide Basic Anonymisation Personal Data Protection Commission (PDPC) Singapore ซึ่งมี 5 กระบวนการ ดังที่กล่าวในบทที่ 2.2 โดยสำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ และสำนักคณะกรรมการคุ้มครองข้อมูลส่วนบุคคลได้ร่วมกันแปลเอกสารเป็นภาษาไทย ฉบับเต็มสามารถอ่านเพิ่มเติมได้ที่https://www.nstda.or.th/nstdaxpdpc/privacytools/

3.1. การพิจารณาข้อมูล และ การขจัดข้อมูลระบุตัวตน

1.การพิจารณาข้อมูล

1.1 การพิจารณาตามคุณลักษณะข้อมูล

การบันทึกข้อมูลส่วนบุคคลประกอบด้วยคุณลักษณะข้อมูล ที่สามารถจัดระดับการระบุตัวตนและความละเอียดอ่อนต่อบุคคลที่แตกต่างกันได้ การทำให้ข้อมูลไม่สามารถระบุตัวตนได้โดยทั่วไปประกอบด้วยการลบตัวบ่งชี้ทางตรง (Direct identifiers) และการปรับเปลี่ยนตัวบ่งชี้ทางอ้อม (Indirect identifiers) โดยคุณลักษณะเป้าหมาย (Target attributes) มักจะถูกเว้นไว้ไม่เปลี่ยนแปลง ยกเว้นในกรณีที่วัตถุประสงค์คือ การสร้างข้อมูลสังเคราะห์ ขึ้นมาเพื่อใช้งาน (Singapore, 2022)

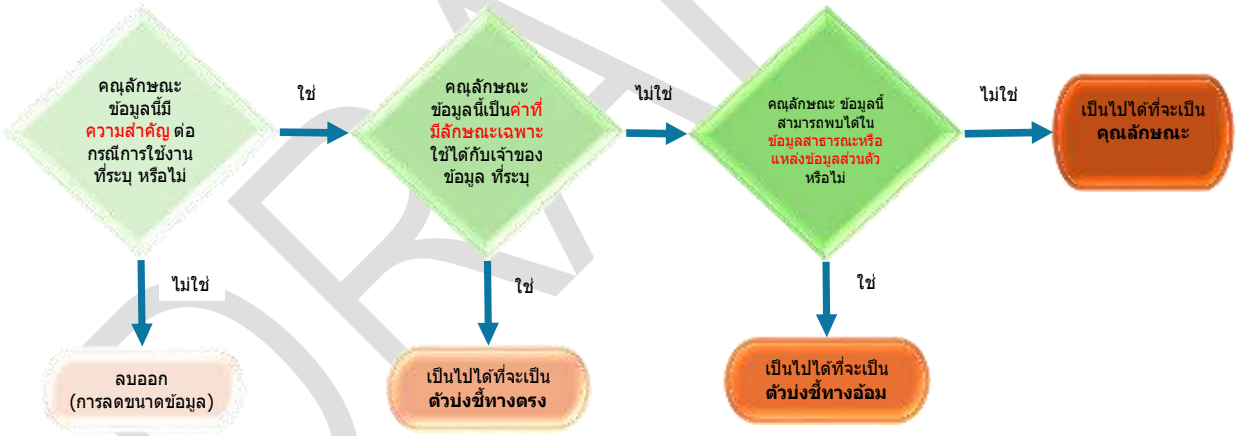
ตารางและตัวอย่างด้านล่างนี้แสดงให้เห็นว่าคุณลักษณะข้อมูลโดยทั่วไปจะถูกจัดประเภทอย่างไรในบันทึกข้อมูล

302 **ตารางที่ 2: การพิจารณาตามคุณลักษณะข้อมูล**

ระดับการระบุตัวตนของชุดข้อมูล	คุณลักษณะของข้อมูล	การเข้าถึงข้อมูล	ตัวอย่างในชุดข้อมูล
1.ตัวบ่งชี้ทางตรง (Direct identifiers)	เป็นคุณลักษณะข้อมูลที่เฉพาะเจาะจงแต่ละบุคคล และสามารถใช้เป็นคุณลักษณะข้อมูลหลักในการระบุตัวตนของบุคคลนั้นได้อีกครั้ง	คุณลักษณะข้อมูลเหล่านี้มักจะเป็นข้อมูลสาธารณะหรือข้อมูลที่เข้าถึงได้ง่าย	• ชื่อ นามสกุล • ที่อยู่อีเมล • หมายเลขโทรศัพท์มือถือ • หมายเลขหนังสือเดินทาง • หมายเลขบัญชี • หมายเลขสูติบัตร • หมายเลขใบอนุญาตทำงาน • ชื่อผู้ใช้งานโซเชียลมีเดีย
2. ตัวบ่งชี้ทางอ้อม (Indirect identifiers)	เป็นคุณลักษณะข้อมูลที่ไม่เฉพาะเจาะจงแต่ละบุคคล แต่อาจทำให้สามารถระบุตัวตนของบุคคลนั้นได้เมื่อรวมกับข้อมูลอื่น (เช่น การรวมกันของอายุ, เพศ และรหัสไปรษณีย์).	คุณลักษณะข้อมูลเหล่านี้มักจะเป็นข้อมูลสาธารณะหรือข้อมูลที่เข้าถึงได้ง่าย	• อายุ • เพศ • เชื้อชาติ • วันเดือนปีเกิด • ที่อยู่ • รหัสไปรษณีย์ • ตำแหน่งงาน • ชื่อบริษัท • สถานภาพการสมรส • ส่วนสูง • น้ำหนัก • ที่อยู่อินเทอร์เน็ตโปรโตคอล (IP Address) • ลขทะเบียนรถ • ตำแหน่งพิกัดบนพื้นโลก (GPS)

3. คุณลักษณะเป้าหมาย (Target attributes)	คุณลักษณะของข้อมูลนี้อาจจะเป็นลักษณะละเอียดอ่อน และอาจส่งผลให้เกิดผลเสียต่อบุคคลได้สูงเมื่อถูกนำไปเปิดเผย	คุณลักษณะข้อมูลเหล่านี้มักจะไม่เป็นข้อมูลสาธารณะหรือไม่สามารถเข้าถึงได้ ซึ่งข้อมูลเหล่านี้ไม่สามารถใช้เพื่อระบุตัวตนของบุคคลนั้นอีกครั้งได้ เนื่องจากโดยทั่วไปแล้วจะเป็นข้อมูลที่มีกรรมสิทธิ์	• ธุรกรรม (เช่น การซื้อของ) • เงินเดือน • อัตราเครดิต • กรมธรรม์ประกัน • การวินิจฉัยทางการแพทย์ • สถานะการฉีดวัคซีน

แนวทางการลดขนาดของข้อมูล โดยเริ่มจากการพิจารณาคุณลักษณะของข้อมูลใดใด ที่ไม่จำเป็นในชุดข้อมูลผลลัพธ์ควรถูกลบออก โดยแผนภาพด้านล่างเพื่อช่วยให้สามารถจำแนกคุณลักษณะของข้อมูลได้อย่างเหมาะสม



รูปที่ 12: ความเชื่อมโยงระหว่างข้อมูลที่ไม่ระบุตัวตนกับข้อมูลที่ระบุตัวตน

1.2 การพิจารณาสถานการณ์ของข้อมูล

นอกเหนือจากการพิจารณาข้อมูลตามคุณลักษณะข้างต้นแล้ว เพื่อให้เป็นไปตาม“พระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562 ส่วนที่ 2 การเก็บรวบรวมข้อมูลส่วนบุคคล” นั้น ผู้จัดทำข้อมูลนิรนามจะต้องสามารถจัดทำผังการเคลื่อนที่ข้อมูล (Data Flowchart) โดยระบุถึงสิ่งแวดล้อมทั้งหมดที่ข้อมูลอาจมีการเคลื่อนย้ายโดยอาจระบุถึง

- บุคคลที่มีส่วนเกี่ยวข้องกับข้อมูลในสิ่งแวดล้อมนั้น
- การกระทำอันเกี่ยวข้องกับข้อมูล
- วิธีการในการเคลื่อนย้าย

- ระบุลักษณะของข้อมูลที่เคลื่อนย้ายดังกล่าวว่าเป็นข้อมูลดั้งเดิม หรือเป็นข้อมูลที่มีการเปลี่ยนแปลงประการใด

ตัวอย่าง บริษัท ก เก็บข้อมูลของผู้ใช้บริการทั้งหมด สมมติว่ามีกฎหมายบังคับให้บริษัท ก นั้นเปิดเผยข้อมูลดังกล่าว กับกระทรวงกลาโหม เพื่อประโยชน์ในด้านความมั่นคง อย่างไรก็ตามข้อมูลดังกล่าวนั้นอาจมีประโยชน์ใน ด้านการวิจัย จึงมีการนำข้อมูลที่ได้ถูกลบตัวบ่งชี้ออกหมดแล้ว (de-identified data) เพื่อให้นักวิจัย ข ที่ ได้รับการรับรองจากสถาบันที่กฎหมายกำหนด ใช้ภายใต้ระบบที่ป้องกันการนำข้อมูลไปใช้เกินขอบเขตของวัตถุประสงค์ในการวิจัยที่ขอไว้ล่วงหน้า หลังจากนั้นนักวิจัย ข ที่มาขออนุญาตจึงได้นำข้อมูลไปวิเคราะห์ และ ตีพิมพ์ผลการวิจัยเพื่อเปิดเผยต่อสาธารณชนต่อไป สถานการณ์ดังกล่าวอาจเขียนเป็นผังการเคลื่อนที่ของ ข้อมูลได้ดังภาพ

รูปที่ 13: ตัวอย่างการเข้าถึงข้อมูล

1.2.1 การพิจารณาความรับผิดทางกฎหมาย ผู้ควบคุมข้อมูลส่วนบุคคล ต้องพิจารณาดังต่อไปนี้

(1) ข้อมูลที่อยู่ในความครอบครองนั้นเป็นข้อมูลส่วนบุคคลหรือไม่?

(2) ตนมีหน้าที่เป็นผู้ควบคุม หรือผู้ประมวลผลข้อมูลหรือไม่ อย่างไร?

1.2.2 การพิจารณาตัวข้อมูล ผู้ควบคุมข้อมูลส่วนบุคคล ต้องพิจารณาถึงคุณสมบัติหลัก ที่เกี่ยวข้องกับข้อมูลดังต่อไปนี้

(1) ใครเป็นผู้เป็นเจ้าของข้อมูล? *(เป็นบุคคลธรรมดา หรือ เป็นกลุ่มบุคคล)*

(2) ข้อมูลเป็นข้อมูลประเภทใด? *(เป็นข้อมูลตัวเลข ตัวอักษร ข้อมูลมาตรสัดส่วน ข้อมูลรายบุคคล/ข้อมูลรวมกลุ่ม หรือ ข้อมูลอ่อนไหว)*

(3) ตัวแปรในข้อมูลเป็นตัวแปรประเภทใดบ้าง? *(ตัวแปรบ่งชี้ทางตรง , ตัวแปรบ่งชี้ทางอ้อม)*

(4) คุณสมบัติของชุดข้อมูล (คุณภาพของการวัด, อายุของข้อมูล, โครงสร้างของข้อมูล เป็นข้อมูลประชากร หรือ กลุ่มตัวอย่าง)

1.2.3. การพิจารณาการใช้งานของข้อมูล ผู้ครอบครองข้อมูลหรือผู้จัดทำข้อมูล จะต้องพิจารณาว่าข้อมูลนั้น อาจจะนำไปใช้ได้ในกรณีใดบ้าง โดยตั้งคำถามดังต่อไปนี้

(1) **ทำไม?** ต้องมีคำตอบที่ชัดเจนว่าทำไมถึงอยากที่จะเปิดเผยข้อมูล หรือเปิดเผยข้อมูลให้กับผู้อื่นหรือสาธารณะ

- เพื่อให้ข้อมูลกับผู้มีส่วนได้เสีย
- เพื่อให้ข้อมูลอันเฉพาะเจาะจงที่เกี่ยวกับเรื่องใดเรื่องหนึ่ง
- เพื่อเอื้อประโยชน์ให้กับผู้มีสิทธิเข้าถึงข้อมูล
- จำเป็นต้องทำด้วยผลของกฎหมาย อาทิ กฎหมายที่ว่าด้วยการเปิดเผยข้อมูลของรัฐ

(2) **ใคร?** ต้องระบุให้ชัดเจนว่าใครบ้างที่จะมีสิทธิเข้าถึงข้อมูล (บุคคล , องค์กร/หน่วยงาน , กลุ่มบุคคล หรือกลุ่มองค์กร)

(3) **อย่างไร?** ต้องอธิบายให้ได้อย่างละเอียดว่า ผู้ที่จะเข้าถึงข้อมูลจะนำข้อมูลไปใช้อย่างไรบ้าง (โดยการสอบถาม หรือ ศึกษาจากการใช้งานข้อมูลจำลอง/ข้อมูลตัวอย่าง)

1.2.4. การพิจารณาการขอใช้ข้อมูลโดยชอบ แม้ในกรณีที่ข้อมูลนั้นถูกจัดทำเป็นข้อมูลนิรนามแล้ว แต่จำเป็นต้องมี มาตรฐานในการขอความยินยอม การแสดงความโปร่งใสในการใช้ข้อมูล และการมีระบบธรรมาภิบาลในด้านข้อมูลที่ดี มาตรฐานดังที่กล่าวเหล่านี้ก็ควรเป็นข้อปฏิบัติที่ผู้ควบคุมข้อมูล หรือประมวลผลข้อมูล ควรที่จะปฏิบัติตาม

2. การขจัดข้อมูลระบุตัวตน

ขั้นตอนนี้เป็นส่วนหนึ่งของกระบวนการทำให้ข้อมูลไม่ระบุตัวตน

โดยขั้นตอนแรก คือการลบตัวบ่งชี้ทางตรงทั้งหมด ในตัวอย่างต่อไปนี้ ชื่อทั้งหมดถูกลบออก หากชุดข้อมูลรวมถึงตัวบ่งชี้ทางตรงนอกเหนือจาก รูปที่ 13 เช่น เลขที่บัตรประชาชน ที่อยู่อีเมล สิ่งเหล่านี้ก็ควรจะถูกลบออกเช่นกัน

ชื่อ – นามกุล	ตำแหน่ง	รหัสพนักงาน	เพศ	อายุงาน/ปี	เบอร์ติดต่อ
แดง เป็นคนไทย	ผู้บริหาร	11000	ชาย	10	083 5675675
ไข่ ขยันทำงาน	พนักงาน 1	110011	ชาย	2	0653333133
สวย ใจดี	พนักงาน 2	110012	หญิง	3	0891116622
น้ำใจ รักงาน	พนักงาน 3	110013	หญิง	2	0879871234

รูปที่ 14: การลบตัวบ่งชี้ทางตรง

อีกทางเลือกหนึ่ง คือ สามารถกำหนดนามแฝงให้กับแต่ละรายการ หากมีความจำเป็นที่จะเชื่อมโยงบันทึกกลับไปยังบุคคลที่เป็นเอกลักษณ์หรือไปยังบันทึกเดิมสำหรับกรณีการใช้งานเช่น :

a. การรวมข้อมูล

b. การวิเคราะห์จากหลายรายการที่เกี่ยวข้องกับบุคคลที่เป็นเอกลักษณ์/ลักษณะเฉพาะ หรือ

c. การสร้างชุดข้อมูลสังเคราะห์ที่ต้องการค่าตัวบ่งชี้โดยตรงสำหรับการพัฒนาและทดสอบแอปพลิเคชัน สำหรับกรณีการใช้งานนี้ให้แทนที่ตัวบ่งชี้โดยตรงที่จำเป็นทั้งหมดด้วยนามแฝง

นามแฝงควรจะเป็นเอกลักษณ์สำหรับแต่ละตัวบ่งชี้ทางตรง (ตามที่แสดงด้านล่าง) การกำหนดนามแฝงควรจะต้องมีมั่นคงปลอดภัย (กล่าวคือ บุคคลที่ได้รับอนุญาต ไม่สามารถย้อนกลับได้โดยผ่านการเดาหรือคำนวณค่าตัวบ่งชี้โดยตรงเดิมจากนามแฝงได้)

ชื่อ – นามกุล	โทเค็น	ตำแหน่ง	รหัสพนักงาน	เพศ	อายุงาน/ปี	เบอร์ติดต่อ
แดง เป็นคนไทย	a896	ผู้บริหาร	11000	ชาย	10	083 5675675
ไข่ ขยันทำงาน	345f	พนักงาน 1	110011	ชาย	2	0653333133
สวย ใจดี	b123	พนักงาน 2	110012	หญิง	3	0891116622
น้ำใจ รักงาน	96a5	พนักงาน 3	110013	หญิง	2	0879871234

