กรมการแพทย์แผนไทยและการแพทย์ทางเลือก
Department of Thai Traditional and Alternative Medicine

แนวทางการประยุกต์ใช้ปัญญาประดิษฐ์อย่างมีธรรมาภิบาล

กรมการแพทย์แผนไทยและการแพทย์ทางเลือก

ประจำปีงบประมาณ พ.ศ. ๒๕๖๙

DTAM next

แนวทางการประยุกต์ใช้ปัญญาประดิษฐ์อย่างมีธรรมาภิบาล

กรมการแพทย์แผนไทยและการแพทย์ทางเลือก

ประจำปีงบประมาณ พ.ศ. ๒๕๖๙

แนวทางการประยุกต์ใช้ปัญญาประดิษฐ์อย่างมีธรรมาภิบาล

ที่ปรึกษา

นายแพทย์พงศธร พอกเพิ่มดี อธิบดีกรมการแพทย์แผนไทยและการแพทย์ทางเลือก

นายสมศักดิ์ กรีชัย รองอธิบดีกรมการแพทย์แผนไทยและการแพทย์ทางเลือก

นายนันทศักดิ์ โชติชนะเดชาวงศ์ ผู้อำนวยการกองวิชาการและแผนงาน

บรรณาธิการ

นายจักรกฤษณ์ สิงห์บุตร เภสัชกรชำนาญการ กองวิชาการและแผนงาน

นางสาวสุวิมล สุมลตรี แพทย์แผนไทยชำนาญการ กองวิชาการและแผนงาน

ออกแบบปกและเล่ม

นายธนาวัฒน์ พลศิลป์ เจ้าหน้าที่วิเคราะห์นโยบายและแผน กองวิชาการและแผนงาน

รวบรวมและจัดทำโดย

กลุ่มงานบริหารจัดการข้อมูลขนาดใหญ่และธรรมาภิบาลข้อมูล กองวิชาการและแผนงาน
กรมการแพทย์แผนไทยและการแพทย์ทางเลือก

Website

http://www.dtam.moph.go.th

บทสรุปผู้บริหาร

เทคโนโลยีปัญญาประดิษฐ์ (Artificial Intelligence: AI) เป็นนวัตกรรมที่ถูกพัฒนาขึ้นเพื่อให้คอมพิวเตอร์มีพฤติกรรมใกล้เคียงมนุษย์ ทั้งในด้านการเรียนรู้ การให้เหตุผล และการแก้ไขปัญหา โดยเฉพาะปัญญาประดิษฐ์ประเภทสร้างสรรค์ (Generative AI) ที่สามารถสร้างเนื้อหาใหม่ได้หลากหลายรูปแบบ เช่น ข้อความ ภาพ วิดีโอ และซอร์สโค้ด ผ่านการประมวลผลบนโครงข่ายประสาทเทียมที่มีความซับซ้อน แม้ว่าเทคโนโลยีดังกล่าวจะสร้างประโยชน์ในการเพิ่มประสิทธิภาพการทำงานและการสร้างสรรค์นวัตกรรม แต่ขณะเดียวกันยังเผชิญกับข้อจำกัดและความเสี่ยงที่สำคัญ เช่น อาการหลอนของข้อมูล (Hallucination) ความเอนเอียงทางจริยธรรม ประเด็นด้านทรัพย์สินทางปัญญา ตลอดจนความเสี่ยงต่อความเป็นส่วนตัวและความมั่นคงปลอดภัยทางไซเบอร์

เพื่อให้การประยุกต์ใช้ปัญญาประดิษฐ์ภายในองค์กรเป็นไปอย่างมีความรับผิดชอบและบรรลุเป้าหมายตามหลักธรรมาภิบาล องค์กรจำเป็นต้องมีแนวทางการบริหารจัดการความเสี่ยงและโครงสร้างการกำกับดูแลที่ชัดเจน โดยควรกำหนดกรอบการทำงานหลักที่ประกอบด้วย ๔ หลักสำคัญ ได้แก่ การจัดตั้งโครงสร้างการกำกับดูแล (AI Governance Structure) ผ่านคณะกรรมการที่มีอำนาจหน้าที่ในการกำหนดนโยบายและติดตามผล การกำหนดกลยุทธ์ (AI Strategy) ที่สอดคล้องกับเป้าหมายทางธุรกิจและความซับซ้อนของเทคโนโลยี การบริหารจัดการการปฏิบัติงาน (AI Operation) เพื่อควบคุมความเสี่ยงให้อยู่ในระดับที่ยอมรับได้ และการกำกับดูแลวัฏจักรชีวิตของปัญญาประดิษฐ์ (AI Lifecycle) ตั้งแต่กระบวนการเตรียมข้อมูลที่มีคุณภาพ การคุ้มครองข้อมูลส่วนบุคคล ไปจนถึงการทดสอบและการประเมินผลอย่างต่อเนื่อง

ทั้งนี้ การดำเนินงานด้านธรรมาภิบาลต้องยึดถือหลักการจริยธรรมปัญญาประดิษฐ์ ๖ ประการตามแนวทางสากลและสำนักงานคณะกรรมการดิจิทัลเพื่อเศรษฐกิจและสังคมแห่งชาติ (สดช.) เป็นสำคัญ ประกอบด้วย ความสามารถในการแข่งขันและการพัฒนาอย่างยั่งยืน ความสอดคล้องกับกฎหมายและสิทธิมนุษยชน ความโปร่งใสและความรับผิดชอบต่อผลของการกระทำ ความมั่นคงปลอดภัยและความเป็นส่วนตัว ความเท่าเทียมและความเป็นธรรม และความน่าเชื่อถือของระบบ การบูรณาการหลักจริยธรรมเหล่านี้เข้ากับกระบวนการทำงานร่วมระหว่างมนุษย์และปัญญาประดิษฐ์จะช่วยลดความผิดพลาดและสร้างความเชื่อมั่นในการใช้งานเทคโนโลยีอย่างยั่งยืน

(นายพงศธร พอกเพิ่มดี)
อธิบดีกรมการแพทย์แผนไทยและการแพทย์ทางเลือก

๑. คำนิยาม

๑.๑ Artificial Intelligence (AI) ปัญญาประดิษฐ์เป็นเทคโนโลยีที่ถูกพัฒนาขึ้นเพื่อให้คอมพิวเตอร์มีคุณสมบัติหรือพฤติกรรมใกล้เคียงมนุษย์ เช่น การเรียนรู้ การรับรู้และตอบสนองต่อสภาพแวดล้อม การให้เหตุผล และการแก้ไขปัญหา เป็นต้นตามวัตถุประสงค์ที่มนุษย์กำหนด

๑.๒ Machine Learning (ML) เทคโนโลยี AI ประเภทหนึ่งที่ทำงานหรือสร้างผลลัพธ์บนพื้นฐานของข้อมูลที่ได้รับจากการฝึกฝนหรือจากสภาพแวดล้อม

๑.๓ Deep Learning (DL) Machine Learning ประเภทหนึ่งที่ประมวลผลผ่านโครงข่ายประสาทเทียม (Artificial Neural Network: ANN) จำนวนหลายชั้น (Layer) ที่ถูกสร้างขึ้นจากข้อมูลที่ได้รับการฝึกฝน เพื่อให้สามารถทำงานหรือสร้างผลลัพธ์ที่มีประสิทธิภาพดียิ่งขึ้น

๑.๔ Artificial Neural Network (ANN) โครงข่ายของเซลล์ประสาทเทียม (Artificial Neuron) ที่คล้ายกับการเชื่อมต่อเซลล์ประสาท (Neuron) ในสมองมนุษย์ โดยในแต่ละเซลล์ประสาทเทียมนั้นมีหน้าที่ในการรับข้อมูลและนำไปประมวลผลเพื่อสร้างเป็นผลลัพธ์ จากนั้นจึงส่งต่อผลลัพธ์ไปยังเซลล์ประสาทเทียมในชั้น (Layer) ถัดไป เพื่อประมวลผลต่อ

๑.๕ Generative AI เทคโนโลยี AI ประเภทหนึ่งที่มีความสามารถในการสร้างเนื้อหาใหม่ในหลากหลายรูปแบบ เช่น ข้อความ ภาพ วิดีโอ ซอซอร์สโค้ด หรือรูปแบบอื่น เป็นต้น ตามข้อความหรือคำสั่ง (Prompt) ที่มนุษย์เป็นผู้กำหนด

๑.๖ Prompt Engineering การสร้างและปรับแต่งข้อความหรือคำสั่ง เพื่อให้ Generative AI สร้างผลลัพธ์ (Output) ที่ดีที่สุดและตรงตามความต้องการ

