



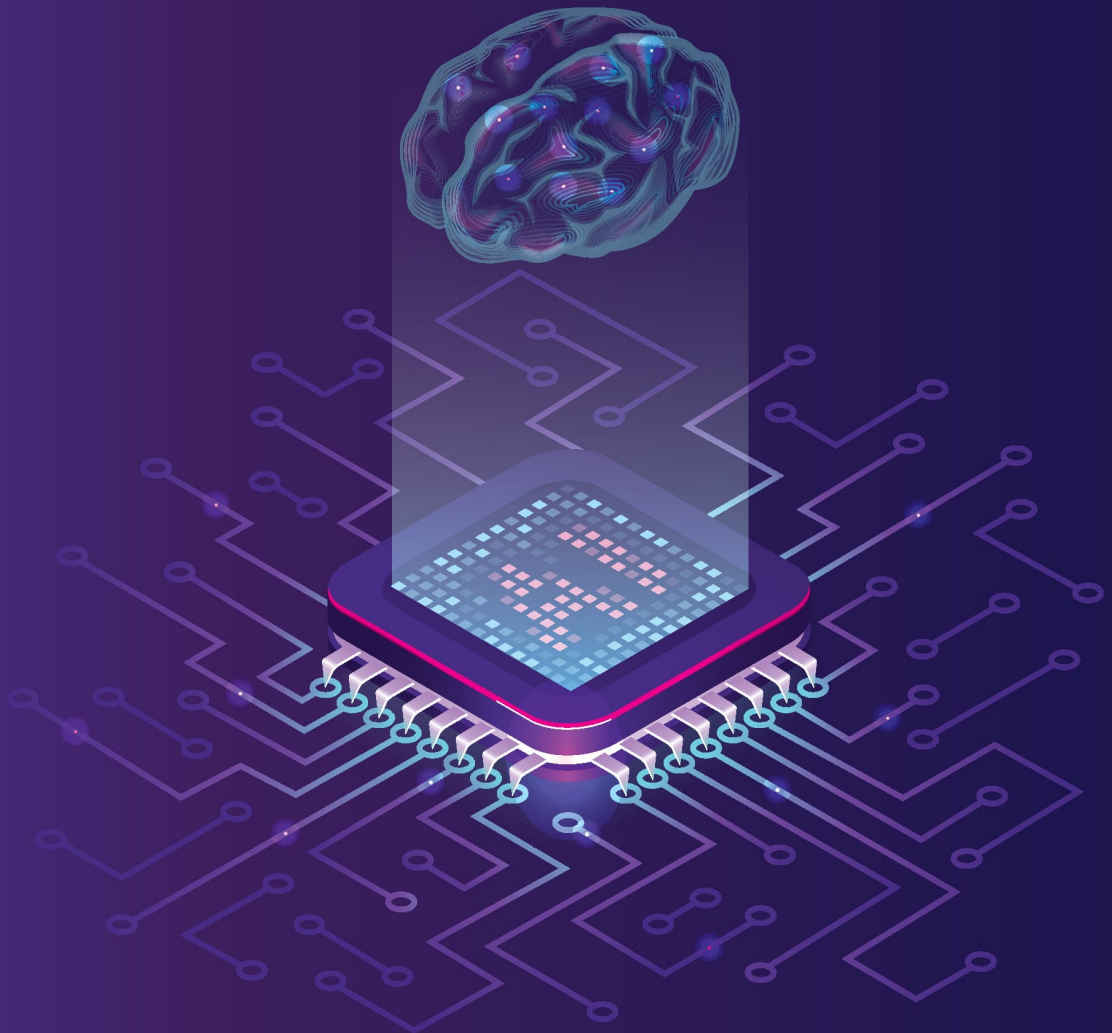
Artificial Intelligence Governance Clinic (AIGC)
Electronic Transactions Development Agency

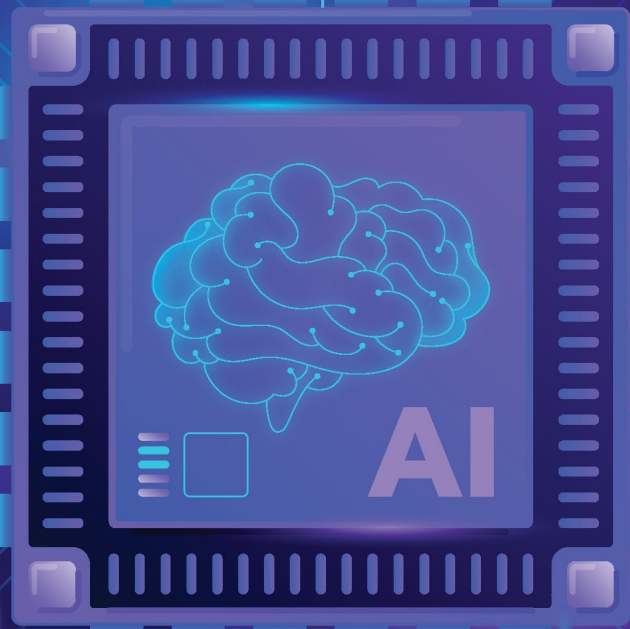


AI

GOVERNANCE GUIDELINE FOR EXECUTIVES

แนวทางการประยุกต์ใช้ปัญญาประดิษฐ์อย่างมีธรรมาภิบาล
สำหรับผู้บริหารองค์กร





สารบัญ

Table of Contents

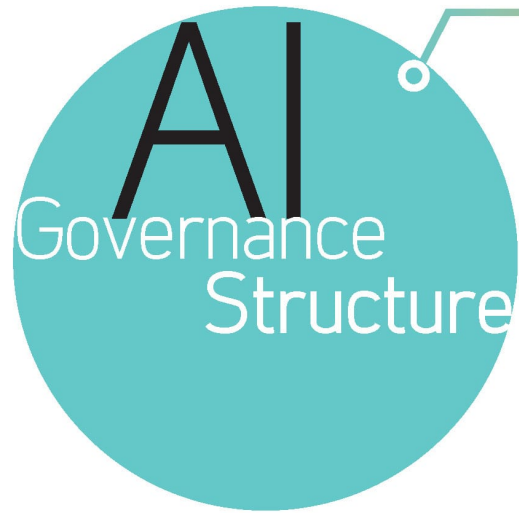
1.	สารบัญ	1
2.	บทสรุปผู้บริหาร	2
3.	ที่มา	4
4.	ขอบเขตและข้อจำกัด	5
5.	บทนำ	6
5.1.	ทำไมจึงต้องมีธรรมาภิบาลในการประยุกต์ใช้ AI	6
5.2.	ธรรมาภิบาลในการประยุกต์ใช้ AI คืออะไร	7
5.3.	หลักการจริยธรรมปัญญาประดิษฐ์	7
5.4.	กรอบการทำงานเพื่อสนับสนุนธรรมาภิบาลในการประยุกต์ใช้ AI	10
5.5.	คำนิยาม	11
6.	การกำหนดโครงสร้างการกำกับดูแล (AI Governance Structure)	12
6.1.	คณะกรรมการกำกับดูแลการประยุกต์ใช้ AI (AI Governance Council)	12
6.2.	หน้าที่และความรับผิดชอบ (Role and Responsibility)	18
6.3.	การพัฒนาศักยภาพบุคลากร (Competency Building)	20
7.	การกำหนดกลยุทธ์ในการประยุกต์ใช้ AI (AI Strategy)	21
7.1.	การกำหนดกลยุทธ์ในการประยุกต์ใช้ AI อย่างมีความรับผิดชอบ (Responsible AI Strategy)	21
7.2.	การบริหารจัดการความเสี่ยงจากการประยุกต์ใช้ AI (AI Risk Management)	24
8.	การกำกับดูแลการปฏิบัติงานที่เกี่ยวข้องกับ AI (AI Operation)	28
8.1.	การกำกับดูแลตลอดวงจรชีวิตของ AI (AI Lifecycle)	28
8.2.	การให้บริการ AI (AI Service)	42
9.	ภาคผนวก ก - ตารางแสดงความสัมพันธ์ระหว่างหลักการจริยธรรมปัญญาประดิษฐ์ และองค์ประกอบหลักในการสนับสนุนธรรมาภิบาลในการประยุกต์ใช้ AI	43
10.	ภาคผนวก ข - ตารางสรุปแนวทางในการพัฒนาโมเดลและนำโมเดลมาปรับใช้	44
11.	เอกสารอ้างอิง	46

2. บทสรุปผู้บริหาร

การประยุกต์ใช้งานปัญญาประดิษฐ์ (Artificial Intelligence หรือ AI) เข้ามามีบทบาทในชีวิตประจำวันของมนุษย์มากขึ้น จึงทำให้ทั่วโลกเกิดความกังวลถึงผลกระทบที่อาจเกิดขึ้นกับมนุษย์ในด้านต่าง ๆ อาทิ ด้านจริยธรรม ด้านความเป็นธรรม ด้านความเป็นส่วนตัว และด้านความปลอดภัย รวมถึงความรับผิดชอบจากผลกระทบอันเกิดจากการนำ AI มาช่วยเหลือหรือตัดสินใจแทนมนุษย์ เป็นต้น

ด้วยเหตุนี้ ธรรมชาติของการประยุกต์ใช้ปัญญาประดิษฐ์ (AI Governance) หรือหลักการในการกำกับดูแลการประยุกต์ใช้ AI ที่ดีนั้น จึงเป็นเรื่องสำคัญที่องค์กรจำเป็นต้องดำเนินการตามบริบทขององค์กรอย่างเหมาะสม เพื่อให้การประยุกต์ใช้ AI สามารถบรรลุตามเป้าหมายที่องค์กรกำหนดและมีความรับผิดชอบต่อบุคคลที่เกี่ยวข้อง องค์กร และสังคมโดยกว้าง (Responsible Use of AI) โดยคำนึงถึงหลักการจริยธรรมปัญญาประดิษฐ์ (AI Ethics Principles) ความสอดคล้องตามกฎหมายและข้อกำหนดที่เกี่ยวข้อง (Laws and Regulations) รวมถึงการควบคุมความเสี่ยงและผลกระทบที่อาจเกิดขึ้นจากการประยุกต์ใช้ AI

โดยเอกสาร "แนวทางการประยุกต์ใช้ปัญญาประดิษฐ์อย่างมีธรรมาภิบาล สำหรับผู้บริหารองค์กร (AI Governance Guideline for Executives)" นี้ มีวัตถุประสงค์เพื่อนำเสนอแนวปฏิบัติสำหรับการกำกับดูแลการประยุกต์ใช้ AI ในมุมมองของผู้บริหารองค์กร พร้อมทั้งนำเสนอตัวอย่างแนวปฏิบัติทั้งจากภายในและต่างประเทศ รวมถึงแนวปฏิบัติตามมาตรฐานสากล เพื่อเป็นแนวทางให้การประยุกต์ใช้ AI ในองค์กรบรรลุตามเป้าหมายที่กำหนดอย่างมีความรับผิดชอบ ซึ่งจะนำไปสู่ความเชื่อมั่นและการยอมรับจากผู้ที่เกี่ยวข้องในทุกภาคส่วน โดยเอกสารนี้ได้นำเสนอแนวทางในการกำกับดูแลซึ่งประกอบด้วย 3 องค์ประกอบหลัก ได้แก่ การกำหนดโครงสร้างการกำกับดูแล (AI Governance Structure) การกำหนดกลยุทธ์ในการประยุกต์ใช้ AI (AI Strategy) และการกำกับดูแลการปฏิบัติงาน (AI Operation)



1. การกำหนดโครงสร้างการกำกับดูแล (AI Governance Structure)

เป็นการนำเสนอแนวทางการกำหนดโครงสร้างการกำกับดูแลการประยุกต์ใช้ AI ขององค์กร โดยมีคณะกรรมการกำกับดูแลการประยุกต์ใช้ AI (AI Governance Council) ซึ่งมีหน้าที่ในการกำหนดกลยุทธ์และนโยบายที่เกี่ยวข้องเพื่อขับเคลื่อนการประยุกต์ใช้ AI ภายในองค์กร รวมถึงมีหน้าที่ในการเฝ้าติดตาม และประเมินผลการประยุกต์ใช้ AI อย่างต่อเนื่อง เพื่อให้มั่นใจได้ว่าองค์กรมีการกำกับดูแลการประยุกต์ใช้ AI อย่างเหมาะสม

นอกจากการกำหนดคณะกรรมการเพื่อกำกับดูแลโดยภาพรวมแล้ว องค์กรยังจำเป็นต้องกำหนดหน้าที่และความรับผิดชอบ (Role and Responsibility) ของบุคลากรและผู้มีส่วนได้เสียที่เกี่ยวข้อง (Stakeholders) ตามโครงสร้างการกำกับดูแล พร้อมทั้งสนับสนุนทรัพยากรที่จำเป็น ส่งเสริมทักษะความรู้ รวมถึงสร้างความตระหนักรู้ (Awareness) ต่อความรับผิดชอบ (Responsibility) และความรับผิดชอบต่อผลของการกระทำ (Accountability) ของแต่ละหน้าที่อีกด้วย



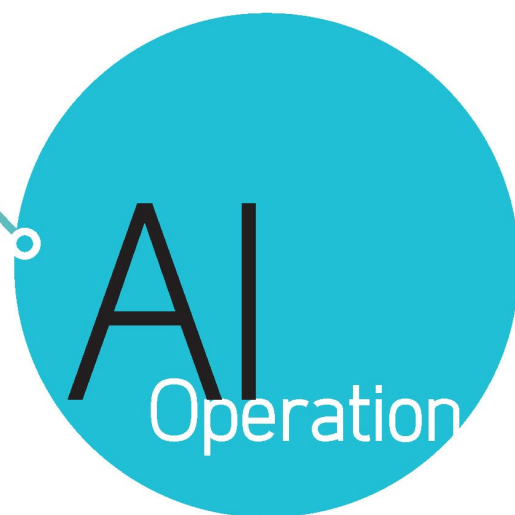
2. การกำหนดกลยุทธ์ในการประยุกต์ใช้ AI (AI Strategy)

เป็นการมองหาโอกาสหรือประโยชน์ที่องค์กรจะได้รับจากการนำ AI มาประยุกต์ใช้ที่ตอบสนองเป้าหมายขององค์กร พร้อมทั้งวิเคราะห์เป้าหมายความสำเร็จจากการนำ AI มาประยุกต์ใช้ ความเป็นไปได้ในการดำเนินการ ความพร้อมขององค์กร และทรัพยากรที่จำเป็น เพื่อให้มั่นใจได้ว่าการประยุกต์ใช้ AI จะบรรลุตามเป้าหมายและประสบความสำเร็จตามที่กำหนด

นอกจากนี้ ในการกำหนดกลยุทธ์ยังจำเป็นต้องวิเคราะห์ถึงความเสี่ยงและผลกระทบจากการนำ AI มาประยุกต์ใช้ ควบคู่ไปกับความสอดคล้องกับหลักการจริยธรรมปัญญาประดิษฐ์ และความสอดคล้องตามกฎหมายและข้อกำหนดที่เกี่ยวข้อง เพื่อแสดงให้เห็นถึงความมุ่งมั่นขององค์กรที่จะประยุกต์ใช้ AI อย่างมีความรับผิดชอบต่อบุคคลที่เกี่ยวข้อง องค์กร และสังคม โดยกว้าง (Responsible use of AI)

3. การกำกับดูแลการปฏิบัติงานเกี่ยวกับ AI (AI Operation)

เป็นการอธิบายถึงแนวปฏิบัติตลอดวงจรชีวิตของ AI (AI Lifecycle) เพื่อให้มั่นใจได้ว่า AI ได้รับการออกแบบ พัฒนาและนำไปใช้งานได้อย่างสอดคล้องตามเป้าหมายขององค์กร รวมถึงสอดคล้องตามหลักการจริยธรรมปัญญาประดิษฐ์ กฎหมายและข้อกำหนดที่เกี่ยวข้อง นอกจากนี้ ยังอธิบายถึงแนวปฏิบัติในการจัดเตรียมข้อมูล (Data Preparation) เพื่อให้ได้ข้อมูลที่มีคุณภาพและเหมาะสมสำหรับนำไปใช้สอนและทำงานร่วมกับ AI รวมถึงแนวปฏิบัติการรับมือกับความเสียหายและผลกระทบที่อาจเกิดขึ้น การเฝ้าติดตามและการประเมินผลการประยุกต์ใช้งาน AI เพื่อปรับปรุงการดำเนินงานให้มีประสิทธิภาพต่อไปในอนาคต



จากแนวทางในการกำกับดูแลทั้ง 3 องค์ประกอบหลักข้างต้น เป็นการนำเสนอแนวทางการกำกับดูแลการประยุกต์ใช้เทคโนโลยี AI อย่างมีธรรมาภิบาล โดยมีคณะกรรมการกำกับดูแลการประยุกต์ใช้ AI ซึ่งมีหน้าที่หลักในการกำกับดูแลในภาพรวม ตั้งแต่การกำหนดกลยุทธ์และนโยบาย การควบคุมความเสี่ยงและผลกระทบที่อาจเกิดขึ้น ไปจนถึงการกำกับดูแลการปฏิบัติงานตลอดวงจรชีวิตของ AI พร้อมทั้ง กำหนดหน้าที่และความรับผิดชอบของบุคลากรและผู้มีส่วนได้ส่วนเสียที่เกี่ยวข้องอย่างชัดเจน เพื่อให้มั่นใจว่าการประยุกต์ใช้ AI ขององค์กร จะบรรลุตามเป้าหมายที่กำหนดและมีความรับผิดชอบต่อบุคคลที่เกี่ยวข้อง องค์กร และสังคม โดยกว้าง ซึ่งจะทำให้เกิดความเชื่อมั่นและได้รับการยอมรับจากทุกภาคส่วน

3. ที่มา

ในยุคโลกาภิวัตน์ที่ปัญญาประดิษฐ์ได้เข้ามามีบทบาทสนับสนุนและอำนวยความสะดวกกับมนุษย์มากขึ้น สำนักงานคณะกรรมการดิจิทัลเพื่อเศรษฐกิจและสังคมแห่งชาติ (สดช.) กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม จึงได้มีการจัดทำเอกสารหลักการแนวทางจริยธรรมปัญญาประดิษฐ์ (Thailand AI Ethics Guideline) และนำเสนอต่อคณะรัฐมนตรีให้ความเห็นชอบให้หน่วยงานราชการใช้เป็นแนวทางปฏิบัติในการพัฒนา ส่งเสริม และนำไปใช้ในทางที่ถูกต้องและมีจริยธรรม เมื่อวันที่ 2 กุมภาพันธ์ พ.ศ. 2564 รวมทั้งเพื่อสนับสนุนการดำเนินการตามยุทธศาสตร์ชาติ 20 ปี (พ.ศ. 2561- 2580) และยุทธศาสตร์กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม (พ.ศ. 2563 - 2567) โดยเอกสารดังกล่าวได้มีการกำหนดหลักการจริยธรรมปัญญาประดิษฐ์ (AI Ethics Principles) ไว้ 6 หลักการ (ดังรายละเอียดในหัวข้อ 4.3) ซึ่งเป็นหลักการจริยธรรมพื้นฐานที่สอดคล้องกับหลักการจริยธรรมปัญญาประดิษฐ์ที่ได้รับการยอมรับจากองค์กรที่มีชื่อเสียงด้านเทคโนโลยีในหลายประเทศ เพื่อเป็นหลักการพื้นฐานที่จะนำไปสู่การพัฒนาและส่งเสริมให้เกิดธรรมาภิบาลในการประยุกต์ใช้ AI ภายในประเทศ

ต่อมา สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์ (สพธอ.) ในฐานะหน่วยงานของรัฐ ซึ่งมีการทํากิจในการส่งเสริมและสนับสนุนการทำธุรกรรมทางอิเล็กทรอนิกส์ให้มีความน่าเชื่อถือและมั่นคงปลอดภัย ได้เล็งเห็นว่าเทคโนโลยี AI มีพัฒนาการแบบก้าวกระโดดและมีผลกระทบต่อตรงกับการทำธุรกรรมอิเล็กทรอนิกส์ในปัจจุบัน จึงควรส่งเสริมและสนับสนุนให้ผู้ประกอบธุรกิจบริการภายใต้กฎหมายธุรกรรมทางอิเล็กทรอนิกส์ทั้งภาครัฐและเอกชน นำหลักการจริยธรรมปัญญาประดิษฐ์ (AI Ethics Principles) ไปปรับใช้เพื่อให้เกิดธรรมาภิบาลในการประยุกต์ใช้ AI

ด้วยเหตุนี้ สพธอ. จึงได้จัดทำเอกสาร "แนวทางการประยุกต์ใช้ปัญญาประดิษฐ์อย่างมีธรรมาภิบาล สำหรับผู้บริหารองค์กร (AI Governance Guideline for Executives)" เพื่อนำเสนอภาพรวมของการกำกับดูแลการประยุกต์ใช้ AI ในมุมมองของผู้บริหารองค์กร พร้อมทั้ง นำเสนอตัวอย่างแนวปฏิบัติทั้งจากภายในและต่างประเทศ รวมถึงแนวปฏิบัติตามมาตรฐานสากล เพื่อเป็นแนวปฏิบัติให้แก่ผู้ประกอบธุรกิจบริการทั้งภาครัฐและเอกชน ในการกำกับดูแลการประยุกต์ใช้ AI เพื่อให้เกิดธรรมาภิบาลในการประยุกต์ใช้ AI ซึ่งจะนำไปสู่ความเชื่อมั่นและการยอมรับจากบุคคลที่เกี่ยวข้องในทุกภาคส่วนต่อไป

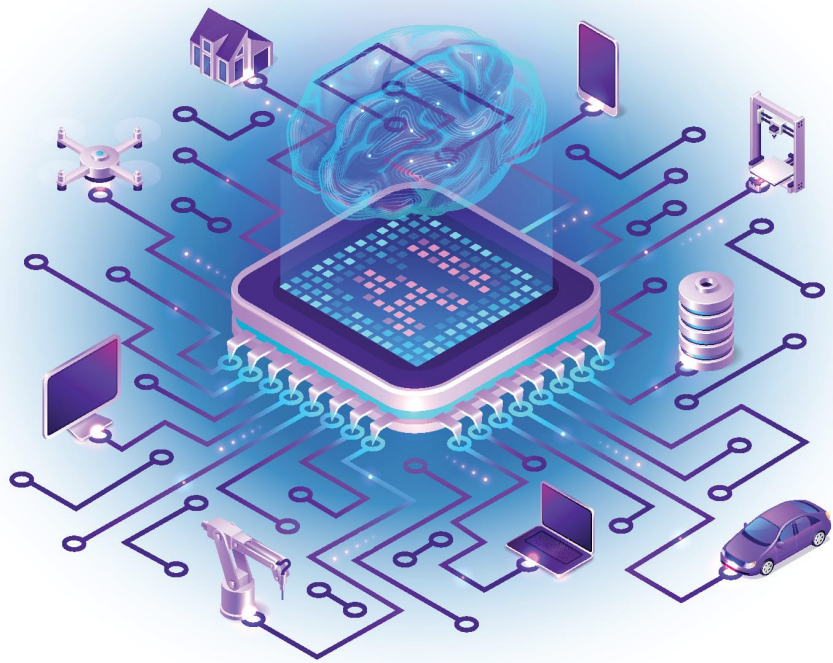


4. ขอบเขตและข้อจำกัด

เอกสารนี้จัดทำโดยมีวัตถุประสงค์เพื่อนำเสนอแนวทางการกำกับดูแลประยุกต์ใช้ AI ในภาพรวม โดยไม่มีการระบุเจาะจงประเภทของเทคโนโลยีและประเภทขององค์กรที่นำแนวปฏิบัติไปปรับใช้ ด้วยเหตุนี้ในการนำแนวปฏิบัติหรือข้อเสนอแนะในเอกสารนี้ไปปรับใช้ จึงอาจไม่เหมาะสมกับบางบริบทของการประยุกต์ใช้ AI ดังนั้น องค์กรจึงควรพิจารณาถึงความเหมาะสมกับบริบทของการนำ AI ไปประยุกต์ใช้ เทคโนโลยีที่ใช้และการเปลี่ยนแปลงของเทคโนโลยีในอนาคต ความสอดคล้องตามกฎหมายและข้อกำหนดที่เกี่ยวข้อง และความเสี่ยงที่เกี่ยวข้องกับบริบทของการนำ AI ไปประยุกต์ใช้ เช่น ภัยคุกคามทางไซเบอร์ ความเสี่ยงในด้านอาชีวอนามัยและความปลอดภัย เป็นต้น