รูปที่ 15: การกำหนดนามแฝง

หากต้องการรักษาความสามารถในการเชื่อมโยงบันทึกข้อมูลที่ไม่ระบุตัวตนกลับไปยังบันทึกเดิมในภายหลัง จำเป็นจะต้องจัดเก็บรักษา โทเค็น (ตัวจับคู่) ให้มีความมั่นคงและปลอดภัย เนื่องจากโทเค็นจะตัวที่ทำให้ระบุตัวตนได้อีกครั้ง

3.2. หลักเกณฑ์การจัดทำข้อมูลนิรนาม

วิธีการจัดทำข้อมูลนิรนามตามบทที่ 2.3 สามารถสรุปได้ 7 วิธี (สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ, 2023) ซึ่งแต่ละวิธีก็จะมีความเหมาะสมข้อมูลที่มีลักษณะไม่เหมือนกัน โดยการจัดทำข้อมูลนิรนามมากกว่า 1 วิธี ในบทนี้จะกล่าวถึงหลักเกณฑ์การจัดทำข้อมูลนิรนาม เพื่อเป็นแนวทางให้แก่เจ้าของข้อมูลใช้ในการประกอบการพิจารณาเพื่อจัดทำข้อมูลนิรนาม พร้อมข้อเสนอแนะและการแสดงตัวอย่างในการจัดทำข้อมูลนิรนาม

ทั้งนี้ เนื้อหาที่กล่าวมาได้อ้างอิงจาก The Personal Data Protection Commission ('PDPC'), Guide to Basic Data Anonymisation (31 March 2022) และแนวทางสำหรับการจัดทำข้อมูลนิรนามขั้นพื้นฐาน โดยสำนักงานคณะกรรมการคุ้มครองข้อมูลส่วนบุคคล (สคส.) หรือ PDPC ร่วมกับ สวทช. เพื่อให้หน่วยงานสามารถเลือกใช้วิธีการจัดทำข้อมูลนิรนามให้เหมาะสมกับข้อมูลของตนเองได้ โดยพิจารณาจากวัตถุประสงค์การใช้งานข้อมูลและลักษณะของข้อมูล ประกอบด้วย ข้อมูลบ่งชี้ทางตรง (Direct identifiers) ข้อมูลทางอ้อม (Indirect identifiers) ข้อมูลที่มีคุณลักษณะในการเชื่อมโยงไปยังข้อมูลบ่งชี้ (Target attributes) ตามข้อที่ 3.1 กล่าวไว้ รายละเอียดมีดังนี้

1. การลบคุณลักษณะข้อมูล (Suppression) คือ การลบหรือซ่อนบันทึกข้อมูลหรือข้อมูลบางส่วนออกจากชุดข้อมูล โดยส่วนมากใช้วิธีนี้กับข้อมูลบ่งชี้ทางตรง (Direct identifiers) เพื่อลดความเสี่ยงในการระบุตัวตน โดยมี 2 วิธีย่อย ได้แก่ 1) การลบคุณลักษณะเฉพาะ (Attribute Suppression) เป็นการลบข้อมูลทั้งคอลัมน์/ฟิลด์ (ต่อไปจะใช้คำว่า คอลัมน์) ออกจากตารางที่เป็นแนวตั้ง/ชุดข้อมูล (ต่อไปจะใช้คำว่า ชุดข้อมูล) และ 2) การลบข้อมูลรายบันทึก (Record Suppression) เป็นการลบข้อมูลบันทึกออกจากชุดข้อมูล ดังนี้

- **การลบคุณลักษณะเฉพาะ (Attribute Suppression)**

ตารางที่ 3: การทำข้อมูลนิรนามด้วยการลบคุณลักษณะเฉพาะ

คำอธิบาย	การลบ/ซ่อนข้อมูลบ่งชี้ทางตรง (Direct identifiers หรือ Formal Identifier) โดยลบเป็นการลบข้อมูลรายคอลัมน์ ซึ่งในหนึ่งชุดข้อมูลจะใช้วิธีการลบคุณลักษณะเฉพาะหนึ่งคอลัมน์หรือหลายคอลัมน์ก็ได้ เช่น ข้อมูลบ่งชี้ทางตรง ได้แก่ ชื่อ-นามสกุล หมายเลขบัตรประชาชน หมายเลขหนังสือเดินทาง อีเมล์ วันเดือนปีเกิด
หลักเกณฑ์การพิจารณา	● เป็นข้อมูลบ่งชี้ทางตรง หรือข้อมูลที่มีความอ่อนไหว ใช่หรือไม่ ● เป็นข้อมูลที่ไม่จำเป็นต่อการนำไปวิเคราะห์ ใช่หรือไม่ ● เป็นข้อมูลที่ลบไปแล้วไม่ส่งผลต่อการใช้ประโยชน์ ใช่หรือไม่ ● เป็นข้อมูลที่สามารถแทนที่ด้วยตัวอักษรหรือตัวเลข หรือไม่ ● เป็นข้อมูลที่ไม่ต้องการเชื่อมโยงกับข้อมูลชุดอื่น ใช่หรือไม่
ข้อแนะนำ	● ควรเป็นข้อมูลที่มีลักษณะตาราง ● ควรเลือกใช้กับข้อมูลบ่งชี้ทางตรง หรือข้อมูลที่มีความอ่อนไหว ซึ่งมีผลกระทบกับหน่วยงานหากมีการเข้าถึงข้อมูลโดยไม่ได้รับอนุญาต ● ควรเลือกใช้กับข้อมูลที่ไม่กระทบกับการใช้งาน ● ควรใช้เพื่อลบค่าผิดปกติออกจากชุดข้อมูล เช่น ค่า k-anonymity

	• หากใช้วิธีการลบบันทึกข้อมูลออกไปอาจไม่สามารถกู้คืนมาได้ และหากใช้วิธีการซ่อนคอลัมน์ควรมีการเข้ารหัสคอลัมน์เพิ่มขึ้น เพื่อให้เก็บรักษาข้อมูลมีความปลอดภัยมากขึ้น (รายละเอียดดูได้ที่ภาคผนวก) • ควรใช้การจัดทำข้อมูลนิรนามวิธีอื่นประกอบกัน เพื่อป้องกันการเชื่อมโยงหรืออนุมานกลับการระบุตัวตน • สามารถประยุกต์กับข้อมูลประเภท 1) ข้อมูลใช้ภายใน 2) ข้อมูลที่มีการแบ่งปันให้กับบุคคลภายนอกองค์กร ซึ่งเจ้าของข้อมูล/ผู้ควบคุมข้อมูลพิจารณาได้ตามความเหมาะสม

การลบคุณลักษณะเฉพาะ (Attribute Suppression)

ตัวอย่างข้อมูล**ก่อน**การทำ Attribute Suppression

ชื่อ - นามสกุล	ตำแหน่ง	รหัสพนักงาน	เพศ	อายุงาน/ปี
แดง เป็นคนไทย	ผู้บริหาร	11000	ชาย	10
ไข่ ขยันทำงาน	พนักงาน 1	110011	ชาย	2
สวย ใจดี	พนักงาน 2	110012	หญิง	3
น้ำใจ รักงาน	พนักงาน 3	110013	หญิง	2

ตัวอย่างข้อมูล**หลัง**การทำ Attribute Suppression

ตำแหน่ง	รหัสพนักงาน	เพศ	อายุงาน/ปี
ผู้บริหาร	11000	ชาย	10
พนักงาน 1	110011	ชาย	2
พนักงาน 2	110012	หญิง	3
พนักงาน 3	110013	หญิง	2

รูปที่ 16: ตัวอย่างการลบคุณลักษณะเฉพาะ

- **การลบข้อมูลรายบันทึก (Record Suppression)**

ตารางที่ 4: การทำข้อมูลนิรนามด้วยการลบข้อมูลรายบันทึก

คำอธิบาย	การลบข้อมูลเป็นรายบันทึก เป็นรายแถว สำหรับข้อมูลที่มีลักษณะโดดเด่น หรือมีค่าที่ผิดปกติออกจากชุดข้อมูล เช่น ข้อมูลผู้บริหาร ซึ่งมีลักษณะที่แปลกแยกจากกลุ่มข้อมูล (Singling out) เนื่องจากข้อมูลมีความโดดเด่นแม้ว่าจะลบข้อมูลบ่งชี้ทางตรงไปแล้ว หรือแปลงข้อมูลบางส่วนออกไปก็ยังมีโอกาสในการเชื่อมโยงหรืออนุมานไปข้อมูลบ่งชี้ทางตรงได้ จึงต้องลบข้อมูลรายบันทึกไปทั้งแถว ซึ่งอาจส่งผลต่อคุณลักษณะของชุดข้อมูลในแง่ของสถิติได้ เช่น ค่าเฉลี่ย
หลักเกณฑ์การพิจารณา	• เป็นข้อมูลบ่งชี้ทางตรง หรือข้อมูลที่มีความอ่อนไหว ใช่หรือไม่ • เป็นข้อมูลที่มีลักษณะโดดเด่นจากชุดข้อมูล ใช่หรือไม่ • เป็นข้อมูลที่ไม่จำเป็นต่อการนำไปวิเคราะห์ ใช่หรือไม่ • เป็นข้อมูลที่ลบไปแล้วไม่ส่งผลต่อการใช้ประโยชน์หรือค่าเฉลี่ยของชุดข้อมูล ใช่หรือไม่ • เป็นข้อมูลที่สามารถแทนที่ด้วยตัวอักษรหรือตัวเลข หรือไม่ • เป็นข้อมูลที่ไม่ต้องการเชื่อมโยงกับข้อมูลชุดอื่น ใช่หรือไม่
ข้อแนะนำ	• ควรเป็นข้อมูลที่มีลักษณะตาราง • ควรเลือกใช้กับข้อมูลบ่งชี้ทางตรง หรือข้อมูลที่มีลักษณะโดดเด่นจากชุดข้อมูล และเป็นข้อมูลที่ไม่ต้องนำมาวิเคราะห์

	• ควรใช้เพื่อลบค่าผิดปกติออกจากชุดข้อมูลที่สามารถเชื่อมโยงไปยังข้อมูลบ่งชี้ทางตรง เช่น ค่า k-anonymity และการลบข้อมูลนี้อาจไม่สามารถกู้คืนมาได้ • ข้อมูลในชุดข้อมูลควรมีปริมาณมากพอเพื่อป้องกันไม่ให้เกิดการเชื่อมโยงกลับไปยังเจ้าของข้อมูลส่วนบุคคลได้ • ควรใช้การจัดทำข้อมูลนิรนามวิธีอื่นประกอบกัน เพื่อป้องกันการเชื่อมโยงหรืออนุมานกลับการระบุตัวตน • สามารถประยุกต์ใช้วิธีนี้กับข้อมูลประเภท 1) ข้อมูลใช้ภายใน 2) ข้อมูลที่มีการแบ่งปันให้กับบุคคลภายนอกองค์กร ซึ่งเจ้าของข้อมูล/ผู้ควบคุมข้อมูลสามารถพิจารณาได้ตามความเหมาะสม

รูปที่ 17: ตัวอย่างการลบข้อมูลรายบันทึก

2. **การปิดทับลักษณะข้อมูล (Character Masking)** คือ การปกปิดหรือปิดบังข้อมูล ด้วยการเปลี่ยนส่วนใดส่วนหนึ่งของข้อมูล โดยการใช้กลุ่มของตัวอักษรที่ได้จากการสุ่ม และนำมาเรียงอักษรในข้อมูลใหม่แบบไม่เป็นระบบ เช่น 234246 อาจจะสลับเป็น 464232 เป็นต้น เพื่อทำให้ข้อมูลนั้นแสดงเป็นข้อมูลหลอกหรือนามแฝงเพื่อปกปิดข้อมูลจริง **ซึ่งการปิดทับลักษณะข้อมูลเป็นวิธีที่ได้รับความนิยมอย่างมาก เนื่องจากมีการปกปิดข้อมูลที่ระบุตัวตน โดยที่ข้อมูลยังคงคุณค่าจากการใช้ประโยชน์ของข้อมูล (Value) ที่ต้องการไว้ได้ ซึ่งผลลัพธ์ภายหลังการปิดทับข้อมูลยังคงเหมือนกับชุดข้อมูลจริงต้นฉบับ เช่น ชุดข้อมูลผู้ป่วยมีการปิดทับชื่อผู้ป่วยแล้ว แต่นักวิเคราะห์ยังสามารถใช้ประโยชน์จากข้อมูลในคอลัมน์อื่น ได้ครบถ้วน และยังหาความสัมพันธ์อื่น โดยเชื่อมโยงกับชุดข้อมูลภายนอกได้ เช่น การรับรู้ว่าเป็นบุคคลเดียวกัน แต่ไม่สามารถระบุตัวตนได้**

ตารางที่ 5: การทำข้อมูลนิรนามด้วยการปิดทับลักษณะข้อมูล

คำอธิบาย	การปิดทับลักษณะข้อมูล คือ การเปลี่ยนส่วนใดส่วนหนึ่งของข้อมูลโดยการใช้กลุ่มของตัวอักษรที่ได้จากการสุ่ม หรือข้อมูลอื่น โดยการสุ่มข้อความโดยเรียงอักษรในข้อมูลนั้นใหม่แบบไม่เป็นระบบ เช่น ลบข้อมูลที่เป็นชื่อ แล้วจึงเอาข้อมูลตัวอักษรดังกล่าวมาแทนที่ชื่อในข้อมูลปัจจุบันแทน เช่น การปิดทับข้อมูลบ่งชี้ทางตรง ได้แก่ ชื่อ อีเมล์ วันเดือนปีเกิด หมายเลขบัตรประชาชน หมายเลขหนังสือเดินทาง
หลักเกณฑ์การพิจารณา	• เป็นข้อมูลบ่งชี้ทางตรง หรือข้อมูลที่มีความอ่อนไหว ใช่หรือไม่ • เป็นข้อมูลที่สามารถแทนที่ด้วยตัวอักษรหรือตัวเลขหรือไม่

	● เป็นข้อมูลที่ต้องการเชื่อมโยงกับข้อมูลอื่น ในชุดข้อมูลเดียวกัน ใช่หรือไม่ ● เป็นข้อมูลที่ปิดทับไปแล้วไม่ส่งผลต่อการใช้ประโยชน์ข้อมูลอื่น ในชุดข้อมูล ใช่หรือไม่
ข้อแนะนำ	● ควรเลือกใช้กับข้อมูลบ่งชี้ทางตรง ● สามารถปิดทับได้มากกว่า 1 คอลัมน์ในชุดข้อมูล และสามารถใช้การจัดทำข้อมูลนิรนามวิธีอื่น มาประกอบกัน เพื่อป้องกันการเชื่อมโยงหรืออนุมานกลับการระบุตัวตน ● ข้อมูลบันทึก (แถว) ในชุดข้อมูลควรมีปริมาณมากพอเพื่อป้องกันไม่ให้เกิดการเชื่อมโยงกลับไปยังเจ้าของข้อมูลส่วนบุคคลได้ ● สามารถประยุกต์ใช้กับข้อมูลประเภท 1) ข้อมูลใช้ภายใน 2) ข้อมูลที่มีการแบ่งปันให้กับบุคคลภายนอกองค์กร ซึ่งเจ้าของข้อมูล/ผู้ควบคุมข้อมูลพิจารณาได้ตามความเหมาะสม

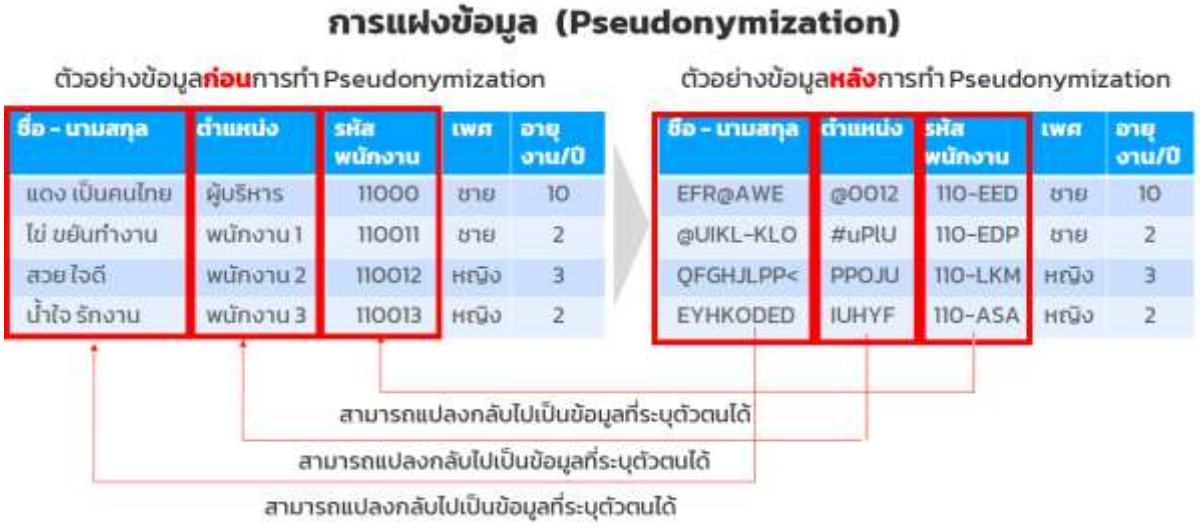
รูปที่ 18: ตัวอย่างการปิดทับลักษณะข้อมูล