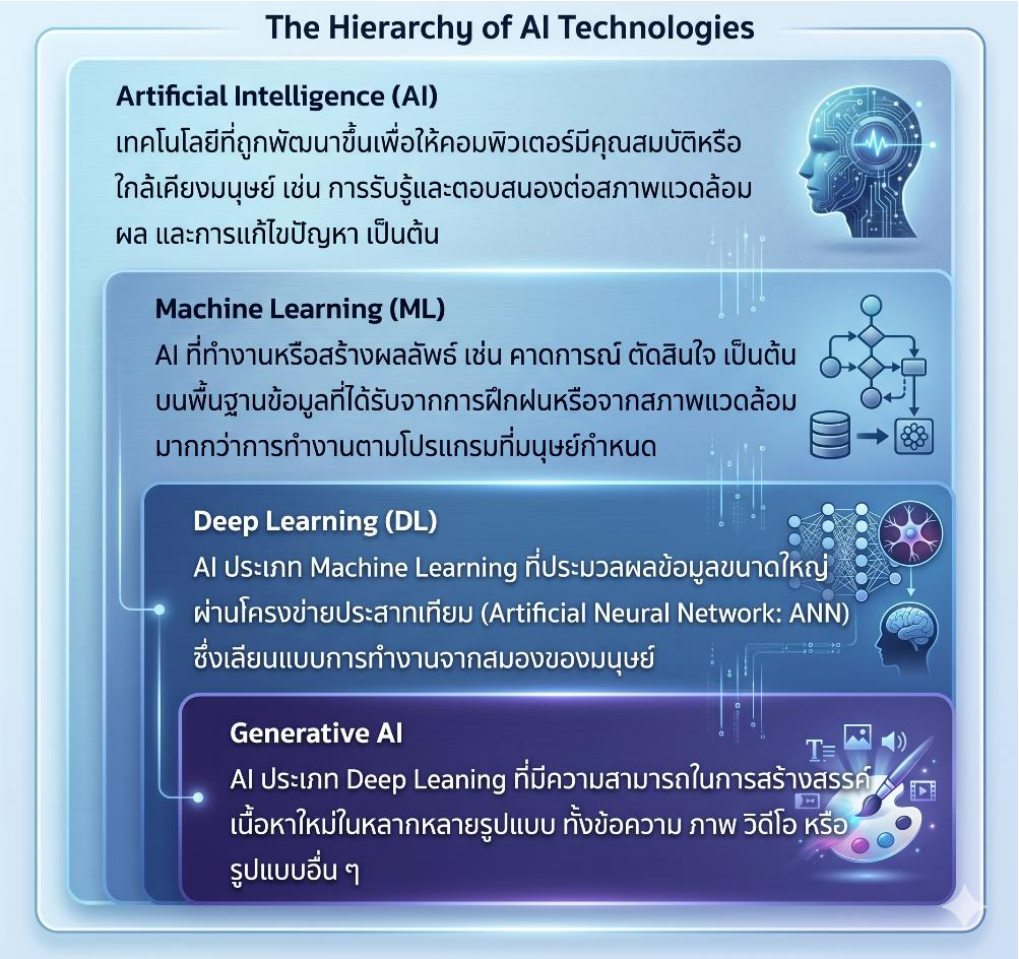
๑.๗ Foundation Model โมเดล AI ประเภท Generative AI ที่ได้รับการฝึกฝนด้วยข้อมูลขนาดใหญ่ โดยมีวัตถุประสงค์เพื่อให้สามารถสร้างเนื้อหาใหม่ที่คล้ายคลึงกับข้อมูลที่ได้รับการฝึกฝน

๑.๘ Large Language Model (LLM) โมเดลภาษาขนาดใหญ่ที่รับข้อความหรือคำสั่ง (Input) ในรูปแบบภาษา และนำไปสร้างผลลัพธ์ (Output) ที่มีความสามารถในด้านภาษาที่หลากหลาย เช่น การสร้างข้อความใหม่ การแปลภาษา การสรุปความการวิเคราะห์ข้อความ เป็นต้น

๒. ความหมายของ Generative AI

Generative AI เป็น AI ประเภทหนึ่งที่สามารถสร้างเนื้อหาได้หลากหลายรูปแบบ เช่น ข้อความ ภาพ วิดีโอ ซอร์สโค้ด หรือรูปแบบอื่น เป็นต้น ด้วยการสั่งการผ่านข้อความหรือคำสั่ง (Prompt) ที่มนุษย์เป็นผู้กำหนดเมื่อ Generative AI ได้รับ Prompt จากผู้ใช้งานแล้ว Generative AI จะทำการสร้างเนื้อหาใหม่ที่คล้ายคลึงกับข้อมูลที่ได้รับการฝึกฝนมา โดยเลือกนำเสนอเนื้อหาที่สอดคล้องและเหมาะสมกับ Prompt ที่ได้รับมามากที่สุด โดยพิจารณาจากหลักการความน่าจะเป็นจากความสามารถและกระบวนการทำงานของ Generative AI ข้างต้น จึงทำให้ Generative AI มีความแตกต่างจาก AI แบบเดิม (Traditional AI) กล่าวคือ Traditional AI ถูกออกแบบมาเพื่อคาดการณ์ (Prediction) การตัดสินใจ (Decision) หรือการให้คำแนะนำ (Recommendation) บนพื้นฐานของข้อมูลที่ได้รับการฝึกฝน (Train) มาก่อน การใช้ AI วินิจฉัยโรคปอดจากภาพเอกซเรย์ปอด และการใช้ AI แนะนำสินค้าและบริการ เป็นต้น แต่ AI แบบนี้ไม่ได้ถูกออกแบบมาเพื่อสร้างเนื้อหาใหม่ที่มีความคล้ายคลึงกับข้อมูลต้นฉบับ

จากรูปที่แสดงลำดับถัดไปจะพบว่า Generative AI เป็น AI ประเภทหนึ่งของ Deep Learning เนื่องจาก Generative AI สร้างเนื้อหาใหม่บนพื้นฐานของข้อมูลที่ได้รับการฝึกฝนมา โดยทำการประมวลผลเพื่อสร้างเนื้อหาใหม่ผ่านโครงข่ายประสาทเทียม (Artificial Neural Network) ที่มีความซับซ้อน



๓. ความสามารถของ Generative AI

๑. ข้อความ (Text) - ความสามารถในการสร้างและปรับปรุงข้อความในหลายรูปแบบ เช่น การเขียนบทความ การสร้างคำตอบ การปรับปรุงข้อความ และอื่น ๆ ตัวอย่างเครื่องมือ เช่น ChatGPT, Claude, Copilot, Gemini เป็นต้น ซึ่งเป็นตัวอย่างของ Generative Al ประเภท Large Language Model

๒. ซอร์สโค้ด (Source Code) - ความสามารถในการสร้างและปรับปรุง Code โปรแกรมในหลายภาษาโปรแกรม ตัวอย่างเครื่องมือเช่น ChatGPT, Claude, Copilot, Gemini เป็นต้น

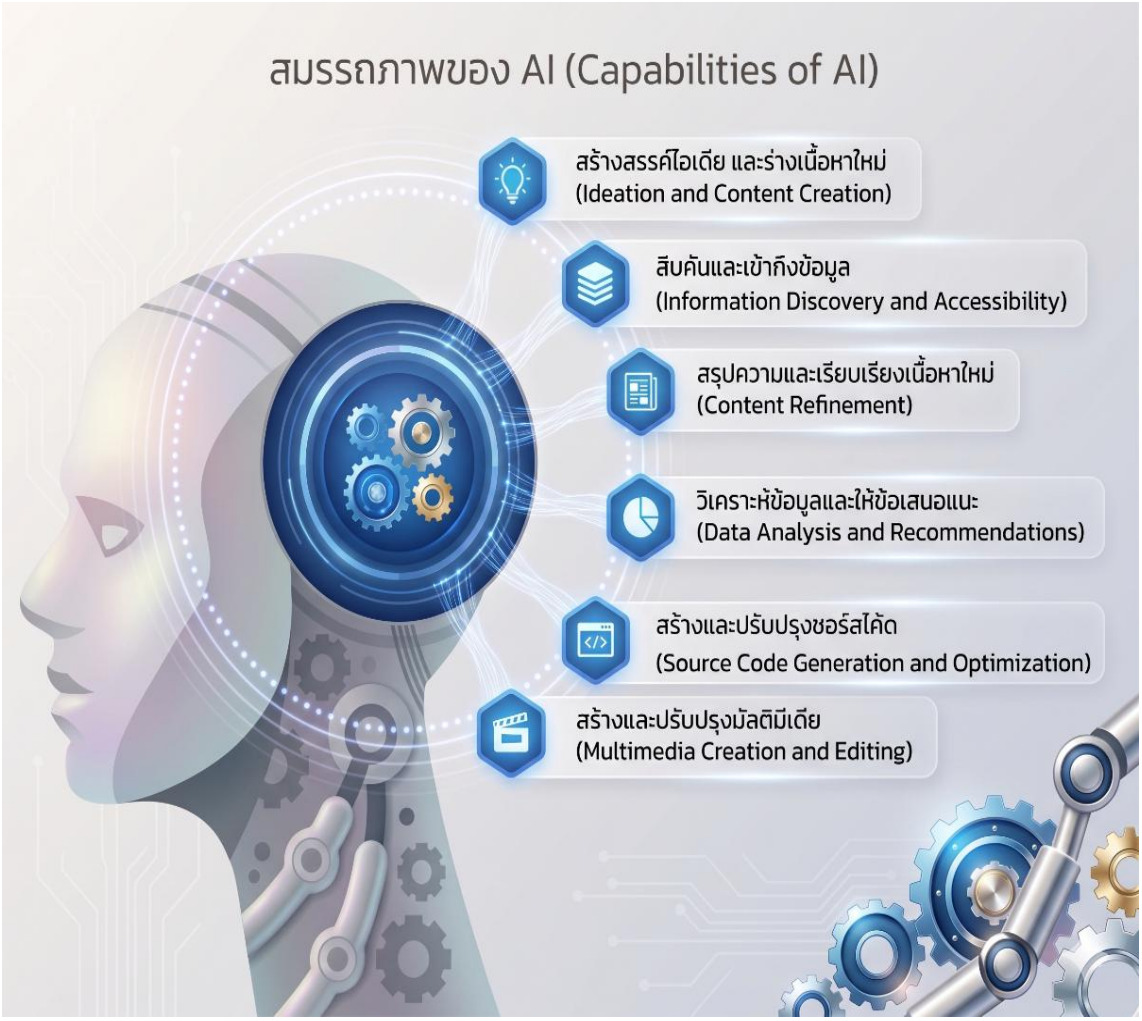
๓. รูปภาพ (Image) - ความสามารถในการสร้างและปรับปรุงรูปภาพต่าง ๆ ได้ไม่ว่าจะเป็นการออกแบบภาพใหม่ การปรับแต่งภาพที่มีอยู่แล้ว หรือการสร้างภาพจากข้อความ ตัวอย่างเครื่องมือเช่น DALL-E, Midjourney เป็นต้น