นอกจากนี้ แนวปฏิบัติหรือข้อเสนอแนะในเอกสารนี้เป็นเพียงการนำเสนอผลการศึกษาตัวอย่างแนวปฏิบัติทั้งจากภายในและต่างประเทศ รวมถึงแนวปฏิบัติตามมาตรฐานสากล เพื่อส่งเสริมและสนับสนุนให้องค์กรทั้งรัฐและเอกชนสามารถประยุกต์ใช้ AI ได้อย่างมีธรรมาภิบาล ดังนั้น การไม่ปฏิบัติตามแนวปฏิบัติหรือข้อเสนอแนะดังกล่าว จึงไม่เป็นการฝ่าฝืนหรือไม่ปฏิบัติตามกฎหมายและข้อกำหนดที่เกี่ยวข้องกับปัญญาประดิษฐ์ของประเทศไทย

5. บทนำ



5.1. ทำไมจึงต้องมีธรรมาภิบาลในการประยุกต์ใช้ AI

ปัญญาประดิษฐ์ หรือ AI เป็นเทคโนโลยีที่มีความแตกต่างเฉพาะเมื่อเปรียบเทียบกับเทคโนโลยีอื่นในด้านความสามารถในการตัดสินใจได้โดยอัตโนมัติ ความสามารถในการประมวลผลข้อมูลที่มีความซับซ้อนและยากที่มนุษย์จะสามารถทำได้ อีกทั้งยังมีความสามารถในการเรียนรู้และปรับปรุงประสิทธิภาพในการตัดสินใจของตนเองได้ตามสถานการณ์หรือสภาพแวดล้อมที่เปลี่ยนแปลงได้ด้วยการนำข้อมูลจากบริบทแวดล้อมและข้อเสนอแนะจากผู้ใช้งานมาเรียนรู้เพิ่มเติม ซึ่งจากความสามารถดังกล่าวของ AI ทำให้หลายองค์กรตัดสินใจที่จะนำ AI มาช่วยในการตัดสินใจ แก้ปัญหา หรือทำงานแทนมนุษย์ รวมถึงช่วยในการหาโอกาสใหม่ทางธุรกิจด้วยการประมวลผลข้อมูลที่มีความซับซ้อน

แต่จากการศึกษาการประยุกต์ใช้เทคโนโลยี AI ในหลายประเทศที่ผ่านมาพบว่า การทำงานของ AI อาจไม่เป็นไปตามที่ออกแบบไว้ตั้งแต่ต้น อีกทั้งยังอาจก่อให้เกิดความเสียหายโดยไม่คาดคิด ซึ่งในหลายกรณีพบว่า AI มีการตัดสินใจที่ผิดพลาดและมีอคติ ส่งผลกระทบต่อเชิงลบทั้งในระดับบุคคล องค์กร และสังคมโดยกว้าง ตัวอย่างเช่น การเกิดอุบัติเหตุรถชนพร้อมกันหลายคัน เนื่องจากการใช้ระบบขับอัตโนมัติเต็มรูปแบบ (Full Self-Driving Mode) หรือในกรณีที่คนผิวขาวในสหรัฐอเมริกาได้รับ

สิทธิประโยชน์ในการรักษาพยาบาลที่ดีกว่าคนผิวดำ เนื่องจาก AI ถูกสอนด้วยข้อมูลตั้งต้นที่มีข้อมูลของคนผิวขาวเป็นส่วนใหญ่ ทำให้การตัดสินใจของ AI มีความเอนเอียงและมีอคติต่อคนผิวดำ เป็นต้น ด้วยเหตุนี้ การนำเทคโนโลยี AI มาประยุกต์ใช้จึงจำเป็นต้องคำนึงถึงปัจจัยต่าง ๆ ที่อาจก่อให้เกิดทั้งประโยชน์ และผลกระทบเชิงลบที่มนุษย์ไม่เคยคาดคิดมาก่อน

ด้วยเหตุผลดังกล่าวข้างต้น จึงทำให้เกิดการพัฒนาหลักการธรรมาภิบาลในการประยุกต์ใช้ AI หรือหลักการในการกำกับดูแลการประยุกต์ใช้ AI ที่ดี โดยเป้าหมายของหลักการธรรมาภิบาลนั้น ไม่ได้มีวัตถุประสงค์เพียงเพื่อสนับสนุนให้องค์กรประสบความสำเร็จในการประยุกต์ใช้ AI เท่านั้น แต่ยังหมายรวมถึงการประยุกต์ใช้ AI อย่างมีความรับผิดชอบ โดยนำหลักการจริยธรรม ปัญญาประดิษฐ์มาปรับใช้ เช่น การตัดสินใจด้วยความเป็นธรรม ไม่มีอคติ มีความโปร่งใส สามารถตรวจสอบและอธิบายเหตุผลในการตัดสินใจของ AI ได้ เป็นต้น และคำนึงถึงความสอดคล้องตามกฎหมายและข้อกำหนดที่เกี่ยวข้อง พร้อมทั้งมีการควบคุมความเสี่ยงจากการประยุกต์ใช้ AI และผลกระทบเชิงลบที่อาจเกิดขึ้นต่อบุคคลที่เกี่ยวข้อง องค์กร และสังคมโดยกว้าง เพื่อนำไปสู่ความเชื่อมั่นและการยอมรับจากบุคคลที่เกี่ยวข้อง

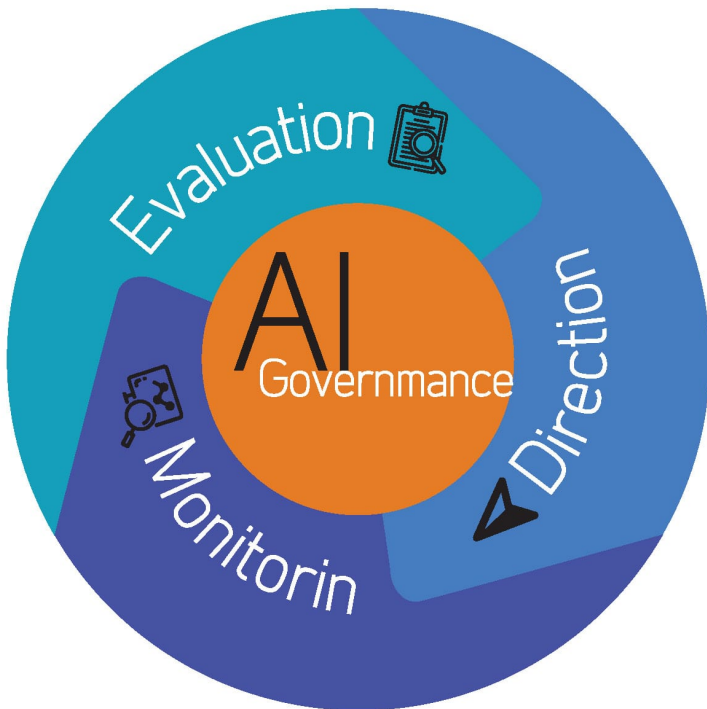
5.2. ธรรมชาติของการประยุกต์ใช้ AI คืออะไร

ธรรมชาติของการประยุกต์ใช้ AI คือหลักการกำกับดูแลการปฏิบัติงานในทุกกระบวนการที่เกี่ยวข้องกับการประยุกต์ใช้ AI โดยจัดให้มีมาตรการในการกำกับดูแลผ่านการกำหนดนโยบาย ขั้นตอนปฏิบัติ และเครื่องมือในการปฏิบัติงาน เพื่อให้มั่นใจได้ว่าการประยุกต์ใช้ AI นั้น สามารถบรรลุตามเป้าหมายขององค์กรอย่างมีความรับผิดชอบ โดยคำนึงถึงความสอดคล้องตามหลักการจริยธรรมปัญญาประดิษฐ์ ความสอดคล้องตามกฎหมายและข้อกำหนดที่เกี่ยวข้อง และมีการควบคุมความเสี่ยงที่อาจส่งผลกระทบต่อบุคคลที่เกี่ยวข้อง องค์กร และสังคมโดยกว้าง โดยมีคณะกรรมการกำกับดูแลการประยุกต์ใช้ AI ซึ่งมีหน้าที่หลัก ดังนี้

1) การกำหนดทิศทางดำเนินการ (Direction) โดยกำหนดกลยุทธ์ในการประยุกต์ใช้ AI และนโยบายที่เกี่ยวข้อง

2) การเฝ้าติดตาม (Monitoring) ประสิทธิภาพ (Performance) ของการประยุกต์ ใช้ AI รวมถึงการปฏิบัติงานตามนโยบายและข้อกำหนดต่าง ๆ (Conformance) ทั้งภายในและภายนอกองค์กร

3) การประเมินผล (Evaluation) การประยุกต์ใช้งาน AI ในปัจจุบันและอนาคต โดยพิจารณาจากปัจจัยที่เกี่ยวข้องทั้งภายในและ/หรือภายนอก เช่น เป้าหมายในการประยุกต์ใช้ AI ภัยคุกคามและโอกาสจากการประยุกต์ใช้ AI ความเปลี่ยนแปลงด้านเทคโนโลยี ประสิทธิภาพและประสิทธิผลในการกำกับดูแล เป็นต้น



5.3. หลักการจริยธรรมปัญญาประดิษฐ์

เพื่อให้การประยุกต์ใช้ AI เป็นไปอย่างมีความรับผิดชอบนั้น จำเป็นต้องมีการนำหลักการจริยธรรมปัญญาประดิษฐ์ที่เกี่ยวข้องมาปรับใช้ โดยองค์กรจะต้องกำหนดกลยุทธ์ในการประยุกต์ใช้ AI (AI Strategy) และนโยบายให้สอดคล้องตามหลักการจริยธรรมที่เกี่ยวข้อง ซึ่งในการพิจารณานำหลักการจริยธรรมปัญญาประดิษฐ์ใดมาปรับใช้นั้นขึ้นอยู่กับบริบทขององค์กรที่นำ AI ไปประยุกต์ใช้

ทั้งนี้ องค์กรอาจพิจารณานำหลักการจริยธรรมปัญญาประดิษฐ์ตามแนวทางของสำนักงานคณะกรรมการดิจิทัลเพื่อเศรษฐกิจและสังคมแห่งชาติ (สดช.) มาปรับใช้ได้ทั้งหมด หรือบางส่วน หรืออาจนำหลักการจริยธรรมปัญญาประดิษฐ์อื่นที่เกี่ยวข้องกับบริบทของการประยุกต์ใช้ AI มาปรับใช้ตามความเหมาะสม

หลักการจริยธรรมปัญญาประดิษฐ์ของไทย ตามแนวทางของ สดช.

ความสามารถ ในการแข่งขัน และการพัฒนาอย่างยั่งยืน (Competitiveness and Sustainable Development)

AI ควรถูกสร้างและใช้งานเพื่อสร้างประโยชน์และคุณค่าให้แก่มนุษย์ สังคม เศรษฐกิจและสิ่งแวดล้อมอย่างยั่งยืน รวมถึงเพิ่มความสามารถในการแข่งขันและสร้างความเจริญให้กับทุกภาคส่วนในโลกอย่างเป็นธรรม นอกจากนี้ AI ยังควรได้รับการวิจัยและพัฒนาอย่างต่อเนื่อง เพื่อสร้างนวัตกรรมและอุตสาหกรรมใหม่

1

ความสอดคล้อง กับกฎหมาย จริยธรรม และมาตรฐานสากล (Laws, Ethics, and International Standards)

AI ควรได้รับการออกแบบ พัฒนา ให้บริการและใช้งาน โดยสอดคล้องกับกฎหมาย บรรทัดฐาน จริยธรรม คุณธรรม ของมนุษย์ และ มาตรฐานสากล โดยเคารพต่อความเป็นส่วนตัว เกียรติ สิทธิเสรีภาพ และสิทธิมนุษยชน นอกจากนี้ AI ยังควรได้รับการออกแบบโดยมีมนุษย์เป็นศูนย์กลางและเป็นผู้ตัดสินใจ (Human-Centered Design) และไม่ควรถูกออกแบบเพื่อใช้ในการกำหนดชะตาชีวิตของมนุษย์

2

ความน่าเชื่อถือ (Reliability)

AI ควรได้รับการสนับสนุนให้มีความน่าเชื่อถือและความมั่นใจในการใช้งานต่อสาธารณะ โดยสนับสนุนให้มีการพัฒนาด้วยความยึดมั่นในความถูกต้อง (Accuracy) ความน่าเชื่อถือ (Reliability) สามารถทนทานต่อเหตุการณ์ที่อาจเกิดความผิดพลาด (Robustness) และสามารถสร้างผลลัพธ์ได้เหมือนเดิม (Reproducibility) นอกจากนี้ ควรมีการควบคุมคุณภาพของข้อมูล (Quality of data) รวมถึงกำหนดกระบวนการและช่องทางรับความคิดเห็น (Feedback) จากผู้ใช้งาน ให้ข้อมูลเพิ่มเติม รับเรื่องร้องเรียน แจ้งปัญหาที่พบ และมีการตอบสนองหรือดำเนินการแก้ไขปัญหาที่พบได้ทันที



ความเท่าเทียม หลากหลาย ครอบคลุม และเป็นธรรม (Fairness)

ในการส่งเสริมการประยุกต์ใช้ AI อย่างเป็นธรรมและลดความมีอคติ (Bias) โดยระบบควรถูกออกแบบและพัฒนาโดยคำนึงความหลากหลาย (Diversity) ลดการเอนเอียง แบ่งแยก และเลือกปฏิบัติ (Discrimination) ต่อบุคคลหรือกลุ่มคนที่มีคุณลักษณะที่ต่างกัน (อายุ เพศ ลักษณะทางกายภาพ เชื้อชาติ) โดยเฉพาะกลุ่มคนผู้ด้อยโอกาสในสังคม รวมถึงสามารถพิสูจน์ถึงความเป็นธรรมสำหรับทุกฝ่ายที่เกี่ยวข้อง



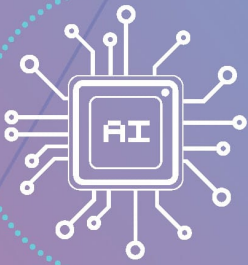
ความมั่นคงปลอดภัย และเป็นส่วนตัว (Security and Privacy)

AI ควรถูกออกแบบและพัฒนาโดยให้ความสำคัญถึงความมั่นคงปลอดภัยและการรักษาความเป็นส่วนตัวโดยจัดให้มีมาตรการป้องกันการโจมตีทางไซเบอร์ และปฏิบัติตามกฎหมายคุ้มครองข้อมูลส่วนบุคคล รวมถึงมีกลไกที่ให้นุชนสามารถแทรกแซงเพื่อควบคุมการทำงานของ AI ในกรณีที่เกิดเหตุการณ์หรือความเสี่ยงที่มีผลกระทบต่อมนุษยชาติ



ความโปร่งใส และความรับผิดชอบ ต่อผลของการกระทำ (Transparency and Accountability)

AI ควรได้รับการออกแบบ พัฒนา ให้บริการ และใช้งานด้วยความโปร่งใส โดยจะต้องสามารถอธิบาย (Explainability) และคาดการณ์ผลลัพธ์จากการทำงานได้ อีกทั้งยังควรมีการเฝ้าติดตามความผิดปกติ มีความสามารถในการสืบย้อนกลับ (Traceability) และสามารถวินิจฉัยปัญหาและความผิดพลาดได้ (Diagnosability) นอกจากนี้ องค์กรยังควรมีการกำหนดหน้าที่และความรับผิดชอบ (Role and Responsibility) รวมถึงความรับผิดชอบต่อผลของการกระทำ (Accountability) ต่อผลกระทบที่เกิดขึ้นจาก AI ตามหน้าที่ที่รับผิดชอบได้



5.4. กรอบการทำงานเพื่อสนับสนุนธรรมาภิบาลในการประยุกต์ใช้ AI

กรอบการทำงานเพื่อสนับสนุนให้เกิดธรรมาภิบาลในการประยุกต์ใช้ AI ประกอบด้วย 3 องค์ประกอบหลัก ได้แก่



1 การกำหนดโครงสร้างการกำกับดูแล (AI Governance Structure)

- จัดตั้งคณะกรรมการกำกับดูแล (AI Governance Council) เพื่อกำหนดทิศทาง การประยุกต์ใช้งาน AI ผ่านการกำหนดกลยุทธ์และนโยบาย รวมถึง ฝ่าฝืนติดตาม และประเมินผล การประยุกต์ใช้ AI อย่างต่อเนื่องเพื่อสนับสนุนให้เกิดธรรมาภิบาล (ตามข้อ 2 และ 3)
- กำหนดหน้าที่ของบุคลากรและผู้มีส่วนได้เสียที่เกี่ยวข้องกับการประยุกต์ใช้ AI พร้อมทั้ง สร้างความตระหนักรู้ในด้านความรับผิดชอบและความรับผิดชอบต่อผลของการกระทำ ของแต่ละหน้าที่
- พัฒนาศักยภาพของบุคลากรและหรือผู้มีส่วนได้ส่วนเสียที่เกี่ยวข้อง เพื่อให้สามารถ ปฏิบัติงานได้อย่างเหมาะสมตามหน้าที่ที่ได้รับมอบหมาย



2 การกำหนดกลยุทธ์ในการประยุกต์ใช้ AI (AI Strategy)

- มองหาโอกาสในการนำ AI มาประยุกต์ใช้ เพื่อสนับสนุนให้บรรลุเป้าหมายขององค์กร หรือเป้าหมายทางธุรกิจ
- กำหนดเป้าหมายในการประยุกต์ใช้ AI ตามลำดับความสำคัญ โดยพิจารณาจาก ประโยชน์ที่จะได้รับ ความพร้อมขององค์กร หลักการจริยธรรมปัญญาประดิษฐ์ กฎหมายและข้อกำหนดที่ต้องดำเนินการให้สอดคล้อง รวมถึงความซับซ้อนและเวลาที่ จำเป็นต้องใช้ในการดำเนินการ
- กำหนดกลยุทธ์ในการบริหารจัดการข้อมูล
- กำหนดแผนปฏิบัติงาน (Roadmap) ในการประยุกต์ใช้ AI
- วิเคราะห์ความเสี่ยงและผลกระทบที่อาจเกิดขึ้นจากการประยุกต์ใช้ AI รวมถึงการกำหนด ระดับการมีส่วนร่วมของมนุษย์ในการทำงานของ AI และมาตรการในการควบคุม ความเสี่ยงที่เหมาะสม เพื่อควบคุมความเสี่ยงให้อยู่ในขอบเขตที่ยอมรับได้



3 การกำกับดูแลการปฏิบัติงานที่เกี่ยวข้องกับ AI (AI Operation)

- จัดทำข้อกำหนดความต้องการในการพัฒนาระบบ AI (AI Requirement) พร้อมทั้ง ออกแบบโซลูชันที่เหมาะสมกับข้อกำหนดดังกล่าว
- จัดเตรียมข้อมูลที่มีคุณภาพสำหรับการสอน ตรวจสอบ และทดสอบ โมเดลปัญญา ประดิษฐ์ และลดความเอนเอียงที่อาจเกิดจากข้อมูล (Data Bias)
- สร้างโมเดลปัญญาประดิษฐ์โดยนำหลักการจริยธรรมปัญญาประดิษฐ์มาปรับใช้ และ ควบคุมความเสี่ยงที่อาจเกิดขึ้น
- ฝ่าฝืนติดตามผลประสิทธิภาพการประยุกต์ใช้ AI รวมถึงการปฏิบัติงานตามนโยบาย หลักการจริยธรรมปัญญาประดิษฐ์ กฎหมายและข้อกำหนดที่เกี่ยวข้อง
- ประเมินผลการประยุกต์ใช้งาน AI ในปัจจุบันและกำหนดแนวทางการดำเนินงานอนาคต



5.5. คำนิยาม

ปัญญาประดิษฐ์ (Artificial Intelligence: AI) - เทคโนโลยีที่ถูกพัฒนาขึ้นเพื่อให้คอมพิวเตอร์มีคุณสมบัติหรือพฤติกรรมใกล้เคียงมนุษย์ เช่น การเรียนรู้ การรับรู้และตอบสนองต่อสภาพแวดล้อม การให้เหตุผล และการแก้ไขปัญหา เป็นต้น ตามวัตถุประสงค์ที่มนุษย์กำหนด

แมชชีนเลิร์นนิง (Machine Learning) - เทคโนโลยี AI ประเภทหนึ่งที่มีความสามารถในการเรียนรู้หรือปรับปรุงประสิทธิภาพการทำงานของตน โดยใช้อัลกอริทึมในการวิเคราะห์และเรียนรู้จากข้อมูลที่ได้รับจากการสอนหรือสภาพแวดล้อม

โมเดลปัญญาประดิษฐ์ (AI Model) - โปรแกรมที่ถูกสร้างขึ้นจากอัลกอริทึมและข้อมูลที่สอนโดยมนุษย์ เพื่อให้คอมพิวเตอร์มีคุณสมบัติหรือพฤติกรรมใกล้เคียงมนุษย์

หมายเหตุ: โมเดลปัญญาประดิษฐ์ประเภทแมชชีนเลิร์นนิง (Machine Learning) เท่านั้น ที่จำเป็นต้องมีการเรียนรู้จากข้อมูลที่สอนโดยมนุษย์

หลักการจริยธรรมปัญญาประดิษฐ์ (AI Ethics Principles) - หลักการในการออกแบบและพัฒนาให้พฤติกรรมหรือผลลัพธ์จากการทำงานของ AI สอดคล้องตามหลักการจริยธรรมหรือหลักการอันดีที่พึงปฏิบัติตามบริษัทที่ AI ถูกนำไปประยุกต์ใช้