3. การแฝงข้อมูล (Pseudonymization) คือ การแทนที่สิ่งที่ระบุตัวตนของเจ้าของข้อมูลโดยตรงด้วยชื่อหรือรหัสที่สร้างขึ้นมา หรือด้วยวิธีการใดวิธีการหนึ่งอันเป็นเอกลักษณ์ โดยแยกเก็บเป็นชุดข้อมูลจริงและชุดข้อมูลที่มีการแทนค่า เพื่อป้องกันการเชื่อมโยงระหว่างชุดข้อมูล

ตารางที่ 6: การทำข้อมูลนิรนามด้วยการแฝงข้อมูล

คำอธิบาย	การแฝงข้อมูล โดยการเปลี่ยนค่าข้อมูลระบุตัวตนเป็นค่าที่กำหนดขึ้น เพื่อลดทอนหรือจำกัดความสามารถในการเชื่อมโยงข้อมูล เช่น การเข้ารหัสข้อมูล ซึ่งเป็นการแปลงข้อมูลให้อยู่ในรูปที่ไม่สามารถอ่านทำความเข้าใจได้ แต่สามารถแปลงกลับเป็นข้อมูลเดิมได้ผ่านการใช้กุญแจ (key)
หลักเกณฑ์การพิจารณา	● เป็นข้อมูลบ่งชี้ทางตรง หรือข้อมูลที่มีความอ่อนไหว ใช่หรือไม่ ● เป็นข้อมูลที่มีโอกาสระบุตัวตนได้ใช่หรือไม่ ● เป็นข้อมูลที่สามารถแทนที่ด้วยตัวอักษรหรือตัวเลข ใช่หรือไม่ ● เป็นข้อมูลสำคัญที่ต้องการใช้ประโยชน์ โดยไม่จำเป็นต้องรับรู้ข้อมูลบ่งชี้ทางตรง และไม่สามารถลบได้ ใช่หรือไม่ ● เป็นข้อมูลที่ต้องการเชื่อมโยงกับข้อมูลอื่น ในชุดข้อมูลเดียวกัน และเชื่อมโยงกับชุดข้อมูลอื่น ใช่หรือไม่
ข้อแนะนำ	● ควรเลือกใช้กับข้อมูลบ่งชี้ทางตรง หรือข้อมูลที่มีความอ่อนไหว ● ควรใช้การจัดทำข้อมูลนิรนามวิธีอื่นประกอบกัน เพื่อป้องกันการเชื่อมโยงหรืออนุมานกลับการระบุตัวตน

	• ควรมีการเก็บรักษาการเข้ารหัส หรือค่ากุญแจ (key) ที่มีความปลอดภัย โดยกำหนดสิทธิในการเข้าถึงรหัสหรือค่ากุญแจ (key) เนื่องจากข้อมูลแฝงสามารถย้อนกลับมาเป็นข้อมูลเดิมได้ • ควรแยกการจัดเก็บข้อมูลเดิม และข้อมูลผ่านการทำข้อมูลแฝง เพื่อป้องกันการเชื่อมโยงระหว่างชุดข้อมูล • ข้อมูลบันทึก (แถว) ในชุดข้อมูลควรมีปริมาณมากพอเพื่อป้องกันไม่ให้เกิดการเชื่อมโยงกลับไปยังเจ้าของข้อมูลส่วนบุคคลได้ • สามารถประยุกต์ใช้กับข้อมูลที่มีการแบ่งปันให้กับบุคคลภายนอกองค์กร ซึ่งเจ้าของข้อมูล/ผู้ควบคุมข้อมูลสามารถพิจารณาได้ตามความเหมาะสม



รูปที่ 19: ตัวอย่างการทำข้อมูลแฝง

4. การทำให้ข้อมูลเป็นสามัญ (Generalization) คือ การลดความละเอียดข้อมูลที่ระบุตัวตนได้โดยการสรุปข้อมูลหรือจับกลุ่มข้อมูลที่มีลักษณะคล้ายคลึงกันให้อยู่ในกลุ่มเดียวกัน เพื่อตัดความสัมพันธ์ระหว่างบุคคลกับชิ้นข้อมูลโดยไม่เสียคุณค่าข้อมูล

ตารางที่ 7: การทำข้อมูลนิรนามด้วยการทำให้ข้อมูลเป็นสามัญ

คำอธิบาย	การทำให้ข้อมูลเป็นสามัญ เป็นการลดความละเอียดข้อมูล สามารถประยุกต์ใช้กับข้อมูลบ่งชี้ทางตรง ข้อมูลบ่งชี้ทางอ้อม ข้อมูลที่มีคุณลักษณะในการเชื่อมโยงไปยังข้อมูลบ่งชี้ทางตรง (Target attributes) โดยการสรุปข้อมูลหรือจับกลุ่มข้อมูลที่มีลักษณะใกล้เคียงกันให้อยู่ในกลุ่มเดียวกัน ซึ่งอาจแทนที่ด้วยลำดับ แทนตัวเลขจริงได้ เพื่อจัดให้ข้อมูลเป็นกลุ่ม เช่น เลขบัตรประชาชน วันเดือนปีเกิด อายุ ส่วนสูง น้ำหนัก เงินเดือน
หลักเกณฑ์การพิจารณา	• เป็นข้อมูลที่มีการลบข้อมูลบ่งชี้ทางตรงออกจากชุดข้อมูลแล้ว ใช่หรือไม่ • เป็นข้อมูลที่สามารถจับกลุ่มข้อมูลได้ ใช่หรือไม่ • เป็นข้อมูลบ่งชี้ทางอ้อม ใช่หรือไม่ • เป็นข้อมูลที่สามารถแทนที่ด้วยตัวอักษรหรือตัวเลขหรือไม่
ข้อแนะนำ	• ควรมีการลบข้อมูลบ่งชี้ทางตรงออกจากชุดข้อมูล

	• ข้อมูลสามารถเป็นตัวเลขหรือตัวอักษรก็ได้ • สามารถทำให้ข้อมูลเป็นสามัญมากกว่า 1 คอลัมน์ในชุดข้อมูล และสามารถใช้การจัดทำข้อมูลนิรนามวิธีอื่นประกอบกัน เพื่อป้องกันการเชื่อมโยงหรืออนุมานกลับการระบุตัวตน • ข้อมูลบันทึก (แถว) ในชุดข้อมูลควรมีปริมาณมากพอเพื่อป้องกันไม่ให้เกิดการเชื่อมโยงกลับไปยังเจ้าของข้อมูลส่วนบุคคลได้ • สามารถประยุกต์ใช้วิธีนี้กับข้อมูลที่มีการแบ่งปันให้กับบุคคลภายนอกองค์กร ซึ่งเจ้าของข้อมูล/ผู้ควบคุมข้อมูลสามารถพิจารณาได้ตามความเหมาะสม

รูปที่ 20: ตัวอย่างการทำข้อมูลให้เป็นสามัญ

5. การสลับข้อมูล (Swapping/Shuffling/Permutation) คือ การสับเปลี่ยนข้อมูล โดยใช้วิธีการสลับข้อมูล โดยจัดเรียงข้อมูลในชุดข้อมูล ซึ่งจะเรียงแบบใดก็ได้ โดยที่ยังคงข้อมูลเดิมไว้ เพื่อป้องกันการเชื่อมโยงข้อมูลต่างตัวแปรภายในชุดข้อมูลได้

ตารางที่ 8: การทำข้อมูลนิรนามด้วยการสลับข้อมูล

คำอธิบาย	การสลับข้อมูล โดยจัดเรียงข้อมูลตามคอลัมน์ใหม่แบบสุ่ม เพื่อไม่ให้ข้อมูลที่เรียงใหม่ตรงกับข้อมูลเริ่มต้นในชุด เพื่อลดความเสี่ยงในการระบุตัวตน โดยยังคงคุณลักษณะข้อมูลไว้ทั้งหมด
หลักเกณฑ์การพิจารณา	• เป็นข้อมูลที่มีการลบข้อมูลบ่งชี้ทางตรงออกจากชุดข้อมูลแล้ว ใช่หรือไม่ • เป็นข้อมูลบ่งชี้ทางอ้อม ใช่หรือไม่ • เป็นข้อมูลที่สามารถแทนที่ด้วยตัวอักษรหรือตัวเลข ใช่หรือไม่ • เป็นข้อมูลสำคัญที่ต้องการใช้ประโยชน์ โดยไม่จำเป็นต้องรับรู้ข้อมูลบ่งชี้ทางตรง และไม่สามารถลบได้ ใช่หรือไม่ • สามารถรับความเสี่ยงที่อาจส่งผลต่อความแม่นยำและความน่าเชื่อถือชุดข้อมูลได้ ใช่หรือไม่ • การสลับข้อมูลในชุดข้อมูลไม่ตรงกับค่าในชุดข้อมูลเดิม ใช่หรือไม่
ข้อแนะนำ	• ควรมีการลบข้อมูลบ่งชี้ทางตรงออกจากชุดข้อมูล • ข้อมูลสามารถเป็นตัวเลขหรือตัวอักษรก็ได้

	• สามารถใช้การสลับข้อมูลได้มากกว่า 1 คอลัมน์ในชุดข้อมูล และสามารถใช้วิธีการจัดทำข้อมูลนิรนามวิธีอื่นมาประกอบกัน เพื่อป้องกันการเชื่อมโยงมาที่ข้อมูลระบุตนตัว • ข้อมูลภายหลังการสลับต้องไม่ตรงกับข้อมูลเดิม • ข้อมูลบันทึก (แถว) ในชุดข้อมูลควรมีปริมาณมากพอเพื่อป้องกันไม่ให้เกิดการเชื่อมโยงกลับไปยังเจ้าของข้อมูลส่วนบุคคลได้ • สามารถประยุกต์ใช้กับข้อมูลสังเคราะห์ หรือข้อมูลที่มีการแบ่งปันให้กับบุคคลภายนอกองค์กร ซึ่งเจ้าของข้อมูล/ผู้ควบคุมข้อมูลสามารถพิจารณาได้ตามความเหมาะสม

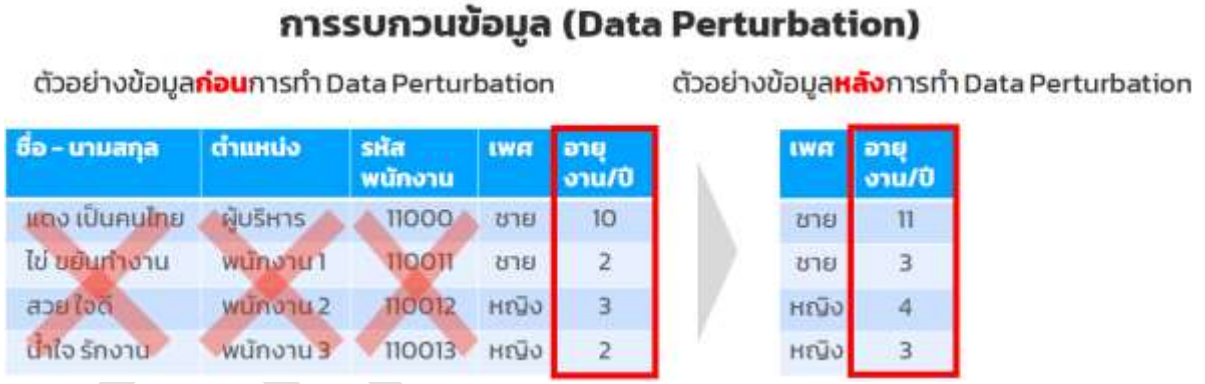
รูปที่ 21: ตัวอย่างการสลับข้อมูล

6. การรบกวนข้อมูล (Data Perturbation) คือ การรบกวนข้อมูล โดยใช้วิธีการปัดเศษและเพิ่มการรบกวนตัวเลข หรือแก้ไขข้อมูลตัวเลขเล็กน้อย เพื่อทำให้ความแม่นยำของข้อมูลมีค่าลดลงและป้องกันการเชื่อมโยงข้อมูลไปยังข้อมูลจริง

ตารางที่ 9: การทำข้อมูลนิรนามด้วยการรบกวนข้อมูล

คำอธิบาย	การรบกวนข้อมูล โดยการปัดข้อมูลที่เป็นตัวเลขในชุดข้อมูล โดยจะปัดให้ข้อมูลมีค่าที่น้อยลงหรือมีค่ามากขึ้น โดยให้พิจารณาเลขโดดตัวที่ถัดจากตำแหน่งที่ต้องการไปทางขวามือตัวเดียว เช่น หากเลขโดดตัวนั้นมีค่าต่ำกว่า 5 ให้ปัดลดลง ตั้งแต่ 5 ขึ้นไปให้ปัดขึ้น อย่างไรก็ดี การใช้วิธีนี้จะคงคุณลักษณะข้อมูลในภาพรวมและไม่มีผลกระทบอย่างมีนัยสำคัญต่อข้อสรุปที่ได้จากการวิเคราะห์ทางสถิติ
หลักเกณฑ์การพิจารณา	• เป็นข้อมูลที่มีการลบข้อมูลบ่งชี้ทางตรงออกจากชุดข้อมูลแล้ว ใช่หรือไม่ • เป็นข้อมูลบ่งชี้ทางอ้อม ใช่หรือไม่ • เป็นข้อมูลที่สามารถแทนที่ด้วยตัวเลข ใช่หรือไม่ • เป็นข้อมูลสำคัญที่ต้องการใช้ประโยชน์ โดยไม่จำเป็นต้องรับรู้ข้อมูลบ่งชี้ทางตรง และไม่สามารถลบได้ ใช่หรือไม่ • สามารถรับความเสี่ยงที่ส่งผลต่อความแม่นยำและความน่าเชื่อถือชุดข้อมูลได้ ใช่หรือไม่ • การปัดเศษตัวเลขในชุดข้อมูลไม่ตรงกับค่าในชุดข้อมูลเดิม ใช่หรือไม่
ข้อแนะนำ	• ควรมีการลบข้อมูลบ่งชี้ทางตรงออกจากชุดข้อมูล

	• ควรใช้การจัดทำข้อมูลนิรนามวิธีอื่นประกอบกัน เพื่อป้องกันการเชื่อมโยงหรืออนุมานกลับการระบุตัวตน • ข้อมูลสามารถเป็นตัวเลขเท่านั้น • สามารถใช้การปัดเศษตัวเลขได้มากกว่า 1 คอลัมน์ในชุดข้อมูล และสามารถใช้การจัดทำข้อมูลนิรนามวิธีอื่น มาประกอบกัน เพื่อป้องกันการเชื่อมโยงหรืออนุมานกลับการระบุตัวตน • ข้อมูลภายหลังการปัดเศษตัวเลขต้องไม่ตรงกับข้อมูลเดิม • ข้อมูลบันทึก (แถว) ในชุดข้อมูลควรมีปริมาณมากพอเพื่อป้องกันไม่ให้เกิดการเชื่อมโยงกลับไปยังเจ้าของข้อมูลส่วนบุคคลได้ • สามารถประยุกต์ใช้วิธีนี้กับข้อมูลสังเคราะห์ หรือข้อมูลที่มีการแบ่งปันให้กับบุคคลภายนอกองค์กร ซึ่งเจ้าของข้อมูล/ผู้ควบคุมข้อมูลสามารถพิจารณาได้ตามความเหมาะสม • การใช้วิธีนี้เป็นการรบกวนข้อมูลซึ่งอาจส่งผลต่อความแม่นยำและความน่าเชื่อถือชุดข้อมูล เช่น ค่าเบี่ยงเบนมาตรฐาน ที่อาจมีการผันแปรจากการปัดเศษตัวเลข



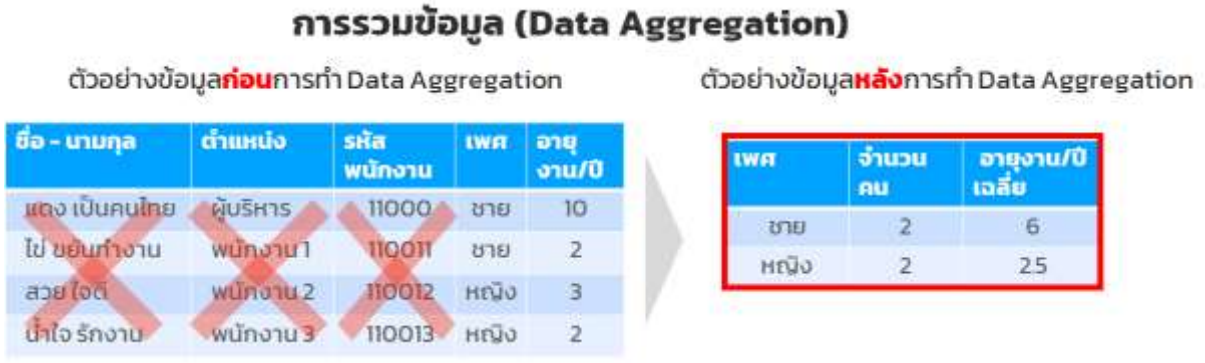
รูปที่ 22: ตัวอย่างการรบกวนข้อมูล

7. การรวมข้อมูล (Data Aggregation) คือ การแปลงข้อมูลให้อยู่ในค่าผลรวม ค่าเฉลี่ย หรือข้อมูลที่สรุปในภาพรวม