๔. เสียง (Audio) - ความสามารถในการสร้างและปรับปรุงเสียง เช่น การสร้างเสียงพูด การปรับแต่งเสียง การสร้างดนตรี และอื่น ๆ ตัวอย่างเครื่องมือ เช่น Suno และ Udio เป็นต้น

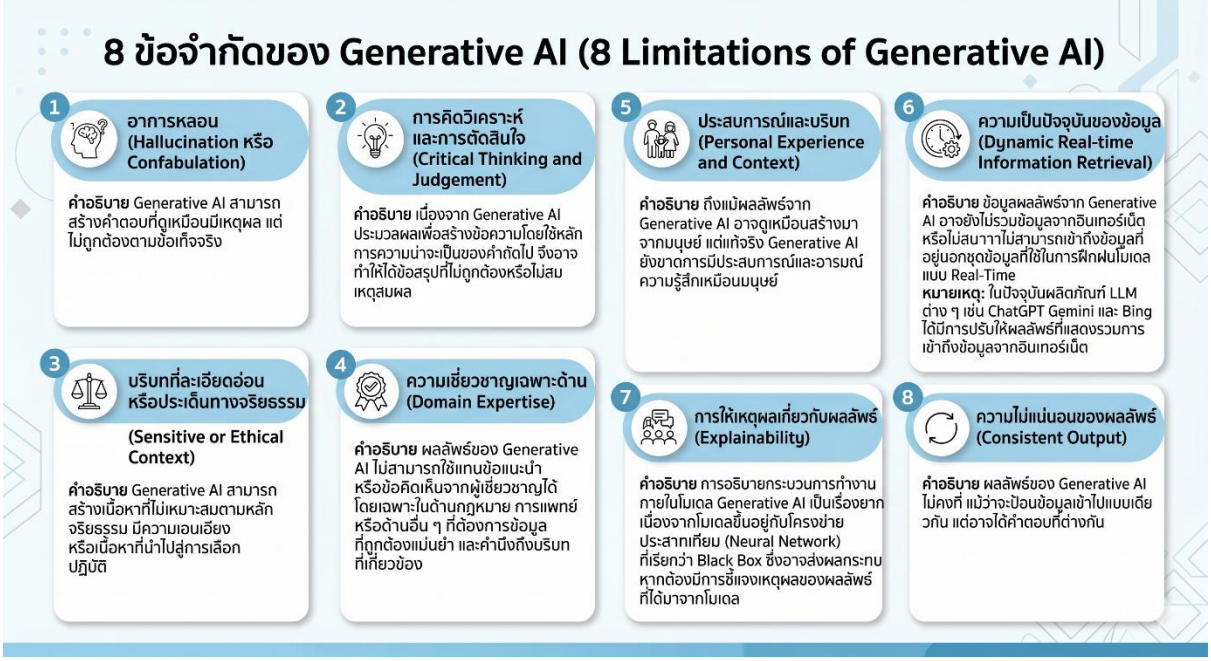
๕. วิดีโอ (Video) - ความสามารถในการสร้างและปรับปรุงวิดีโอ เช่น การตัดต่อวิดีโอ การสร้างวิดีโอจากข้อความ การสร้างวิดีโอแบบแอนิเมชัน และอื่น ๆ ตัวอย่าง เครื่องมือเช่น Runway, Adobe Firefly เป็นต้น

๖. โมเดล ๓ มิติ (3D Model) – ความสามารถในการสร้างและปรับปรุงโมเดล ๓ มิติได้ไม่ว่าจะเป็นการสร้างโมเดล ๓ มิติใหม่ การปรับแต่งโมเดลที่มีอยู่แล้ว หรือการสร้างโมเดลจากข้อความ

๔. ประโยชน์จากการประยุกต์ใช้ Generative AI



๕. ข้อจำกัดของ Generative AI



จากข้อจำกัดในการนำ Generative AI มาประยุกต์ใช้ข้างต้น องค์กรจึงควรคำนึงถึงประเด็นด้านต่าง ๆ รวมถึงความสอดคล้องกับกฎระเบียบหรือข้อบังคับที่เกี่ยวข้อง นอกจากนี้ ควรเพิ่มการมีส่วนร่วมของมนุษย์ในกระบวนการทำงานร่วมกับ Generative AI และการอัปเดตการเปลี่ยนแปลงทางด้านเทคโนโลยี เพื่อให้มั่นใจว่าการประยุกต์ใช้มีความสอดคล้องกับความสามารถของเทคโนโลยีที่มีการพัฒนาอย่างต่อเนื่อง ทั้งนี้ เพื่อเป็นการเตรียมพร้อมรับมือกับข้อจำกัดต่าง ๆ ของ Generative AI ได้อย่างมีประสิทธิภาพ

๖. ความเสี่ยงที่อาจเกิดขึ้นจากการประยุกต์ใช้ Generative AI

การทำความเข้าใจประเด็นความเสี่ยงที่อาจเกิดขึ้นเมื่อนำเทคโนโลยี Generative AI มาประยุกต์ใช้ในองค์กรจะช่วยให้องค์กรหาสมดุลระหว่างประโยชน์และความเสี่ยงทำให้สามารถตัดสินใจนำ Generative AI มาประยุกต์ใช้งานได้อย่างเหมาะสม Generative AI นำมาซึ่งความเสี่ยงรูปแบบใหม่ ดังนั้น องค์กรจึงควรมีวิธีการวิเคราะห์และจัดการความเสี่ยงเพิ่มเติมอย่างเหมาะสม โดยจากเอกสาร "Artihcial Intelligence Risk Management Framework: Generative Artifcial Intelligence Profle" ของ National Institute of Standards and Technology (NIST) ซึ่งมีการระบุประเด็นความเสี่ยงที่ควรให้ความสำคัญ ซึ่งประกอบด้วย

๑. ความเสี่ยงด้านข้อมูลที่เป็นอันตรายต่อการผลิตอาวุธ เคมี รังสี หรือนิวเคลียร์ (Chemical, Biological, Radiological, or Nuclear (CBRN) Weapons) Generative AI อาจถูกใช้เป็นเครื่องมือในการสร้างเนื้อหาที่เกี่ยวข้องกับการผลิตอาวุธเคมี รังสี หรือนิวเคลียร์ และอาจถูกนำไปใช้ในทางที่ไม่เหมาะสม

๒. ความเสี่ยงด้านเนื้อหาที่น่าเชื่อถือแต่ไม่ถูกต้อง (Confabulation) การผลิตเนื้อหาที่ดูน่าเชื่อถือแต่ไม่ถูกต้องตามข้อเท็จจริง อาจถูกเรียกว่า "Hallucination" หรือ "Fabrication" โดย Generative AI อาจสร้างผลลัพธ์ที่ผิดไปจากข้อเท็จจริง หรือขัดแย้งกับข้อความที่สร้างขึ้นก่อนหน้าทั้งที่อยู่ในบริบทเดียวกัน เช่น สับสนเกี่ยวกับบุคคล สถานที่ หรือรายละเอียดเหตุการณ์ทางประวัติศาสตร์ เป็นต้น ซึ่งอาจทำให้เกิดการนำเนื้อหาที่ผิดไปใช้งาน