หลักการในการประยุกต์ใช้ AI อย่างมีความรับผิดชอบ (Responsible AI) - หลักการในการนำ AI มาประยุกต์ใช้เพื่อประโยชน์ของบุคคลที่เกี่ยวข้องและองค์กร อย่างมีความรับผิดชอบต่อผลกระทบที่อาจเกิดขึ้นกับบุคคล องค์กร และสังคมโดยกว้างเพื่อนำไปสู่การสร้างเชื่อมั่น และการยอมรับจากบุคคลที่เกี่ยวข้องในทุกภาคส่วนที่เกี่ยวข้อง

องค์กร - องค์กร หน่วยงาน หรือกลุ่มบุคคล ในรูปแบบใดรูปแบบหนึ่งที่น่าเทคโนโลยี AI มาประยุกต์ใช้เพื่อสนับสนุนการปฏิบัติงานหรือบริการให้บรรลุตามเป้าหมายที่กำหนด

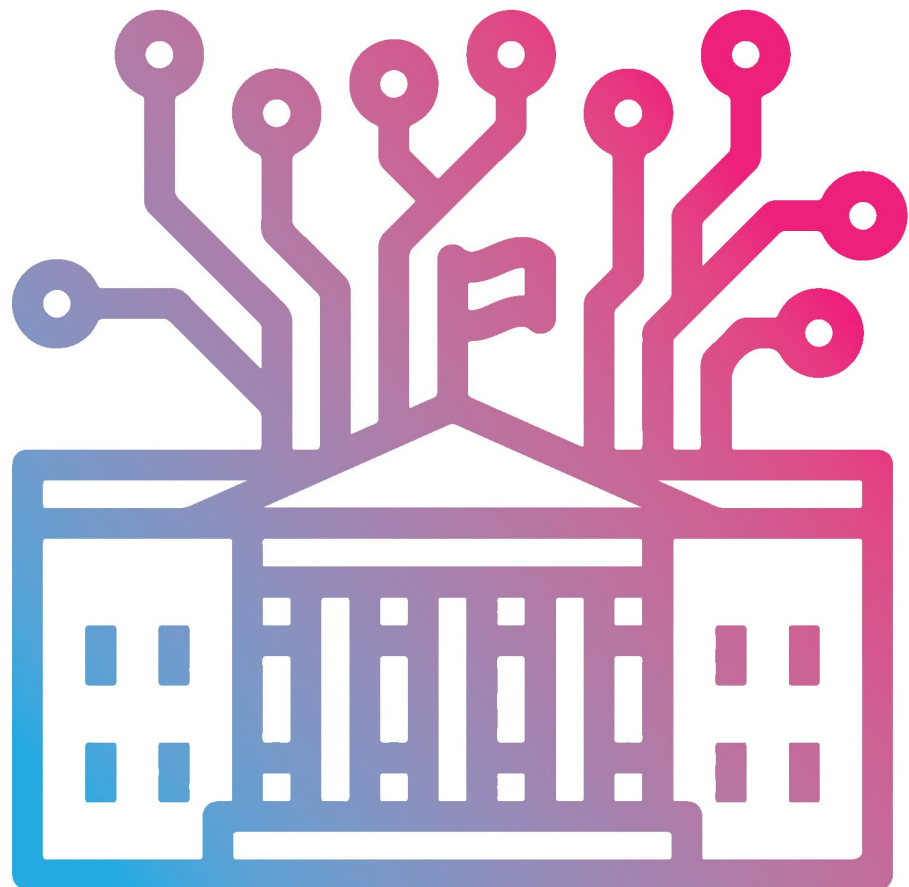
ผู้ให้บริการโซลูชันด้าน AI (AI Solution Providers) - ผู้ให้บริการที่นำเทคโนโลยี AI มาพัฒนาเป็นโซลูชันหรือซอฟต์แวร์ เพื่อให้บริการแก่บุคคลหรือองค์กรที่ต้องการนำเทคโนโลยี AI ไปประยุกต์หรือปรับใช้ตามวัตถุประสงค์ที่ต้องการ



6. การกำหนดโครงสร้างการกำกับดูแล (AI Governance Structure)

6.1 คณะกรรมการกำกับดูแลการประยุกต์ใช้ AI (AI Governance Council)

ในการประยุกต์ใช้ AI อย่างมีธรรมาภิบาลนั้น องค์กรจำเป็นต้องมีการกำหนดโครงสร้างการกำกับดูแลภายในองค์กร โดยมีการกำหนดหน้าที่และความรับผิดชอบ (Role and Responsibility) รวมถึงความรับผิดชอบต่อผลของการกระทำ (Accountability) ของบุคลากรและผู้มีส่วนได้เสียที่เกี่ยวข้อง (Stakeholders) ตามหน้าที่ที่ได้รับมอบหมาย เพื่อให้มั่นใจได้ว่าการประยุกต์ใช้ AI จะเป็นไปตามเป้าหมายอย่างมีความรับผิดชอบ โดยคำนึงถึงความสอดคล้องตามหลักการจริยธรรมปัญญาประดิษฐ์ ความสอดคล้องตามกฎหมาย และข้อกำหนดที่เกี่ยวข้อง และมีการควบคุมความเสี่ยงที่อาจเกิดขึ้น โดยมีคณะกรรมการกำกับดูแลการประยุกต์ใช้ AI (AI Governance Council) เป็นศูนย์กลางในการกำหนดทิศทาง การดำเนินการ (Direction) การเฝ้าติดตาม (Monitoring) และการประเมินผล (Evaluation) การประยุกต์ใช้ AI



กำหนดทิศทางการดำเนินการ (Direction)

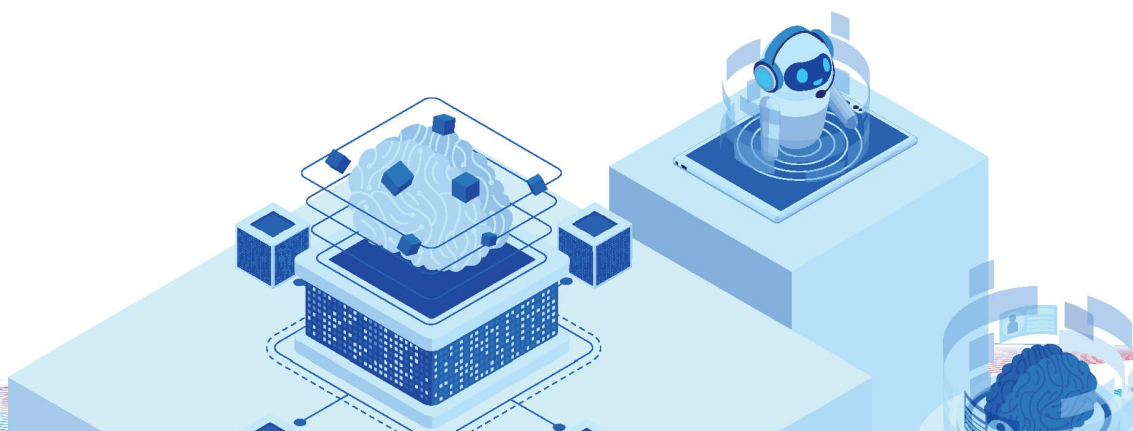
ในด้านการกำหนดทิศทางนั้น คณะกรรมการกำกับดูแล ทำหน้าที่ในการพิจารณากำหนดกลยุทธ์การประยุกต์ใช้ AI (AI Strategy) รวมถึงกำหนดนโยบายที่เกี่ยวข้องกับการประยุกต์ใช้ AI และการกำกับดูแล (AI Governance Policies) ในด้านต่าง ๆ เช่น

- โครงสร้างการกำกับดูแล
- หน้าที่และความรับผิดชอบ (Role and Responsibility) และความรับผิดชอบต่อผลของการกระทำ (Accountability)
- การบริหารจัดการและการใช้งานข้อมูล
- การออกแบบโซลูชัน สร้าง ทดสอบ และนำ AI ไปประยุกต์ใช้งาน
- การมีส่วนร่วมของมนุษย์ในการควบคุมการทำงานหรือตัดสินใจของ AI
- การประเมินความเสี่ยงและผลกระทบที่อาจเกิดขึ้นจากการประยุกต์ใช้ AI
- การขออนุมัติ พิจารณา หรือตัดสินใจด้านต่าง ๆ เช่น แผนปฏิบัติงานในการประยุกต์ใช้ AI (AI Roadmap) การนำโมเดลปัญญาประดิษฐ์ (AI Model) ไปใช้งานความสอดคล้องตามหลักการจริยธรรมปัญญาประดิษฐ์ การแก้ไขประเด็นปัญหาและแนวทางการดำเนินงานที่เกี่ยวข้องกับการประยุกต์ใช้ AI เป็นต้น

พร้อมทั้ง สื่อสารนโยบายและทิศทางของการกำกับดูแลให้แก่บุคลากรและผู้มีส่วนได้เสียที่เกี่ยวข้องเพื่อให้เกิดความเข้าใจและรับทราบร่วมกัน

โดยในการกำหนดกลยุทธ์และนโยบายดังกล่าว จำเป็นต้องคำนึงถึงความสอดคล้องตามเป้าหมายขององค์กรหรือเป้าหมายในการดำเนินธุรกิจ ความสอดคล้องตามหลักการจริยธรรมปัญญาประดิษฐ์ และความสอดคล้องตามกฎหมายและข้อกำหนดที่เกี่ยวข้อง พร้อมทั้งมีการควบคุมความเสี่ยงและผลกระทบที่อาจเกิดขึ้นตามบริบทของการนำ AI ไปประยุกต์ใช้

นอกจากนี้ ในด้านการควบคุมความเสี่ยงและผลกระทบที่อาจเกิดขึ้นนั้น คณะกรรมการกำกับดูแล ยังมีหน้าที่ในการพิจารณาว่าความเสี่ยงใดเป็นความเสี่ยงที่ยอมรับได้ (Risk Appetite) พร้อมทั้งกำกับดูแลให้มีการบริหารจัดการความเสี่ยง (Risk Management) อย่างเหมาะสม เพื่อควบคุมความเสี่ยงให้อยู่ในขอบเขตที่ยอมรับได้





การเฝ้าติดตาม (Monitoring)

การเฝ้าติดตาม (Monitoring) การประยุกต์ใช้ AI นั้น เป็นกระบวนการที่คณะกรรมการกำกับดูแล ต้องดำเนินการอย่างต่อเนื่อง เพื่อเฝ้าติดตามประสิทธิภาพ (Performance) ของการประยุกต์ใช้ AI ในด้านความสำเร็จตามเป้าหมายและประสิทธิภาพในการทำงานของ AI ตามที่กำหนด รวมถึงเฝ้าติดตามการปฏิบัติงานตามนโยบายขององค์กร หลักการจริยธรรมปัญญาประดิษฐ์ กฎหมายและข้อกำหนดที่เกี่ยวข้อง (Conformance) และเฝ้าติดตามความเสี่ยงว่ายังอยู่ในขอบเขตที่ยอมรับได้

โดยในกระบวนการเฝ้าติดตามนี้ คณะกรรมการกำกับดูแล จะทำการเฝ้าติดตามจากรายงานผลการปฏิบัติงานของบุคลากรและผู้มีส่วนได้เสียที่เกี่ยวข้องตามหน้าที่ที่ได้รับมอบหมาย โดยรายงานอาจประกอบด้วยข้อมูล เช่น

- ผลการดำเนินงานตามตัวชี้วัดที่องค์กรกำหนด (Key Performance Indicator: KPI)
- รายงานข้อมูลเกี่ยวกับความถูกต้องของการตัดสินใจของ AI
- ความสอดคล้องตามหลักการจริยธรรมปัญญาประดิษฐ์ที่เกี่ยวข้อง
- ข้อคิดเห็น (Feedback) ประเด็นปัญหา (Issue) และความผิดพลาด (Error) ที่พบจากประยุกต์ใช้ AI

ทั้งนี้ การรายงานผลการปฏิบัติงานอย่างทันทั่วถึงและมีข้อมูลที่เพียงพอต่อการตัดสินใจ หรือให้ความเห็นโดยคณะกรรมการกำกับดูแล จะช่วยให้เกิดการประยุกต์ใช้ AI อย่างมีความรับผิดชอบ อีกทั้งยังเป็นการสร้างความเชื่อมั่น และการยอมรับจากบุคคลที่เกี่ยวข้อง

การประเมินผล (Evaluation)

ในด้านการประเมินผลการประยุกต์ใช้ AI นั้น คณะกรรมการกำกับดูแล ควรมีการประเมินผลการปฏิบัติงานและประเมินผลผลลัพธ์จากการประยุกต์ใช้งาน AI ที่ผ่านมา เพื่อพิจารณาสิ่งที่ต้องดำเนินการปรับปรุง รวมถึงกำหนดแนวทางการดำเนินงานในอนาคต เพื่อลดผลกระทบที่อาจเกิดขึ้น โดยในการประเมินผลการประยุกต์ใช้ AI นั้น องค์กรอาจจัดให้มีการประเมินผลโดยผู้ตรวจประเมินภายใน (Internal Auditor) หรือผู้ตรวจประเมินภายนอก (External Auditor) ตามความเหมาะสม เพื่อช่วยเพิ่มความน่าเชื่อถือ (Reliability) และสร้างการยอมรับจากบุคคลที่เกี่ยวข้อง

โดยในการประเมินผลการประยุกต์ใช้งานนั้น อาจมีการพิจารณาในด้านต่าง ๆ ที่เกี่ยวข้องดังตัวอย่างด้านล่าง แต่อย่างไรก็ตามหลักเกณฑ์ในการประเมินผลของแต่ละองค์กรนั้น อาจมีความแตกต่างกันตามบริบทของการประยุกต์ใช้งาน AI ได้ด้วยเช่นกัน ตัวอย่างเช่น

- ประเมินประสิทธิภาพในการทำงานของ AI และผลลัพธ์ความสำเร็จจากการประยุกต์ใช้ AI ที่ผ่านมา เมื่อเทียบกับเป้าหมายที่กำหนด
- พิจารณาปรับเปลี่ยนเป้าหมายในการประยุกต์ใช้ AI หากพบว่าการประยุกต์ใช้ AI ไม่สามารถช่วยให้องค์กรบรรลุเป้าหมายตามที่กำหนดได้

- พิจารณาปรับปรุงข้อกำหนดความต้องการในการพัฒนาระบบ AI (AI Requirements) หากพบว่าข้อกำหนดความต้องการไม่ถูกต้อง หรือยกเลิกข้อกำหนดในบางเรื่องหากพบว่าไม่มีความจำเป็น
- นำความคิดเห็น (Feedback) ประเด็นปัญหา (Issue) และความผิดพลาด (Error) ที่พบ มาพิจารณาปรับปรุงเป้าหมายในการประยุกต์ใช้หรือข้อกำหนดความต้องการในการพัฒนาระบบ
- ประเมินมาตรการในการกำกับดูแลเพื่อให้การประยุกต์ใช้ AI สอดคล้องตามหลักการจริยธรรมปัญญาประดิษฐ์ รวมถึงความสอดคล้องตามกฎหมายและข้อกำหนดที่เกี่ยวข้อง
- ทบทวนความเสี่ยงจากการดำเนินงานที่ผ่านมา และวิเคราะห์ความเสี่ยงที่อาจเกิดขึ้นในอนาคต เพื่อระบุความเสี่ยงที่ยอมรับได้ (Risk Appetite) รวมถึงปรับปรุงมาตรการในการควบคุมความเสี่ยง และแผนการจัดการความเสี่ยง (Risk Treatment Plan) ให้เหมาะสม





การจัดตั้งคณะกรรมการกำกับดูแลการประยุกต์ใช้ AI (Setting up AI Governance Council)

คณะกรรมการกำกับดูแลฯ ซึ่งเป็นศูนย์กลางในการกำกับดูแลการประยุกต์ใช้ AI มีหน้าที่และความรับผิดชอบหลักในการกำกับดูแล ได้แก่

- กำกับดูแลการปฏิบัติงานเพื่อให้การประยุกต์ใช้ AI ประสบความสำเร็จและบรรลุตามเป้าหมายที่องค์กรกำหนด ผ่านการกำหนดกลยุทธ์ในการประยุกต์ใช้ AI (AI Strategy) และการกำหนดนโยบายที่เกี่ยวข้อง พร้อมทั้งเฝ้าติดตามประสิทธิภาพ (Performance) และประเมินผลลัพธ์จากการประยุกต์ใช้ AI
- เฝ้าติดตามและประเมินผลการปฏิบัติงานให้สอดคล้อง (Conformance) ตามนโยบายขององค์กร หลักการจริยธรรมปัญญาประดิษฐ์ กฎหมายและข้อกำหนดที่เกี่ยวข้อง
- กำกับดูแลและควบคุมความเสี่ยงให้อยู่ในขอบเขตที่ยอมรับได้
- อนุมัติ พิจารณา หรือตัดสินใจด้านต่าง ๆ ที่เกี่ยวข้องกับการประยุกต์ใช้ AI

จากหน้าที่และความรับผิดชอบข้างต้น สมาชิกของคณะกรรมการกำกับดูแลฯ จึงควรประกอบด้วยผู้บริหารและบุคลากรด้านต่าง ๆ ที่เกี่ยวข้องกับการประยุกต์ใช้ AI ทั้งจากภายในและภายนอกองค์กร (หากจำเป็น) แต่อย่างไรก็ตาม การคัดเลือกสมาชิกของคณะกรรมการกำกับดูแลฯ ในแต่ละองค์กรอาจแตกต่างกันตามขนาดขององค์กร ข้อจำกัดด้านทรัพยากรบุคคล กฎหมายและข้อกำหนดที่เกี่ยวข้อง บริบทของการประยุกต์ใช้ AI และความเสี่ยงที่อาจเกิดขึ้น เป็นต้น

ตัวอย่างโครงสร้างคณะกรรมการกำกับดูแล

- ผู้บริหารระดับสูง เช่น ผู้บริหารในระดับ C-level เป็นต้น ที่มีอำนาจในการกำหนดทิศทางในการดำเนินงาน การตัดสินใจ และสนับสนุนทรัพยากรที่จำเป็น

- ผู้บริหารหรือผู้แทนจากทีมงานด้านต่าง ๆ ที่เกี่ยวข้องกับการประยุกต์ใช้ AI เช่น กฎหมาย/กำกับดูแล ทีมงานด้านธุรกิจ และทีมงานด้านเทคนิค

- บุคคลภายนอก (หากจำเป็น) เช่น หน่วยงานกำกับดูแลและผู้เชี่ยวชาญในด้านที่จะนำ AI ไปประยุกต์ใช้

ซึ่งจากโครงสร้างดังกล่าว องค์กรอาจพิจารณา นำโครงสร้างการกำกับดูแลภายในองค์กรที่มีอยู่แล้ว มาปรับใช้เพื่อสนับสนุนหรือทำงานร่วมกับคณะกรรมการกำกับดูแล เช่น

- มอบหมายให้คณะกรรมการจริยธรรม (Ethics Board/Committee) กำกับดูแลเรื่องความสอดคล้องตามหลักการจริยธรรมปัญญาประดิษฐ์

- มอบหมายให้คณะกรรมการธรรมาภิบาลข้อมูล (Data Governance Council) กำกับดูแลเรื่องการบริหารจัดการข้อมูลให้มีคุณภาพและเหมาะสมกับการใช้งานร่วมกับ AI รวมถึงกำกับดูแลมาตรการในการการคุ้มครองข้อมูลส่วนบุคคลอย่างเหมาะสม

- มอบหมายให้คณะกรรมการตามโครงสร้าง Governance, Risk, and Compliance (GRC) บริหารจัดการและควบคุมความเสี่ยงที่เกี่ยวกับการประยุกต์ใช้ AI รวมถึงกำกับดูแลในด้านความสอดคล้องของการปฏิบัติงานในด้านต่าง ๆ



6.2. หน้าที่และความรับผิดชอบ (Role and Responsibility)

องค์กรที่จะนำ AI มาประยุกต์ใช้งานจำเป็นต้องมีการกำหนดหน้าที่และความรับผิดชอบ (Role and Responsibility) รวมถึงกำหนดความรับผิดชอบต่อผลของการกระทำ (Accountability) ของบุคลากรและผู้มีส่วนได้ส่วนเสียที่เกี่ยวข้องในทุกกระบวนการที่เกี่ยวข้องกับการประยุกต์ใช้ AI โดยอาจสรุปได้ดังนี้

ระดับนโยบาย (Strategic Level)

อาจประกอบด้วย:

- ผู้บริหารระดับสูง
- คณะกรรมการจริยธรรม*
- คณะกรรมการตามโครงสร้าง GRC*
- คณะกรรมการธรรมาภิบาลข้อมูล*
- ผู้บริหารหรือผู้แทนจากทีมงานภายในองค์กรที่เกี่ยวข้องกับการประยุกต์ใช้ AI
- ผู้เชี่ยวชาญหรือหน่วยงานกำกับดูแล**

หมายเหตุ:

- * โครงสร้างการกำกับดูแลภายในองค์กรที่อาจนำมาปรับใช้เพื่อสนับสนุนหรือร่วมดำเนินงาน
- ** บุคคลหรือกลุ่มบุคคลภายนอกที่อาจเข้าร่วมดำเนินงาน

ด้านกำกับดูแลการปฏิบัติงานเพื่อบรรลุเป้าหมายที่กำหนด

- กำหนดกลยุทธ์และเป้าหมายในการประยุกต์ใช้ AI ภายในองค์กร
- กำหนดนโยบายที่เกี่ยวข้องกับการประยุกต์ใช้ AI และนโยบายในการกำกับดูแลการปฏิบัติงานของบุคลากรและผู้มีส่วนได้ส่วนเสียที่เกี่ยวข้อง
- อนุมัติ พิจารณา หรือตัดสินใจด้านต่าง ๆ ที่เกี่ยวข้องกับการประยุกต์ใช้ AI เช่น
 - แผนปฏิบัติงานในการประยุกต์ใช้ AI (AI Roadmap)
 - การนำโมเดลปัญญาประดิษฐ์ (AI Model) ไปใช้งานจริง
 - ความสอดคล้องตามหลักการจริยธรรมปัญญาประดิษฐ์
 - การแก้ไขประเด็นปัญหาและกำหนดแนวทางการดำเนินงาน
- เฝ้าติดตามประสิทธิภาพการทำงานของ AI และระดับความสำเร็จในการประยุกต์ใช้ AI เมื่อเทียบกับเป้าหมายที่กำหนด
- ประเมินผลประยุตใช้งาน AI ในปัจจุบัน และกำหนดแนวทางการดำเนินงานในอนาคต

ด้านการปฏิบัติงานให้สอดคล้อง (Conformance) ตามข้อกำหนดภายในและภายนอกองค์กร

- กำหนดนโยบายในการกำกับดูแลการปฏิบัติงานของบุคลากรและผู้มีส่วนได้ส่วนเสียที่เกี่ยวข้อง รวมถึงเฝ้าติดตามและประเมินผลการปฏิบัติงานตามนโยบายองค์กร หลักการจริยธรรมปัญญาประดิษฐ์ กฎหมายและข้อกำหนดที่เกี่ยวข้อง

ด้านการควบคุมความเสี่ยงและผลกระทบที่อาจเกิดขึ้น

- กำหนดความเสี่ยงที่ยอมรับได้ (Risk Appetite) และพิจารณาความเหมาะสมของมาตรการเพื่อจำกัดความเสี่ยงให้อยู่ในขอบเขตที่ยอมรับได้