ตารางที่ 10: การทำข้อมูลนิรนามด้วยการรวมข้อมูล

คำอธิบาย	การรวมข้อมูลเป็นการสรุปค่าสามารถรวบรวมข้อมูลที่ได้มาจากหลายแหล่ง ให้อยู่ร่วมกันเป็นกลุ่มข้อมูลเดียวกัน ความมีคล้ายเคียงกับการทำให้ข้อมูลเป็นสามัญ (Generalization) แต่วิธีของการผสมข้อมูลมีการสรุปข้อมูลให้บันทึกมีจำนวนที่น้อยลง ในขณะที่การทำให้ข้อมูลเป็นสามัญ (Generalization) จะคงจำนวนบันทึกไว้คงเดิม
หลักเกณฑ์การพิจารณา	• เป็นข้อมูลที่มีการลบข้อมูลบ่งชี้ทางตรงออกจากชุดข้อมูลแล้ว ใช่หรือไม่ • เป็นข้อมูลบ่งชี้ทางอ้อม ใช่หรือไม่ • เป็นข้อมูลสำคัญที่ต้องการใช้ประโยชน์ โดยไม่จำเป็นต้องรับรู้ข้อมูลบ่งชี้ทางตรง และไม่สามารถลบได้ ใช่หรือไม่

	• สามารถรับความเสี่ยงที่อาจส่งผลต่อความแม่นยำและความน่าเชื่อถือชุดข้อมูลได้ ใช่หรือไม่ • ปริมาณข้อมูลในชุดข้อมูลมีเพียงพอต่อการรวมข้อมูล ใช่หรือไม่
ข้อแนะนำ	• ควรมีการลบข้อมูลบ่งชี้ทางตรงออกจากชุดข้อมูล • ข้อมูลบันทึก (แถว) ในชุดข้อมูลควรมีปริมาณมากพอเพื่อป้องกันไม่ให้เกิดการเชื่อมโยงกลับไปยังเจ้าของข้อมูลส่วนบุคคลได้ • สามารถประยุกต์ใช้วิธีนี้กับข้อมูลสังเคราะห์ หรือข้อมูลที่มีการแบ่งปันให้กับบุคคลภายนอกองค์กร ซึ่งเจ้าของข้อมูล/ผู้ควบคุมข้อมูลสามารถพิจารณาได้ตามความเหมาะสม • การใช้วิธีนี้เป็นการรบกวนข้อมูลซึ่งอาจส่งผลต่อความแม่นยำและความน่าเชื่อถือชุดข้อมูล เนื่องจากเป็นการสรุปข้อมูล • ข้อมูลอาจใช้ประโยชน์ได้ไม่เต็มที่เนื่องจากความละเอียดที่แปลงมาในรูปแบบสรุป



รูปที่ 23: ตัวอย่างการรวมข้อมูล

สรุปข้อเสนอแนะในการเลือกใช้วิธีการจัดทำข้อมูลนิรนาม

การในประยุกต์ใช้วิธีการจัดทำข้อมูลนิรนามทั้ง 7 วิธี โดยในชุดข้อมูลแต่ละชุดข้อมูลควรทำข้อมูลนิรนามได้มากกว่า 1 วิธี ซึ่งจะพิจารณาได้จากประเภทข้อมูลและการนำข้อมูลไปใช้ประโยชน์ เช่น การแบ่งปันข้อมูลให้แก่หน่วยงานภายนอก โดยข้อมูลที่บ่งชี้ทางตรงควรมีการลบหรือปิดทับข้อมูล หรือใช้การแฝงข้อมูลเพื่อให้เชื่อมโยงกับชุดข้อมูลอื่นได้ ในขณะที่ข้อมูลที่บ่งชี้ทางอ้อมไม่จำเป็นต้องลบข้อมูล แต่อาจลดความละเอียดข้อมูลหรือเพิ่มการรบกวนข้อมูลเข้าไป โดยข้อพิจารณาถึงความเหมาะสม สามารถสรุปได้ดังนี้

- วิธีการจัดทำข้อมูลนิรนามสามารถประยุกต์ใช้กับข้อมูลหลากหลายประเภทได้ โดยมักยกเว้นสำหรับข้อมูลที่บ่งชี้ทางตรง ซึ่งไม่เหมาะกับการใช้วิธีการ 1) การทำให้ข้อมูลเป็นสามัญ 2) การสลับข้อมูล และ 3) การรบกวนข้อมูล เนื่องจากเป็นวิธีที่ยังคงรักษาข้อมูลที่บ่งชี้ทางตรงเดิมไว้ ทำให้ข้อมูลสามารถระบุตัวตนได้ จึงควรมีการใช้วิธีการลบคุณลักษณะข้อมูลเพื่อกำหนดข้อมูลที่บ่งชี้ทางตรงก่อน และจึงใช้วิธีการจัดทำข้อมูลนิรนามด้วยวิธีต่างๆ กับข้อมูลอื่น ในชุดข้อมูลเดียวกัน เช่น ข้อมูลบ่งชี้ทางอ้อม เพื่อป้องกันการระบุตัวตน

- นอกเหนือจากการพิจารณาการใช้วิธีการจัดทำข้อมูลร่วมกับประเภทข้อมูลบ่งชี้แล้ว ยังควรพิจารณาถึงการใช้ประโยชน์จากข้อมูลร่วมด้วย เพื่อให้ข้อมูลที่ผ่านการจัดทำเป็นข้อมูลนิรนามสามารถนำไปใช้ประโยชน์ได้ โดยเจ้าของข้อมูล/ผู้ควบคุมข้อมูลส่วนบุคคลเป็นผู้พิจารณาการเลือกใช้วิธีการจัดทำข้อมูลนิรนาม

วิธีการจัดทำข้อมูลนิรนาม / ข้อเสนอแนะวิธีการจัดทำข้อมูลนิรนาม	ข้อเสนอแนะสำหรับการเลือกวิธีการจัดทำข้อมูลนิรนามให้เหมาะสมกับประเภทข้อมูล			ข้อเสนอแนะเพื่อการเลือกวิธีการจัดทำข้อมูลนิรนามให้เหมาะสมกับการใช้ประโยชน์		
	ข้อมูลบ่งชี้ทางตรง (Direct identifiers)	ข้อมูลบ่งชี้ทางอ้อม (Indirect identifiers)	ข้อมูลเชื่อมโยงไปข้อมูลบ่งชี้ได้ (Target attributes)	ข้อมูลนำมาวิเคราะห์ได้	ข้อมูลแปลงย้อนกลับได้	ข้อมูลสามารถเชื่อมโยงกับชุดข้อมูลอื่นได้
การลบคุณลักษณะข้อมูล (Suppression)	✓	✓	✓	!	!	!
การปิดทับลักษณะข้อมูล (Character Masking)	✓	✓	✓	!	!	!
การแฝงข้อมูล (Pseudonymization)	✓	✓	✓	!	✓	✓
การทำให้ข้อมูลเป็นสามัญ (Generalization)	!	✓	✓	✓	✓	!
การสลับข้อมูล (Swapping/Shuffling/ Permutation)	!	✓	✓	✓	!	!
การรบกวนข้อมูล (Data Perturbation)	!	✓	✓	✓	!	!
การรวมข้อมูล (Data Aggregation)	✓	✓	✓	✓	!	!

รูปที่ 24: ข้อเสนอแนะเพื่อพิจารณาการเลือกใช้วิธีการจัดทำข้อมูลนิรนาม

3.3. การวิเคราะห์ความเสี่ยงในการเปิดเผยข้อมูล

ตามประกาศกระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม เรื่อง มาตรฐานการรักษาความมั่นคงปลอดภัยของข้อมูลส่วนบุคคล พ.ศ. 2563 ระบุไว้ว่า ความมั่นคงปลอดภัยของข้อมูลส่วนบุคคล หมายถึง การธำรงไว้ซึ่งความลับ (confidentiality) ความถูกต้องครบถ้วน (integrity) และสภาพพร้อมใช้งาน (availability) ของข้อมูลส่วนบุคคล ทั้งนี้ เพื่อป้องกันการสูญหาย เข้าถึง ใช้ เปลี่ยนแปลง แก้ไข หรือเปิดเผยข้อมูลส่วนบุคคลโดยมิชอบ แต่ปัจจุบันยังคงเห็นข่าวการละเมิดข้อมูลส่วนบุคคลเพื่อนำไปใช้ประโยชน์ในทางมิชอบจากการโจรกรรมข้อมูลการสวมรอย ทำให้เกิดความเสียหายต่อตัวบุคคล ดังนั้นการจัดการความเสี่ยงและความปลอดภัยของข้อมูล เช่น การคุ้มครองข้อมูลส่วนบุคคล การเข้าถึงข้อมูล จึงเป็นสิ่งที่หน่วยงานภาครัฐควรตระหนักและให้ความสำคัญเป็นอย่างมาก ซึ่งเป็นส่วนสำคัญที่จะช่วยให้หน่วยงานภาครัฐสามารถพิจารณาประเมินโอกาสที่จะเกิดความเสี่ยง (Likelihood) และ ความร้ายแรง (Severity) เพื่อกำหนดระดับความเสี่ยง (Level of Risk) (จุฬาลงกรณ์มหาวิทยาลัย, 2023)

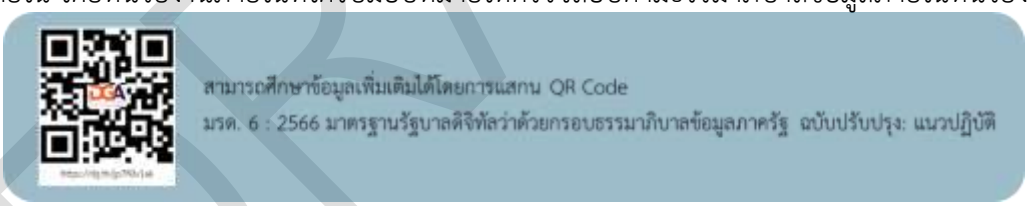
ความรุนแรง				
	ร้ายแรงมาก	ระดับต่ำ	ระดับสูง	ระดับสูง
	ร้ายแรงปานกลาง	ระดับต่ำ	ระดับกลาง	ระดับสูง
	ร้ายแรงน้อย	ระดับต่ำ	ระดับต่ำ	ระดับสูง
		โอกาสต่ำ	**โอกาสพอสมควร**	**โอกาสสูง**
		ความน่าจะเป็นของโอกาสที่จะเกิดขึ้น		

รูปที่ 25: ระดับความเสี่ยง

จากภาพจะเห็นว่าผลกระทบที่มีความร้ายแรงมากไม่จำเป็นต้องมีความเสี่ยงสูงเสมอไป ในทำนองเดียวกันหากความร้ายแรงน้อยแต่มีโอกาสเกิดขึ้นสูงก็ถือเป็นความเสี่ยงสูงได้เช่นกัน การประเมินความเสี่ยงจึงเป็นขั้นตอนที่ต้องการข้อมูลที่ค่อนข้างชัดเจนและเป็นระบบ ทั้งนี้ผู้ควบคุมข้อมูลส่วนบุคคลต้องประเมินความเสี่ยงของผลกระทบจากการประมวลผลข้อมูลดังกล่าวที่จะมีต่อเจ้าของข้อมูลส่วนบุคคล ทั้งในเชิงร่างกาย จิตใจ และทรัพย์สิน โดยควรคำนึงถึงประเด็นเฉพาะว่ามีผลกระทบต่อเจ้าของข้อมูลหรือไม่ เช่น

- ทำให้ไม่สามารถใช้สิทธิได้ตามสมควร ทั้งที่เป็นสิทธิความเป็นส่วนตัว และสิทธิอื่นที่มี
- ทำให้ไม่สามารถเข้าถึงบริการ หรือเสียโอกาสบางอย่าง
- ทำให้ไม่สามารถควบคุมการใช้งานข้อมูลส่วนบุคคลของตนได้
- ทำให้ถูกเลือกปฏิบัติ
- ทำให้ถูกสวมรอยบุคคล (identity theft) หรือหลอกลวงได้
- ทำให้เกิดความเสียหายทางการเงิน ชื่อเสียง หรือร่างกาย
- ทำให้สูญเสียความลับ
- ทำให้ข้อมูลส่วนบุคคลที่ผ่านกระบวนการแฝงข้อมูล (pseudonymization) สามารถระบุตัวบุคคลได้

การควบคุมความเสี่ยง ถือเป็นองค์ประกอบหนึ่งของการดำเนินงานด้านธรรมาภิบาลข้อมูลภาครัฐ ซึ่งการละเมิดข้อมูล การเข้าถึงข้อมูลโดยไม่ได้รับอนุญาต หรือการละเมิดมาตรการ รักษาความมั่นคงปลอดภัยที่นำไปสู่ การทำลาย การสูญหาย การแก้ไข การเปิดเผยข้อมูลที่มีความอ่อนไหว ซึ่งเป็นความเสี่ยงที่ต้องมีการจัดการอย่างเป็นระบบ โดยการประเมินความเสี่ยงและกำหนดการควบคุมเพื่อจัดการความเสี่ยงที่สามารถทำได้ด้วยตัวบุคคลหรือการใช้เทคโนโลยีเข้ามาช่วย ทั้งนี้ควรมีการตรวจประเมินการบริหารจัดการความเสี่ยงด้านข้อมูลภายใน โดยหน่วยงานภายในที่ได้รับมอบหมายให้ตรวจสอบตามธรรมาภิบาลข้อมูลภายในหน่วยงาน



ตาม Thailand Data Protection Guidelines 3.0 แนวปฏิบัติเกี่ยวกับการคุ้มครองข้อมูลส่วนบุคคลได้อธิบายไว้ว่าปัจจัยที่ก่อให้เกิดความเสี่ยงในการระบุตัวตนเจ้าของข้อมูลนิรนามนั้น ประกอบด้วยปัจจัย 2 ประการ ได้แก่ ข้อมูล และ สภาพแวดล้อมข้อมูล

รูปที่ 26: ปัจจัยความเสี่ยงของข้อมูลนิรนาม

ดังนั้นการรักษาความลับเพื่อคุ้มครองข้อมูลส่วนบุคคลนั้น จึงเกิดจากการลดความเสี่ยงของการเปิดเผยข้อมูล โดยมีปัจจัยสำคัญ คือ ลักษณะของข้อมูล เช่น เป็นข้อมูลที่มีความอ่อนไหวหรือไม่ (Sensitive Data)

เป็นต้น และ สิ่งแวดล้อมของข้อมูล เช่น มีข้อมูลสาธารณะจำนวนมากที่อาจนำมาเทียบเคียงเพื่อระบุตัวตนของเจ้าของข้อมูลได้ หรือ จำนวนผู้ที่สามารถเข้าถึงข้อมูลได้ ยิ่งทำให้ต้องมีการทำข้อมูลให้เป็นนิรนามมากยิ่งขึ้น โดยกระบวนการในการจัดทำข้อมูลนิรนามอาจแบ่งออกได้เป็น 2 ขั้นตอน คือ

(1) การพิจารณาสถานการณ์ของข้อมูล

อ้างอิงตาม 3.1 การพิจารณาข้อมูล และ การขจัดข้อมูลระบุตัวตน โดยทั้งนี้ผู้ควบคุมข้อมูลส่วนบุคคลต้องพิจารณาถึงคุณสมบัติหลัก ที่เกี่ยวข้องกับข้อมูล ดังนี้

1. ใครเป็นผู้เป็นเจ้าของข้อมูล เป็นบุคคลธรรมดา หรือเป็นหน่วยข้อมูลที่สามารถระบุบุคคลธรรมดาหรือ เป็นกลุ่มบุคคลที่มีความเป็นไปได้ว่าจะถูกละเมิดสิทธิในข้อมูลส่วนบุคคล
2. ข้อมูลเป็นข้อมูลประเภทใด เป็นข้อมูลอ่อนไหว เป็นข้อมูลตัวเลข ตัวอักษร หรือรูปภาพ เป็นต้น
3. ประเภทของตัวแปรของข้อมูล เป็นตัวแปรที่ระบุตัวตนของเจ้าของข้อมูลได้โดยตรงหรืออาจจะระบุได้ทางอ้อม
4. คุณสมบัติของชุดข้อมูล เช่น คุณภาพของข้อมูล อายุของข้อมูล รวมทั้งโครงสร้างของข้อมูล