๓. ความเสี่ยงด้านเนื้อหาอันตรายหรือรุนแรง (Dangerous or Violent Recommendations) Generative Al อาจให้คำแนะนำที่ยั่วยุ ปลุกปั่นหรือคุกคาม ที่จะนำไปสู่ความรุนแรง โดยอาจ สร้างภาพ, วิดีโอหรือเสียง เพื่อให้เกิดความเข้าใจผิดในตัวองค์กร หรือบุคคล และอาจนำไปสู่การกระทำผิดทางกฎหมาย

๔. ความเสี่ยงด้านความเป็นส่วนตัวของข้อมูล (Data Privacy) ข้อมูลที่ใช้ในการฝึกฝนโมเดล Generative Al อาจถูกเก็บรวบรวม ใช้ หรือเปิดเผยข้อมูลส่วนบุคคลที่ส่งผลกระทบต่อความเป็นส่วนตัวได้ เช่น ข้อมูลบัตรประชาชน ประวัติการรักษา ที่อยู่ หรือข้อมูลที่สามารถระบุตัวตนได้ เป็นต้น ซึ่งข้อมูลดังกล่าวอาจรั่วไหลจากการโจมตี หรือเป็นส่วนหนึ่งของผลลัพธ์ นอกจากนี้ ผลลัพธ์ที่ปรากฏข้อมูลส่วนบุคคลอาจทำให้เกิดผลลัพธ์ที่มีอคติและการเลือกปฏิบัติที่เป็นอันตรายต่อตัวบุคคลได้

๕. ความเสี่ยงด้านสิ่งแวดล้อม (Environmental) ความเสี่ยงด้านสิ่งแวดล้อมจากการใช้ทรัพยากรจำนวนมากในการฝึกโมเดล อาจสร้างผลกระทบด้านพลังงานและการปล่อยคาร์บอนของ Generative Al ซึ่งขึ้นอยู่กับประเภทของโมเดล, รูปแบบ, ฮาร์ดแวร์ และประเภทของแอปพลิเคชัน

๖. ความเสี่ยงด้านการทำงานร่วมกันระหว่างมนุษย์และ Generative AI (Human-AI Configuration) การกำหนดระดับการมีส่วนร่วมหรือปฏิสัมพันธ์ระหว่างมนุษย์และระบบ AI อาจก่อให้เกิดความเสี่ยง เช่น การอคติหรือไม่เชื่อผลลัพธ์ การเชื่อผลลัพธ์มากเกินไปโดยไม่ตรวจสอบ ความไม่สอดคล้องกับเป้าหมายและ/หรือผลลัพธ์ที่ต้องการ การใช้ในทางที่ผิด การใช้อย่างไม่ถูกต้อง และการใช้อย่างไม่ปลอดภัยกับมนุษย์ เป็นต้น

๗. ความเสี่ยงด้านการนำข้อมูลที่ไม่ครบถ้วนถูกต้อง หรือบิดเบือน (Information Integrity) Generative AI อาจถูกใช้เป็นเครื่องมือในการสร้างและเผยแพร่ข้อมูลที่ไม่ถูกต้อง (misinformation) หรือข้อมูลที่ถูกทำให้บิดเบือน (disinformation) ซึ่งอาจทำลายความเชื่อมั่นต่อบุคคล องค์กร และสังคมในวงกว้าง

๘. ความเสี่ยงด้านความปลอดภัยของข้อมูล (Information Security) ผู้ไม่หวังดีอาจใช้ Generative Al ในการหาช่องโหว่ของระบบ การเขียนโปรแกรมในการเจาะระบบขององค์กร นอกจากนี้ Generative Al เองยังเป็นเป้าหมายในการถูกโจมตี เช่น การโจมตีที่โมเดลโดยตรงโดยเฉพาะการโจมตีแบบ Prompt injection หรือการแก้ไขข้อมูลที่ใช้ฝึกฝนโมเดล (Data Poisoning) ซึ่งส่งผลให้ Generative AI ทำงานผิดพลาด

๙. ความเสี่ยงด้านทรัพย์สินทางปัญญา (Intellectual Property) ผลลัพธ์ที่มาจาก Generative Al อาจละเมิดทรัพย์สินทางปัญญา เนื่องจากการจำข้อมูลหรือการสร้างเนื้อหาที่คล้ายกับผลงานที่ได้รับความคุ้มครองลิขสิทธิ์ รวมถึงการใช้อัตลักษณ์หรือลักษณะเด่นของบุคคลที่ไม่ได้รับอนุญาตอาจเป็นปัญหาที่ไม่ได้รับการคุ้มครองจากกฎหมายทรัพย์สินทางปัญญา

๑๐. ความเสี่ยงด้านเนื้อหาลามก คุกคามหรือล่วงละเมิดทางเพศ (Obscene, Degrading, and/or Abusive Content) Generative Al มีโอกาสสร้างเนื้อหาลามก อนาจาจาร คุกคามหรือล่วงละเมิดทางเพศ เนื่องจากโมเดล AI ถูกฝึกฝนด้วยชุดข้อมูลเปิดในอินเกอร์เน็ตที่อาจมีเนื้อหาลักษณะดังกล่าว รวมถึงข้อมูลที่ไม่ได้รับอนุญาต ดังนั้น ผลลัพธ์ที่ถูกสร้างขึ้นอาจส่งผลกระทบทางจิตใจและร่างกายของบุคคลที่เกี่ยวข้อง

๑๑. ความเสี่ยงด้านเนื้อหาที่มีความคิดเชิงลบ อคติแบ่งแยก (Toxicity, Bias, and Homogenization)Generative AI มีโอกาสสร้างเนื้อหาที่มีความคิดเชิงลบ อคติ หรือแบ่งแยก ที่อาจถูกเผยแพร่เป็นวงกว้างและไม่สามารถควบคุมการแพร่กระจายได้ง่าย ซึ่งอาจก่อให้ความเสียหายแก่ชื่อเสียงและจิตใจของบุคคลที่เกี่ยวข้องได้

๑๒. ความเสี่ยงโดยรวมของห่วงโซ่อุปทาน (Value Chain and Component Integration) โมเดล Generative AI อาจถูกฝึกด้วยเนื้อหาที่ไม่ได้รับการตรวจสอบจากแหล่งที่มาของบุคคลที่สามซึ่งอาจทำให้ผลลัพธ์ของโมเดลไม่สามารถตรวจสอบได้ และอาจจะก่อให้เกิดความเสี่ยงต่อผู้ที่เกี่ยวข้องในห่วงโซ่อุปทาน

การวิเคราะห์ประเด็นความเสี่ยงของการประยุกต์ใช้ Generative Al ของแต่ละองค์กรนั้นอาจมีความแตกต่างกัน จึงควรพิจารณาโอกาสความเป็นไปได้ (Likelihood) และผลกระทบ (Impact) ที่อาจเกิดขึ้นในมิติต่าง ๆ ดังนั้น องค์กรควรหมั่นติดตามข้อมูลที่เกี่ยวข้องกับเทคโนโลยีและการพัฒนาของ Generative AI เพื่อนำมาทบทวนแนวทางจัดการความเสี่ยงที่อาจเกิดขึ้นจากการประยุกต์ใช้ Generative AI อย่างเหมาะสมต่อไป

๗. แนวทางการบริหารจัดการความเสี่ยง

Generative Al มีความสามารถที่แตกต่างไปจาก AI ประเภทอื่นจึงนำมาซึ่งความเสี่ยงรูปแบบใหม่ ดังนั้น ในการประยุกต์ใช้ Generative AI จึงควรคำนึงถึงประเด็นความเสี่ยงที่เกี่ยวข้องและกำหนดมาตรการเพื่อป้องกันหรือลดความเสี่ยงที่อาจจะเกิดขึ้นอย่างเหมาะสม โดยแนวทางในการจัดการความเสี่ยงในการประยุกต์ใช้ Generative AI นั้นมีหลากหลายแนวทาง โดยยกตัวอย่าง ดังนี้

๑. กำหนดกรอบแนวทางการประยุกต์ใช้ Generative AI อย่างมีธรรมาภิบาล (Establish Generative Al Governance Structure) องค์กรควรกำหนดกรอบแนวทางในการประยุกต์ใช้ Generative AI ที่ครอบคลุมรูปแบบและลักษณะการประยุกต์ใช้งานจริงภายในองค์กร ทั้งนี้ โดยควรพิจารณาประเด็นความเสี่ยงและผลกระทบที่อาจจะเกิดขึ้น และกำหนดหน้าที่และความรับผิดชอบ (Role and Responsibility) รวมถึงกำหนดความรับผิดชอบต่อผลของการกระทำ (Accountability) ในการประยุกต์ใช้ Generative Al