อาจประกอบด้วย:

- ผู้บริหาร/ผู้จัดการที่เกี่ยวข้องกับการประยุกต์ใช้ AI
- ทีมงานด้านธุรกิจที่เกี่ยวข้องกับการประยุกต์ใช้ AI
- ทีมงานด้านเทคนิคที่เกี่ยวข้องกับข้อมูลและปัญญาประดิษฐ์ เช่น นักพัฒนาปัญญาประดิษฐ์ (AI Developers) นักวิเคราะห์ข้อมูล (Data Analyst) นักวิทยาศาสตร์ข้อมูล (Data Sciences) วิศวกรข้อมูล (Data Engineers)
- ทีมงานด้านเทคนิคอื่นที่เกี่ยวข้องกับการนำ AI ไปใช้งานจริง เช่น วิศวกรระบบ (System Engineers) วิศวกรซอฟต์แวร์ (Software Developers)
- หน่วยงานภายนอก* เช่น หน่วยงานที่รับจ้างพัฒนาระบบ (Vendor) ผู้ให้บริการโซลูชันด้าน AI (AI Solution Providers)

หมายเหตุ:

* บุคคลหรือกลุ่มบุคคลภายนอกที่อาจเข้าร่วมดำเนินงานด้านการปฏิบัติงานตามกลยุทธ์ในการประยุกต์ใช้ AI

ด้านการปฏิบัติงานตามกลยุทธ์ในการประยุกต์ใช้ AI

- จัดเตรียมข้อมูลสำหรับนำเสนอคณะกรรมการกำกับดูแล เพื่อประกอบการกำหนดกลยุทธ์ในการประยุกต์ใช้ AI เช่น กระบวนการ/ขั้นตอนปฏิบัติที่สามารถนำ AI มาประยุกต์ใช้ ประโยชน์ที่จะได้รับระดับความสำเร็จที่คาดหวัง ความพร้อมขององค์กร ความซับซ้อนในการดำเนินการและเวลาที่ใช้ในการดำเนินการ หลักการจริยธรรมปัญญาประดิษฐ์ กฎหมายและข้อกำหนดที่เกี่ยวข้อง เป็นต้น
- จัดทำแผนปฏิบัติงานในการประยุกต์ใช้ AI (AI Roadmap) ตามกลยุทธ์ที่คณะกรรมการกำกับดูแล กำหนด
- จัดทำสถาปัตยกรรมที่เกี่ยวข้องกับการประยุกต์ใช้ AI (AI Architecture) และจัดเตรียมทรัพยากรที่จำเป็นในการดำเนินงานตามแผนปฏิบัติงานในการประยุกต์ใช้ AI

ด้านการออกแบบ พัฒนา และนำ AI ไปประยุกต์ใช้

- ออกแบบโซลูชัน พัฒนา ทดสอบ และนำ AI ไปใช้งาน พร้อมทั้งจัดทำเอกสารที่เกี่ยวข้อง
- จัดเตรียมข้อมูลและบริหารจัดการข้อมูลสำหรับทำงานร่วมกับ AI เพื่อให้ได้ข้อมูลที่มีคุณภาพในการใช้งานร่วมกับ AI
- จัดทำมาตรการในการรักษาความมั่นคงปลอดภัยและการคุ้มครองความเป็นส่วนตัว
- จัดให้มีช่องทางรับข้อคิดเห็น (Feedback) ประเด็นปัญหา (Issue) ความผิดพลาด (Error) จากการปฏิบัติงานและการให้บริการ พร้อมทั้งจัดให้มีมาตรการในการแก้ไขปัญหา และรับมือกับเหตุการณ์ผิดปกติ (Incident) ได้ทันทั่วถึงเพื่อลดผลกระทบที่อาจเกิดขึ้น
- เฝ้าติดตามประสิทธิภาพการทำงานของ AI และระดับความสำเร็จในการประยุกต์ใช้ AI และนำเสนอคณะกรรมการกำกับดูแล
- ประเมินผลประยุกต์ใช้งาน AI ในปัจจุบัน จัดทำข้อเสนอในการดำเนินงานในอนาคต และนำเสนอคณะกรรมการกำกับดูแล

ด้านการปฏิบัติตามข้อกำหนดภายในและภายนอกองค์กร

- จัดทำขั้นตอนปฏิบัติ มาตรการ และเครื่องมือเพื่อควบคุมการปฏิบัติงานให้สอดคล้องตามนโยบายขององค์กร หลักการจริยธรรมปัญญาประดิษฐ์ กฎหมายและข้อกำหนดที่เกี่ยวข้อง
- เฝ้าติดตามผลการปฏิบัติงานตามข้อกำหนดทั้งภายในและภายนอกองค์กร และนำเสนอคณะกรรมการกำกับดูแล
- ประเมินผลมาตรการในการกำกับดูแล จัดทำข้อเสนอในการดำเนินงานในอนาคต และนำเสนอคณะกรรมการกำกับดูแล

ด้านการควบคุมความเสี่ยงและผลกระทบที่อาจเกิดขึ้น

- ประเมินความเสี่ยงและผลกระทบที่อาจเกิดขึ้นจากการประยุกต์ใช้ AI
- จัดทำมาตรการเพื่อจำกัดความเสี่ยงให้อยู่ในขอบเขตที่ยอมรับได้
- กำหนดระดับการมีส่วนร่วมของมนุษย์ในการทำงานของ AI เพื่อลดผลกระทบที่อาจเกิดขึ้น
- นำเสนอผลการดำเนินงานเกี่ยวกับการควบคุมความเสี่ยงต่อคณะกรรมการกำกับดูแล

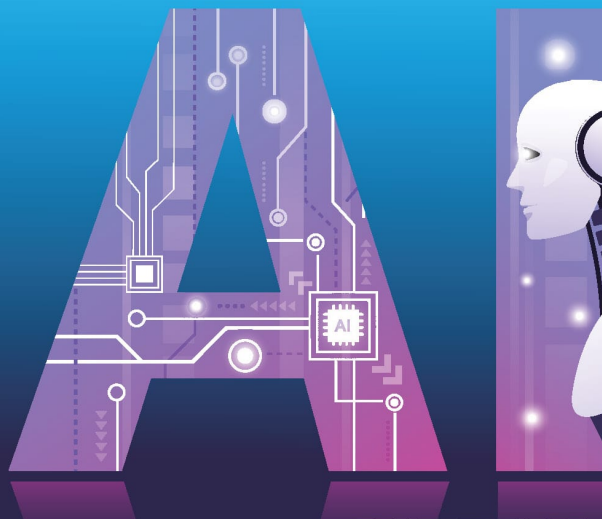
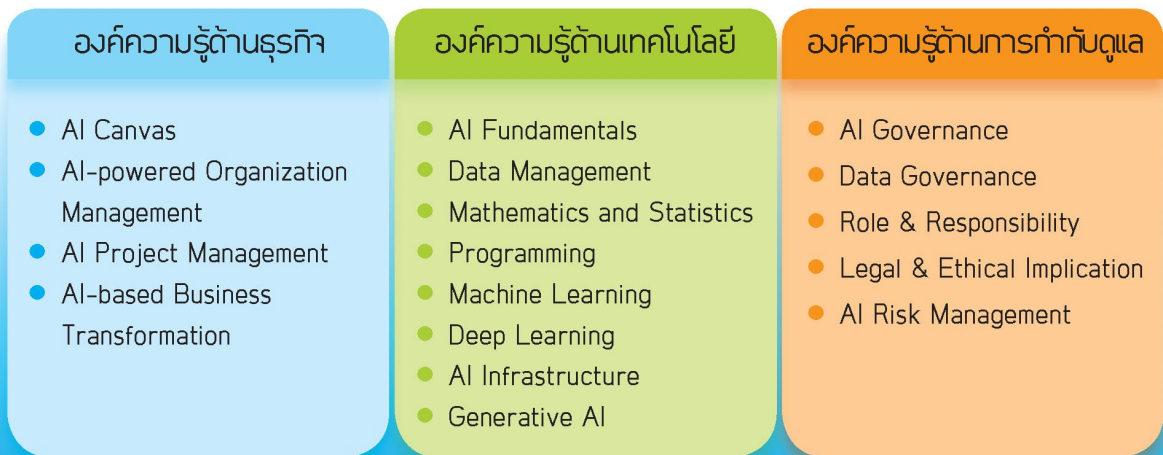
นอกจากนี้ องค์กรยังจำเป็นต้องมีการสื่อสารนโยบายและขั้นตอนปฏิบัติที่เกี่ยวข้อง พร้อมทั้งสร้างความตระหนักรู้ (Awareness) ต่อความรับผิดชอบ (Responsibility) และความรับผิดชอบต่อผลของการกระทำ (Accountability) ของแต่ละหน้าที่

ในกรณีที่มีการว่าจ้างหน่วยงานภายนอก (Outsource) เช่น หน่วยงานที่รับจ้างพัฒนาระบบ (Vendor) ผู้ให้บริการโซลูชันด้าน AI (AI Solution Providers) เป็นต้น องค์กรยังคงต้องมีความรับผิดชอบต่อผลของการกระทำ (Accountability) อันเกิดจากการทำงานของ AI ด้วยเหตุนี้ องค์กรจึงควรมีการกำกับดูแลการปฏิบัติงานของ หน่วยงานภายนอกอย่างเหมาะสม พร้อมทั้งกำหนดหน้าที่และความรับผิดชอบในข้อตกลงว่าจ้างหรือข้อตกลงการใช้บริการ

6.3. การพัฒนาศักยภาพบุคลากร (Competency Building)

เพื่อให้องค์กรประสบความสำเร็จในการประยุกต์ใช้ AI คณะกรรมการกำกับดูแล รวมถึงบุคลากรและผู้มีส่วนได้ส่วนเสียที่เกี่ยวข้อง ควรได้รับการส่งเสริมให้มีความรู้ความเข้าใจเกี่ยวกับ AI ในมิติต่าง ๆ อย่างเหมาะสมตามหน้าที่ที่ได้รับมอบหมาย เพื่อให้สามารถปฏิบัติหน้าที่ได้อย่างถูกต้องและสอดคล้องตามนโยบายขององค์กร หลักการจรรยาบรรณ ปัญญาประดิษฐ์ กฎหมายและข้อกำหนดที่เกี่ยวข้อง รวมถึงสามารถรับมือกับความเสี่ยงและเหตุการณ์ผิดปกติได้อย่างเหมาะสม

ตัวอย่างองค์ความรู้ที่เกี่ยวกับการประยุกต์ใช้ AI



7. การกำหนดกลยุทธ์ในการประยุกต์ใช้ AI (AI Strategy)

7.1. การกำหนดกลยุทธ์ในการประยุกต์ใช้ AI อย่างมีความรับผิดชอบ (Responsible AI Strategy)

ปัจจุบันเทคโนโลยี AI ถูกนำมาประยุกต์ใช้ในรูปแบบที่หลากหลายแตกต่างกันตามวัตถุประสงค์หรือเป้าหมายขององค์กร เพื่อให้การประยุกต์ใช้ AI ประสบความสำเร็จและบรรลุตามเป้าหมาย องค์กรจึงควรกำหนดกลยุทธ์ในการประยุกต์ใช้ AI และแผนปฏิบัติงานอย่างเหมาะสม โดยพิจารณาถึงโอกาสหรือประโยชน์ที่องค์กรจะได้รับ และพิจารณาความเป็นไปได้ในการประยุกต์ใช้ AI โดยในการพิจารณาความเป็นไปได้นั้น องค์กรควรพิจารณาถึงปัจจัยในด้านต่าง ๆ เช่น ความสามารถของเทคโนโลยี และ/หรือผลิตภัณฑ์ในปัจจุบันที่สามารถตอบโจทย์ความต้องการขององค์กร ความพร้อมของข้อมูล และทรัพยากรที่เกี่ยวข้อง เช่น บุคลากรที่มีความรู้ความสามารถ งบประมาณ ตลอดจน ความซับซ้อน และเวลาที่ใช้ดำเนินการ ข้อจำกัดและความท้าทายในการดำเนินงาน เป็นต้น

นอกจากนี้ เพื่อให้เกิดการประยุกต์ใช้งาน AI อย่างมีความรับผิดชอบ ในการกำหนดกลยุทธ์ยังจำเป็นต้องมีการพิจารณาและคาดการณ์ถึงผลกระทบเชิงลบที่อาจส่งผลกระทบต่อบุคคลที่เกี่ยวข้อง องค์กร หรือสังคม โดยกว้าง เช่น การตัดสินใจของ AI ที่มีอคติจนนำไปสู่การเลือกปฏิบัติ ความเสียหายต่อชื่อเสียง ชีวิต หรือทรัพย์สินซึ่งเกิดจากการทำงานที่ผิดพลาดของ AI เป็นต้น ด้วยเหตุนี้ องค์กรจึงควรมีการกำหนดกลยุทธ์และจัดทำแผนปฏิบัติงาน เพื่อให้ระบบหรือบริการที่เกี่ยวข้องกับ AI ทำงานได้อย่างสอดคล้องตามหลักการจริยธรรมปัญญาประดิษฐ์ กฎหมายและข้อกำหนดที่เกี่ยวข้อง พร้อมทั้งลดความเสี่ยงและผลกระทบที่อาจเกิดขึ้นดังกล่าว



มองหาโอกาสในการนำ AI มาประยุกต์ใช้



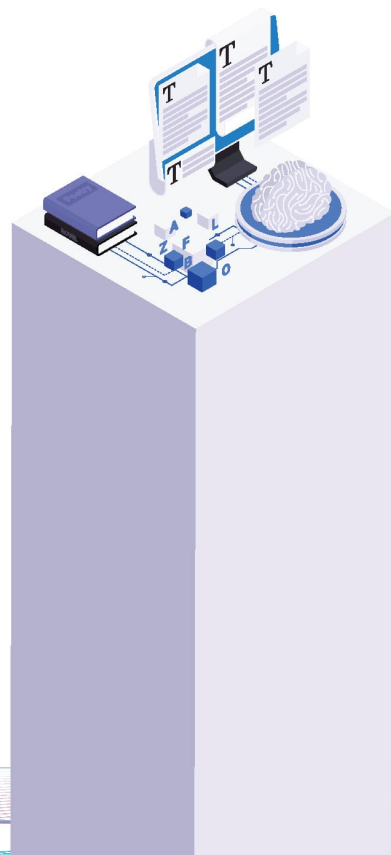
กำหนดเป้าหมายในการประยุกต์ใช้ AI ตามลำดับความสำคัญ และความพร้อมขององค์กร



กำหนดกลยุทธ์ในการบริหารจัดการข้อมูลเพื่อสนับสนุนการประยุกต์ใช้ AI



จัดทำแผนปฏิบัติงาน (Roadmap) ในการประยุกต์ใช้ AI



มองหาโอกาสในการนำ AI มาประยุกต์ใช้

เพื่อให้เห็นโอกาสในการนำ AI มาประยุกต์ใช้ องค์กรจึงควรเริ่มต้นจากการทำความเข้าใจภาพรวม วิสัยทัศน์ พันธกิจ รวมถึงเป้าหมายระยะสั้นและระยะยาวขององค์กร เพื่อเข้าใจเป้าหมายขององค์กรหรือเป้าหมายในการดำเนินธุรกิจ จากนั้นจึงเริ่มทำการวิเคราะห์กระบวนการหรือขั้นตอนปฏิบัติต่าง ๆ เพื่อมองหาโอกาสที่ AI จะเข้ามาเป็นองค์ประกอบหนึ่งในการทำงาน ซึ่งจะช่วยให้องค์กรบรรลุตามเป้าหมายหรือตอบโจทย์ความต้องการทางธุรกิจได้อย่างมีประสิทธิภาพมากยิ่งขึ้น โดยในการวิเคราะห์นั้นอาจพิจารณาในประเด็นดังต่อไปนี้ ตัวอย่างเช่น

- ประโยชน์ที่องค์กรจะได้รับจากการนำ AI มาประยุกต์ใช้ เช่น เพิ่มรายได้ให้แก่องค์กร การยกระดับความพึงพอใจของลูกค้า ปรับปรุงประสิทธิภาพ และผลิตผล ลดค่าใช้จ่ายในการปฏิบัติงาน เป็นต้น
- ระดับความสำเร็จที่คาดหวังว่า AI จะเข้ามาช่วยให้องค์กรบรรลุเป้าหมายในการดำเนินธุรกิจ หรือความสำเร็จตามตัวชี้วัดที่องค์กรกำหนด (Key Performance Indicator: KPI) เช่น สามารถผลิตสินค้าได้มากขึ้น 500 ชิ้น/วัน สามารถช่วยลดค่าใช้จ่ายลงได้ 20% เป็นต้น
- เทคโนโลยีและวิธีการที่จะนำ AI มาประยุกต์ใช้
- ขอบเขตการดำเนินงานที่ AI จะเข้าไปเกี่ยวข้องกับการปฏิบัติงาน
- ข้อมูลและทรัพยากรที่จำเป็นต้องใช้ในการดำเนินการ เช่น งบประมาณ บุคลากร โครงสร้างพื้นฐาน เทคโนโลยีสารสนเทศ เป็นต้น
- ผลกระทบจากการนำ AI ไปประยุกต์ใช้ เช่น การปรับเปลี่ยนกระบวนการหรือขั้นตอนปฏิบัติ ผลกระทบต่อผู้มีส่วนได้ส่วนเสีย ความเสี่ยงและผลกระทบที่อาจเกิดขึ้นจากการประยุกต์ใช้

AI และมาตรการในการควบคุมความเสี่ยง และผลกระทบดังกล่าว เป็นต้น

- ข้อจำกัดหรือความท้าทายในการประยุกต์ใช้ AI เช่น กฎหมายและข้อกำหนดที่ต้องปฏิบัติตาม ข้อจำกัดด้านทรัพยากร ปริมาณและคุณภาพของข้อมูลที่มีอยู่ในปัจจุบัน เป็นต้น

โดยในการวิเคราะห์กระบวนการหรือขั้นตอนปฏิบัติต่าง ๆ นั้น ควรมีการทำงานร่วมกันระหว่างผู้เชี่ยวชาญด้าน AI และทีมงานด้านธุรกิจหรือผู้ที่เกี่ยวข้องกับการปฏิบัติงานโดยตรง เพื่อสะท้อนมุมมองในทางปฏิบัติ และวิเคราะห์ความเป็นไปได้ในการที่จะนำ AI มาปรับใช้ เพื่อให้องค์กรบรรลุตามเป้าหมายที่กำหนด

กำหนดเป้าหมายในการประยุกต์ใช้ AI

ภายหลังจากทำการวิเคราะห์กระบวนการหรือขั้นตอนปฏิบัติเพื่อมองหาโอกาสและความเป็นไปได้ที่จะนำ AI มาประยุกต์ใช้นั้น องค์กรอาจพบว่ามีกระบวนการหรือขั้นตอนปฏิบัติจำนวนมาก หลากหลาย และอาจไม่สามารถดำเนินการไปพร้อมกันได้ทั้งหมด ด้วยเหตุนี้ องค์กรจึงจำเป็นต้องจัดลำดับความสำคัญในการดำเนินงาน โดยอาจพิจารณาจากปัจจัยต่าง ๆ ได้แก่

- ประโยชน์ที่องค์กรจะได้รับจากการนำ AI มาประยุกต์ใช้ในกระบวนการหรือขั้นตอนปฏิบัตินั้น
- ระดับความสำเร็จที่คาดหวังว่า AI จะเข้ามาช่วยให้บรรลุตามเป้าหมายขององค์กร เป้าหมายในการดำเนินธุรกิจ หรือความสำเร็จตามตัวชี้วัดที่องค์กรกำหนด (Key Performance Indicator: KPI)
- หลักการจริยธรรมปัญญาประดิษฐ์ กฎหมายและข้อกำหนดที่ต้องออกแบบและพัฒนาระบบให้สอดคล้อง
- ความพร้อมของทรัพยากรที่จำเป็นต้องใช้ในการดำเนินการ เช่น งบประมาณ ข้อมูล บุคลากร โครงสร้างพื้นฐานเทคโนโลยีสารสนเทศ เป็นต้น

- ความพร้อมในการกำกับดูแลและปฏิบัติงานเพื่อเปลี่ยนผ่านไปสู่การประยุกต์ใช้งาน AI รวมถึงการรับมือความเสี่ยงและผลกระทบที่อาจเกิดขึ้น
- ความซับซ้อนในการดำเนินการและเวลาที่ใช้ในการดำเนินการ

ซึ่งการจัดลำดับความสำคัญในการดำเนินงานดังกล่าว จะช่วยให้องค์กรสามารถเลือกเป้าหมายในการประยุกต์ใช้งาน AI ที่มีความสำคัญมาดำเนินการก่อน และสามารถจัดสรรทรัพยากรเพื่อมาดำเนินงานได้อย่างมีประสิทธิภาพ

กำหนดกลยุทธ์ในการบริหารจัดการข้อมูลเพื่อสนับสนุนการประยุกต์ใช้ AI

เนื่องจาก AI จำเป็นต้องใช้ข้อมูลในการประมวลผลด้วยเหตุนี้ องค์กรจึงจำเป็นต้องมีการกำหนดกลยุทธ์ในการบริหารจัดการข้อมูลที่จะใช้ทำงานร่วมกับ AI อย่างเหมาะสม ตั้งแต่กระบวนการสรรหาแหล่งข้อมูลที่เหมาะสม การเตรียมข้อมูลให้อยู่ในรูปแบบที่พร้อมใช้งานและเพียงพอสำหรับ AI ในการประมวลผล ไปจนถึงการพัฒนาอัลกอริทึมเพื่อวิเคราะห์คาดการณ์ ให้คำแนะนำ ตัดสินใจ หรือดำเนินการใด ๆ แทนมนุษย์ โดยอยู่บนพื้นฐานของข้อมูลที่มีอยู่ได้อย่างมีประสิทธิภาพ

เพื่อให้มั่นใจได้ว่าองค์กรจะมีข้อมูลที่เหมาะสมและเพียงพอต่อการสนับสนุนการประยุกต์ใช้ AI ให้บรรลุตามเป้าหมายที่กำหนด องค์กรควรมีการสรรหาแหล่งข้อมูลที่เหมาะสม โดยอาจพิจารณาจากคุณภาพของข้อมูล ความน่าเชื่อถือในการเก็บรวบรวมข้อมูล ความสามารถในการเข้าถึงข้อมูลได้โดยสะดวก ข้อมูลอยู่ในรูปแบบ