โดยข้อมูลที่มีลักษณะต่างกันย่อมมีความเสี่ยงต่อการเปิดเผยข้อมูลส่วนบุคคลต่างกัน โดยหากมีข้อมูลหลายชุดในความควบคุม ผู้ควบคุมข้อมูลส่วนบุคคลหรือผู้ประมวลผลข้อมูลส่วนบุคคล ก็ควรให้ความสำคัญกับข้อมูลที่อาจมีความเสี่ยงสูง

ตารางที่ 11: การพิจารณาสถานการณ์ของข้อมูล

คุณสมบัติ	ความเสี่ยงต่ำ	ความเสี่ยงสูง
คุณภาพของข้อมูล	ต่ำ	สูง
อายุของข้อมูล	เก่า	ใหม่
ระดับของข้อมูล	ข้อมูลรวมกลุ่ม	ข้อมูลรายบุคคลหรือรายหน่วยย่อย
โครงสร้างของข้อมูล	มีมิติเดียว	มีหลายมิติ
ความครบถ้วนข้อมูล	ข้อมูลตัวอย่าง	ข้อมูลประชากร
ข้อมูลที่มีความอ่อนไหว	น้อย	มาก
จำนวนตัวแปรหลัก	น้อย	มาก

จากตารางจะเห็นได้ว่าสามารถเลือกใช้ข้อมูลที่มีความเสี่ยงต่ำได้ โดยไม่กระทบต่อวัตถุประสงค์ของการเก็บข้อมูลหรือการใช้ข้อมูล ซึ่งหากเลือกใช้ข้อมูลที่มีความเสี่ยงสูงจะก่อให้เกิดความเสี่ยงในการเปิดเผยข้อมูลส่วนบุคคลมากขึ้น

นอกจากนี้ยังมีการพิจารณาการใช้งานของข้อมูลเพื่อกำหนดว่าชุดข้อมูลนั้นสามารถนำไปใช้ในกรณีใดได้บ้าง หรือใครมีสิทธิเข้าถึงข้อมูล ซึ่งต้องมีการระบุให้มีความชัดเจน

(2) การวิเคราะห์ความเสี่ยง และมาตรการจัดการความเสี่ยง

เป็นพิจารณาภาพรวมของข้อมูล ข้อมูลต่างลักษณะย่อมมีความเสี่ยงต่อการเปิดเผยข้อมูลส่วนบุคคลต่างกัน โดยการวิเคราะห์สถานการณ์ของข้อมูล เป็นการวิเคราะห์ว่าถ้าหากข้อมูลชุดหนึ่ง นั้นถูกเปิดเผยหรือรั่วไหลออกไป จะมีความเสี่ยงมากน้อยเพียงใดที่ข้อมูลชุดอื่น จะสามารถถูกนำมาใช้ในการระบุตัวตนย้อนกลับได้ โดยวิธีในการละเมิดข้อมูลมีด้วยกันหลากหลายวิธี และแต่ละวิธีก็มีความซับซ้อนที่แตกต่างกันไป ซึ่งส่วนใหญ่มักเกิดจากการนำข้อมูลภายนอกมาเทียบเคียงเพื่อหาความสัมพันธ์จนสามารถนำไปสู่การระบุตัวตนของ

เจ้าของข้อมูลได้ในที่สุด ใช้ความเชื่อมโยงของข้อมูลหลายชุดผ่านตัวแปรหลัก หรือการสรุปคุณลักษณะร่วมกันของคนกลุ่มหนึ่ง รวมถึงการตัดกรณีที่เป็นไปไม่ได้ออกไป ดังนั้นการกำหนดมาตรการในการควบคุมความเสี่ยงจึงต้องกำหนดให้สอดคล้องกับสถานการณ์ของข้อมูล ซึ่งอาจทำได้สองวิธี

1. การเปลี่ยนข้อมูล ต้องคำนึงถึง 2 ปัจจัย คือ ความง่ายต่อการเปิดเผยข้อมูลส่วนบุคคลและความอ่อนไหวของข้อมูลส่วนบุคคลในชุดข้อมูลนั้น วิธีการนี้ไม่เกิดผลกระทบกับเนื้อข้อมูลแต่เป็นการลดความเสี่ยงในการเปิดเผยชุดข้อมูล โดยมีมาตรฐานที่นิยมใช้ในการตรวจสอบข้อมูลว่ามีความปลอดภัยจากการระบุตัวตนมากน้อยเพียงใด คือ k-anonymization ซึ่งเป็นการใช้หลักการเรื่อง การยืนยันตัวตนในวิชาพีชคณิตเชิงเส้น กล่าวคือหากมีแถวของข้อมูลที่เป็นอิสระในเชิงเส้นจากกันน้อยกว่าจำนวนตัวแปร ย่อมเป็นกรณีที่อาจเป็นข้อมูลของใครก็ได้ เช่น หากมีผู้ทราบว่าคนที่ป่วยนั้นมีผลรวมของอายุกับสี่เท่าของวันเกิดมีผลรวมเป็น 100 ดังสมการ

$$x + 4y = 100$$

จะมีความน่าจะเป็นมากมายที่จะมีข้อมูลของคู่ตัวแปร x หรือ y ได้ไม่จำกัดจำนวนที่เป็นไปตามข้อมูลดังกล่าว แต่ถ้าเกิดมีข้อมูลที่เป็นอิสระในเชิงเส้นจากกันเท่ากับจำนวนของคู่ตัวแปร เช่น

$$x + 2y = 7$$

$$3x-y = 7$$

กรณีนี้สามารถสรุปได้โดยง่ายว่า $x = 3$ และ $y = 2$ และหาเจ้าของข้อมูลที่มีลักษณะดังกล่าวได้ทันทีโดยการพิจารณาปัจจัยที่ส่งผลต่อระดับที่เหมาะสมของการจัดทำข้อมูลนิรานาม ผู้จัดทำข้อมูลนิรนามอาจพิจารณาปัจจัยหลักได้ 2 ประการ คือ

(1) ความเสี่ยงในการถูกเปิดเผยของข้อมูล (Data disclosiveness)

(2) ความอ่อนไหวของข้อมูล (Data sensitivity)

โดยเฉพาะในเรื่องที่ความเสี่ยงในการถูกเปิดเผยของข้อมูลนั้นขึ้นอยู่กับปัจจัยอื่นเป็นจำนวนมาก ทั้งตัวข้อมูลเอง และสิ่งแวดล้อมของข้อมูลที่ รวมถึง ขนาดของข้อมูล (Data Size) จำนวนตัวแปรหลัก ความยากง่ายในการหาข้อมูลภายนอกที่มีตัวแปรหลักเพื่อเทียบเคียง จำนวนคนที่อาจเข้าถึงทั้งข้อมูลของผู้จัดทำข้อมูลนิรนาม เป็นต้น ดังนั้นจึงจำเป็นต้องมีการกำหนดปัจจัยสำคัญที่สุด 3 ปัจจัยที่จะส่งผลต่อความเสี่ยงในการถูกเปิดเผยข้อมูล โดยควรเป็นทั้งปัจจัยที่เป็นตัวข้อมูลเองและสิ่งแวดล้อมของข้อมูล *สามารถศึกษากรอบแนวคิดและขั้นตอนการดำเนินการได้ที่ G4. การตัดสินใจถึงระดับของการจัดทำข้อมูลนิรนาม (หน้า 293.) Thailand Data Protection Guidelines 3.0 ศูนย์วิจัยกฎหมายและการพัฒนา คณะนิติศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย*

2. การปรับสิ่งแวดล้อม คือการควบคุมการเข้าถึงข้อมูล ไม่ว่าจะบุคคลที่สามารถเข้าถึงข้อมูลได้ วิธีการในการเข้าถึงข้อมูล และวัตถุประสงค์ของการเข้าถึงข้อมูล ต้องมีการกำหนดสิทธิการเข้าถึงข้อมูลและจะต้องปฏิบัติตามกฎระเบียบอย่างเคร่งครัด ทั้งนี้หากต้องการลดความเสี่ยงผู้ควบคุมข้อมูลต้องมีการกำหนดมาตรฐานหรือเงื่อนไขในก่อนการเข้าถึงข้อมูล โดยวิธีในการเปิดเผยข้อมูลให้แก่บุคคลภายนอก มี 4 วิธี ซึ่งเรียงลำดับในการควบคุมการเข้าถึงและใช้ข้อมูลตามความจำเป็นจากน้อยไปมาก ดังนี้

2.1 การเปิดให้ใช้ข้อมูลโดยทั่วไป (open access)

2.2 การจัดส่งข้อมูลให้เป็นรายกรณี (delivered access)

2.3 การใช้ข้อมูล ณ สถานที่ที่จัดเตรียมไว้ (on-site safe settings)

2.4 การใช้ใบอนุญาต (Licenses)

อย่างไรก็ตามแม้ข้อมูลที่ได้ผ่านการดำเนินการโดยใช้เทคนิคหรือวิธีการ อันหลากหลาย เพื่อทำให้ข้อมูลส่วนบุคคลนั้นอยู่ในรูปของข้อมูลนิรนาม ซึ่งถือเป็นเครื่องมือสำคัญในการคุ้มครองข้อมูลส่วนบุคคลเพื่อลดความเสี่ยงและผลกระทบที่อาจเกิดขึ้นกับบุคคลลงได้ แต่ทั้งนี้ก็ยังมีโอกาสที่จะเกิดขึ้นได้ ดังนั้น ผู้ควบคุมข้อมูลส่วนบุคคลยังคงต้องคำนึงถึงความเสี่ยง 3 ประการดังนี้ (Article 29 Data Protection Working Party (European Commission), 2014)

ประการที่ 1 ข้อมูลแฝงไม่เทียบเท่ากับข้อมูลนิรนาม เนื่องจากยังสามารถเชื่อมโยงข้อมูลทำให้ระบุตัวตนบุคคลได้ ดังนั้นถือว่ายังคงอยู่ภายใต้กฎหมายคุ้มครองข้อมูลส่วนบุคคล เช่น การวิจัยทางวิทยาศาสตร์ สถิติ หรือประวัติศาสตร์

ประการที่ 2 ข้อมูลนิรนามที่ไม่ปฏิบัติตามข้อกำหนด เงื่อนไข และเกณฑ์ที่กำหนดไว้ได้ เช่น มาตรา 5(3) ของ e-Privacy Directive ห้ามการจัดเก็บและการเข้าถึงข้อมูลทุกประเภท รวมถึงข้อมูลที่ไม่ใช่ข้อมูลส่วนบุคคล บนอุปกรณ์ปลายทางโดยไม่ได้รับความยินยอมจากสมาชิกหรือผู้ใช้งาน ซึ่งเป็นส่วนหนึ่งของการรักษาความลับของการสื่อสาร

ประการที่ 3 ความประมาทที่เกิดจากการนำข้อมูลนิรนามไปใช้งานหรือเผยแพร่โดยไม่พิจารณาผลกระทบต่อบุคคลซึ่งอาจทำให้สูญเสียความเป็นส่วนตัวของเจ้าของข้อมูล ควรมีการกำหนดวัตถุประสงค์ตามความจำเป็นในการใช้ข้อมูล เพื่อประมวลผลตามปัจจัยที่เกี่ยวข้อง เช่น ลักษณะของความสัมพันธ์ระหว่างเจ้าของข้อมูลและผู้ควบคุมข้อมูล ภาระผูกพันทางกฎหมายที่เกี่ยวข้อง ความโปร่งใสในการประมวลผล ความชัดเจนในการดำเนินงาน

ดังนั้นการใช้ประโยชน์จากข้อมูลส่วนบุคคลแม้ว่าจะอยู่ในรูปของข้อมูลนิรนามแล้วก็ตาม ควรมีการใช้งานอย่างระมัดระวัง ไม่ละเมิดสิทธิ์และความเป็นส่วนตัวของบุคคล ใช้ข้อมูลตามวัตถุประสงค์หรืออำนาจหน้าที่ที่ได้รับ เพื่อเป็นการป้องกันไม่ให้เกิดข้อมูลส่วนบุคคลรั่วไหลและนำไปใช้ในการแสวงหาผลประโยชน์จากกลุ่มคนผู้ไม่หวังดี ซึ่งอาจทำให้สูญเสียความเป็นส่วนตัวของเจ้าของข้อมูลและเกิดผลกระทบที่รุนแรงตามมาได้

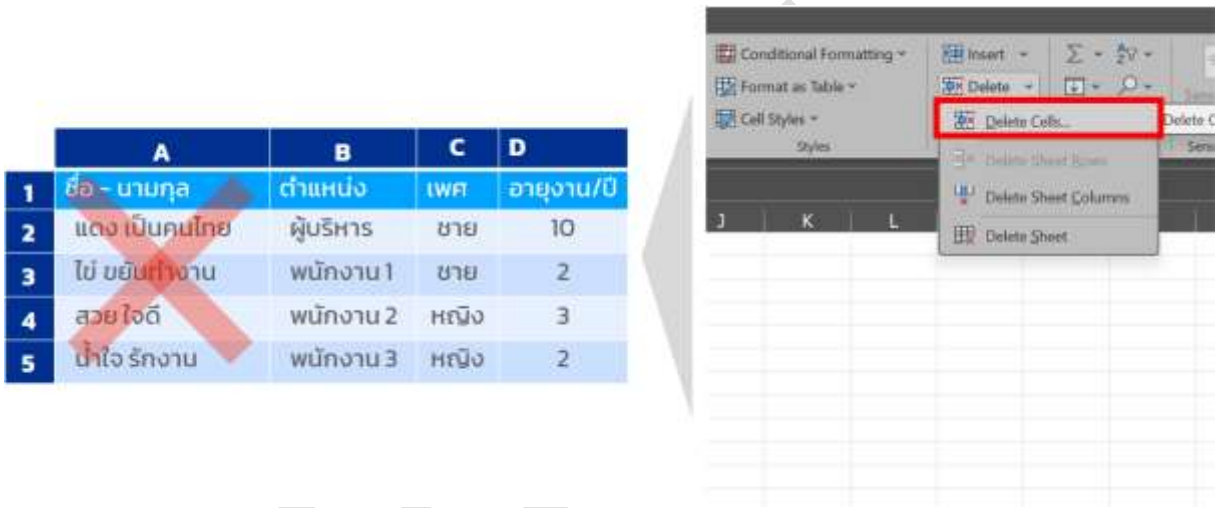
4. ภาคผนวก

4.1. เครื่องมือการจัดทำข้อมูลนิรนาม

1. ตัวอย่างการจัดทำข้อมูลนิรนามเบื้องต้น โดยการใช้โปรแกรม Microsoft Excel

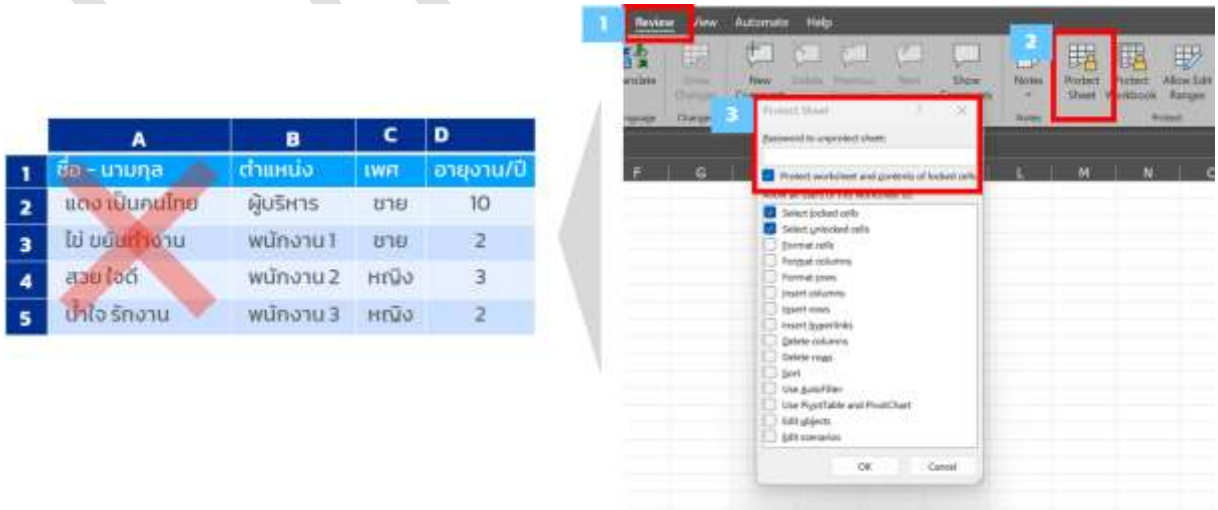
จากหัวข้อ 3.2 หลักเกณฑ์การจัดทำข้อมูลนิรนาม จะเห็นได้ว่า วิธีการจัดทำ ข้อมูลนิรนามที่ควรใช้กับข้อมูลบ่งชี้ทางตรง เพื่อป้องกันการนำข้อมูลไปใช้ประโยชน์ ประกอบด้วย การลบข้อมูลและการปิดทับข้อมูล ซึ่งมีแนวทางเบื้องต้นได้ ดังนี้