๒. หมั่นตรวจสอบและทบทวนความสอดคล้องตามข้อกำหนด หรือข้อกฎหมาย (Ensure Regulatory and Legal Compliance) ที่เที่เกี่ยวข้องกับการประยุกต์ใช้ Generative Al ว่ายังเป็นไปตามข้อกำหนดหรือข้อกฎหมายหรือไม่ โดยอาจจำเป็นต้องปรับปรงกฎระเบียบภายในองค์กรให้กับสมัยอยู่เสมอ ทั้งนี้ ในบางกรณีอาจต้องมีการปรับปรุงระบบให้สอดคล้องกับกฎ ระเบียบ หรือกฎหมายที่เกี่ยวข้อง

๓. ส่งเสริมการประยุกต์ใช้ Generative AI อย่างมีจริยธรรม (Foster a Culture of Ethical Generative Al) พัฒนาองค์ความรู้บุคลากรในการใช้งาน Generative AI อย่างรู้เท่าทัน สร้างความเข้าใจเกี่ยวกับประเด็นด้านจริยธรรมและแนวปฏิบัติที่ดีในการประยุกต์ใช้ Generative AI พร้อมทั้งมีมาตรการตรวจสอบว่ามีการปฏิบัติตามแบวทางที่ดีที่องค์กรกำหนดไว้

๔. กำหนดแนวทางการทำงานร่วมกันระหว่างมนุษย์และ Generative Al (Ensure Human Oversight) สร้างแนวทางการทำงานร่วมกันระหว่างมนุษย์และ Generative AI เพื่อหลีกเลี่ยงการพึ่งพา Generative AI มากเกินไป โดยเพิ่มบทบาทการมีส่วนร่วมของมนุษย์ในกระบวนการทำงานกับ Generative Al ซึ่งจะช่วยลดความเสี่ยงจากความผิดพลาดของ Generative Al

๕. สร้างความร่วมมือระหว่างฝ่ายงานต่าง ๆ ในองค์กร (Promote Interdisciplinary Collaboration) ส่งเสริมการบูรณาการ การทำงานระหว่างฝ่ายต่าง ๆ ในการพัฒนาและใช้งาน Generative Al ตั้งแต่ประเมินความเสี่ยงไปจนถึงการออกแบบกลยุทธ์บริหารจัดการความเสี่ยง

๖. พัฒนาแนวทางกำกับดูแลด้านข้อมูลขององค์กร (Enhance Data Governance) จัดทำกรอบแนวทางการกำกับดูแลการนำข้อมูลไปใช้กับ Generative AI อย่างเหมาะสม โดยคำนึงถึงความสอดคล้องกับหลักจริยธรรม AI การจัดการข้อมูลที่เป็นระบบ การตรวจสอบความถูกต้อง การปกป้องความเป็นส่วนตัว การทำความเข้าใจบริบทของข้อมูลก่อนนำไปฝึกฝนโมเดลการนำข้อมูลไปใช้งานกับ Generative AI ทั้งนี้ควรตรวจสอบการกำกับดูแลข้อมูลในประเด็นต่าง ๆ ข้างต้นอย่างสม่ำเสมอ

๗. เฝ้าติดตาม ประเมินผล และปรับปรุงการใช้งาน (Monitor, Evaluate, and Improve) ควรมีการประเมินผลทั้งก่อนและหลังการนำ Generative Al ไปใช้จริงในองค์กร เพื่อให้แน่ใจว่าการใช้งาน

เป็นไปตามเป้าหมายในบริบทสถานการณ์จริงและควรมีกลไกช่องทางการรับฟังความคิดเห็น (Feedback) จากผู้ใช้งานและนำมาปรับปรุงระบบให้มีประสิทธิภาพ

๘. ประเมินและตรวจสอบผลิตภัณฑ์หรือบริการที่เกี่ยวข้องจากหน่วยงานภายนอก (Monitor and Evaluate Products/Services by External Parties) ประเมินความเสี่ยงและตรวจสอบผลิตภัณฑ์หรือบริการ (เช่น เครื่องมือโมเดล และชุดข้อมูล) จากหน่วยงานภายนอก เพื่อให้แน่ใจว่ามีความสอดคล้องตามนโยบายการจัดการความเสี่ยงขององค์กร รวมถึงมีการติดตามประสิทธิภาพอย่างสม่ำเสมอ เพื่อให้สามารถรับรู้ถึงความเสี่ยงใหม่ที่อาจจะเกิดและปรับแผนจัดการความเสี่ยงตามความจำเป็นเหมาะสม

๙. กำหนดมาตรการและเฝ้าระวังความมั่นคงปลอดภัยทางไซเบอร์ (Establish Cyber Security Mechanisms) จัดทำมาตรการการรักษาความปลอดภัยทางไซเบอร์ เพื่อป้องกันการเข้าถึงข้อมูลและระบบโดยไม่ได้รับอนุญาต (Unauthorized Access) การเจาะระบบ (Hacking) การละเมิดข้อมูล (Data Breaches) การรั่วไหลข้อมูล (Data Leakage) เข้ารหัสข้อมูลสำคัญ อัปเดตโปรโตคอลความปลอดภัยและตรวจสอบสิทธิ์การเข้าถึงเป็นประจำ

๘. ธรรมาภิบาลการใช้ปัญญาประดิษฐ์

ธรรมาภิบาลในการใช้ AI คือ หลักการกำกับดูแลการปฏิบัติงานในทุกกระบวนการที่เกี่ยวข้องกับการประยุกต์ใช้ AI โดยจัดให้มีมาตรการในการกำกับดูแลผ่านการกำหนดนโยบาย ขั้นตอนปฏิบัติและเครื่องมือในการปฏิบัติงาน เพื่อให้มั่นใจได้ว่าการประยุกต์ใช้ AI นั้น สามารถบรรลุตามเป้าหมายขององค์กรอย่างมีความรับผิดชอบ โดยคำนึงถึงความสอดคล้องตามหลักการจริยธรรมปัญญาประดิษฐ์ความสอดคล้องตามกฎหมายและข้อกำหนดที่เกี่ยวข้อง และมีการควบคุมความเสี่ยงที่อาจส่งผลกระทบต่อบุคคลที่เกี่ยวข้ององค์กรและสังคมโดยกว้าง โดยมีคณะกรรมการกำกับดูแลการประยุกต์ใช้ใช้ AI ซึ่งมีหน้าที่หลัก ดังนี้

๑. กำหนดทิศทางการดำเนินการ (Direction) โดยกำหนดกลยุทธ์ในการใช้ AI และนโยบายที่เกี่ยวข้อง

๒. เฝ้าติดตาม (Monitoring) ประสิทธิภาพ (Performance) ของการใช้ AI รวมถึงการปฏิบัติงานตามนโยบายและข้อกำหนดต่าง ๆ (Conformance) ทั้งภายในและภายนอกองค์กร

๓. ประเมินผล (Evaluation) การใช้งาน AI ในปัจจุบันและอนาคต โดยพิจารณาจากปัจจัยที่เกี่ยวข้องทั้งภายในและภายนอก เช่น เป้าหมายในการประยุกต์ใช้ AI ภัยคุกคามและโอกาสจากการประยุกต์ใช้ AI ความเปลี่ยนแปลงด้านเทคโนโลยี ประสิทธิภาพและประสิทธิผลในการกำกับดูแล เป็นต้น

๙. หลักการจริยธรรมปัญญาประดิษฐ์

เพื่อให้การประยุกต์ใช้ AI เป็นไปอย่างมีความรับผิดชอบนั้น จำเป็นต้องมีการนำหลักการจริยธรรมปัญญาประดิษฐ์ที่เกี่ยวข้องมาปรับใช้ โดยองค์กรจะต้องกำหนดกลยุทธ์ในการประยุกต์ใช้ AI (AI Strategy) และนโยบายให้สอดคล้องตามหลักการจริยธรรมที่เกี่ยวข้อง ซึ่งในการพิจารณานำหลักการจริยธรรมปัญญาประดิษฐ์ใดมาปรับใช้นั้นขึ้นอยู่กับบริบทขององค์กรที่นำ AI ไปประยุกต์ใช้ ทั้งนี้ องค์กรอาจพิจารณานำหลักการจริยธรรมปัญญาประดิษฐ์ตามแนวทางของ สำนักงานคณะกรรมการดิจิทัลเพื่อเศรษฐกิจและสังคมแห่งชาติ (สดช.) มาปรับใช้ทั้งหมดหรือบางส่วน หรืออาจนำหลักการจริยธรรมปัญญาประดิษฐ์อื่น ๆ ที่เกี่ยวข้องกับบริบทของ การประยุกต์ใช้ AI มาปรับใช้ตามความเหมาะสม

๑๐. หลักการจริยธรรมปัญญาประดิษฐ์ของไทย ตามแนวทางของ สดช.