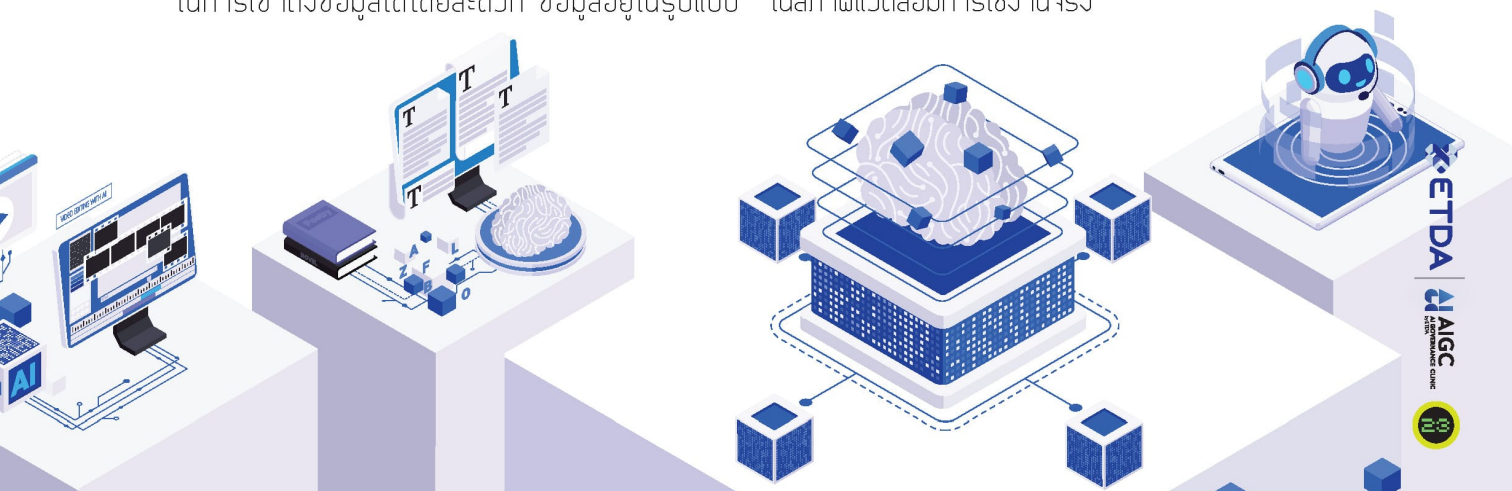
หรือประเภทที่ตรงตามความต้องการใช้งาน ข้อมูลมีความหลากหลายและครอบคลุมตัวอย่างที่ต้องการ มีขนาดของข้อมูลหรือขนาดของกลุ่มตัวอย่างที่เหมาะสม ข้อมูลมีการปรับปรุงให้ทันสมัยอยู่เสมอ เป็นต้น

นอกจากนี้องค์กรยังจำเป็นต้องมีการเตรียมความพร้อมในด้านบุคลากร เครื่องมือ และกระบวนการอย่างเหมาะสม เพื่อจัดเตรียมข้อมูลที่เหมาะสมสำหรับการพัฒนา ทดสอบ และใช้งาน AI (ซึ่งจะกล่าวถึงในหัวข้อ 7.1) รวมถึงควรมีการกำกับดูแลโดยนำหลักการธรรมาภิบาลข้อมูล (Data Governance) มาปรับใช้ เพื่อให้มั่นใจได้ว่าข้อมูลมีคุณภาพตรงตามความต้องการในการใช้งาน

จัดทำแผนปฏิบัติงาน (Roadmap) ในการประยุกต์ใช้ AI

ภายหลังจากที่องค์กรกำหนดเป้าหมายในการประยุกต์ใช้งาน AI ที่จะนำมาดำเนินการ รวมถึงแนวทางการบริหารจัดการข้อมูลที่สอดคล้องตามกลยุทธ์ที่วางไว้แล้ว องค์กรควรจัดทำแผนปฏิบัติงานตามลำดับความสำคัญที่ได้วิเคราะห์ไว้ พร้อมทั้งกำหนดขั้นตอนการดำเนินงานของแต่ละเป้าหมายของการประยุกต์ใช้ AI เช่น การจัดเตรียมข้อมูล การพัฒนา ทดสอบ และการนำ AI ไปใช้งาน เป็นต้น

นอกจากนี้ องค์กรอาจกำหนดให้มีการจัดทำระบบต้นแบบ (Prototype) ก่อนนำไปขยายผลเพื่อใช้งานจริง เพื่อทดสอบและประเมินประสิทธิภาพในการทำงานของ AI พร้อมทั้งประเมินความคุ้มค่าในการลงทุน วิเคราะห์ปัญหาและอุปสรรคที่อาจเกิดขึ้นเมื่อนำ AI ไปประยุกต์ใช้ในสภาพแวดล้อมการใช้งานจริง



7.2. การบริหารจัดการความเสี่ยงจากการประยุกต์ใช้ AI (AI Risk Management)

เนื่องจากการประยุกต์ใช้ AI อาจนำมาซึ่งโอกาสใหม่ในการดำเนินธุรกิจหรือสร้างประโยชน์ให้แก่องค์กร แต่ในขณะเดียวกันการประยุกต์ใช้ AI ก็อาจก่อให้เกิดผลกระทบเชิงลบต่อบุคคลที่เกี่ยวข้อง องค์กร และสังคมโดยกว้างได้เช่นกัน ด้วยเหตุนี้ ในระหว่างการกำหนดกลยุทธ์และเป้าหมายในการประยุกต์ใช้ AI นั้น องค์กรจำเป็นต้องมีการประเมินความเสี่ยงเพื่อให้เห็นถึงความไม่แน่นอนหรือโอกาสที่การประยุกต์ใช้ AI จะไม่เป็นไปตามเป้าหมายที่กำหนด รวมถึงเห็นผลกระทบเชิงลบที่อาจเกิดขึ้น

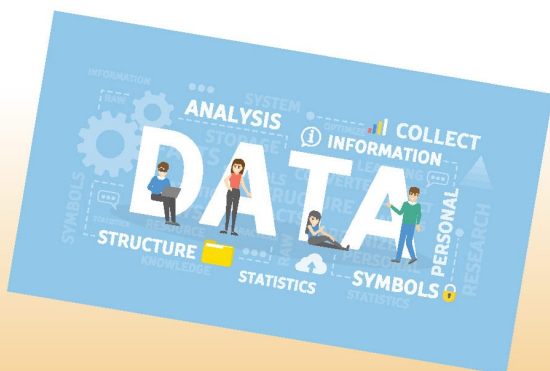
เนื่องจากความเสี่ยงจากการประยุกต์ใช้ AI นั้น มีความหลากหลายและแตกต่างกันไปตามบริบทของการประยุกต์ใช้ AI ตัวอย่างเช่น

- ความเสี่ยงด้านคุณภาพข้อมูล (Data Quality)
 - เป็นความเสี่ยงที่ข้อมูลขององค์กรมีคุณภาพไม่เพียงพอจนส่งผลให้การทำงานของ AI ไม่มีประสิทธิภาพตามเป้าหมายที่กำหนด หรือทำงานผิดพลาด เป็นต้น
- ความเสี่ยงด้านความไม่เป็นธรรมและการเลือกปฏิบัติ (Unfairness and Discrimination)
 - เป็นความเสี่ยงที่ผลลัพธ์จากการทำงานของ AI นำไปสู่ความไม่เป็นธรรมและการเลือกปฏิบัติ ซึ่งอาจเกิดจากอคติที่มาจากกรออกแบบและสร้างโมเดล (Bias Introduced by Engineering Decisions) หรือเกิดจากอคติที่มาจากข้อมูล (Data Bias) เช่น การใช้ข้อมูลที่ไม่มีความหลากหลายหรือเอนเอียงมาสอน AI เป็นต้น

- ความเสี่ยงด้านภัยคุกคามทางไซเบอร์ (Cyber Attack) - เป็นความเสี่ยงที่ AI ถูกโจมตีโดยผู้ประสงค์ร้าย โดยอาศัยช่องโหว่ (Vulnerability) เพื่อโจมตี AI หรือโจมตีข้อมูลที่ใช้ในการสอน AI โดยมีวัตถุประสงค์ให้ AI ทำงานผิดพลาดหยุดการทำงาน หรือเกิดการรั่วไหลของข้อมูล เป็นต้น
- ความเสี่ยงด้านการคุ้มครองความเป็นส่วนตัว (Privacy) - เป็นความเสี่ยงที่ข้อมูลส่วนบุคคลถูกละเมิด โดยอาจเกิดจากการมีมาตรการในการคุ้มครองข้อมูลส่วนบุคคลที่ไม่เหมาะสม ความผิดพลาดในการปฏิบัติงาน หรือถูกโจมตีโดยผู้ประสงค์ร้าย เป็นต้น
- ความเสี่ยงด้านการไม่ปฏิบัติตามกฎหมายและข้อกำหนดที่เกี่ยวข้อง (Non-compliance)
 - เป็นความเสี่ยงที่ผลลัพธ์จากการทำงานของ AI นำไปสู่ความการละเมิดกฎหมายหรือข้อกำหนดที่เกี่ยวข้อง
- ความเสี่ยงด้านความน่าเชื่อถือหรือชื่อเสียง (Trust and Reputation) - เป็นความเสี่ยงที่ AI ทำงานผิดพลาดหรือก่อให้เกิดผลกระทบเชิงลบ เช่น AI ทำงานเมื่ออยู่ในสถานการณ์ที่ไม่เคยถูกทดสอบ หรือต้องตัดสินใจบนพื้นฐานข้อมูลที่ไม่เคยได้รับการสอนมาก่อน เป็นต้น

เพื่อลดความเสี่ยงจากการประยุกต์ใช้ AI ดังกล่าว องค์กรจึงจำเป็นต้องมีการบริหารจัดการความเสี่ยง (Risk Management) เพื่อควบคุมความเสี่ยงในทุกกิจกรรมตลอดวงจรชีวิตของ AI (AI Lifecycle) ให้อยู่ในขอบเขตที่ยอมรับได้

ซึ่งในการบริหารจัดการและความคุมความเสี่ยงนั้น องค์กรจำเป็นต้องมีการกำหนดนโยบาย ขั้นตอนปฏิบัติ และมาตรการเพื่อควบคุมความเสี่ยงที่อาจเกิดขึ้น



ในแต่ละกิจกรรมตลอดวงจรชีวิตของ AI รวมถึงจัดให้มีกระบวนการในการบริหารจัดการความเสี่ยง ซึ่งประกอบด้วย

1. การสื่อสารและหารือร่วมกัน (Communication and Consultation) ระหว่างบุคลากรภายในองค์กรและผู้มีส่วนได้เสียที่เกี่ยวข้องกับการประยุกต์ใช้ AI
2. การทำความเข้าใจขอบเขตและบริบทของการประยุกต์ใช้ AI รวมถึงกำหนดเกณฑ์ในการประเมินความเสี่ยง (Risk Criteria)
3. การประเมินความเสี่ยง (Risk Assessment) เพื่อให้องค์กรมองเห็นความเสี่ยงและผลกระทบที่อาจเกิดขึ้นจากการประยุกต์ใช้ AI รวมถึงมาตรการในการควบคุมความเสี่ยงดังกล่าว
4. การกำหนดมาตรการที่เหมาะสมในการควบคุมและแก้ไขความเสี่ยง (Risk Treatment) รวมถึงจัดทำแผนการดำเนินการเพื่อควบคุมและแก้ไขความเสี่ยง (Risk Treatment Plan) ดังกล่าวเพื่อควบคุมความเสี่ยงให้อยู่ในขอบเขตที่ยอมรับได้ นอกจากนี้การกำหนดมาตรการในการควบคุมและแก้ไขความเสี่ยงจะต้องครอบคลุมถึงการปฏิบัติงานของหน่วยงานภายนอก เช่น หน่วยงานที่รับจ้างพัฒนาระบบ (Vendor) ผู้ให้บริการโซลูชันด้าน AI (AI Solution Providers) เป็นต้น และครอบคลุมถึงโซลูชัน/ผลิตภัณฑ์ AI ที่ใช้งาน เพื่อควบคุมความเสี่ยงที่อาจเกิดขึ้นจากบุคคลหน่วยงานภายนอก
5. การเฝ้าติดตามและทบทวน (Monitoring and Review) ประสิทธิภาพในการควบคุมและแก้ไขความเสี่ยงที่เกิดขึ้น
6. การบันทึกและรายงานผลการบริหารจัดการความเสี่ยง (Recording & Reporting) ต่อคณะกรรมการกำกับดูแล บุคลากร และผู้มีส่วนได้เสียที่เกี่ยวข้อง เพื่อประเมินผลการปฏิบัติงาน (Evaluation) และปรับปรุงประสิทธิภาพในการบริหารจัดการและความคุมความเสี่ยง

กรอบแนวทางการบริหารจัดการความเสี่ยงจากการประยุกต์ใช้ AI (AI Risk Management Framework)

จากกระบวนการในการบริหารจัดการความเสี่ยงข้างต้น องค์กรอาจพิจารณานำกรอบแนวทางการบริหารจัดการความเสี่ยงที่ประกาศโดยหน่วยงานด้านมาตรฐานสากลมาปรับใช้ โดยในปี 2023 องค์กร International Organization for Standardization (ISO) ซึ่งเป็นองค์กรระหว่างประเทศด้านมาตรฐาน และสถาบัน National Institute of Standards and Technology (NIST) ซึ่งเป็นสถาบันมาตรฐานและเทคโนโลยีแห่งชาติของสหรัฐอเมริกา ได้มีการประกาศมาตรฐาน

1. ISO/IEC 23894:2023 Information technology – Artificial intelligence – Guidance on risk management ซึ่งเป็นส่วนขยาย (Extension) จากมาตรฐาน ISO 31000:2018 Risk management – Guidelines ที่อธิบายรายละเอียดเพิ่มเติมในส่วนของการบริหารจัดการความเสี่ยงจากการประยุกต์ใช้ AI

2. Artificial Intelligence Risk Management Framework (AI RMF 1.0) โดย NIST

ซึ่งทั้ง 2 มาตรฐานดังกล่าว ได้มีการวางแผนปฏิบัติในการบริหารจัดการความเสี่ยงจากการประยุกต์ใช้ AI เป็นการเฉพาะ เพื่อให้องค์กรมีความเข้าใจและตระหนักถึงความเสี่ยงรูปแบบใหม่ที่เกิดจากการประยุกต์ใช้ AI รวมถึงใช้เป็นแนวทางในการรับมือกับความเสี่ยงที่จะมาเปลี่ยนแปลงสังคมและบริบทของการดำรงชีวิตของมนุษย์ในอนาคต



การเข้ามามีส่วนร่วมของมนุษย์ในการควบคุมการทำงานหรือตัดสินใจของ AI นั้น อาจแบ่งออกได้เป็น 3 ระดับ ได้แก่

1. **Human-in-the-loop** - เป็นระดับการมีส่วนร่วมที่มนุษย์ควบคุมการทำงานหรือตัดสินใจทั้งหมด โดยมี AI ทำหน้าที่ในการให้คำแนะนำหรือข้อมูล แต่ไม่สามารถทำงานหรือตัดสินใจในการดำเนินการใด ๆ ได้ โดยปราศจากมนุษย์

ตัวอย่างเช่น

ในกรณีของการวินิจฉัยโรคนั้น AI จะช่วยวิเคราะห์โอกาสการเป็นโรค พร้อมทั้งนำเสนอข้อมูลที่เป็นประโยชน์แก่แพทย์ในการวินิจฉัยและกำหนดแนวทางการรักษาต่อไป

ซึ่งจากกรณีดังกล่าว หาก AI วินิจฉัยโรคผิดพลาด การวินิจฉัยของแพทย์จะช่วยลดโอกาสและผลกระทบที่อาจนำไปสู่การเสียชีวิต ซึ่งเป็นความเสียหายที่ไม่อาจแก้ไขให้กลับคืนมาเหมือนเดิมได้

2. **Human-over-the-Loop** - เป็นระดับการมีส่วนร่วมที่ AI สามารถทำงานหรือตัดสินใจได้โดยอัตโนมัติ (ไม่ต้องรอการตัดสินใจจากมนุษย์) แต่ยังคงจำเป็นต้องมีมนุษย์ในการกำกับดูแล อีกทั้ง มนุษย์ยังสามารถเข้าควบคุมหรือระงับการทำงานได้ หากพบความผิดพลาดหรือเกิดเหตุการณ์ที่อาจก่อให้เกิดผลกระทบเชิงลบ

ตัวอย่างเช่น

กรณีของระบบควบคุมคุณภาพสินค้านั้น AI สามารถคัดแยกสินค้าที่ไม่ผ่านคุณภาพได้ โดยไม่ต้องรอการตัดสินใจจากมนุษย์ แต่ระบบยังคงอนุญาตให้มนุษย์สามารถระงับการทำงานได้ หากพบว่าระบบไม่สามารถคัดแยกสินค้าที่ไม่ผ่านคุณภาพได้ตามที่กำหนด

โดยในกรณีดังกล่าว การทำงานของ AI ยังคงอยู่ภายใต้การกำกับดูแลของมนุษย์ อีกทั้ง ยังสามารถระงับเหตุเพื่อแก้ไขปัญหาและคัดแยกสินค้าที่ไม่ได้คุณภาพก่อนส่งถึงมือลูกค้าได้

3. **Human-out-of-the-loop** - เป็นระดับการมีส่วนร่วมที่ AI ควบคุมการทำงานหรือตัดสินใจทั้งหมดได้ โดยไม่จำเป็นต้องมีมนุษย์เข้ามาเกี่ยวข้อง อีกทั้งมนุษย์ยังไม่สามารถเข้าควบคุมหรือระงับการทำงานได้

ตัวอย่างเช่น

ในกรณีของระบบช่วยตอบคำถามลูกค้า (Chatbot) และการแนะนำสินค้าบนแพลตฟอร์มขายของออนไลน์นั้น องค์กรมีความประสงค์ที่จะประยุกต์ใช้ AI เพื่อตอบสนองความต้องการของลูกค้าได้อย่างทันก่วงที โดยไม่จำเป็นต้องรอการตัดสินใจจากมนุษย์

โดยทั้ง 2 กรณีข้างต้น มนุษย์อาจไม่สามารถเข้าไปแทรกแซงการทำงานของ AI และไม่สามารถติดต่อกับลูกค้าหรือผู้ซื้อโดยตรงได้ ดังนั้น เพื่อลดผลกระทบจากความไม่พอใจต่อระบบช่วยตอบคำถาม องค์กรจึงควรมีการเปิดให้มีช่องทางอื่นเพื่อให้ลูกค้าสามารถติดต่อพนักงานที่เป็นมนุษย์ได้โดยตรง หรือในกรณีที่ผู้ซื้อไม่สามารถเห็นสินค้าประเภทอื่นที่สนใจได้นั้น องค์กรอาจลดผลกระทบดังกล่าวได้โดย อนุญาตให้ผู้ซื้อปิดฟังก์ชันการนำเสนอสินค้าได้ด้วยตนเอง

จากตัวอย่างของการมีส่วนร่วมของมนุษย์ทั้ง 3 ระดับ จะพบว่าปัจจัยในการพิจารณาเลือกระดับการมีส่วนร่วมของมนุษย์นั้น มีความหลากหลายแตกต่างกันไปตามบริบทของการประยุกต์ใช้ AI ซึ่งองค์กรจำเป็นต้องมีการพิจารณาเลือกระดับการมีส่วนร่วมของมนุษย์ฯ อย่างเหมาะสม โดยอาจพิจารณาจากปัจจัยดังต่อไปนี้ เช่น

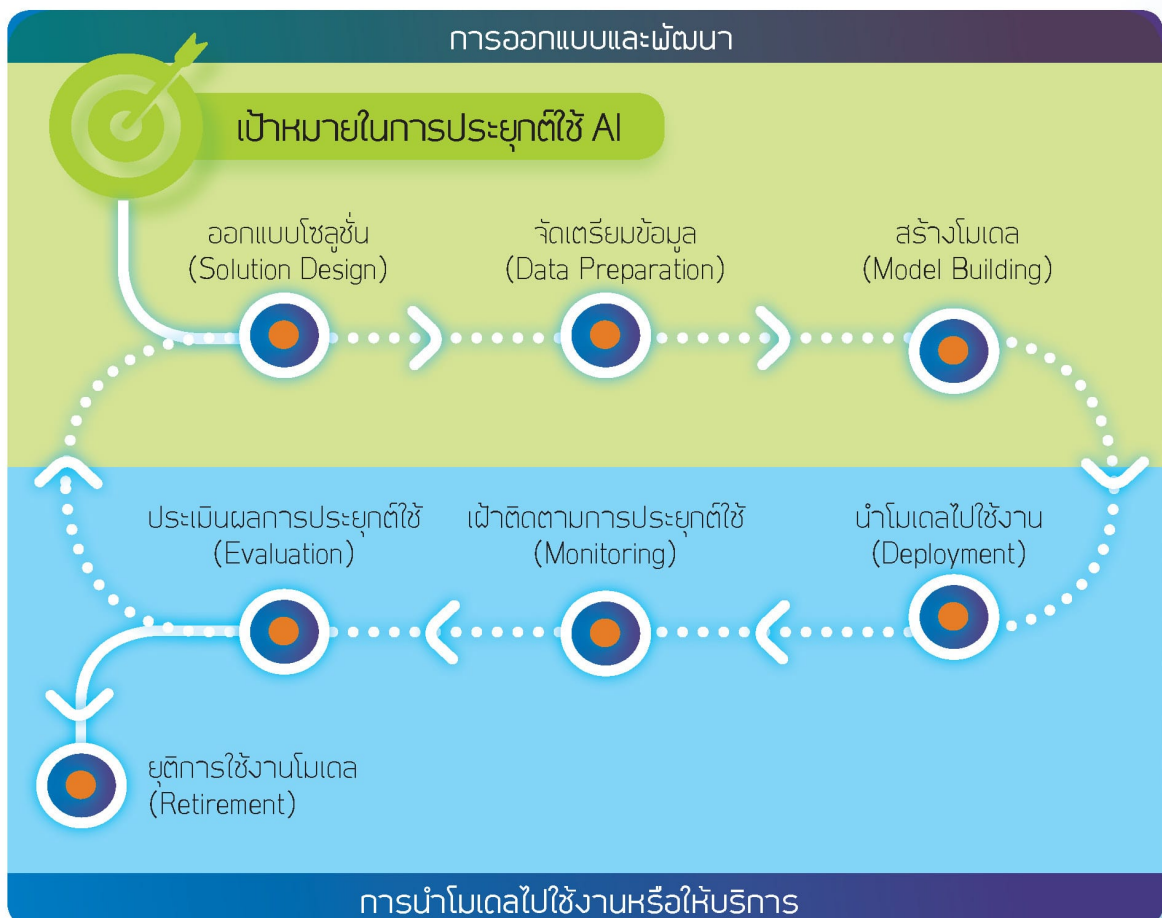
1. ระดับความเสี่ยงที่อาจส่งผลกระทบเชิงลบ
2. ความรุนแรงของผลกระทบเชิงลบที่อาจเกิดขึ้น
3. ผลกระทบที่เกิดขึ้นนั้นสามารถย้อนกลับไปได้แก้ไขได้หรือไม่
4. ความเป็นไปได้หรือความเหมาะสมในการเข้าไปแทรกแซงการทำงานโดยมนุษย์

8. การกำกับดูแลการปฏิบัติงานที่เกี่ยวข้องกับ AI (AI Operation)

8.1. การกำกับดูแลตลอดวงจรชีวิตของ AI (AI Lifecycle)