- **กรณีต้องการลบคุณลักษณะเฉพาะ (Attribute Suppression)** ให้ 1) คลิกขวาในเซลล์คอลัมน์ (Column) ที่ต้องการลบ และ 2) คลิกบนแถบเครื่องมือด้านบน คลิก Delete Cell



รูปที่ 27: ตัวอย่างการลบคุณลักษณะเฉพาะ

หากต้องการใช้วิธีการซ่อนคอลัมน์ สามารถกดไปที่ 1) คลิกขวาในเซลล์คอลัมน์ (Column) ที่ต้องการซ่อน และกดซ่อน (Hide) 2) จากนั้นมาที่ เมนูทบทวน (Review) 3) เลือก Protect Sheet และใส่รหัสการเข้าถึงคอลัมน์



รูปที่ 28: ตัวอย่างการซ่อนคอลัมน์

- **กรณีต้องการลบข้อมูลรายบันทึก (Record Suppression)** ให้ 1) คลิกขวาในเซลล์แถว (Row) ที่ต้องการลบ และ 2) คลิกบนแถบเครื่องมือด้านบน ให้คลิก Delete Row (ลบแถว)

	A	B	C	D
1	ชื่อ - นามสกุล	ตำแหน่ง	เพศ	อายุงาน/ปี
2	แดง เป็นคนไทย	ผู้บริหาร	ชาย	10
3	ไข่ ขยันทำงาน	พนักงาน 1	ชาย	2
4	สวย ใจดี	พนักงาน 2	หญิง	3
5	น้ำใจ รักงาน	พนักงาน 3	หญิง	2

รูปที่ 29: ตัวอย่างการลบข้อมูลรายบันทึก

- **การปิดทับลักษณะข้อมูลเบื้องต้น** โดยการใช้โปรแกรม Microsoft Excel

วิธีนี้จะใช้ฟังก์ชัน RIGHT เพื่อแสดงตัวเลขสุดท้าย โดยต้องเลือกคอลัมน์ (Column) ที่ต้องการปิดทับข้อมูล และใส่ฟังก์ชัน RIGHT เพื่อปิดทับข้อมูลในคอลัมน์นั้น

การใช้ฟังชั่น	ข้อมูลก่อนการ Masking	ข้อมูลหลังการ Masking
=("**** ") & RIGHT (C2, 2))	11001	**** 01

คอลัมน์ที่ต้องการแสดง

จำนวนอักขระที่ต้องการปิดทับ ซึ่งควรมีจำนวนเท่ากันกับข้อมูลเดิม

จำนวนอักขระตัวท้ายที่ต้องการแสดง (จำนวนหลักเลข)

ตัวอย่างการใช้ฟังก์ชัน RIGHT

	A	B	C	D	
1	เพศ	อายุงาน /ปี	รหัสพนักงาน ก่อนการ Masking	รหัสพนักงาน หลังการ Masking	ตัวอย่างการใช้ฟังชั่น
2	ชาย	10	515-43-51101	****-**-51101	=("****-**-"&RIGHT(C2,5))
3	ชาย	2	515-43-51102	@@@-@@-51102	=("@@@-@@-"&RIGHT(C3,5))
4	หญิง	3	515-43-51103	XXX-XX-51103	=("XXX-XX-"&RIGHT(C4,5))

รูปที่ 30: ตัวอย่างการใช้ฟังก์ชัน RIGHT

ดังตัวอย่างข้างต้น จะเห็นได้ว่า จะต้องมีการคัดลอกข้อมูลจากคอลัมน์ C เพื่อปิดทับข้อมูลในคอลัมน์ D ซึ่งหากข้อมูลในคอลัมน์ D มีการปิดทับโดยใช้ฟังก์ชัน RIGHT เรียบร้อย ควร 1) ลบข้อมูลในคอลัมน์ C โดยการคัดลอก (Copy) ข้อมูลและนำไปวาง (Paste) แบบ Value ในคอลัมน์ใหม่ เพื่อแสดงแค่ข้อมูลที่ผ่าน

การปิดทับ หรือ 2) ซ่อนคอลัมน์ข้อมูลก่อนการปิดทับ (คอลัมน์ C) โดยการเข้ารหัสตามวิธีที่กล่าวไว้ด้านบน (**หัวข้อ กรณีต้องการลบคุณลักษณะเฉพาะ**)

วิธีนี้จะใช้ฟังก์ชัน REPT เพื่อปิดทับข้อมูล โดยต้องเลือกคอลัมน์ (Column) ที่ต้องการปิดทับข้อมูล และใส่ฟังก์ชัน REPT และเลือกตัวอักขระเพื่อปิดทับข้อมูลในคอลัมน์นั้น และเมื่อมีการปิดทับข้อมูลเรียบร้อยแล้ว ควรมีการซ่อนคอลัมน์ข้อมูลก่อนการปิดทับ (คอลัมน์ C) โดยการเข้ารหัสตามวิธีที่กล่าวไว้ด้านบน

การใช้ฟังชัน	ข้อมูลก่อนการ Masking	ข้อมูลหลังการ Masking
ฟังชัน =(REPT("*", LEN (C2))) — อักขระที่ต้องการแสดงเมื่อมีการปิดทับ; คอลัมน์ที่ต้องการปิดทับเนื้อหา	11001	*****

ตัวอย่างการใช้ฟังก์ชัน REPT

	A	B	C	D	
1	เพศ	อายุงาน /0	รหัสพนักงาน ก่อนการ Masking	รหัสพนักงาน หลังการ Masking	ตัวอย่างการใช้ฟังชัน
2	ชาย	10	515-43-51101	**********	=(REPT("*",LEN(C2)))
3	ชาย	2	51102	WWWWW	=(REPT("W",LEN(C3)))
4	หญิง	3	001	@@@	=(REPT(@",LEN(C2)))

รูปที่ 31: ตัวอย่างการใช้ฟังก์ชัน REPT

2. ตัวอย่างรายการเครื่องมือการจัดทำข้อมูลนิรนามเชิงพาณิชย์หรือโอเพ่นซอร์ส (Open Source)

อ้างอิงจาก The Personal Data Protection Commission , Guide to Basic Data Anonymisation (31 March 2022) และแนวทางสำหรับการจัดทำข้อมูลนิรนามขั้นพื้นฐาน โดยสำนักงานคณะกรรมการคุ้มครองข้อมูลส่วนบุคคล (สคส.) หรือ PDPC ร่วมกับ สวทช. สรุปได้ดังนี้

ตารางที่ 12: รายการเครื่องมือการจัดทำข้อมูลนิรนามเชิงพาณิชย์หรือโอเพ่นซอร์ส

รายการ	คำอธิบาย	ที่มา
PDPC	เป็นเครื่องมือเพื่อตามกระบวนการจัดทำข้อมูลนิรนามในการแปลงให้เป็นข้อมูลนิรนาม ตั้งแต่การการพิจารณาว่าข้อมูล การเลือกใช้เทคนิคข้อมูลนิรนามกับข้อมูลบ่งชี้ทางตรง ไปจนถึงการการประเมินความเสี่ยง	https://www.pdpc.gov.sg/help-and-resources/2018/01/basic-anonymisation

รายการ	คำอธิบาย	ที่มา
Amnesia	เป็นเครื่องมือในการแปลงให้เป็นข้อมูลนิรนาม โดยลบข้อมูลบ่งชี้ทางตรงและข้อมูลอ่อนไหว โดยมีการประเมินโดยใช้ k-anonymity และ km-anonymity	https://amnesia.openaire.eu/
ARGUS	เครื่องมือนี้ใช้วิธีการลบข้อมูลระบุตัวบุคคลทางสถิติ โดยการจัดทำข้อมูลนิรนามด้วยวิธี 1) การทำข้อมูลให้เป็นสามัญ 2) การรบกวนข้อมูล และ 3) การรวมข้อมูล	https://research.cbs.nl/casc/mu.htm
ARX	เป็นแบบจำลอง (Model) เพื่อแสดงการจัดทำข้อมูลนิรนามในหลากหลายมิติ เช่น การเลือกใช้วิธีการจัดทำข้อมูลนิรนาม การวิเคราะห์การใช้ประโยชน์ของข้อมูล และ การวิเคราะห์ความเสี่ยงในการระบุตัวตัวตน	https://arx.deidentifier.org/

4.2. กรณีศึกษาการจัดทำข้อมูลนิรนามของสถาบันข้อมูลขนาดใหญ่ (องค์การมหาชน)

ในยุคดิจิทัล ข้อมูลกลายเป็นทรัพยากรที่มีค่ามหาศาล การเข้าถึงและวิเคราะห์ข้อมูลอย่างมีประสิทธิภาพช่วยขับเคลื่อนการพัฒนาหน่วยงาน องค์กร หรือประเทศได้ในหลากหลายมิติ แต่ทว่าข้อมูลที่ถูกเก็บรวบรวมโดยหน่วยงานและองค์กรต่าง ๆ มักถูกนำมาใช้ประโยชน์ภายในหน่วยงานและองค์กรเท่านั้น การแบ่งปันข้อมูล (Data Sharing) จึงเป็นแนวทางสำคัญในการเพิ่มศักยภาพของข้อมูล ให้หน่วยงาน องค์กร หรือบุคคลอื่น ๆ สามารถเข้าถึงและใช้ประโยชน์จากข้อมูลที่มีอยู่ร่วมกันได้ เพื่อประโยชน์ในหลายภาคส่วน อาทิ สนับสนุนการตัดสินใจเชิงนโยบาย ส่งเสริมการวิจัยและพัฒนา เพิ่มประสิทธิภาพในการทำงาน พัฒนาสินค้าและบริการใหม่ ๆ หรือ กระตุ้นเศรษฐกิจ เป็นต้น อย่างไรก็ดี แม้การแบ่งปันข้อมูลจะมีประโยชน์มากมาย แต่ก็ยังมีความกังวลและความท้าทายที่ต้องพิจารณาโดยเฉพาะประเด็นความเป็นส่วนตัวและความปลอดภัยของข้อมูล การนิรนามข้อมูล (Data Anonymization) เป็นหนึ่งวิธีการที่ช่วยลดความกังวลและความเสี่ยงที่เกี่ยวข้องกับการแบ่งปันข้อมูลได้

การแบ่งปันข้อมูลด้านการท่องเที่ยวเป็นหนึ่งกลไกสำคัญที่ช่วยขับเคลื่อนเศรษฐกิจ ส่งเสริมการพัฒนาการท่องเที่ยวอย่างยั่งยืน สร้างพัฒนาประโยชน์ให้ทุกภาคส่วน และร่วมสร้างอนาคตการท่องเที่ยวที่ขับเคลื่อนด้วยข้อมูล คณะทำงานโครงการ Travel Link ภายใต้หน่วยงานสถาบันข้อมูลขนาดใหญ่ (องค์การมหาชน) ร่วมกับหน่วยงานพันธมิตร อาทิ สำนักงานปลัดกระทรวงการท่องเที่ยวและกีฬา การท่องเที่ยวแห่งประเทศไทย สำนักงานตรวจคนเข้าเมือง กรมการปกครอง ฯลฯ จึงได้ร่วมมือกันเชื่อมโยงข้อมูลด้านการท่องเที่ยวเพื่อนำมาสร้างประโยชน์ต่อภาครัฐ เอกชน ประชาชน และประเทศ โดยมุ่งเน้นการรักษาความเป็น

ส่วนตัว ความปลอดภัย และความสมบูรณ์ของข้อมูลเป็นหลัก การนิรนามข้อมูลจึงถูกนำมาใช้ก่อนการเชื่อมโยงข้อมูลเพื่อสร้างความปลอดภัยสูงสุดในการเชื่อมโยงข้อมูล ในบทนี้ คณะทำงานได้สรุปปัจจัยหลักในการนิรนามข้อมูล การนิรนามฟิลด์ข้อมูล พร้อมทั้งการเข้ารหัสไฟล์เพื่อใช้ในการเชื่อมโยงข้อมูล

ปัจจัยที่ส่งผลต่อการนิรนามข้อมูล

การนิรนามข้อมูลมีหลากหลายวิธีขึ้นโดยอาจเลือกตามความเหมาะสมจากปัจจัยดังต่อไปนี้

- **ความปลอดภัยของข้อมูล** ในแต่ละกระบวนการนิรนามข้อมูลนั้น ย่อมมีความปลอดภัยของข้อมูลที่แตกต่างกัน เช่น การปิดทับข้อมูลบางส่วนนั้นย่อมมีความปลอดภัยน้อยกว่าการปิดบังข้อมูลทั้งหมด เนื่องจากการปิดทับข้อมูลอาจเดาได้ง่ายกว่าเมื่อประกอบกับข้อมูลอื่น ๆ
- **ความเร็วในการประมวลผล** เมื่อข้อมูลมีปริมาณมาก เวลาที่ใช้ในการประมวลผลข้อมูลในแต่ละวิธีก็อาจแตกต่างกัน ซึ่งขึ้นอยู่กับความซับซ้อนของวิธีการประมวลผลและทรัพยากรที่ใช้ ส่งผลให้มีระยะเวลารอคอยการประมวลผล หากระยะเวลาการรอคอยไม่สามารถเป็นที่ยอมรับได้ ก็อาจต้องเปลี่ยนเป็นวิธีการอื่นหรือเพิ่มทรัพยากรที่ใช้ในการประมวลผล
- **การนำข้อมูลไปใช้ต่อ** การใช้ประโยชน์จากข้อมูลถือเป็นเรื่องสำคัญในวางแผนและตัดสินใจในการดำเนินกิจการต่าง ๆ การนิรนามข้อมูลด้วยวิธีการที่ยังคงไว้ซึ่งประโยชน์จากการใช้ข้อมูลจึงสำคัญด้วย เช่น การจัดกลุ่มนักท่องเที่ยวในประเทศไทย ซึ่งอายุอาจมีผลต่อการจัดกลุ่มดังกล่าว อาจใช้ช่วงอายุแทนอายุที่มีความเสี่ยงในการถูกระบุตัวตนย้อนกลับได้ง่ายกว่า หรือกรณีที่ต้องการแยกว่าข้อมูลนั้นเป็นข้อมูลของคนละบุคคลกันก็อาจใช้การเข้ารหัสข้อมูลทางเดียว (Hashing) กับข้อมูลส่วนบุคคลในการสร้างรหัสจำแนกตัวตนได้

การนิรนามฟิลด์ข้อมูล

ด้วยปัจจัยในการเลือกวิธีการในการนิรนามข้อมูลที่ได้กล่าวมา จึงได้พิจารณาการเลือกใช้วิธีการต่าง ๆ กับประเภทข้อมูล และรูปแบบการนำไปใช้งานต่อที่แตกต่างกันดังนี้

ตารางที่ 13: ตารางแสดงวิธีการเลือกการนิรนามฟิลด์ข้อมูล

ประเภทฟิลด์ข้อมูล	รูปแบบการนำไปใช้งาน	วิธีการ
ข้อมูลส่วนบุคคล	ต้องการแยกแยะตัวตนแต่ละบุคคล	เข้ารหัสข้อมูลทางเดียวขั้นสูง
ข้อมูลส่วนบุคคล	เผยแพร่สู่สาธารณะ	ลบฟิลด์ข้อมูล
ข้อมูลที่เชื่อมโยงกับข้อมูลส่วนบุคคล	เผยแพร่สู่สาธารณะ	ลบฟิลด์ข้อมูล หรือ ลดความละเอียดข้อมูล
ข้อมูลที่ไม่ใช่ข้อมูลสาธารณะ	เผยแพร่สู่สาธารณะ	ผสมข้อมูล

การเข้ารหัสข้อมูลทางเดียวขั้นสูงนั้นเป็นการต่อยอดจากการเข้ารหัสข้อมูลทางเดียวเพื่อให้เกิดความปลอดภัยจากการถูกเดาสุ่มค่าข้อมูลด้วยอัลกอริทึมการเข้ารหัสทางเดียวที่มีอยู่ในปัจจุบัน เช่น หากผู้ไม่ประสงค์ดีต้องการเดาข้อมูลวันเดือนปีเกิด ด้วยอัลกอริทึม SHA-512 ก็อาจทำการสร้างตารางข้อมูลวันเดือนปีเกิดด้วยอัลกอริทึมดังกล่าว แล้วนำไปเทียบกับข้อมูลที่ถูกเข้ารหัสไว้แล้วเพื่อแปลงกลับเป็นวันเดือนปีเกิด เนื่องจากการเข้ารหัสหากเป็นค่าเดิมและใช้อัลกอริทึมเดิมทุกครั้งจะได้ค่าเดิมเสมอ