๑๐.๑ ความสามารถในการแข่งขันและการพัฒนาอย่างยั่งยืน (Competitiveness and Sustainable Development) AI ควรถูกสร้างและใช้งานเพื่อสร้างประโยชน์และคุณค่าให้แก่มนุษย์ สังคม เศรษฐกิจและสิ่งแวดล้อมอย่างยั่งยืน รวมถึงเพิ่มความสามารถในการแข่งขันและสร้างความเจริญให้กับทุกภาคส่วนในโลกอย่างเป็นธรรม นอกจากนี้ AI ยังควรได้รับการวิจัยและพัฒนาอย่างต่อเนื่องเพื่อสร้างสรรค์นวัตกรรมและอุตสาหกรรมใหม่

๑๐.๒ ความสอดคล้องกับกฎหมาย จริยธรรม และมาตรฐานสากล (Laws, Ethics, and International Standards) AI ควรได้รับการออกแบบ พัฒนาให้บริการและใช้งาน โดยสอดคล้องกับกฎหมาย บรรทัดฐาน จริยธรรม คุณธรรมของมนุษย์ และ มาตรฐานสากลโดยเคารพต่อความเป็นส่วนตัว เกียรติ สิทธิเสรีภาพ และสิทธิมนุษยชน นอกจากนี้ AI ยังควรได้รับการออกแบบโดยมีมนุษย์เป็นศูนย์กลางและเป็นผู้ตัดสินใจ (Human-Centered Design) และไม่ควรถูกออกแบบเพื่อใช้ในการกำหนดชะตาชีวิตของมนุษย์

๑๐.๓ ความโปร่งใสและความรับผิดชอบต่อผลของการกระทำ (Transparency and Accountability) AI ควรได้รับการออกแบบ พัฒนาให้บริการ และใช้งานด้วยความโปร่งใส โดยจะต้องสามารถอธิบาย (Explainability) และคาดการณ์ผลลัพธ์จากการทำงานได้ อีกทั้งยังควรมีการเฝ้าติดตามความผิดปกติ มีความสามารถในการสืบย้อนกลับ (Traceability) และสามารถวินิจฉัยปัญหาและความผิดพลาดได้ (Diagnosability) นอกจากนี้ องค์กรยังควรมีการกำหนดหน้าที่และความรับผิดชอบ (Role and Responsibility) รวมถึงความรับผิดชอบต่อผลของการกระทำ (Accountability) ต่อผลกระทบที่เกิดขึ้นจาก AI ตามหน้าที่ที่รับผิดชอบได้

๑๐.๔ ความมั่นคงปลอดภัยและความเป็นส่วนตัว (Security and Privacy) AI ควรถูกออกแบบและพัฒนาโดยให้ความสำคัญถึงความมั่นคงปลอดภัยและการรักษาความเป็นส่วนตัว โดยจัดให้มีมาตรการป้องกันการโจมตีทางไซเบอร์ และปฏิบัติตามกฎหมายคุ้มครองข้อมูลส่วนบุคคล รวมถึงมีกลไกที่ให้มนุษย์สามารถแทรกแซงเพื่อควบคุมการทำงานของ AI ในกรณีที่เกิดเหตุการณ์หรือความเสี่ยงที่มีผลกระทบกับมนุษย์ได้

๑๐.๕ ความเท่าเทียม หลากหลาย ครอบคลุม และเป็นธรรม (Fairness)ในการส่งเสริมการประยุกต์ใช้ AI อย่างเป็นธรรมและลดความมีอคติ (Bias) ระบบควรถูกออกแบบและพัฒนาโดยคำนึงความหลากหลาย (Diversity) ลดการเอนเอียง แบ่งแยก และเลือกปฏิบัติ (Discrimination) ต่อบุคคลหรือกลุ่มคนที่มีคุณลักษณะที่ต่างกัน (อาทิอายุ เพศ ลักษณะทางกายภาพ เชื้อชาติ) โดยเฉพาะกลุ่มคนผู้ด้อยโอกาสในสังคม รวมถึงสามารถพิสูจน์ถึงความเป็นธรรมสำหรับทุกฝ่ายที่เกี่ยวข้อง

๑๐.๖ ความน่าเชื่อถือ (Reliability) AI ควรได้รับการสนับสนุนให้มีความน่าเชื่อถือและความมั่นใจในการใช้งานต่อสาธารณะ โดยสนับสนุนให้มีการพัฒนาด้วยความยึดมั่นในความถูกต้อง (Accuracy) ความน่าเชื่อถือ (Reliability) สามารถทนทานต่อเหตุการณ์ที่อาจเกิดความผิดพลาด (Robustness) และสามารถสร้างผลลัพธ์ได้เหมือนเดิม (Reproducibility) นอกจากนี้ควรมีการควบคุมคุณภาพของข้อมูล (Quality of data) รวมถึงกำหนดกระบวนการและช่องทางรับความคิดเห็น (Feedback) จากผู้ใช้งานให้ข้อมูลเพิ่มเติม รับเรื่องร้องเรียนแจ้งปัญหาที่พบ และมีการตอบสนองหรือดำเนินการแก้ไขปัญหาที่พบได้ทันท่วงที

๑๑. แนวทางปฏิบัติทางจริยธรรมปัญญาประดิษฐ์

๑๑.๑ ความสามารถในการแข่งขันและการพัฒนาอย่างยั่งยืน (Competitiveness and Sustainability Development)

o ควรตรวจสอบและประเมินงานวิจัย พัฒนาและการนำปัญญาประดิษฐ์ไปใช้งานว่าสามารถสร้างให้เกิดประโยชน์และความผาสุกต่อมนุษย์ สังคม และสิ่งแวดล้อม

o ควรสนับสนุนการวิจัยและพัฒนาในการนำปัญญาประดิษฐ์มาใช้ให้เกิดประโยชน์

o ควรส่งเสริมและสนับสนุนการทำงานร่วมกันระหว่างมนุษย์และปัญญาประดิษฐ์ที่ทำให้เกิดประโยชน์ต่อมนุษย์

o ควรส่งเสริมและสนับสนุนให้หน่วยงานต่าง ๆ สามารถเข้าถึงองค์ความรู้ทางปัญญาประดิษฐ์ด้วยเทคโนโลยีโครงสร้างพื้นฐานทางดิจิทัลและกลไกต่าง ๆ เพื่อให้เกิดการแลกเปลี่ยนข้อมูลความรู้และเทคโนโลยีในการวิจัย ออกแบบและพัฒนา

o ควรเฝ้าระวังและติดตามการใช้งานปัญญาประดิษฐ์ต่อผู้ใช้งาน ว่าก่อให้เกิดประโยชน์ได้จริงและไม่สร้างผลกระทบในเชิงลบ

หน่วยงานรัฐ ควรออกนโยบายเพื่อสนับสนุนให้หน่วยงานทั้งภาครัฐและเอกชนดำเนินการวิจัยและพัฒนา เพื่อกระตุ้นให้เกิดนวัตกรรมปัญญาประดิษฐ์ที่ช่วยในการสร้างอุตสาหกรรมใหม่ควรส่งเสริมและสนับสนุนด้านการศึกษา อบรมประชาชนทั่วไปเพื่อให้มีความรู้ และทักษะเกี่ยวกับปัญญาประดิษฐ์และผลกระทบที่อาจเกิดขึ้นจากปัญญาประดิษฐ์สร้างทัศนคติที่ถูกต้องในการยอมรับการใช้งานเทคโนโลยีปัญญาประดิษฐ์ที่ก่อให้เกิดประโยชน์ และทัศนคติของการตระหนักรู้ภัยคุกคามจากปัญญาประดิษฐ์ ควรสร้างความร่วมมือกับองค์กร ภายในประเทศ ภูมิภาคและนานาชาติ เพื่อพัฒนาโครงการปัญญาประดิษฐ์ที่ก่อให้เกิดประโยชน์ต่อภาพรวมในระดับภูมิภาคและระดับโลก

๑๑.๒ ความสอดคล้องกับกฎหมาย จริยธรรม และมาตรฐานสากล (Laws Ethics and International Standards)

o ควรสนับสนุนการให้การศึกษาเพื่อสร้างความตระหนักรู้และความเข้าใจในปัญญาประดิษฐ์และผลกระทบของปัญญาประดิษฐ์ให้กับผู้ใช้งาน สนับสนุนการอบรมผู้วิจัย และพัฒนาปัญญาประดิษฐ์ เพื่อให้ตระหนักและเข้าใจผลกระทบของระบบปัญญาประดิษฐ์ที่มีต่อบุคคลและสังคม และให้การสนับสนุนงานวิจัยเกี่ยวกับปัญญาประดิษฐ์กับสิทธิมนุษยชน

o การจัดซื้อจัดจ้างควรกำหนดให้ผู้พัฒนา และผู้ให้บริการปัญญาประดิษฐ์ต้องดำเนินการสอดคล้องกับหลักเกณฑ์ของความโปร่งใส มีการประเมินผลกระทบ จากการประมวลผลข้อมูลต่อสิทธิมนุษยชนและความเป็นส่วนตัวพร้อมกำหนดภาระความรับผิดชอบต่อผลกระทบเชิงลบที่เกิดขึ้นกับปัญญาประดิษฐ์