ความสามารถของ AI ที่หลากหลายนั้นมีเบื้องหลังมาจากความสามารถของโมเดลปัญญาประดิษฐ์ (AI Model) (ซึ่งต่อไปในเอกสารนี้จะเรียกว่า "โมเดล") ที่สร้างขึ้นจากอัลกอริทึมและข้อมูลที่สอนโดยมนุษย์ (เฉพาะในกรณีของ AI ประเภท Machine Learning เท่านั้น) ที่จำเป็นต้องมีการเรียนรู้จากข้อมูลที่สอนโดยมนุษย์) ด้วยเหตุนี้ ประสิทธิภาพในการประมวลผลของ AI เช่น ระดับความถูกต้อง (Accuracy) ความแม่นยำ (Precision) และระดับความน่าเชื่อถือในการตัดสินใจ (Level of Confidence) เป็นต้น จึงขึ้นอยู่กับอัลกอริทึม และคุณภาพของข้อมูลที่ใช้ในการสอน AI เป็นสำคัญ

ด้วยเหตุนี้ การกำกับดูแลการประยุกต์ใช้ AI ตลอดวงจรชีวิต (AI Lifecycle) จึงมีความสำคัญเป็นอย่างยิ่งในการทำให้การประยุกต์ใช้ AI เป็นไปตามเป้าหมายที่กำหนดและมีความรับผิดชอบต่อบุคคลที่เกี่ยวข้ององค์กร และสังคมโดยกว้าง ซึ่งในการกำกับดูแลนั้นควรมีการกำกับดูแลตั้งแต่ระยะเริ่มต้นของการออกแบบและพัฒนาไปจนถึงการนำโมเดลไปใช้งานหรือให้บริการ ซึ่งประกอบด้วย กระบวนการออกแบบโซลูชัน จัดเตรียมข้อมูล สร้างโมเดล นำโมเดลไปใช้งาน ฝ้าติดตาม และประเมินผลการประยุกต์ใช้ ไปจนถึงยุติการใช้งานโมเดล



ออกแบบโซลูชัน (Solution Design)

เพื่อให้การประยุกต์ใช้ AI บรรลุตามเป้าหมายอย่างมีความรับผิดชอบต่อบุคคลที่เกี่ยวข้อง องค์กร และสังคม โดยกว้าง องค์กรจึงจำเป็นต้องวิเคราะห์และนำเป้าหมายในการประยุกต์ใช้ AI หลักการจริยธรรมปัญญาประดิษฐ์ กฎหมายและข้อกำหนดที่เกี่ยวข้อง และมาตรการในการควบคุม และแก้ไขความเสี่ยง มาเปลี่ยนเป็นข้อกำหนดความต้องการในการพัฒนาระบบ AI (AI Requirements) พร้อมทั้งเลือกใช้เครื่องมือหรือผลิตภัณฑ์ที่สามารถตอบสนองความต้องการดังกล่าวได้

แต่เนื่องจากบริบทในการประยุกต์ใช้ AI และความพร้อมด้านบุคลากรที่แตกต่างกันในแต่ละองค์กร การเลือกใช้เครื่องมือหรือผลิตภัณฑ์ในแต่ละองค์กรจึงมีความแตกต่างกัน บางองค์กรอาจเลือกใช้ผลิตภัณฑ์ AI ที่พร้อมใช้งาน (Off-the-shelf AI Product) หรือว่าจ้างหน่วยงานภายนอกมาดำเนินการทั้งหมด แต่ในขณะที่บางองค์กรอาจมีความต้องการที่จะสร้างและนำโมเดลไปใช้งานด้วยบุคลากรของตนเอง จึงทำให้แนวทางในการสร้างและนำโมเดลไปใช้งานในปัจจุบัน อาจแบ่งออกเป็น 4 แนวทางหลัก ได้แก่

1. สร้างและนำโมเดลไปใช้งานด้วยบุคลากรภายในองค์กร

สำหรับองค์กรที่มีความพร้อมด้านบุคลากรในด้าน AI อาจทำการสร้างโมเดลและนำโมเดลไปใช้งานกับระบบหรือบริการด้วยบุคลากรภายในองค์กรเอง โดยในกรณีดังกล่าวองค์กรจำเป็นต้องดำเนินการตั้งแต่กระบวนการสร้าง สอน ตรวจสอบ และทดสอบโมเดล รวมถึงจัดเตรียมข้อมูลสำหรับกระบวนการดังกล่าวด้วยตนเองทั้งหมด ซึ่งแนวทางการพัฒนาและนำโมเดลไปใช้งานในรูปแบบนี้ จะช่วยให้องค์กรได้ระบบหรือบริการที่ตรงตามความต้องการ แต่จำเป็นต้องลงทุนด้านบุคลากรและเวลามากกว่าแนวทางอื่น

2. นำโมเดลในรูปแบบโอเพนซอร์ส (Open-source Model) มาปรับใช้

การนำโมเดลในรูปแบบโอเพนซอร์สมาปรับใช้จะช่วยลดเวลาในการสร้างโมเดล แต่อย่างไรก็ตาม องค์กรยังคงมีหน้าที่ในการสอน ตรวจสอบ และทดสอบโมเดล รวมถึงจัดเตรียม

ข้อมูลสำหรับกระบวนการดังกล่าว อีกทั้งยังต้องทำการทดสอบประสิทธิภาพและความสอดคล้องตามข้อกำหนดความต้องการในการพัฒนาระบบ AI (AI Requirement)

3. เลือกใช้โมเดลที่ได้รับการสอนแล้ว (Pre-trained Model) มาปรับใช้

การเลือกใช้โมเดลที่ได้รับการสอนแล้ว (Pre-trained Model) จากผู้ให้บริการโซลูชันด้าน AI หรือหน่วยงานภายนอกมาปรับใช้นั้น เป็นอีกหนึ่งแนวทางที่ช่วยลดเวลาและภาระในกระบวนการสร้าง สอน และตรวจสอบโมเดล แต่อย่างไรก็ตามองค์กรยังคงมีหน้าที่ในการจัดเตรียมข้อมูล รวมถึงทดสอบประสิทธิภาพและความสอดคล้องตามข้อกำหนดความต้องการในการพัฒนาระบบ AI (AI Requirement)

4. เลือกใช้ผลิตภัณฑ์ AI ที่พร้อมใช้งาน (Off-the-shelf AI Product)

การเลือกใช้ผลิตภัณฑ์ AI ที่พร้อมใช้งานในท้องตลาดจะช่วยลดเวลาและความซับซ้อนในการพัฒนาระบบ อีกทั้งยังช่วยลดภาระขององค์กรในการสร้างบุคลากรด้าน AI แต่อย่างไรก็ตาม องค์กรยังคงมีความจำเป็นที่จะต้องจัดเตรียมข้อมูลและดำเนินการทดสอบเช่นเดียวกันกับแนวทางข้างต้น

จากแนวทางการดำเนินการที่ 2-4 จะพบว่าองค์กรมีการใช้โมเดลหรือผลิตภัณฑ์ AI ที่สร้างโดยหน่วยงานภายนอก รวมถึงอาจมีว่าจ้างหน่วยงานภายนอกเพื่อนำโมเดลหรือผลิตภัณฑ์ AI มาปรับใช้ร่วมกับกระบวนการหรือขั้นตอนปฏิบัติต่าง ๆ ให้แก่องค์กร

ในกรณีที่มีความเสียหายหรือผลกระทบเชิงลบซึ่งเกิดความผิดพลาดของหน่วยงานภายนอก รวมถึงความผิดพลาดจากการทำงานของโมเดลหรือผลิตภัณฑ์ AI (ตามแนวทางที่ 2-4) นั้น องค์กรยังคงต้องรับผิดชอบต่อความเสียหายหรือผลกระทบเชิงลบเช่นเดียวกันกับการดำเนินการตามแนวทางการดำเนินการที่ 1 ด้วยเหตุนี้ องค์กรจึงจำเป็นต้องมีการกำกับดูแลการปฏิบัติงานของหน่วยงานภายนอกให้สอดคล้องตามนโยบายและขั้นตอนปฏิบัติขององค์กร รวมถึงจัดให้มีการทดสอบอย่างเหมาะสม เพื่อให้มั่นใจได้ว่าโมเดลที่ใช้งานมีประสิทธิภาพและสอดคล้องตามข้อกำหนดความต้องการในการพัฒนาระบบ AI (AI Requirement)

จัดเตรียมข้อมูล (Data Preparation)

เนื่องจาก AI จำเป็นต้องใช้ข้อมูลในการเรียนรู้และประมวลผลในการทำงาน ด้วยเหตุนี้ องค์กรจึงควรมีการกำหนดคุณสมบัติและคุณภาพของข้อมูลที่เหมาะสม เพื่อให้ AI สามารถทำงานได้อย่างมีประสิทธิภาพตามเป้าหมายที่กำหนด โดยในการกำหนดคุณสมบัติและคุณภาพนั้น อาจพิจารณาจากประเภทของข้อมูลที่เป็นต้องใช้ในการประมวลผล ขนาดของข้อมูล ความถูกต้องของข้อมูล ข้อมูลมีการจัดเก็บภายในช่วงเวลาที่ต้องการ (Time Series Data) มีการปรับปรุงข้อมูลเป็นประจำสม่ำเสมอ เป็นต้น อีกทั้ง องค์กรควรมีการสรรหาแหล่งข้อมูลที่น่าเชื่อถือและทำการเชื่อมโยงข้อมูลดังกล่าว เพื่อนำมาใช้งานร่วมกับ AI

นอกจากนี้ องค์กรยังควรมีการพิจารณาคุณสมบัติของข้อมูลในด้านอื่นเพื่อลดการเกิดอคติที่มาจากข้อมูล (Data Bias) เช่น ข้อมูลมีความหลากหลาย ข้อมูลมีความครอบคลุมทุกกลุ่มประชากรโดยไม่เอนเอียงไปยังกลุ่มใดกลุ่มหนึ่ง ขนาดของกลุ่มตัวอย่างสามารถสะท้อนหรือเป็นตัวแทนของประชากรได้อย่างเหมาะสมผล เป็นต้น รวมถึงพิจารณาความน่าเชื่อถือของการเก็บรวบรวมข้อมูล เพื่อลดการเกิดอคติที่เกิดจากความคลาดเคลื่อนในการวัดหรือความผิดพลาดจากเครื่องมือวัด (Measurement Bias)

โดยภายหลังจากทำการเชื่อมโยงและรวบรวมข้อมูลจากแหล่งต่าง ๆ (Data Integration) เรียบร้อยแล้ว ไม่ว่าจะจะเป็นแหล่งที่มาจากภายในหรือภายนอกองค์กร ควรมีการจัดทำเอกสารเพื่อแสดงแหล่งที่มาของข้อมูล (Data Provenance) รวมถึงรายละเอียดการประมวลผลหรือการเปลี่ยนแปลงใด ๆ ที่เกิดขึ้นกับข้อมูลตลอดกระบวนการจัดเตรียมข้อมูล (Data Lineage) เพื่อประโยชน์ในการตรวจสอบย้อนกลับ (Traceability) ในกรณีที่ AI ทำงานผิดพลาด



เป้าหมายในการประยุกต์ใช้ AI



กำหนดคุณสมบัติของข้อมูล และคุณภาพของข้อมูลที่ต้องการ



เชื่อมโยงและประมวลผลข้อมูล จากแหล่งข้อมูลที่น่าเชื่อถือ



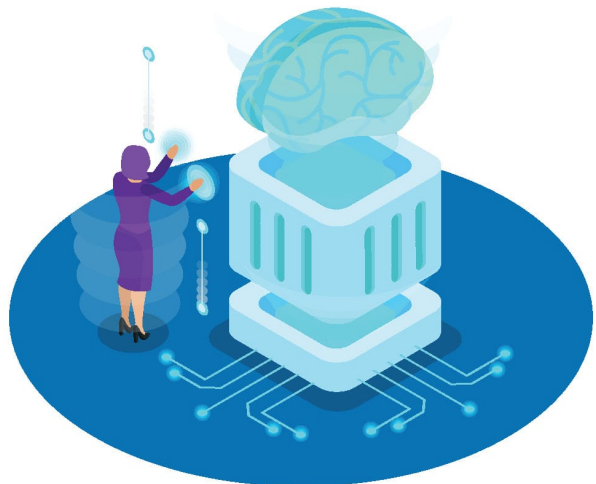
ระบุประเภทข้อมูล



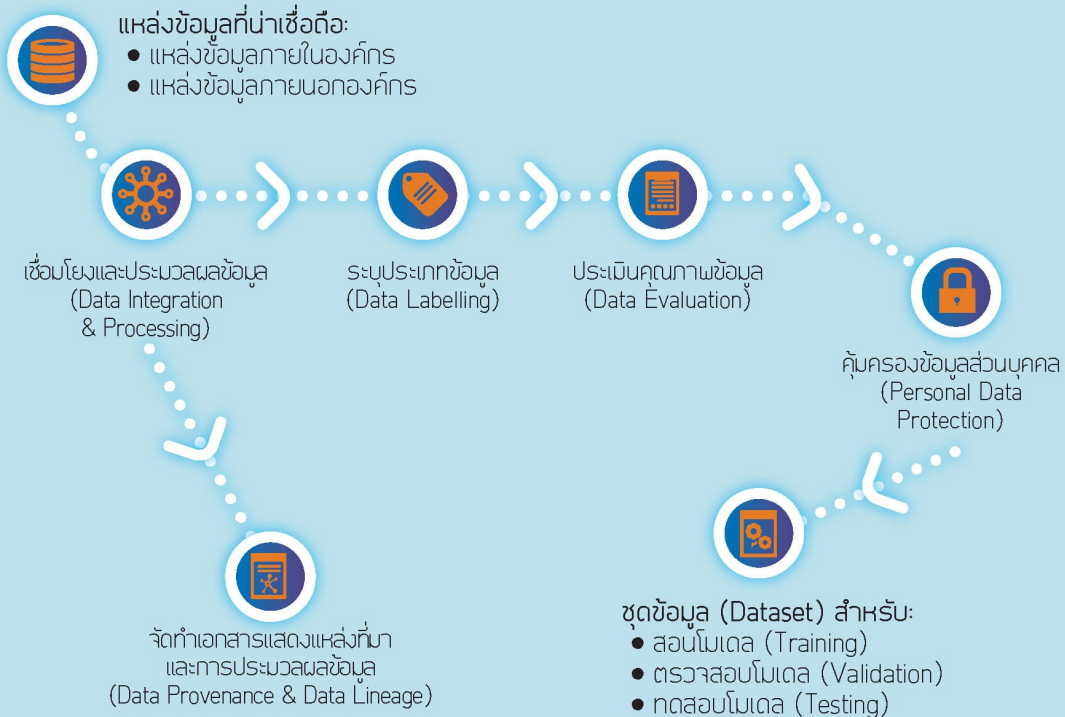
ประเมินคุณภาพข้อมูล



คุ้มครองข้อมูลส่วนบุคคล



ขั้นตอนการดำเนินงานเพื่อจัดเตรียมข้อมูลสำหรับทำงานร่วมกับ AI (Data Pipeline)



เนื่องจากการสร้างโมเดลจำเป็นต้องใช้ข้อมูลในการดำเนินการ โดยเฉพาะอย่างยิ่ง AI ประเภท Supervised Machine Learning จำเป็นต้องเรียนรู้จากข้อมูลที่สอนโดยมนุษย์ เพื่อให้สามารถวิเคราะห์คาดการณ์ให้คำแนะนำ ตัดสินใจ หรือดำเนินการใด ๆ แทนมนุษย์ได้ ด้วยเหตุนี้ ในการจัดเตรียมข้อมูลสำหรับการสอน ตรวจสอบ และทดสอบโมเดล จำเป็นต้องมีการระบุประเภทของข้อมูล (Data Labelling) อย่างถูกต้อง เพื่อให้ AI มีข้อมูลและสามารถรับรู้ความหมายของข้อมูลได้อย่างถูกต้อง ซึ่งในการระบุประเภทของข้อมูลสามารถทำได้ทั้งในรูปแบบรูปภาพ ข้อความ และเสียง ตัวอย่างเช่น การระบุรูปคน สัตว์ และสิ่งของต่าง ๆ เพื่อให้ AI สามารถเข้าใจรูปภาพหรือวิดีโอในแบบเดียวกับที่มนุษย์มองเห็น หรือการระบุอารมณ์และความรู้สึกของแต่ละข้อความหรือเสียง เพื่อให้ AI สามารถเข้าใจได้ว่าผู้ที่แสดงความคิดเห็นดังกล่าวมีความรู้สึกอย่างไร เป็นต้น

โดยก่อนที่จะนำชุดข้อมูล (Dataset) ไปใช้สำหรับการสอน ตรวจสอบ และทดสอบ องค์กรจำเป็นต้องมีการประเมินคุณภาพของข้อมูล (Data Evaluation) และ

อาจทำการปรับปรุงข้อมูลให้มีคุณภาพตามหลักเกณฑ์ที่องค์กรกำหนด เช่น ข้อมูลมีความถูกต้อง (Accuracy) เนื้อหาข้อมูลครบถ้วน (Completeness) ตรงตามความต้องการ (Relevancy) มีรูปแบบของข้อมูล (Data Format) ตามที่กำหนด และมีการปรับปรุงให้เป็นปัจจุบัน (Timeliness) เป็นต้น เพื่อให้มั่นใจได้ว่า AI มีความสามารถและประสิทธิภาพเป็นไปตามเป้าหมายที่กำหนด

นอกจากนี้ ในกรณีที่มีการใช้ข้อมูลส่วนบุคคลในการประมวลผล รวมถึงกรณีที่มีการประมวลผลข้อมูลที่มีความอ่อนไหว (Sensitive Data) ตามกฎหมายคุ้มครองข้อมูลส่วนบุคคล (เช่น เชื้อชาติ เผ่าพันธุ์ ข้อมูลสุขภาพ ข้อมูลชีวภาพ เป็นต้น) องค์กรควรมีการเก็บรักษาข้อมูลส่วนบุคคลอย่างมั่นคงปลอดภัยและจัดให้มีมาตรการในการคุ้มครองข้อมูลส่วนบุคคลที่เหมาะสม เช่น การควบคุมการเข้าถึง (Access Control) การเข้ารหัสลับข้อมูล (Encryption) การทำข้อมูลส่วนบุคคลให้เป็นข้อมูลนิรนาม (Anonymization) เป็นต้น เพื่อลดความเสี่ยงที่จะทำให้ข้อมูลส่วนบุคคลรั่วไหล (Data Breach) และป้องกันผลกระทบที่อาจเกิดขึ้นจากการรั่วไหลของข้อมูล

สร้างโมเดลปัญญาประดิษฐ์ (Model Building)

การสร้างโมเดลนั้น อาจประกอบด้วยขั้นตอนพื้นฐาน 4 ขั้นตอน โดยเริ่มจากการจัดเตรียมอัลกอริทึมสำหรับสร้างโมเดล ซึ่งในขั้นตอนนี้องค์กรอาจพัฒนาอัลกอริทึมขึ้นเองหรืออาจนำอัลกอริทึมที่มีการเปิดเผยซอร์สโค้ด (Open Source) มาใช้งาน โดยในกรณีที่องค์กรสร้างโมเดลประเภท Machine Learning นั้น องค์กรจำเป็นต้องทำการสอนโมเดลด้วยชุดข้อมูลสำหรับสอนโมเดล (Training Dataset) จากนั้นจึงนำโมเดลไปตรวจสอบประสิทธิภาพ โดยใช้ชุดข้อมูลสำหรับตรวจสอบโมเดล (Validation Dataset) ซึ่งหาก AI ยังไม่สามารถทำงานได้ตามประสิทธิภาพที่กำหนด โมเดลดังกล่าวจะต้องถูกปรับแต่ง (Tuning) และนำกลับไปตรวจสอบจนกว่าประสิทธิภาพในการทำงานจะเป็นไปตามเป้าหมายที่กำหนด โดยโมเดลที่ดีที่สุดจะถูกนำไปทดสอบด้วยชุดข้อมูลสำหรับทดสอบโมเดล (Testing Dataset) ในขั้นตอนนี้สุดท้ายก่อนนำไปใช้งาน (Deployment) ซึ่งชุดข้อมูลสำหรับทดสอบโมเดลนี้ จะต้องเป็นชุดข้อมูลที่ไม่เคยถูกใช้ในขั้นตอนการสอนและตรวจสอบโมเดลมาก่อน