ตารางที่ 14: ตัวอย่างตารางข้อมูลเงินเดือนพนักงาน

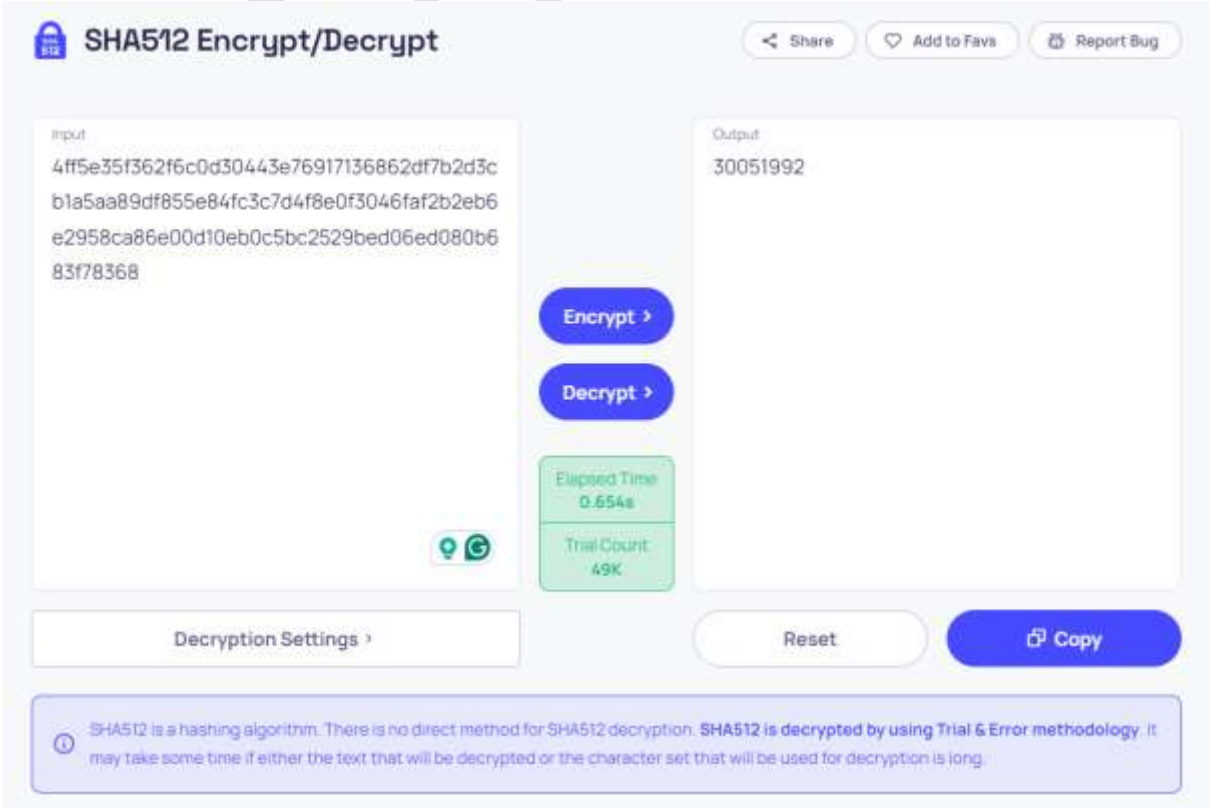
เพศ	วันเกิด	เงินเดือน
ชาย	4ff5e35f362f6c0d30443e76917136862df7b2d3cb1a5aa89df855e84fc3c7d4f8e0f3046faf2b2eb6e2958ca86e00d10eb0c5bc2529bed06ed080b683f78368	30,000

ตารางที่ 15: ตัวอย่างตารางการสุ่มเดาค่าวันเกิด

รหัสที่ได้จาก SHA-512	ค่าวันเกิด
ee5542eee1b6aea5eecd5fcc289202c5ab1edaa6b2e1e050fff1a3b35ce08ffd6ff69bf0a5774af5378c790beacd889a3d8185078feddebe8d0efabd6639de7a	28051992
6b00221214caafb5009d30d8c4fb897ecc27b98bff72baf9cb9709a1058c58433e34911b8ca8d77794de0268c8c26f944966c3184bb0747ef72ef173999ca5c4	29051992
4ff5e35f362f6c0d30443e76917136862df7b2d3cb1a5aa89df855e84fc3c7d4f8e0f3046faf2b2eb6e2958ca86e00d10eb0c5bc2529bed06ed080b683f78368	30051992
a7eb47d1221634fdc5e305d184a780216a531543e062df346337d0ec0d06983963f2540d306f433aadf385a5fb48f9db94f3cc1cb52022f8bf9aab513f2a25f1	31051992

เมื่อทำการเชื่อม 2 ตารางนี้เข้าด้วยกันด้วยรหัสวันเกิด จะทำให้เดาได้ว่าฟิลด์ข้อมูลที่ถูกเข้ารหัสดังกล่าวเป็นวันที่ 30/05/1992 ซึ่งผู้ไม่ประสงค์ดีอาจระบุตัวตนของบุคคลนี้ได้ง่ายขึ้นและทราบว่าใครเป็นผู้มีเงินเดือน 30,000 บาท

นอกจากนี้การเข้ารหัสข้อมูลด้วยปริมาณข้อมูลในฟิลด์ที่น้อยไป ก็อาจถูกสุ่มเดาย้อนกลับด้วยเครื่องประมวลผลที่มีกำลังประมวลผลสูงได้



รูปที่ 32: ภาพจากการใช้งานบนเว็บไซต์ [1]

จากตัวอย่างที่เกิดขึ้นการเข้ารหัสทางเดียวขั้นสูงจึงสำคัญในการเก็บรักษาความลับของข้อมูลให้มั่นคงปลอดภัย โดยสามารถเพิ่มองค์ประกอบต่าง ๆ ประกอบกับข้อมูลตั้งต้นได้ดังนี้

- **ค่าประกอบการเข้ารหัส (salt) เป็นค่าสุ่มเพื่อทำให้การเดานั้นยากขึ้นโดยเป็นค่าที่ควรประกอบด้วยตัวอักษรพิมพ์เล็ก ตัวอักษรพิมพ์ใหญ่ ตัวเลข และเครื่องหมาย เช่น Q65e;=Cyx$hr8+?H**
- **วิธีการสลับหรือเรียงข้อมูล** เช่น การเรียงหลังไปหน้า การสลับตัวอักษร 5 ตัวแรกกับ 5 ตัวสุดท้าย จากนั้นสร้างรูปแบบการผสมค่าข้อมูล ค่าประกอบการเข้ารหัส และวิธีการสลับหรือเรียงข้อมูล เพื่อให้เป็นรูปแบบที่ตายตัวในการเข้ารหัสแต่ละครั้ง

ตัวอย่างกระบวนการ	ใช้ข้อมูลดั้งเดิม	เรียงจากหลังมาหน้า + ค่าประกอบการสุ่ม	สลับ 2 ตัวหน้ากับ 2 ตัวท้าย + ค่าประกอบการสุ่ม
ตัวอย่างข้อมูล	30051992	29915003Ox9udp@d	92051930Ox9udp@d
ผลลัพธ์	4ff5e35f362f6c0d30443e769 17136862df7b2d3cb1a 5aa89df855e84fc3c7d4f8e0f304 6faf2b2eb6e2958ca 86e00d10eb0c5bc2529bed06e d080b683f78368	e0ebd53a34b3f8d0cf3b1610ec e24e62f9d739c6e00c926c3d58 93cdc2b7ab08abb84ebc4cd9b9 d818932af2d119f81f87517ba8642 e8a20f9ed97bb80364c75	8f607ff69696b851ae2f9 0ebb0931c6670d71282 7582d47171ad5bb36e e7eeb29e56d4ea41a1 003773b4882d4338e1c4db b9e681439e3dfd432a30fd521dba62

รูปที่ 33: ตัวอย่างการสร้างรูปแบบการผสมข้อมูลก่อนเข้ารหัส

การเข้ารหัสไฟล์เพื่อใช้ในการรับส่งข้อมูล

คณะทำงานได้ทำการประยุกต์ใช้หลักการจดหมายดิจิทัล (Digital Envelope) ในการเข้ารหัสไฟล์เพื่อใช้ในการแลกเปลี่ยนข้อมูล โดยการเข้ารหัสนั้นจะเป็นการเข้ารหัสที่สามารถถอดกลับมาเป็นค่าก่อนเข้ารหัสได้ (Encryption) โดยจะแบ่งการเข้ารหัสเป็นทั้งหมด 2 รูปแบบ ได้แก่แบบสมมาตร (Symmetric Encryption) และไม่สมมาตร (Asymmetric Encryption) โดยมีความแตกต่างกันดังนี้

ตารางที่ 16: ตารางเปรียบเทียบรูปแบบการเข้ารหัส

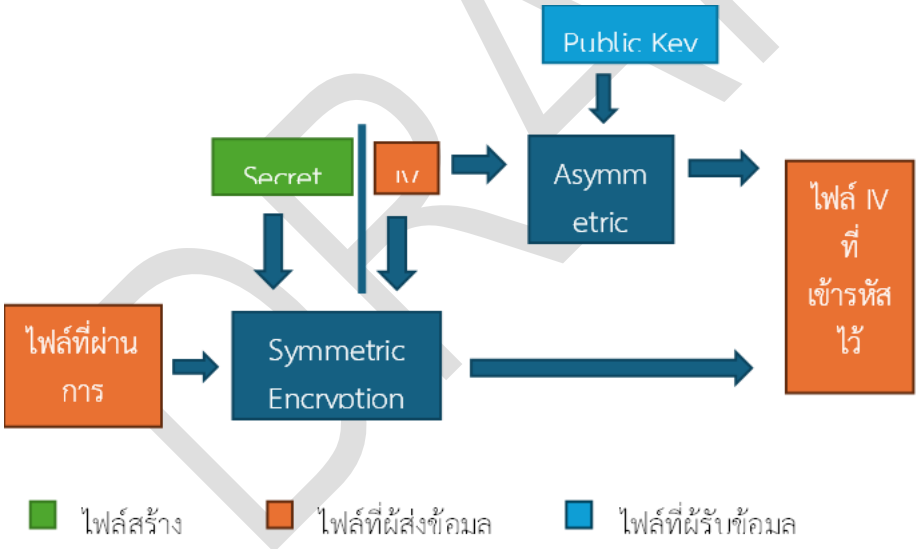
หัวข้อ	Symmetric Encryption	Asymmetric Encryption
ความปลอดภัย	ใช้กุญแจเข้ารหัสดอกเดียวกัน จึงทำให้ความปลอดภัยน้อยกว่า	แยกใช้กุญแจทำให้ปลอดภัยมากขึ้น
จำนวนกุญแจ	1 ดอก	1 คู่ แบ่งเป็น Public key และ Private key
ความเร็ว	เร็วกว่า	ช้ากว่า
ตัวอย่างอัลกอริทึม	AES, DES, 3DES	RSA, Diffie-Hellman

[1] https://10015.io/tools/sha512-encrypt-decrypt

การเข้ารหัสแบบสมมาตรถูกนำมาใช้ในการเข้ารหัสไฟล์ข้อมูลเพราะมีปริมาณที่มาก และใช้กุญแจแบบไม่สมมาตร public key ในการเข้ารหัสไฟล์ที่ใช้ประกอบการถอดรหัสก่อนหน้านี้เพราะมีขนาดเล็กและเพื่อให้มั่นใจว่าจะมีคู่กุญแจ private key เท่านั้นที่จะสามารถถอดออกมาได้

ขั้นตอนการเข้ารหัส

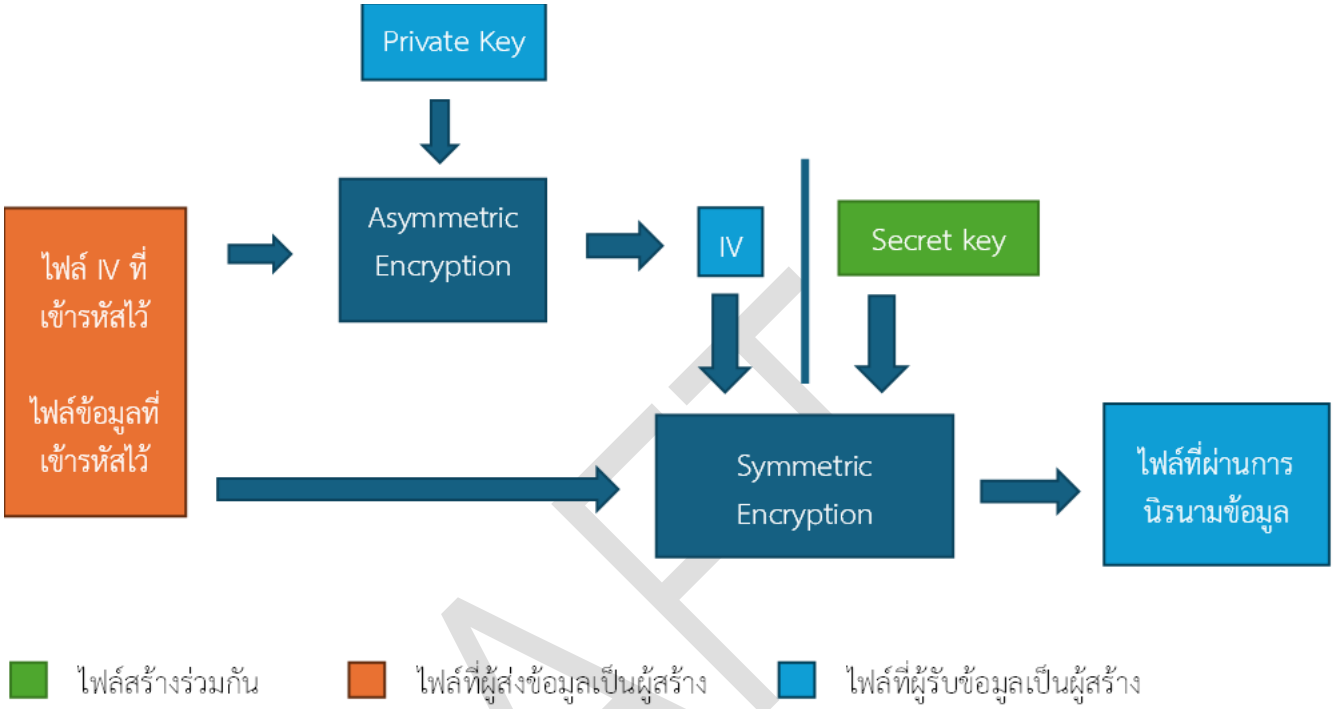
1. ผู้รับข้อมูลสร้างคู่กุญแจด้วยอัลกอริทึมการเข้ารหัสแบบไม่สมมาตร แล้วนำส่ง Public key ให้ผู้ที่จะนำส่งข้อมูล
2. ผู้นำส่งข้อมูลหรือผู้รับข้อมูลสร้างกุญแจ (Secret key) ที่สอดคล้องกับอัลกอริทึมการเข้ารหัสแบบสมมาตรแล้วส่งให้อีกฝ่าย
3. ผู้นำส่งข้อมูลทำการสร้างค่าประกอบการเข้ารหัส (IV: initialization vector) ที่สอดคล้องกับอัลกอริทึมการเข้ารหัสแบบสมมาตร
4. ผู้นำส่งข้อมูลนิรนามฟิลด์ข้อมูลแล้วทำการเข้ารหัสด้วยกุญแจที่ได้จากขั้นตอนที่ 2 และ ค่าประกอบการเข้ารหัสจากขั้นตอนที่ 3
5. ผู้นำส่งข้อมูลเข้ารหัสค่าประกอบการเข้ารหัสจากขั้นตอนที่ 3 ด้วย Public key จากขั้นตอนที่ 1
6. ผู้นำส่งข้อมูลนำส่งข้อมูลที่ได้จากขั้นตอนที่ 4 และ 5 ที่ทำการเข้ารหัสไว้แล้วไปยังผู้รับผ่านช่องทางที่มีการเข้ารหัสระหว่างการส่งข้อมูล



รูปที่ 34: แผนภาพประกอบการเข้ารหัส

ขั้นตอนการถอดรหัส

1. ถอดรหัสค่าประกอบการเข้ารหัส (IV: initialization vector) ด้วย Private key
2. นำค่าประกอบการเข้ารหัสไปถอดรหัสไฟล์ข้อมูลที่เข้ารหัสไว้ร่วมกับ Secret key



รูปที่ 35: แผนภาพประกอบการถอดรหัส

บทสรุป

การนิรนามข้อมูลมีบทบาทสำคัญในการเชื่อมโยงข้อมูลจากหลายหน่วยงานเข้าด้วยกัน ทั้งเพื่อปกป้องข้อมูลส่วนบุคคลทำให้สามารถเชื่อมโยงข้อมูลโดยไม่ต้องกังวลเรื่องความเป็นส่วนตัว เพื่อเพิ่มประสิทธิภาพในการวิเคราะห์ข้อมูลที่ครอบคลุมมากยิ่งขึ้นและแม่นยำขึ้น รวมถึงเพื่อส่งเสริมการมีส่วนร่วมในการแบ่งปันข้อมูล การเข้าถึงข้อมูล และการมีส่วนร่วมในการวิเคราะห์ พัฒนา และตัดสินใจเกี่ยวกับการท่องเที่ยว อย่างไรก็ตาม การนิรนามข้อมูลในการเชื่อมโยงข้อมูลด้านการท่องเที่ยวอย่างมีประสิทธิภาพ โปร่งใส และปลอดภัยนั้น หน่วยงานจำเป็นต้องมีมาตรการป้องกันและแนวทางปฏิบัติที่เหมาะสม เพื่อให้ข้อมูลด้านการท่องเที่ยวจากการเชื่อมโยงข้อมูลจากหน่วยงานและองค์กรต่าง ๆ นั้นสามารถถูกนำไปใช้ เพื่อพัฒนาและส่งเสริมการท่องเที่ยวได้อย่างยั่งยืน