หน่วยงานรัฐ ควรสนับสนุนในการสร้างมาตรฐานให้เป็นที่ยอมรับโดยสากลและแนวทางปฏิบัติที่ดีที่สุด (Best Practice) และกำหนดให้ผู้ออกแบบผู้พัฒนา และผู้ให้บริการต้องปฏิบัติตาม พร้อมกลไกการให้ใบรับรองสำหรับปัญญาประดิษฐ์ เพื่อเพิ่มความน่าเชื่อถือให้กับผู้ออกแบบผู้พัฒนา และผู้ให้บริการรวมถึงผลิตภัณฑ์และบริการที่เกี่ยวข้องและควรสนับสนุนให้มีองค์กรต่าง ๆ มีหน่วยงานกำกับดูแลการพัฒนาและใช้งานปัญญาประดิษฐ์เพื่อให้คำปรึกษาผู้วิจัย ผู้ออกแบบ ผู้พัฒนาและผู้ให้บริการปัญญาประดิษฐ์ เกี่ยวกับ

ผลกระทบด้านกฎหมายจริยธรรม และสิทธิมนุษยชน ประเมินความเสี่ยงที่มีผลระทบเชิงลบของระบบ และดูแลด้านการสร้าง ภาระความรับผิดชอบของผู้วิจัยผู้ออกแบบ ผู้พัฒนา และผู้ให้บริการปัญญาประดิษฐ์

๑๑.๓ ความโปร่งใสและภาระความรับผิดชอบ (Transparency and Accountability)

หน่วยงานกำกับดูแลการพัฒนาและใช้งานปัญญาประดิษฐ์ควรตรวจสอบความโปร่งใสของโมเดลและอัลกอริทึม ทีมที่ผู้วิจัย ผู้ออกแบบ และผู้พัฒนาใช้ในปัญญาประดิษฐ์ โดยใช้หลักการเรื่องความสามารถในการอธิบายได้ (Explainable) ซึ่งอัลกอรีทึมที่นำมาใช้ต้องสามารถอธิบายที่มา และหน้าที่การทำงานในการคาดการณ์ต่าง ๆ รวมถึงสามารถอธิบายวิธีการสอนและคัดเลือกโมเดลได้หน่วยงานกำกับดูแลการพัฒนาและใช้งานปัญญาประดิษฐ์ควรกำหนดนโยบายด้านการให้คำอธิบายปัญญาประดิษฐ์กับผู้ใช้งานขึ้น โดยกำหนดให้ผู้วิจัยออกแบบ และพัฒนาปัญญาประดิษฐ์ สร้างเอกสารทางเทศนิคเพื่อแสดงรายละเอียดออกแบบและการทำงานในหลายมุมมอง เพื่อรองรับผู้ใช้งานที่มีความรู้ความเข้าใจที่แตกต่างกัน และสถานการณ์ที่แตกต่างกันของการนำใช้และกำหนดให้มีช่องทางที่ง่ายและรวดเร็วเพื่อให้ ผู้ที่อาจได้รับผลกระทบส่งการร้องขอคำการอธิบายเหล่านี้ได้

o ควรมีกลไกเพื่อสร้างภาระความรับผิดชอบ (Accountability) ของผู้วิจัยผู้ออกแบบ ผู้พัฒนา ผู้ให้บริการและผู้ใช้งานปัญญาประดิษฐ์ รวมถึงผู้ที่มีส่วนได้เสียอื่น ๆ ตลอดวัฏจักรชีวิตของระบบ โดยมีกระบวนการในการตรวจสอบชุดข้อมูล อัลกอริทึม กระบวนการออกแบบ การนำไปใช้และผลลัพธ์ของการใช้งานปัญญาประดิษฐ์ ทั้งการตรวจสอบภายในและภายนอกอย่างอิสระ การรายงานผลการตรวจสอบ การประเมินความเสียงเชิงลบและการดำเนินการเพื่อลดหรือหลีกเลี่ยง ผลกระทบเชิงลบของปัญญาประดิษฐ์ โดยเฉพาะผลกระทบต่อมนุษย์และสิ่งแวดล้อม

o กำหนดผู้รับผิดชอบในการสืบสวน และแก้ไขสาเหตุของความสูญเสียและเสียหายที่เกิดขึ้นจากปัญญาประดิษฐ์

๑๑.๔ ความมั่นคงปลอดภัยและความเป็นส่วนตัว (Security and Privacy)

o ควรกำหนดนโยบายและมาตรฐานทางเทคนิคด้านความมั่นคงปลอดภัยและการคุ้มครองความเป็นส่วนตัวสำหรับปัญญาประดิษฐ์ เพื่อลดช่องโหว่และป้องกันภัยคุกคามของปัญญาประดิษฐ์ที่ก่อให้เกิดผลกระทบในด้านความลับความครบถ้วนถูกต้อง ความพร้อมใช้ของข้อมูล การคุ้มครองข้อมูลบุคคล รวมถึงผลกระทบด้านจริยธรรม ชีวิตและสิ่งแวดล้อม และให้ผู้วิจัย ออกแบบพัฒนา และให้บริการปฏิบัติตาม

o หน่วยงานกำกับดูแลการพัฒนาและใช้งานปัญญาประดิษฐ์ ควรดำเนินการจัดการความเสี่ยง กำหนดวิธีการจัดการความเสี่ยงและควบคุมภายในเพื่อทำหน้าที่จัดการความเสี่ยงที่เกี่ยวข้องกับปัญญาประดิษฐ์ หาแนวทางในการจัดการความเสี่ยง ตรวจสอบและรายงาน ประสิทธิภาพในการจัดการความเสี่ยง ทบทวนและปรับปรุงแนวทางและกระบวนการจัดการความเสี่ยง ทั้งนี้ควรกำหนดให้มีการดำเนินการจัดการความเสี่ยงอย่างสม่ำเสมอตลอดวัฏจักรชีวิตของระบบ รวมถึงเมื่อจำเป็นต้องรื้อถอนระบบ โดยดำเนินการอย่างน้อยปีละ ๑ ครั้งหรือมีการเปลี่ยนแปลงที่สำคัญเกิดขึ้น

o หน่วยงานกำกับดูแลการพัฒนาและใช้งานปัญญาประดิษฐ์ควรประเมินความเสี่ยงเชิงลบที่กระทบกับบุคคลและสังคมจากการเพิกเฉยข้อมูลหรือสถานะการณ์เฉพาะ (De-contextualised data) และอัลกอริทึมที่เพิกเฉยกับบริบทหรือสถานะการณ์เฉพาะอย่างเพียงพอระหว่างขั้นตอน การพัฒนาและนำปัญญาประดิษฐ์ไปใช้งานหน่วยงานกำกับดูแลการพัฒนาและใช้งานปัญญาประดิษฐ์ควรดำเนินการ

ประเมินระดับการเข้าแทรกแซงปัญญาประดิษฐ์โดยมนุษย์ในกระบวนการต่าง ๆ จากโอกาสและความรุนแรงของผลกระทบที่จะเกิดขึ้นกับผู้ใช้งาน

o หน่วยงานกำกับดูแลการพัฒนาและใช้งานปัญญาประดิษฐ์ควรดำเนินการประเมินระดับการเข้าแทรกแซงปัญญาประดิษฐ์โดยมนุษย์ในกระบวนการต่าง ๆ จากโอกาสและความรุนแรงของผลกระทบที่จะเกิดขึ้นกับผู้ใช้งาน

o หน่วยงานกำกับดูแลการพัฒนาและใช้งานปัญญาประดิษฐ์ ควรให้บุคคลหรือกลุ่มคนที่ได้รับผลกระทบที่สำคัญจากระบบปัญญาประดิษฐ์ ได้มีส่วนร่วมในกระบวนการประเมินความเสี่ยงด้วย

o หน่วยงานกำกับดูแลการพัฒนาและใช้งานปัญญาประดิษฐ์ ควรมีการทบทวนประสิทธิภาพการทำงานของหน่วยงานและเจ้าหน้าที่ในโครงสร้างการกำกับดูแลอย่างสม่ำเสมอ และทุกครั้งที่มีการเปลี่ยนแปลงโครงสร้างหรือบุคคลที่สำคัญ