เป้าหมายในการประยุกต์ใช้ AI



สร้างโมเดลปัญญาประดิษฐ์



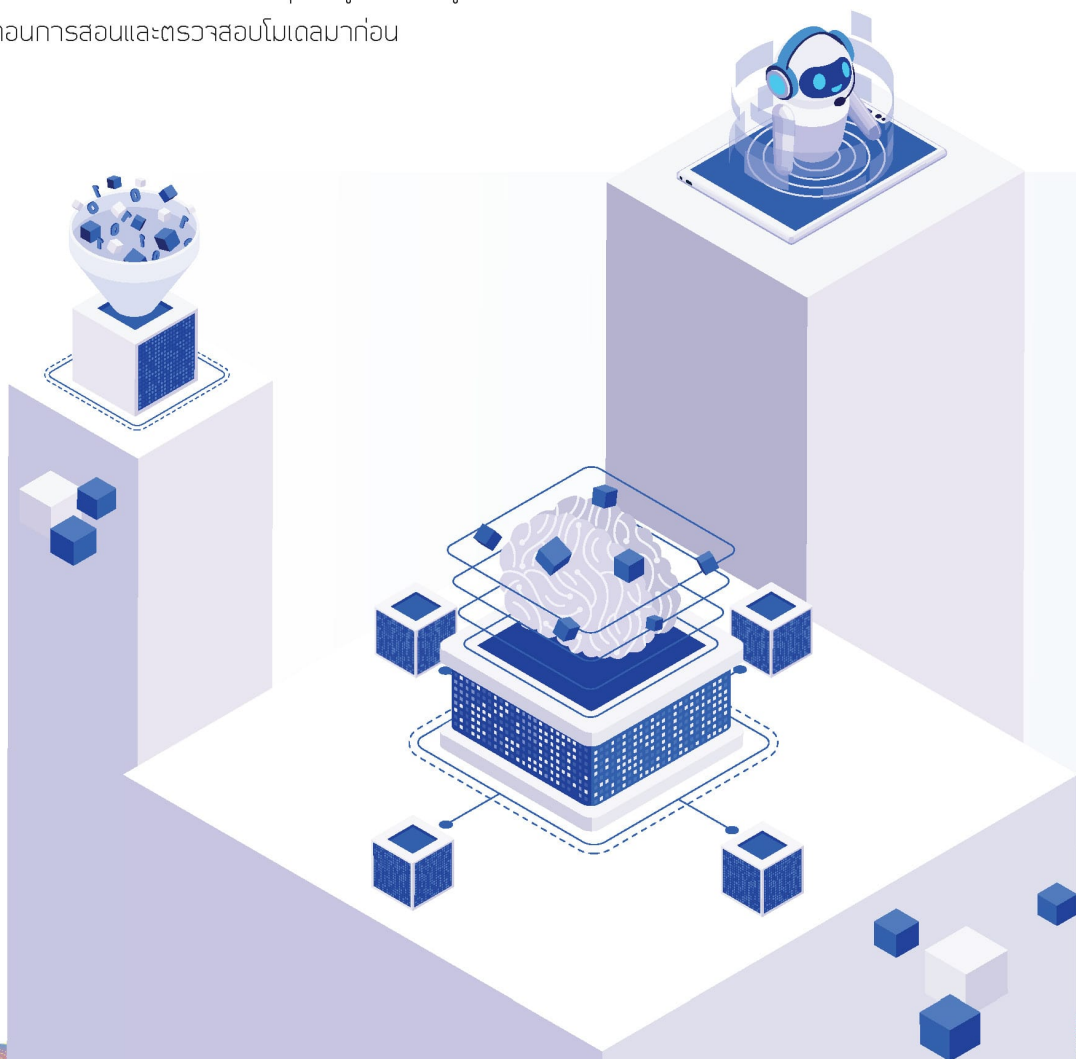
สอนโมเดลปัญญาประดิษฐ์

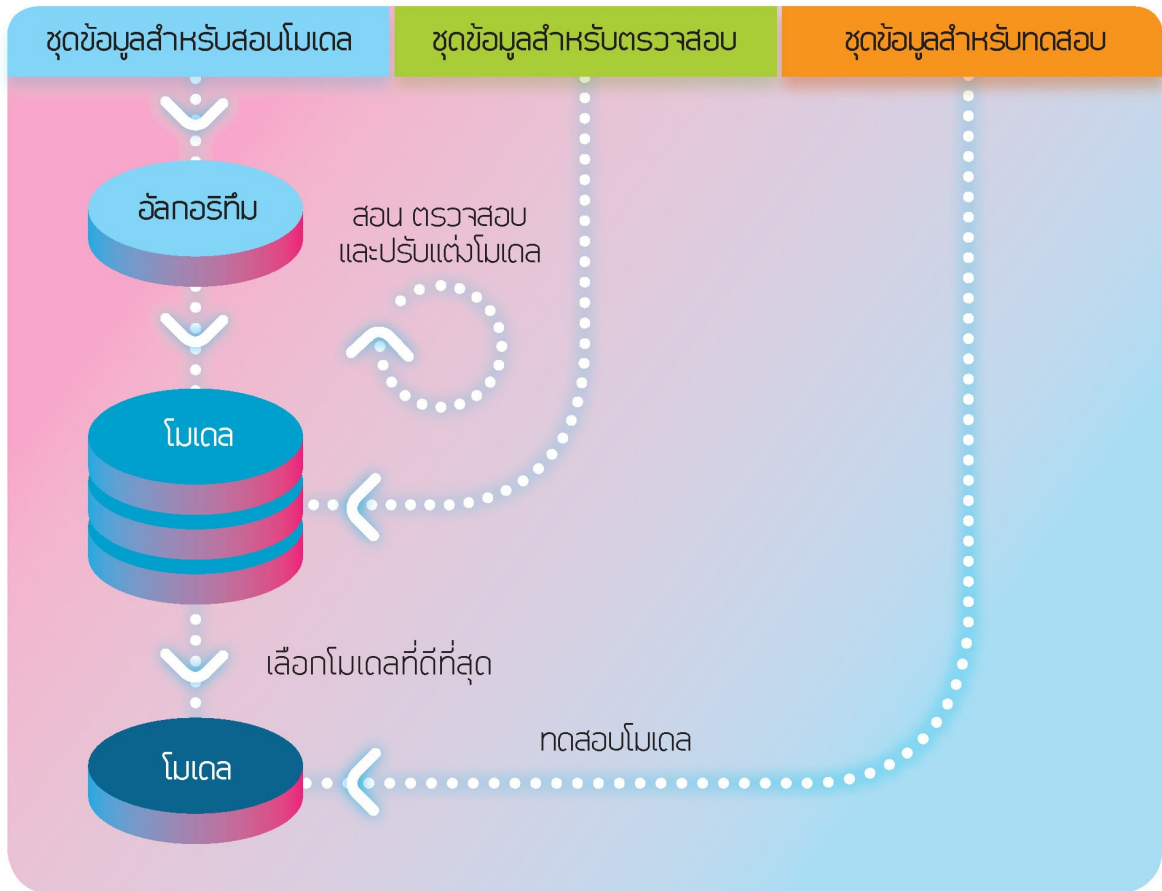


ตรวจสอบและเลือกโมเดลปัญญาประดิษฐ์



ทดสอบโมเดลปัญญาประดิษฐ์





ภาพขั้นตอนพื้นฐานในการสร้างโมเดลปัญญาประดิษฐ์ประเภท Machine Learning ¹

สำหรับองค์กรที่เลือกใช้โมเดลจากผู้ให้บริการโซลูชันด้าน AI (AI Solution Provider) ถึงแม้ว่าองค์กรจะไม่จำเป็นต้องทำการสร้างโมเดลด้วยตนเอง แต่องค์กรยังคงมีความรับผิดชอบ (Responsibility) ในการทดสอบว่าโมเดลสามารถทำงานได้อย่างมีประสิทธิภาพและสามารถบรรลุตามเป้าหมายที่กำหนด พร้อมทั้ง พิจารณาและตรวจสอบความเหมาะสมของกระบวนการสร้างโมเดล เช่น ขั้นตอนและข้อมูลที่เกี่ยวข้องกับกระบวนการสร้าง สอน และทดสอบโมเดล รวมถึงมาตรการต่าง ๆ ที่เกี่ยวข้อง เป็นต้น

นอกเหนือจากการควบคุมกระบวนการสร้างโมเดลข้างต้นเพื่อให้ AI มีประสิทธิภาพและสามารถบรรลุเป้าหมายตามที่กำหนด องค์กรยังจำเป็นต้องจัดให้มีมาตรการควบคุมเพื่อให้มั่นใจได้ว่า การทำงานของ AI สอดคล้องตามหลักการจริยธรรมปัญญาประดิษฐ์ รวมถึงกฎหมายและข้อกำหนดที่เกี่ยวข้อง

โดยในการพิจารณาว่าการทำงานของ AI ควรสอดคล้องตามหลักการจริยธรรมปัญญาประดิษฐ์ใดนั้น องค์กรควรพิจารณาจากบริบทของการนำ AI ไปประยุกต์ใช้ ความเสี่ยง และผลกระทบที่อาจเกิดขึ้น โดยไม่จำเป็นต้องดำเนินการให้สอดคล้องกับทุกหลักการจริยธรรมปัญญาประดิษฐ์ที่กล่าวถึงในเอกสารนี้

¹ Adapt from "Train, Validate, and Test" PRIMO.ai, accessed April 13, 2023, http://primo.ai/index.php?title=Train,_Validate,_and_Test

ความเท่าเทียม หลากหลาย ครอบคลุม และเป็นธรรม (Fairness)

ผลลัพธ์จากการทำงานของ AI ไม่ว่าจะอยู่ในรูปแบบของการวิเคราะห์คาดการณ์ การให้คำแนะนำ การตัดสินใจ หรือการดำเนินการใด ๆ นั้น อาจส่งผลกระทบต่อความไม่เป็นธรรม (Unfairness) จนนำไปสู่การเลือกปฏิบัติ (Discrimination) ซึ่งผลกระทบดังกล่าวนี้อาจเกิดจากอคติที่มาจากข้อมูล (Data Bias) หรืออคติที่มาจากการออกแบบและสร้างโมเดล (Bias Introduced by Engineering Decisions)

เพื่อป้องกันความไม่เป็นธรรมซึ่งเกิดจากอคติที่มาจากการออกแบบและสร้างโมเดล (Bias Introduced by Engineering Decisions) นั้น บุคลากรที่เกี่ยวข้องกับกระบวนการสร้างและการทดสอบโมเดลจึงควรมีความรู้ และมีความหลากหลาย รวมถึงจัดให้มีผู้เชี่ยวชาญในด้านที่เกี่ยวข้องกับการประยุกต์ใช้ AI คอยให้คำปรึกษา เพื่อให้องค์กรมองเห็นอคติที่อาจเกิดขึ้นและกำหนดกลยุทธ์ในการลดอคติดังกล่าว

สำหรับอคติที่มาจากข้อมูล (Data Bias) นั้น อาจเกิดจากการมีข้อมูลที่เอนเอียง ไม่ครอบคลุมทุกกลุ่มประชากร หรือมีขนาดของกลุ่มตัวอย่างที่ไม่เหมาะสม ทำให้เกิดความไม่เป็นธรรม (ซึ่งเรียกว่า "Selection Bias") ตัวอย่างเช่น AI คัดเลือกใบสมัครจากผู้สมัครงานเพศชายมากกว่าเพศหญิง เนื่องจากข้อมูลส่วนใหญ่ที่ใช้สอนโมเดลนั้นเป็นข้อมูลของเพศชาย จึงทำให้ AI ให้น้ำหนักในการคัดเลือกผู้สมัครงานเพศชายเพื่อเข้ารับการสัมภาษณ์มากกว่าผู้สมัครเพศหญิง เป็นต้น อีกทั้งในบางกรณีอาจเกิดจากอคติที่มาจากความคลาดเคลื่อนในการวัดหรือความผิดพลาดจากเครื่องมือวัด (Measurement Bias) เช่น อุปกรณ์วัดระดับมลพิษทางอากาศบางพื้นที่แจ้งระดับมลพิษต่ำกว่าความเป็นจริง ทำให้พื้นที่ดังกล่าวไม่ได้รับการแก้ไขปัญหามา จนนำไปสู่ปัญหาสุขภาพของประชาชนในพื้นที่ เป็นต้น

องค์กรอาจพิจารณาเลือกใช้มาตรการดังต่อไปนี้เพื่อลดอคติที่มาจากข้อมูล ตัวอย่างเช่น



1. เลือกใช้ข้อมูลจากแหล่งข้อมูลที่มีความน่าเชื่อถือ และมีการเก็บรวบรวมข้อมูลอย่างถูกต้อง
2. ข้อมูลมีความหลากหลายและครอบคลุมทุกกลุ่มประชากร อีกทั้งมีการรวบรวมมาจากแหล่งข้อมูลที่หลากหลาย (หากเป็นไปได้)
3. ขนาดของกลุ่มตัวอย่างสามารถสะท้อนหรือเป็นตัวแทนของประชากรได้อย่างสมเหตุสมผล
4. ระมัดระวังในการเลือกใช้ข้อมูลเฉพาะกลุ่มประชากรกลุ่มใดกลุ่มหนึ่ง รวมถึงข้อมูลที่มีความอ่อนไหว (Sensitive Data) ตัวอย่างเช่น เพศ เชื้อชาติ ศาสนา ประวัติการรักษายาเสพติด อาชญากรรม เป็นต้น
5. พิจารณากำหนดชุดข้อมูลสำหรับทดสอบแต่ละชุด ให้มีความแตกต่างกันของกลุ่มประชากร เพื่อตรวจสอบว่า AI ยังคงตัดสินใจได้อย่างถูกต้องไม่เอนเอียงไปตามกลุ่มประชากรกลุ่มใดกลุ่มหนึ่ง
6. อาจพิจารณาสุ่มข้อมูลที่คาดว่าจะส่งผลให้เกิดอคติเพื่อค้นหาความผิดพลาด พร้อมทั้งเก็บรวบรวมข้อมูลดังกล่าวสำหรับใช้ในครั้งต่อไป

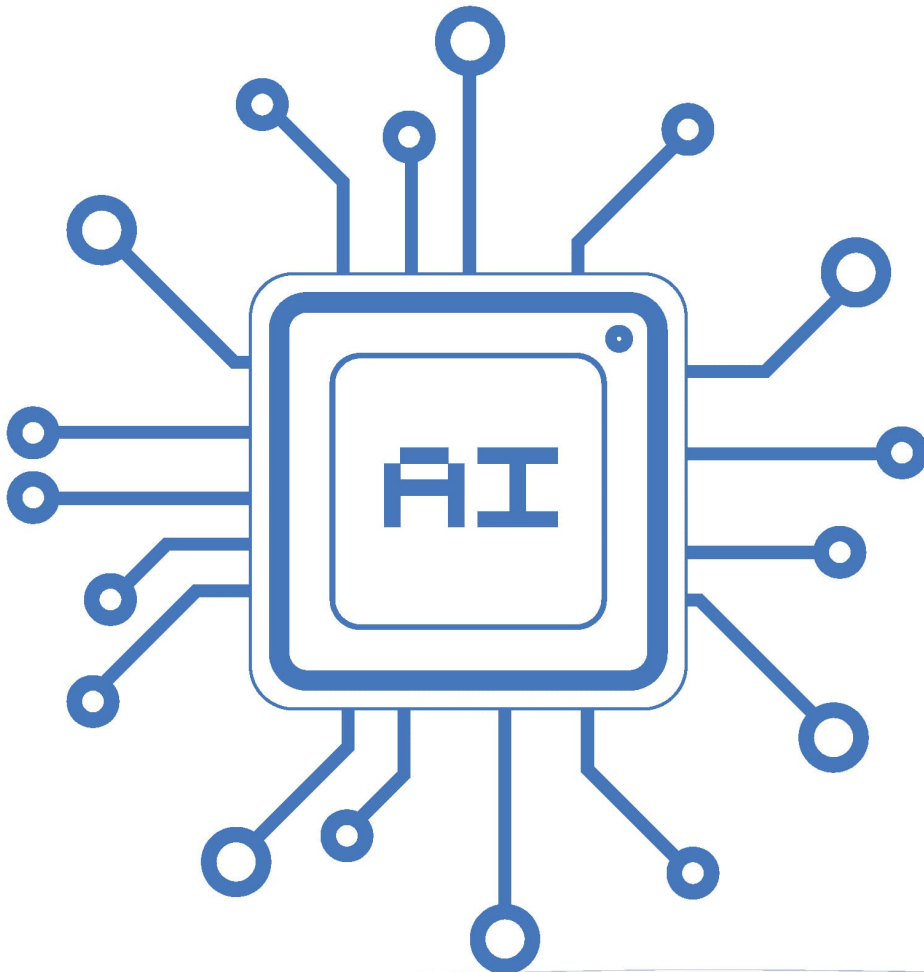
ความโปร่งใส (Transparency)

ความสามารถในการอธิบายรายละเอียดเกี่ยวกับ AI (Explainability) เป็นหนึ่งในมาตรการสำคัญที่จะช่วยให้เกิดความโปร่งใสในการประยุกต์ใช้ AI โดยองค์กรจำเป็นต้องมีการจัดทำเอกสารเพื่ออธิบายพฤติกรรมการทำงานของ AI และเหตุผลเบื้องหลังการทำงาน รวมถึงอธิบายกระบวนการในการสร้างและทดสอบโมเดล ซึ่งจะช่วยให้ผู้ที่เกี่ยวข้องรับทราบและเข้าใจถึงกระบวนการทำงานของ AI อีกทั้งยังเป็นการสร้างความเชื่อมั่นและการยอมรับจากบุคคลที่เกี่ยวข้องอีกด้วย

นอกจากนี้ เพื่อให้เกิดความโปร่งใสในการประยุกต์ใช้ AI องค์กรควรมีการดำเนินการเพื่อให้สามารถตรวจสอบย้อนกลับ (Traceability) ในกรณีที่พบความผิดพลาดหรือความผิดพลาดได้

โดยทำการเก็บรวบรวมข้อมูลและรายละเอียดที่เกี่ยวข้องกับกระบวนการในการสร้างและทดสอบโมเดล เช่น

- แหล่งที่มาของข้อมูล (Data Provenance)
- การประมวลผลหรือการเปลี่ยนแปลงใด ๆ ที่เกิดขึ้นกับข้อมูลตลอดกระบวนการจัดเตรียมข้อมูล (Data Lineage)
- รายละเอียดการออกแบบและการทำงานของอัลกอริทึม
- ชุดข้อมูลที่ใช้และผลลัพธ์ในการสอน ตรวจสอบ และทดสอบโมเดล
- ผลการประเมินความเสี่ยงและมาตรการการรับมือ



ความน่าเชื่อถือ (Reliability)

เพื่อให้การทำงานของ AI มีความน่าเชื่อถือ องค์กรควรมีการกำหนดมาตรการและควบคุมกระบวนการสร้างและทดสอบโมเดล เพื่อให้มั่นใจได้ว่าโมเดลสามารถสร้างผลลัพธ์ได้เหมือนเดิมทุกครั้ง (Reproducibility) เมื่ออยู่ภายใต้สถานการณ์หรือได้รับข้อมูลที่เหมือนกัน อีกทั้งยังสามารถทนทาน (Robustness) ต่อเหตุการณ์ที่อาจเกิดความผิดพลาดหรือสร้างผลกระทบเชิงลบได้

โดยในการทดสอบความทนทานต่อเหตุการณ์ที่อาจเกิดความผิดพลาดหรือสร้างผลกระทบเชิงลบนั้น องค์กรจำเป็นต้องทำการทดสอบโดยใช้ข้อมูลที่อยู่นอกเหนือจากขอบเขตที่ออกแบบไว้ (Exceptions) รวมถึงทำการทดสอบภายใต้สถานการณ์หรือข้อมูลที่ไม่เคยถูกสอนมาก่อน เพื่อให้มั่นใจได้ว่า AI ยังคงสามารถทำงานได้อย่างถูกต้องตามที่ออกแบบไว้ หรือสามารถรับมือกับเหตุการณ์ที่เกิดขึ้นโดยไม่ส่งผลกระทบเชิงลบต่อบุคคล องค์กร และสังคมโดยกว้าง

นอกจากนี้ในการทดสอบความทนทาน องค์กรยังจำเป็นต้องเฝ้าติดตามและนำผลลัพธ์การทำงานของ AI มาปรับปรุงอย่างต่อเนื่อง เพื่อให้ AI สามารถทำงานได้อย่างมีประสิทธิภาพภายใต้สภาพแวดล้อมการใช้งานจริงและอยู่ในขอบเขตที่ยอมรับได้





ความมั่นคงปลอดภัยและความเป็นส่วนตัว (Security and Privacy)

ในด้านการรักษาความมั่นคงปลอดภัยและการคุ้มครองความเป็นส่วนตัวนั้น การปฏิบัติตามแนวปฏิบัติที่ดีในการพัฒนาซอฟต์แวร์ (เช่น การควบคุมเวอร์ชัน การทดสอบ และการควบคุมคุณภาพของซอร์สโค้ด เป็นต้น) อาจจะไม่เพียงพอต่อการรักษาความมั่นคงปลอดภัยและคุ้มครองความเป็นส่วนตัว เนื่องจาก AI อาจมีช่องโหว่ที่เปิดโอกาสให้ผู้ประสงค์ร้ายโจมตีข้อมูลที่ใช้ในการสอนโมเดลและโจมตีที่โมเดลโดยตรง ตัวอย่างเช่น

- Poisoning Attack เป็นการโจมตีโดยทำให้ข้อมูลที่ใช้ในการสอนโมเดลปนเปื้อนด้วยข้อมูลที่ทำให้โมเดลเกิดช่องโหว่ (Adversarial Manipulation) เพื่อให้ผู้ประสงค์ร้าย (Attacker) สามารถบรรลุเป้าหมายที่ต้องการ
- Evasion Attack เป็นการโจมตีโมเดลโดยตรงโดยส่งข้อมูลที่ทำให้โมเดลประมวลผลแล้วทำงานผิดพลาดหรือหยุดการทำงาน (Adversarial Input)

เพื่อป้องกันภัยคุกคามข้างต้นจึงจำเป็นต้องมีความร่วมมือกันระหว่างนักพัฒนา AI (AI Developer) นักวิทยาศาสตร์ข้อมูล (Data Scientist) นักวิเคราะห์ข้อมูล

(Data Analyst) ผู้เชี่ยวชาญด้านความมั่นคงปลอดภัย รวมถึงผู้มีส่วนได้ส่วนเสียที่เกี่ยวข้อง ในการจัดทำมาตรการเพื่อป้องกันภัยคุกคามที่อาจเกิดขึ้น โดยปฏิบัติตามมาตรฐานหรือแนวทางการรักษาความมั่นคงปลอดภัย เช่น มาตรฐานที่ประกาศโดย The ETSI Industry Specification Group on Securing Artificial Intelligence (ISG SAI) และมาตรฐาน ISO/IEC 27090 Cybersecurity – Artificial Intelligence – Guidance for addressing security threats and failures in artificial intelligence systems ซึ่งที่กำลังจะประกาศใช้ โดยหน่วยงาน International Organization for Standardization (ISO) นอกจากนี้ควรติดตามข่าวสารเกี่ยวกับภัยคุกคามจากแหล่งข้อมูลต่าง ๆ เช่น The Open Worldwide Application Security Project (OWASP) และ Adversarial Threat Landscape for Artificial-Intelligence System (MITRE ATLAS)

นอกจากนี้ ในการคุ้มครองความเป็นส่วนตัว ควรมีการป้องกันการเข้าถึงโดยไม่ได้รับอนุญาต (Unauthorized Access) รวมถึงจัดให้มีมาตรการในการคุ้มครองข้อมูลส่วนบุคคลที่เหมาะสม ซึ่งได้กล่าวมาแล้วในหัวข้อการจัดเตรียมข้อมูล (Data Preparation)

นำโมเดลไปใช้งาน (Deployment)

ก่อนนำโมเดลไปใช้งานจริง องค์กรควรมีการนำโมเดลไปติดตั้งและทดสอบการทำงานภายใต้สภาพแวดล้อมเสมือนจริง (Pre-production Environment) เพื่อทดสอบการทำงานร่วมกับกระบวนการหรือขั้นตอนปฏิบัติงานที่เกี่ยวข้อง พร้อมทั้งทดสอบความสามารถในการรองรับปริมาณการใช้งาน (Workload) และความสามารถในการตอบสนองภายในระยะเวลาที่ยอมรับได้ (Acceptable Latency Time) ซึ่งการติดตั้งและทดสอบในสภาพแวดล้อมดังกล่าว จะช่วยให้องค์กรพบปัญหา และสามารถป้องกันปัญหาที่อาจเกิดขึ้นก่อนนำโมเดลไปใช้งานจริง อีกทั้งยังเป็นการสร้างความน่าเชื่อถือ (Reliability) ให้แก่ระบบหรือบริการ

นอกจากนี้ ก่อนนำโมเดลไปใช้งานจริงองค์กรควรมีการออกแบบสถาปัตยกรรมที่เกี่ยวข้องกับการประยุกต์ใช้ AI (AI Architecture) ที่มีประสิทธิภาพและสามารถรองรับปริมาณการใช้งานได้อย่างเพียงพอ เพื่อให้การประยุกต์ใช้ AI บรรลุตามเป้าหมายที่กำหนด พร้อมทั้งอาจพิจารณาออกแบบสถาปัตยกรรมให้มีความยืดหยุ่น เพื่อรองรับความต้องการที่เพิ่มขึ้นหรือการขยายขอบเขตการใช้งานในอนาคตได้