4.3. กรณีศึกษาการจัดทำข้อมูลนิรนามของธนาคารแห่งประเทศไทย

4.3.1. ลักษณะข้อมูลของธนาคารแห่งประเทศไทยที่มีการจัดทำข้อมูลนิรนามเพื่อการใช้งาน

การปฏิบัติงานของธนาคารแห่งประเทศไทย (ธปท.) มีการจัดเก็บข้อมูลจากแหล่งต่าง ๆ ในหลายรูปแบบ ทั้งรูปแบบที่เป็นข้อมูลเชิงสถิติ ข้อมูลรายบุคคลที่ไม่เปิดเผยตัวตนเจ้าของข้อมูล และข้อมูลรายบุคคลที่สามารถระบุตัวตนบุคคลเจ้าของข้อมูลได้ (ทั้งบุคคลธรรมดาและนิติบุคคล) โดยข้อมูลรายบุคคลที่สามารถระบุตัวตนได้ที่ ธปท. ใช้งาน สามารถแบ่งตามแหล่งข้อมูลได้ 3 กลุ่ม ดังนี้

(1) ข้อมูลที่ ธปท. ได้รับจากหน่วยงานภายใต้การกำกับดูแล โดยอาศัยอำนาจตามกฎหมายหรืออาศัยข้อตกลงความร่วมมือระหว่างกันให้ผู้ประกอบธุรกิจจัดส่งข้อมูลดังกล่าวให้ ธปท. เช่น ข้อมูลเงินให้สินเชื่อจากผู้ประกอบธุรกิจภายใต้กฎหมายธุรกิจสถาบันการเงิน ข้อมูลการแลกเปลี่ยนเงินตราต่างประเทศจากผู้ประกอบธุรกิจภายใต้กฎหมายควบคุมการแลกเปลี่ยนเงิน ข้อมูลธุรกรรมการโอนเงินจากผู้ประกอบธุรกิจภายใต้กฎหมายระบบการชำระเงิน

(2) ข้อมูลที่ได้จากการสำรวจ เช่น ข้อมูลหนี้ต่างประเทศของภาคเอกชน ข้อมูลฐานะการลงทุนระหว่างประเทศ

(3) ข้อมูลที่ได้จากหน่วยงานอื่นตามข้อตกลงความร่วมมือการแลกเปลี่ยนข้อมูล เช่น ข้อมูลงบการเงินของนิติบุคคล ข้อมูลการออกตราสารหนี้หรือตราสารทุนในตลาดหลักทรัพย์ เป็นต้น

4.3.2. หลักการและวิธีปฏิบัติในการใช้งานข้อมูลรายบุคคลใน ธปท.

ธปท. มีมาตรการในการรักษาความปลอดภัยของข้อมูลรายบุคคลที่สามารถระบุตัวตนเจ้าของข้อมูล โดยกำหนดเป็นแนวปฏิบัติการกำกับดูแลข้อมูลว่า การใช้งานข้อมูลสำหรับวัตถุประสงค์ที่ไม่จำเป็นต้องทราบตัวตนเจ้าของข้อมูล จะต้องใช้ข้อมูลแบบปกปิดตัวตน หรือแบบข้อมูลรวม (aggregate data) ที่ไม่มีรายละเอียดที่สามารถนำไปสู่หรือชี้แนะให้รู้ตัวตนที่แท้จริงของเจ้าของข้อมูลได้ไม่ว่าทางตรงหรือทางอ้อม

นอกจากนี้ยังมีข้อกำหนดให้ผู้ที่ได้รับอนุญาตให้ใช้ข้อมูลที่ผ่านการจัดทำเป็นข้อมูลนิรนามหรือปกปิดตัวตนแล้ว จะต้องไม่ดำเนินการหรือพยายามดำเนินการเพื่อคาดเดาหรือระบุตัวบุคคลในข้อมูลนั้น รวมถึงไม่ดำเนินการหรือพยายามดำเนินการเพื่อให้เกิดการจับคู่หรือเชื่อมโยงกับข้อมูลแวดล้อมอื่น หรือนำไปเปรียบเทียบกับข้อมูลอื่น ๆ เพื่อนำไปสู่การคาดเดาหรือระบุตัวบุคคลที่เป็นเจ้าของข้อมูลซึ่งได้ถูกปกปิดไว้

โดยหลักการแล้ว ฝ่ายงานที่มีหน้าที่กำกับตรวจสอบผู้ประกอบธุรกิจแต่ละประเภทตามที่กฎหมายบัญญัติไว้จะสามารถใช้งานข้อมูลที่เห็นตัวตนบุคคลได้ เพื่อประโยชน์ในการตรวจสอบการปฏิบัติตามกฎหมายและสามารถดำเนินการทางกฎหมายได้อย่างถูกต้องเมื่อพบการปฏิบัติที่ไม่เป็นไปตามกฎหมาย โดยต้องคำนึงถึงการคุ้มครองข้อมูลให้มีความปลอดภัยอย่างเข้มงวดตามหลักธรรมาภิบาลข้อมูล ตัวอย่างเช่น พนักงานในฝ่ายตรวจสอบของสายกำกับสถาบันการเงิน จะสามารถใช้งานข้อมูลเงินให้สินเชื่อรายสัญญาที่ ธปท. ได้รับจากธนาคารพาณิชย์โดยเห็นทั้งชื่อธนาคารพาณิชย์ผู้ให้สินเชื่อ และชื่อลูกหนี้ที่กู้เงินไปใช้ทำธุรกิจได้ อนึ่ง ธปท. มีนโยบายไม่ให้เปิดเผยตัวตนลูกหนี้บุคคลธรรมดาที่กู้เงินไปใช้ในการอุปโภคบริโภค เช่น ซื้อบ้าน ซื้อรถ บัตรเครดิต ฯลฯ ให้พนักงานใน ธปท. ใช้ในการปฏิบัติงาน เพื่อลดความเสี่ยงจากการละเมิดข้อมูลส่วนบุคคลตามกฎหมายคุ้มครองข้อมูลส่วนบุคคล

ในทางกลับกัน การใช้งานข้อมูลที่กล่าวข้างต้นโดยฝ่ายงานอื่นที่ไม่ใช่ฝ่ายตรวจสอบ เช่น ฝ่ายงานที่ใช้ข้อมูลการปล่อยสินเชื่อของธนาคารพาณิชย์เพื่อประเมินภาวะเศรษฐกิจหรือศึกษาพฤติกรรมของหน่วยเศรษฐกิจในมิติต่าง ๆ จะต้องใช้ข้อมูลดังกล่าวแบบข้อมูลนิรนามทั้งหมด กล่าวคือเป็นข้อมูลที่ถูกปกปิดตัวตนทั้งชื่อธนาคารพาณิชย์และชื่อลูกหนี้ผู้กู้เงิน ในขณะที่ฝ่ายงานที่มีหน้าที่กำหนดนโยบายในการกำกับดูแลสถาบันการเงินจะสามารถใช้ข้อมูลดังกล่าวแบบเห็นชื่อธนาคารพาณิชย์ได้ เพื่อให้สามารถคาดการณ์หรือประเมินผลกระทบต่อธนาคารพาณิชย์แต่ละแห่งได้อย่างถูกต้องในการพิจารณาทางเลือกนโยบายที่เหมาะสมที่สุด แต่ชื่อลูกหนี้จะเป็นถูกปกปิดเป็นข้อมูลนิรนาม

ในทางปฏิบัติ เมื่อฝ่ายงานใดต้องการขอใช้ข้อมูลรายบุคคลของสายงานหนึ่ง ๆ ซึ่งมีอำนาจตามกฎหมายระบุให้จัดเก็บข้อมูลนั้นได้ “ผู้ควบคุมข้อมูล” ซึ่งเป็นหัวหน้าสายงานที่เป็นเจ้าของข้อมูล (ผู้บริหารระดับผู้ช่วยผู้ว่าการ) คือผู้มีอำนาจพิจารณาว่าจะอนุญาตให้ผู้ขอใช้ข้อมูลได้ใช้ข้อมูลรายบุคคลแบบเห็นตัวตนหรือแบบข้อมูลนิรนาม โดยดูวัตถุประสงค์ความจำเป็นในการขอใช้ข้อมูลประกอบการพิจารณาอนุญาต ในกรณีที่อนุญาตให้ใช้ข้อมูลแบบนิรนาม จะพิจารณากำหนดฟิลด์ข้อมูลที่อนุญาตให้ใช้งานอย่างเข้มงวดด้วย เพื่อให้แน่ใจว่า ผู้ใช้ข้อมูลจะไม่เห็นฟิลด์ข้อมูลที่อาจนำไปใช้คาดเดาตัวตนบุคคลได้ (อาจเทียบเคียงได้กับการทำข้อมูลนิรนามด้วยวิธีลบคุณลักษณะข้อมูล (Suppression))

4.3.3. รูปแบบการทำข้อมูลนิรนาม

วิธีปฏิบัติในการทำข้อมูลนิรนามที่ ธปท. ใช้อยู่ในปัจจุบัน นิยมทำใน 2 รูปแบบ ได้แก่

(1) **การเข้าฟังก์ชันแฮช (Hashing)** ตามมาตรฐานสากล SHA256 โดยแปลงข้อมูลให้อยู่ในอักขระรูปแบบอื่นซึ่งจะไม่สามารถแปลงกลับเป็นข้อมูลเดิมได้ (One-way function) โดย ธปท. จะใช้ค่า salt ร่วมกับ key ในการเข้าฟังก์ชันแฮช เพื่อลดความเสี่ยงในการคาดเดาเพื่อระบุตัวตนข้อมูล ทั้งนี้ การใช้ฟังก์ชันแฮชจะต่างจากวิธีการเข้ารหัส (Encryption) ที่สามารถถอดรหัสกลับไปข้อมูลเดิมได้ ทำให้การใช้ฟังก์ชันแฮชมีความปลอดภัยสูงกว่าในการจัดทำข้อมูลนิรนาม

ข้อมูลรายบุคคลชุดใดที่ผู้ควบคุมข้อมูลอนุญาตให้ฝ่ายงานต่าง ๆ มีสิทธิใช้ข้อมูลแบบนิรนามด้วย ข้อมูลจะถูกเข้าฟังก์ชันแฮชเมื่อ ธปท. ได้รับข้อมูลจากหน่วยงานภายนอก และเก็บเป็นฐานข้อมูลแยกต่างหากจากฐานข้อมูลชุดที่เห็นตัวตนบุคคล สำหรับในกรณีที่ผู้ใช้ข้อมูลต้องการใช้ข้อมูลรายบุคคลแบบเชื่อมโยงกันหลายฐานข้อมูล เช่น ต้องการดูข้อมูลสถานะและปริมาณการขอสินเชื่อของลูกหนี้บริษัทในภาคอุตสาหกรรมหนึ่ง เชื่อมโยงกับฐานะการเงินของบริษัทตามงบการเงินของกรมพัฒนาธุรกิจการค้า พฤติกรรมการแลกเปลี่ยนเงินตราต่างประเทศ และการออกตราสารหนี้/ตราสารทุน ข้อมูลต่าง ๆ ของบุคคลเดียวกันจะถูกนำมาเชื่อมโยงกันก่อน แล้วจึงดำเนินการ hashing เพื่อปกปิดตัวตนบุคคล ก่อนส่งให้แก่ผู้ใช้ข้อมูลต่อไป

(2) **การรวมข้อมูล (Data Aggregation)** โดยแสดงข้อมูลด้วยค่าผลรวมตัวเลขแยกตามมิติต่าง ๆ วิธีนี้มักจะใช้ในการแบ่งปันข้อมูลให้ฝ่ายงานที่ไม่จำเป็นต้องเห็นรายละเอียดข้อมูลในระดับรายบุคคล และใช้ในการแบ่งปันข้อมูลกลับให้แก่ผู้ที่จัดส่งข้อมูลให้ ธปท. หรือทำข้อมูลสถิติเผยแพร่ต่อสาธารณชนบนเว็บไซต์ ตัวอย่างเช่น ธนาคารพาณิชย์แต่ละแห่งส่งข้อมูลสินเชื่อรายสัญญาให้ ธปท. ธปท. จะรวมข้อมูลและส่งกลับให้ธนาคารพาณิชย์ทุกแห่งได้เห็นข้อมูลเชิงสถิติภาพรวมของทั้งระบบธนาคารพาณิชย์ เพื่อเป็นการคืน

ประโยชน์กลับแก่ผู้ให้ข้อมูล (data give back) รวมทั้งเผยแพร่ข้อมูลเชิงสถิติภาพรวมของระบบธนาคารพาณิชย์ให้ประชาชนใช้งานได้ทางเว็บไซต์ของ ธปท. โดยอาจมีรายละเอียดน้อยกว่าข้อมูลที่คืนประโยชน์กลับแก่ผู้ให้ข้อมูล

ทั้งนี้ การรวมข้อมูลจะมีการพิจารณาปัจจัย K-anonymity และ L-diversity[21/] ด้วยเพื่อลดความเสี่ยงที่ผู้รับข้อมูลจะคาดเดาและระบุตัวตนบุคคลจากข้อมูลรวมได้ กรณีที่ผลรวมจำนวนข้อมูลในมิติใดมีค่าไม่ผ่านเกณฑ์ K-anonymity และ L-diversity ที่กำหนด ธปท. จะเพิ่มข้อมูลรบกวน (noise) หรือยุบรายละเอียดของมิติ (attribute) นั้นให้มีความละเอียดลดลง เช่น หากการแสดงข้อมูลรวมแยกเป็นรายจังหวัดไม่ผ่านเกณฑ์ดังกล่าว จะยุบความละเอียดลงเป็นการแสดงข้อมูลรายภูมิภาคแทน (เหนือ ตะวันออกเฉียงเหนือ กลาง ใต้) เป็นต้น

5. บรรณานุกรม

Article 29 Data Protection Working Party (European Commission). (2014). *Opinion 05/2014 on Anonymisation Techniques.*

Garfinkel, S. L. (October 2015). *De-Identification of Personal Information.* NISTIR 8053.

PDPA Thailand. (2023). *สรุปเหตุการณ์ "ข้อมูลรั่วไหล" 2561-2566*. Retrieved from https://pdpathailand.com/.

Satori Cyber Ltd. (2021). *Data Anonymization: Use Cases and 6 Common Techniques*. Retrieved from https://satoricyber.com/data-masking/data-anonymization-use-cases-and-6-common-techniques/.

Singapore, P. D. (2022). *Guide To Basic Anonymisation.*

จุฬาลงกรณ์มหาวิทยาลัย, ศ. ค. (2023). *แนวปฏิบัติเกี่ยวกับการคุ้มครองข้อมูลส่วนบุคคล เวอร์ชัน 3.0.*

สถาบันข้อมูลขนาดใหญ่. (2021). *การจัดทำข้อมูลนิรนาม (Data Anonymization)*. Retrieved from https://bdi.or.th/big-data-101/data-anonymization/.

สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ, ส. (2023). *แนวทางสำหรับการจัดทำข้อมูลนิรนามขั้นพื้นฐาน.*

[2] การทำ K-anonymity คือการทำให้มั่นใจว่าข้อมูลในมิตินั้นจะมีจำนวนไม่ต่ำกว่า K เช่น หาก K < 3 ผู้ใช้ข้อมูลอาจคาดเดาได้ง่ายว่าข้อมูลของบุคคลใดบ้างที่ถูกนำมาแสดงอยู่ในมิตินั้น การปรับข้อมูลเพื่อการันตี K-anonymity สามารถทำได้ผ่านการปรับข้อมูลที่ละเอียดเกินไปให้มีสเกลที่หยาบขึ้น เช่น ให้แสดงค่าจังหวัดแทนค่าตำบล

การทำ L-diversity เป็นส่วนขยายของการทำ K-anonymity โดยการันตีว่าในข้อมูลจำนวน K นั้น จะมี L ค่าที่ต่างกันในแต่ละตัวแปร เพื่อป้องกันไม่ให้กลุ่มหนึ่งกลุ่มใดที่ได้จากการทำ K-anonymity มีแต่ค่าข้อมูลอ่อนไหวค่าหนึ่งไปกองรวมกันในกลุ่มเดียว อันอาจทำให้มีการเลือกปฏิบัติกับคนกลุ่มนั้นทั้งกลุ่ม หรือคาดเดาตัวตนบุคคลนั้นได้