เพื่อให้องค์กรมีความพร้อมในการทำงานหรือให้บริการอย่างต่อเนื่อง องค์กรจึงควรมีการบริหารจัดการความเปลี่ยนแปลง (Change Management) เพื่อลดผลกระทบที่อาจเกิดขึ้นจากการนำโมเดลไปใช้งานร่วมกับกระบวนการหรือขั้นตอนปฏิบัติงานจริง นอกจากนี้ในการบริหารจัดการความเปลี่ยนแปลงดังกล่าว ยังจำเป็นต้องครอบคลุมถึงกรณีที่มีการปรับปรุงแก้ไขโมเดลและการยุติการใช้งานโมเดลด้วยเช่นกัน เพื่อไม่ให้เกิดการเปลี่ยนแปลงใด ๆ ที่เกี่ยวข้องกับโมเดลส่งผลกระทบต่อการทำงานหรือการให้บริการ



เป้าหมายในการประยุกต์ใช้ AI



ทดสอบการทำงานของโมเดลใน Pre-production Environment



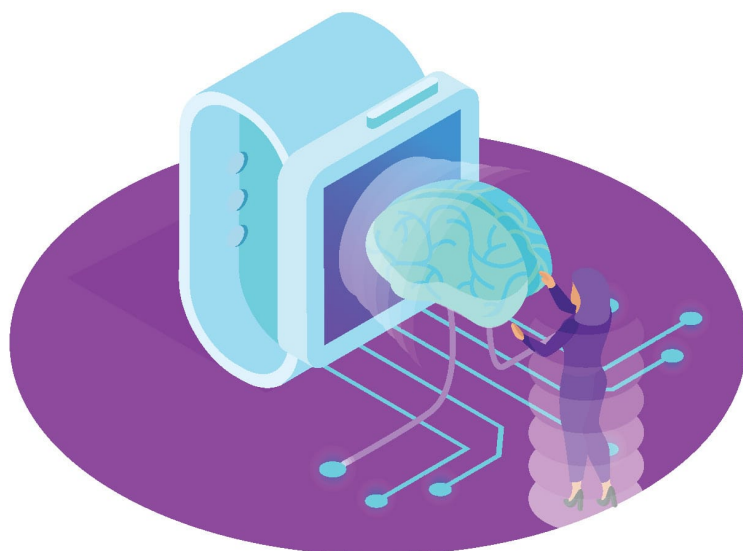
บริหารจัดการความเปลี่ยนแปลง (Change Management)



รักษาความมั่นคงปลอดภัย และคุ้มครองข้อมูลส่วนบุคคล



จัดให้มีมาตรการรับมือ ประเด็นปัญหาและเหตุการณ์ผิดปกติ





เนื่องจาก AI เป็นเทคโนโลยีสารสนเทศประเภทหนึ่งที่ต้องทำงานร่วมกับระบบเทคโนโลยีสารสนเทศที่หลากหลาย เช่น ระบบปฏิบัติการ (Operating System) ระบบเครือข่าย (Network) ระบบบริการคลาวด์ (Cloud Computing) ระบบบริหารจัดการข้อมูล (Data Management System) เป็นต้น ดังนั้น ผลกระทบใด ๆ ที่เกิดขึ้นกับระบบเทคโนโลยีสารสนเทศดังกล่าวย่อมส่งผลกระทบต่อการทำงานของ AI ด้วยเช่นกัน ด้วยเหตุนี้ องค์กรจึงควรมีมาตรการในการรักษาความมั่นคงปลอดภัยสารสนเทศและคุ้มครองข้อมูลส่วนบุคคลอย่างเหมาะสม เพื่อควบคุมความเสี่ยงให้อยู่ในขอบเขตที่ยอมรับได้ และป้องกันผลร้ายที่อาจเกิดขึ้น เช่น การโจรกรรมข้อมูลสำคัญขององค์กร ข้อมูลส่วนบุคคลรั่วไหล การโจมตีช่องโหว่ของโมเดลเพื่อให้ระบบบริการหยุดชะงักหรือใช้ประโยชน์จากความผิดพลาดเพื่อเข้าถึงระบบ เป็นต้น

นอกจากนี้ เพื่อให้องค์กรสามารถรับมือกับเหตุการณ์ผิดปกติและลดผลกระทบที่อาจเกิดขึ้นได้อย่างทันท่วงที องค์กรจึงควรจัดให้มีช่องทางรับข้อคิดเห็น (Feedback) ประเด็นปัญหา (Issue) ความผิดพลาด (Error) จาก การปฏิบัติงานและการให้บริการ พร้อมทั้งจัดให้มี มาตรการในการบริหารจัดการเหตุการณ์ผิดปกติ (Incident Management) เพื่อรับมือและแก้ไขปัญหาก็เกิดขึ้น อีกทั้งยังเป็นการลดผลกระทบที่อาจเกิดขึ้น ต่อบุคคล องค์กร หรือสังคมโดยกว้าง

เฝ้าติดตามการประยุกต์ใช้ (Monitoring)

ภายหลังจากนำโมเดลไปใช้งานในระบบบริการจริง องค์กรควรมีการเฝ้าติดตามประสิทธิภาพ (Performance) ของการประยุกต์ใช้ AI ในด้านความสำเร็จตามเป้าหมายและประสิทธิภาพในการทำงานของ AI พร้อมทั้งปรับแต่งค่า (Model Tuning) เมื่อมีความจำเป็น เพื่อให้มั่นใจได้ว่าโมเดลยังสามารถทำงานได้ตามเป้าหมายที่องค์กรกำหนด นอกจากนี้ องค์กรยังควรพิจารณาให้มีกระบวนการในการเฝ้าติดตามว่ามีข้อมูลใหม่ที่เหมาะสมนำมาใช้ในการสอนโมเดลเพื่อปรับปรุงประสิทธิภาพของ AI หรือไม่ นอกจากนี้ การเฝ้าติดตามยังรวมถึงการปฏิบัติงานตามนโยบายขององค์กร หลักการจริยธรรมปัญญาประดิษฐ์ กฎหมายและข้อกำหนดที่เกี่ยวข้องอย่างสม่ำเสมอ

โดยในการเฝ้าติดตามประสิทธิภาพของการประยุกต์ใช้ AI นั้น อาจดำเนินการโดยพิจารณาจากผลลัพธ์จากการทำงาน ระดับความถูกต้องในการตัดสินใจ (Accuracy Rate) ข้อคิดเห็น (Feedback) ประเด็นปัญหา (Issue) และความผิดพลาดที่เกิดขึ้น (Error) เป็นต้น ทั้งนี้ หากองค์กรมีเครื่องมือที่ช่วยในการเฝ้าติดตามและมีการรายงานผลโดยอัตโนมัติ ไม่ว่าจะอยู่ในรูปแบบของการรายงานผ่านหน้าจอสรุปข้อมูล (Dashboard) หรือมีการแจ้งเตือนผ่านช่องทางต่าง ๆ จะช่วยเพิ่มประสิทธิภาพในการเฝ้าติดตาม อีกทั้งยังช่วยลดภาระในการเฝ้าติดตาม และรายงานผลคณะกรรมการกำกับดูแล เพื่อรับทราบ



ประเมินผลการประยุกต์ใช้ (Evaluation)

การประเมินผลการประยุกต์ใช้ AI เป็นการนำผลลัพธ์จากการประยุกต์ใช้ในปัจจุบันมาพิจารณาเพื่อกำหนดแนวทางการดำเนินการในอนาคต โดยอาจพิจารณาในด้านต่าง ๆ เช่น

- ประสิทธิภาพ (Performance) ของการประยุกต์ใช้ AI ในด้านความสำเร็จตามเป้าหมายและประสิทธิภาพในการทำงานของ AI ที่ผ่านมา เมื่อเทียบกับเป้าหมายที่กำหนด ความสอดคล้องตามหลักการจริยธรรมปัญญาประดิษฐ์ กฎหมายและข้อกำหนดที่เกี่ยวข้อง
- พิจารณาปรับปรุงเป้าหมายในการประยุกต์ใช้ AI หากพบว่าการประยุกต์ใช้ AI ไม่สามารถช่วยให้องค์กรบรรลุตามเป้าหมายที่กำหนดได้ หรือมีเหตุจำเป็นที่ต้องปรับปรุงเป้าหมาย
- ปรับปรุงข้อกำหนดความต้องการในการพัฒนาระบบ AI (AI Requirements) หากพบว่ามีข้อกำหนดความต้องการไม่ถูกต้อง หรือยกเลิกข้อกำหนดในบางเรื่องหากพบว่ามีไม่จำเป็น
- วิเคราะห์และทบทวนความเสี่ยงจากการดำเนินงานที่ผ่านมา รวมถึงวิเคราะห์ความเสี่ยงที่อาจเกิดขึ้นในอนาคต เพื่อปรับปรุงแผนการจัดการความเสี่ยง (Risk Treatment Plan) ให้เหมาะสม

โดยในการประเมินผลการประยุกต์ใช้ AI นั้น อาจจัดให้มีการประเมินผลโดยผู้ตรวจประเมินภายใน (Internal Auditor) หรือผู้ตรวจประเมินภายนอก (External Auditor) ตามความเหมาะสม เพื่อช่วยเพิ่มความน่าเชื่อถือ (Reliability) ให้แก่บริการ พร้อมทั้งแจ้งผลการประเมินให้คณะกรรมการกำกับดูแล รักรบาท เพื่อพิจารณาและกำหนดแนวทางการดำเนินงานในอนาคต พร้อมทั้งสื่อสารให้แก่ผู้มีส่วนได้ส่วนเสียที่เกี่ยวข้องรับทราบและดำเนินการตามแนวทางที่กำหนด



Retirement

ยุคการใช้งานโมเดล (Retirement)

ในกรณีที่โมเดล ผลิตภัณฑ์ ระบบหรือบริการที่เกี่ยวข้องกับ AI ไม่สามารถช่วยให้องค์กรบรรลุตามเป้าหมายหรือตอบโจทย์ความต้องการทางธุรกิจได้อย่างมีประสิทธิภาพ ซึ่งอาจเกิดจากเทคโนโลยีที่ใช้ล้าสมัย (Obsolete) หรือผลลัพธ์จากการทำงานของ AI ในปัจจุบันไม่สามารถตอบสนองต่อความต้องการขององค์กรที่เปลี่ยนแปลงได้ ซึ่งในกรณีดังกล่าว องค์กรอาจพิจารณาดำเนินการใน 2 กรณี กล่าวคือ

1. ยุติการใช้งานโมเดล ระบบ หรือบริการที่เกี่ยวข้องกับ AI ดังกล่าว
2. นำโมเดล ผลิตภัณฑ์ ระบบหรือบริการที่เกี่ยวข้องกับ AI ใหม่ มาใช้งาน เพื่อทดแทนของเดิม (Replace) หากองค์กรพบว่ามีแนวทางใหม่ที่สามารถตอบสนองต่อความต้องการขององค์กรที่เปลี่ยนแปลงไปได้



8.2. การให้บริการ AI (AI Service)

การสื่อสารเพื่อสร้างความสัมพันธ์ระหว่างองค์กรและผู้ใช้งานนั้น มีความสำคัญอย่างยิ่งต่อการสร้างความโปร่งใสในการให้บริการ อีกทั้งยังเป็นการสร้างความเชื่อมั่นของผู้ใช้งานที่มีต่อบริการและองค์กร ด้วยเหตุนี้ องค์กรจึงควรมีการสื่อสารนโยบายในการให้บริการ ข้อกำหนดในการให้บริการ และข้อมูลที่เกี่ยวข้องกับการใช้งาน AI รวมถึงการเปิดให้มีช่องทางการติดต่อสื่อสารกับผู้ใช้งานเพื่อรับฟังเสียงสะท้อนจากการใช้งานจริง ดังตัวอย่างต่อไปนี้

การประกาศนโยบายและข้อมูลทั่วไปเกี่ยวกับการใช้งาน AI (Policy and General Disclosure)

ในการให้บริการที่เกี่ยวข้องกับ AI องค์กรควรมีการแจ้งให้ผู้ใช้งานรับทราบถึงนโยบายด้านต่าง ๆ เกี่ยวกับการให้บริการ ตัวอย่างเช่น นโยบายในการใช้งาน AI (AI Usage Policy) แนวปฏิบัติตามหลักการจริยธรรมปัญญาประดิษฐ์ (AI Ethics Principles) นโยบายด้านความมั่นคงปลอดภัย (Security Policy) นโยบายความเป็นส่วนตัว (Privacy Policy) เป็นต้น

นอกจากนี้ องค์กรยังควรมีการแจ้งผู้ใช้งานทราบว่า กำลังใช้งาน รับบริการ หรือทำงานร่วมกับ AI ผ่านการแจ้งเตือนบนแอปพลิเคชัน (In-app Notification) พร้อมทั้งควรมีการแจ้งถึงวิธีการใช้งาน ข้อห้ามในการใช้งาน ความสามารถ ข้อจำกัด ผลลัพธ์การตัดสินใจของ AI รวมถึงวิธีการและเหตุผลเบื้องหลังการทำงานของ

AI ผ่านทางคู่มือการใช้งาน คำถามที่พบบ่อย (FAQ) และข้อตกลงการให้บริการ (Terms and Conditions) เป็นต้น

การปิดการทำงานของ AI (AI Function Opt-out)

ในกรณีที่ระบบสามารถเปิดโอกาสให้ผู้ใช้งานปิดการทำงานของ AI ได้ด้วยตนเอง เช่น รถยนต์ที่มีระบบการขับเคลื่อนอัตโนมัติด้วยตนเอง อนุญาตให้ผู้ใช้งานสามารถปิดระบบขับเคลื่อนอัตโนมัติด้วยตนเองได้ หากพบว่ารถยนต์ไม่สามารถตัดสินใจได้ถูกต้องในสภาพถนนที่ไม่เหมาะสม เป็นต้น อีกทั้ง องค์กรควรมีการสื่อสารอย่างเหมาะสมเพื่อให้ผู้ใช้งานทราบถึงขั้นตอนการปิดการทำงานดังกล่าว

ช่องทางการติดต่อสื่อสาร (Communication Channel)

การเปิดช่องทางการติดต่อสื่อสารเพื่อเปิดรับความคิดเห็น (Feedback) ประเด็นปัญหา (Issue) และความผิดพลาด (Error) ที่พบโดยผู้ใช้งานนั้นมีความสำคัญมาก เนื่องจากจะช่วยให้องค์กรสามารถปรับปรุง รวมถึงแก้ไขปัญหาและความผิดพลาดที่พบจากการให้บริการจริง อีกทั้งยังช่วยในการปรับปรุงประสบการณ์ในการให้บริการและป้องกันปัญหาหรือผลกระทบที่อาจเกิดขึ้นในอนาคต

9. ภาคผนวก ก - ตารางแสดงความสัมพันธ์ระหว่าง หลักการจริยธรรมปัญญาประดิษฐ์และองค์ประกอบหลัก ในการสนับสนุนธรรมาภิบาล ในการประยุกต์ใช้ AI

หลักการจริยธรรมปัญญาประดิษฐ์ ตามแนวของ สดช. (ONDE's AI Ethics Principles)	องค์ประกอบในการสนับสนุนธรรมาภิบาล ในการประยุกต์ใช้ AI		
	การกำหนดโครงสร้าง การกำกับดูแล (AI Governance Structure)	การกำหนด กลยุทธ์ในการประยุกต์ ใช้ AI (AI Strategy)	การกำกับดูแลการปฏิบัติ งานที่เกี่ยวข้องกับ AI (AI Operation)
1. ความสามารถในการแข่งขัน และการพัฒนาอย่างยั่งยืน (Competitiveness and Sustainable Development)		✓	
2. ความสอดคล้องกับกฎหมาย จริยธรรม และมาตรฐานสากล (Laws, Ethics, and International Standards)		✓	✓
3. ความโปร่งใสและความรับผิดชอบ ต่อผลของการกระทำ (Transparency and Accountability)	✓		✓
4. ความมั่นคงปลอดภัยและ ความเป็นส่วนตัว (Security and Privacy)			✓
5. ความเท่าเทียม หลากหลาย ครอบคลุม และเป็นธรรม (Fairness)			✓
6. ความน่าเชื่อถือ (Reliability)			✓

10. កាតព្វកិច្ច - តារាងឧទ្ទេសវិធានការងារក្នុងការងារកែលម្អនិងការប្រើប្រាស់

វិធានការងារកែលម្អនិងការប្រើប្រាស់	កែលម្អ	សាងសង់	ត្រួតពិនិត្យ	ត្រួតពិនិត្យ	ត្រួតពិនិត្យ	ត្រួតពិនិត្យ	អាយត្យូ
1. កែលម្អនិងការប្រើប្រាស់កម្មវិធីកែលម្អកម្រិតខ្ពស់	✓	✓	✓	✓	✓	✓	<ul style="list-style-type: none"> កែលម្អនិងការប្រើប្រាស់កម្រិតខ្ពស់ រួមមាន ការងារកែលម្អនិងការប្រើប្រាស់ កម្រិតខ្ពស់ ប៉ុន្តែមិនមែនជាការ កែលម្អនិងការប្រើប្រាស់កម្រិតខ្ពស់ ក្នុងការងារកែលម្អនិងការប្រើប្រាស់
2. កែលម្អនិងការប្រើប្រាស់កម្មវិធីកែលម្អកម្រិតទាប (Open-source Model) ប្រើប្រាស់		✓	✓	✓	✓	✓	<ul style="list-style-type: none"> កែលម្អនិងការប្រើប្រាស់កម្រិតទាប
3. កែលម្អនិងការប្រើប្រាស់កម្មវិធីកែលម្អកម្រិតទាប (Pre-trained model) ប្រើប្រាស់					✓	✓	<ul style="list-style-type: none"> កែលម្អនិងការប្រើប្រាស់កម្រិតទាប
4. កែលម្អនិងការប្រើប្រាស់កម្មវិធីកែលម្អកម្រិតទាប (Off-the-shelf AI Product)					✓	✓	<ul style="list-style-type: none"> កែលម្អនិងការប្រើប្រាស់កម្រិតទាប

11. เอกสารอ้างอิง

- C3 AI. n.d. "Best Practices in Developing an Enterprise AI Roadmap." *C3 AI*. Accessed April 2023 ,24. <https://c3.ai/wp-content/uploads/06/2020/Best-Practices-in-Developing-an-Enterprise-AI-Roadmap.pdf>.
- Concordia University and Dawson College. 2021. "ARTIFICIAL INTELLIGENCE COMPETENCY FRAMEWORK." *DawsonAI*. September. Accessed March 2023 ,29. https://www.dawsoncollege.qc.ca/ai/wp-content/uploads/sites/180/Corrected-FINAL_PIA_ConcordiaDawson_AICompetencyFramework.pdf.
- Info-Tech Researc Research Group. n.d. "AI Governance A framework for building responsible, ethical, fair, and transparent AI." *Info-Tech Researc Research Group*. Accessed May 2023 ,20. <https://www.infotech.com/research/ss/ai-governance>.
- International Organization for Standardization. 2022. *ISO/IEC 22989:2022 Information technology – Artificial intelligence – Artificial intelligence concepts and terminology*.
- . 2023. *ISO/IEC 23894:2023 Information technology – Artificial intelligence – Guidance on risk management*.
- . 2015. *ISO/IEC 38500:2015 Information technology – Governance of IT for the organization*.
- . 2022. *ISO/IEC 38507:2022 Information technology – Governance of IT – Governance implications of the use of artificial intelligence by organizations*.
- . 2021. *ISO/IEC TR 24027:2021 Information technology – Artificial intelligence (AI) – Bias in AI systems and AI aided decision making*.
- NASSCOM. n.d. "Responsible AI Architects Guide Prescribing Responsible AI Best Practices Implementation Methods and Tools for Enterprise Adoption." *INDIAai*. Accessed April 2023 ,14. <https://indiaai.gov.in/responsible-ai/pdf/architect-guide.pdf>.
- National Institute of Standards and Technology (NIST). 2023. "Artificial Intelligence Risk Management Framework (AI RMF 1.0)." *AI Risk Management Framework*. January. Accessed 2023 ,30 06. <https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.1-100.pdf>.
- NTT Data Inc. 2022. "AI for Business Value Generation." *NTT Data*. June. Accessed April 2023 ,1. <https://es.nttdata.com/documents/whitepaper-e2e-v02.pdf>.
- Office of the National Digital Economy and Society Commission. 2022. "Thailand AI Ethics Guideline." February 07. [https://www.onde.go.th/assets/portals/1/files/Thailand20%AI20%Ethics20%Guideline20*\(White20*paper\)20*Edit20*Version.pdf](https://www.onde.go.th/assets/portals/1/files/Thailand20%AI20%Ethics20%Guideline20*(White20*paper)20*Edit20*Version.pdf).
- PRIMO.ai. n.d. *Train, Validate, and Test*. Accessed 2023 ,13 04. http://primo.ai/index.php?title=Train,_Validate,_and_Test.
- Scientific American. 2019. *Racial Bias Found in a Major Health Care Risk Algorithm*. October 24. Accessed March 2023 ,22. <https://www.scientificamerican.com/article/racial-bias-found-in-a-major-health-care-risk-algorithm/>.
- The Guardian. 2022. *The Guardian*. December 22. Accessed March 2023 ,22. <https://www.theguardian.com/technology/2022/dec/22/tesla-crash-full-self-driving-mode-san-francisco>.
- The Info-communications Media Development Authority (IMDA) and Personal Data Protection Commission (PDPC). 2020. "Model Artificial Intelligence Governance Framework Second Edition." *Model AI Governance Framework*. January 21. Accessed April 2023 ,13. <https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Resource-for-Organisation/AI/SGModelAIGovFramework2.pdf>.
- The Open Worldwide Application Security Project (OWASP). n.d. "OWASP AI Security and Privacy Guide." *OWASP*. Accessed April 2023 ,22. <https://owasp.org/www-project-ai-security-and-privacy-guide/>.
- The World Economic Forum Centre for the Fourth Industrial Revolution, the Info-communications Media Development Authority (IMDA) and Personal Data Protection Commission (PDPC). 2020. "Implementation and Self Assessment Guide for Organisations (ISAGO)." Accessed 2023 ,13 04. <https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Resource-for-Organisation/AI/SGIsago.pdf>.

ภาพกระบวนการพัฒนาแนวทางการประยุกต์ใช้ปัญญาประดิษฐ์อย่างมีธรรมาภิบาล สำหรับผู้บริหารองค์กร (AI Governance Guideline for Executives)

Timeline (draft) AI Governance Guideline for Executives



ที่ปรึกษา International Policy Advisory Panel

Professor Dr. Urs Gasser

Professor Dr. Christian Fieseler

Professor Dr. Kerstin Vokinger

Professor Dr. Peter Parycek

Professor Dr. Stefaan Verhulst

Professor Albert Bifet

Dr. Sandra Cortesi

Dr. Nawanan Theera-Ampornpunt

Dr. Rom Hiranpruk

ผู้บริหารสำนักงานพัฒนารัฐกรรมทางอิเล็กทรอนิกส์

ดร. ชัยชนะ มิตรพันธ์

ดร. ศักดิ์ เสกขุนทด

ดร. ชลิตดา มัชฌมบุรุษ

คณะทำงาน AIGC

รจนา ล้าเลิศ

ธิตกร ตระกูลศิริศักดิ์

ธงชัย แสงศิริ

หทัยชนก พุทธิรักษา



Artificial Intelligence Governance Clinic (AIGC)
Electronic Transactions Development Agency



จัดทำโดย
สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์ (สพธอ.)
Electronic Transactions Development Agency หรือ ETDA (เอ็ตด้า)