หน่วยงานรัฐ ควรมีการวางแผนเพื่อกำกับดูแล เฝ้าระวังและจัดการความเสี่ยงในระยะยาว โดยเฉพาะเมื่อปัญญาประดิษฐ์มีความฉลาดมากยิ่งขึ้นกว่าในปัจจุบัน เช่น ปัญญาประดิษฐ์ทั่วไป (Ationial General Inteligence (AG)) ซุปเปอร์ปัญญาประดิษฐ์ (Articial Superintelligence) และปัญญาประดิษฐ์ที่มีความสามารถในการปรับปรุงตนเองได้อย่างต่อเนื่อง (Recursive Self-Improving Al) ควรสนับสนุนให้เกิดความร่วมมือเพื่อพัฒนาโครงสร้างการกำกับดูแลปัญญาประดิษฐ์ในแบบบูรณาการ ทั้งในระดับองค์กร ประเทศ ภูมิภาคและให้ความร่วมมือกับนานาชาติ ในการหลีกเลี่ยงการแข่งขันด้านปัญญาประดิษฐ์ที่ไม่พึงประสงค์ รวมถึงการสร้างอาวุธอัตโนมัติจากปัญญาประดิษฐ์ที่ร้ายแรง สนับสนุนให้เกิดการแลกเปลี่ยนความรู้และประสบการณ์ในการกำกับดูแล และร่วมกันกับมือกับผลกระทบของปัญญาประดิษฐ์

๑๑.๕ ความเท่าเทียม หลากหลาย ครอบคลุม และเป็นธรรม (Fairness)

o ควรกำหนด ส่งเสริมและสนับสนุนแนวทางปฏิบัติที่เกี่ยวข้องกับการประยุกต์ใช้ปัญญาประดิษฐ์ที่หลากหลายตามชนิดและสถานการณ์ในการใช้งาน

o หน่วยงานกำกับดูแลการพัฒนาและใช้งานปัญญาประดิษฐ์ ควรมีกระบวนการควบคุมเพื่อวิเคราะห์ ประเมินความเสี่ยง และจัดการปัญหาการเอนเอียงไปสู่ความไม่เป็นธรรมในขั้นตอนการวิจัย ออกแบบ พัฒนา และให้บริการปัญญาประดิษฐ์อย่างชัดเจนและโปร่งใส

o หน่วยงานรัฐควรส่งเสริมให้เกิดแพลตฟอร์มเปิดของปัญญาประดิษฐ์เพื่อหลีกเลี่ยงการผูกขาด ทำให้เกิดการแลกเปลี่ยนความรู้ในการพัฒนาปัญญาประดิษฐ์เพื่อสร้างสรรค์ ต่อยอดองค์ความรู้ และพัฒนาอย่างต่อเนื่องเพื่อใช้งานในระดับอุตสาหกรรมได้

o ควรสนับสนุนให้เกิดโอกาสในการพัฒนาปัญญาประดิษฐ์ที่เท่าเทียมกันและการแข่งขันที่เป็นธรรม ทั้งในระดับภูมิภาคและอุตสาหกรรมที่แตกต่างกัน

o ควรส่งเสริมและสนับสนุนให้เกิดโอกาสที่เท่าเทียมกัน ในการเข้าถึงการศึกษา สินค้า บริการ และเทคโนโลยีที่เกี่ยวข้องด้านปัญญาประดิษฐ์

o ควรส่งเสริมให้ประชาชนมีทักษะด้านปัญญาประดิษฐ์และสนับสนุนกลุ่มคนทำงานเพื่อการเปลี่ยนผ่านที่เป็นธรรม (Fair Transition) หน่วยงานรัฐควรส่งเสริมให้เกิดการวิจัยและพัฒนาเทคนิคการวิเคราะห์เพื่อใช้ตรวจสอบและแก้ไขปัญหาความเอนเอียง แบ่งแยกและไม่เป็นธรรมของปัญญาประดิษฐ์

๑๑.๖ ความน่าเชื่อถือ (Reliability)

o ควรกำหนดนโยบาย หลักเกณฑ์ และกระบวนการในการประเมินคุณภาพของชุดข้อมูลและโมแดลปัญญาประดิษฐ์ ดำเนินการพบทวนและปรับปรุงความน่าเชื่อถือของปัญญาประดิษฐ์และชุดข้อมูลอย่างสม่ำเสมอ โดยนำข้อมูลผลสะท้อนกลับที่ได้รับจากผู้ใช้งานระบบจริงมาใช้ เพื่อให้ระบบสอดรับกับการเปลี่ยนแปลงพฤติกรรมของผู้ใช้งานตามระยะเวลา ควรปรับปรุงโมเดลด้วยชุดข้อมูลที่เป็นปัจจุบัน และควรปรับปรุงโมเดลเมื่อวัตถุประสงค์และความเสี่ยงเปลี่ยนแปลงไป

o ควรกำหนดนโยบายและกระบวนการเพื่อทบทวนช่องทางการสื่อสารกับผู้ใช้งาน และดำเนินการทบทวนอย่างสม่ำเสมอ เพื่อสนับสนุนการเปิดเผยข้อมูลที่จำเป็น และรับผลสะท้อนจากการให้บริการได้อย่างมีประสิทธิภาพ

หน่วยงานรัฐ ควรสนับสนุนการวิจัย ออกแบบ พัฒนา ให้บริการ และใช้งานปัญญาประดิษฐ์ที่มีความน่าเชื่อถือ

๑๒. กรอบการทำงานเพื่อสนับสนุนธรรมาภิบาลในการประยุกต์ใช้ AI

๑๒.๑ Al Governance Structure

o จัดตั้งคณะกรรมการกำกับดูแล (AI Governance Council) เพื่อกำหนดทิศทางการประยุกต์ใช้งาน AI ผ่านการกำหนดกลยุทธ์และนโยบาย รวมถึงเฝ้าติดตาม และประเมินผลกาประยุกต์ใช้ AI อย่างต่อเนื่อง เพื่อสนับสนุนให้เกิดธรรมาภิบาลในการประยุกต์ใช้ AI

o กำหนดหน้าที่ของบุคลากรและผู้มีส่วนได้เสียที่เกี่ยวข้องกับการประยุกต์ใช้ AI พร้อมทั้งสร้างความตระหนักรู้ในด้านความรับผิดชอบ (Responsibility) และความรับผิดชอบต่อผลของการกระทำ (Accountability) ของแต่ละหน้าที่พัฒนาศักยภาพของบุคลากรและผู้มีส่วนได้ส่วนเสียที่เกี่ยวข้อง เพื่อให้สามารถปฏิบัติงานได้อย่างเหมาะสมตามหน้าที่ที่ได้รับมอบหมาย

o พัฒนาศักยภาพของบุคลากรและผู้มีส่วนได้ส่วนเสียที่เกี่ยวข้อง เพื่อให้สามารถปฏิบัติงานได้อย่างเหมาะสมตามหน้าที่ที่ได้รับมอบหมาย

o ควรกำหนดนโยบายที่ช่วยเปิดโอกาสให้เกิดการใช้งานปัญญาประดิษฐ์ที่มีความน่าเชื่อถือ และพัฒนาแนวทางการประเมินความน่าเชื่อถือของปัญญาประดิษฐ์เพื่อใช้ตรวจประเมินผู้พัฒนาและผู้ให้บริการ

- แนวทางการประเมินควรสามารถปรับเปลี่ยนได้ตามลักษณะการนำปัญญาประดิษฐ์ไปใช้งาน
- แนวทางการประเมินควรได้รับการพัฒนาขึ้นด้วยความร่วมมือของผู้ที่มีส่วนได้เสียทั้งภาครัฐและเอกชน

o ควรประเมินและตรวจสอบคุณภาพผู้ให้บริการปัญญาประดิษฐ์ที่ใช้ข้อมูลจากผลลัพธ์ของระบบเพื่อการตัดสินใจที่สำคัญที่เกี่ยวข้องกับมนุษย์อย่างเข้มงวด

๑๒.๒ Al Strategy

o มองหาโอกาสในการนำ AI มาประยุกต์ใช้เพื่อสนับสนุนให้บรรลุเป้าหมายขององค์กรหรือเป้าหมายทางธุรกิจ

o กำหนดเป้าหมายในการประยุกต์ใช้ AI ตามลำดับความสำคัญ โดยพิจารณาจากประโยชน์ที่จะได้รับ ความพร้อมขององค์กร หลักการจริยธรรมปัญญาประดิษฐ์ กฎหมายและข้อกำหนดที่ต้องดำเนินการให้สอดคล้อง รวมถึง ความซับซ้อนและเวลาที่จำเป็นต้องใช้ในการดำเนินการ

o กำหนดกลยุทธ์ในการบริหารจัดการข้อมูล

o กำหนดแผนปฏิบัติงานในการประยุกต์ใช้ AI (AI Roadmap)

o วิเคราะห์ความเสี่ยงและผลกระทบที่อาจเกิดขึ้นจากการประยุกต์ใช้ AI รวมถึงการกำหนดระดับการมีส่วนร่วมของมนุษย์ในการทำงานของ AI และมาตรการในการควบคุมความเสี่ยงที่เหมาะสม เพื่อควบคุมความเสี่ยงให้อยู่ในขอบเขตที่ยอมรับได้

๑๒.๓ AI Operation

o มองหาโอกาสในการนำ AI มาประยุกต์ใช้เพื่อสนับสนุนให้บรรลุเป้าหมายขององค์กรหรือเป้าหมายทางธุรกิจ

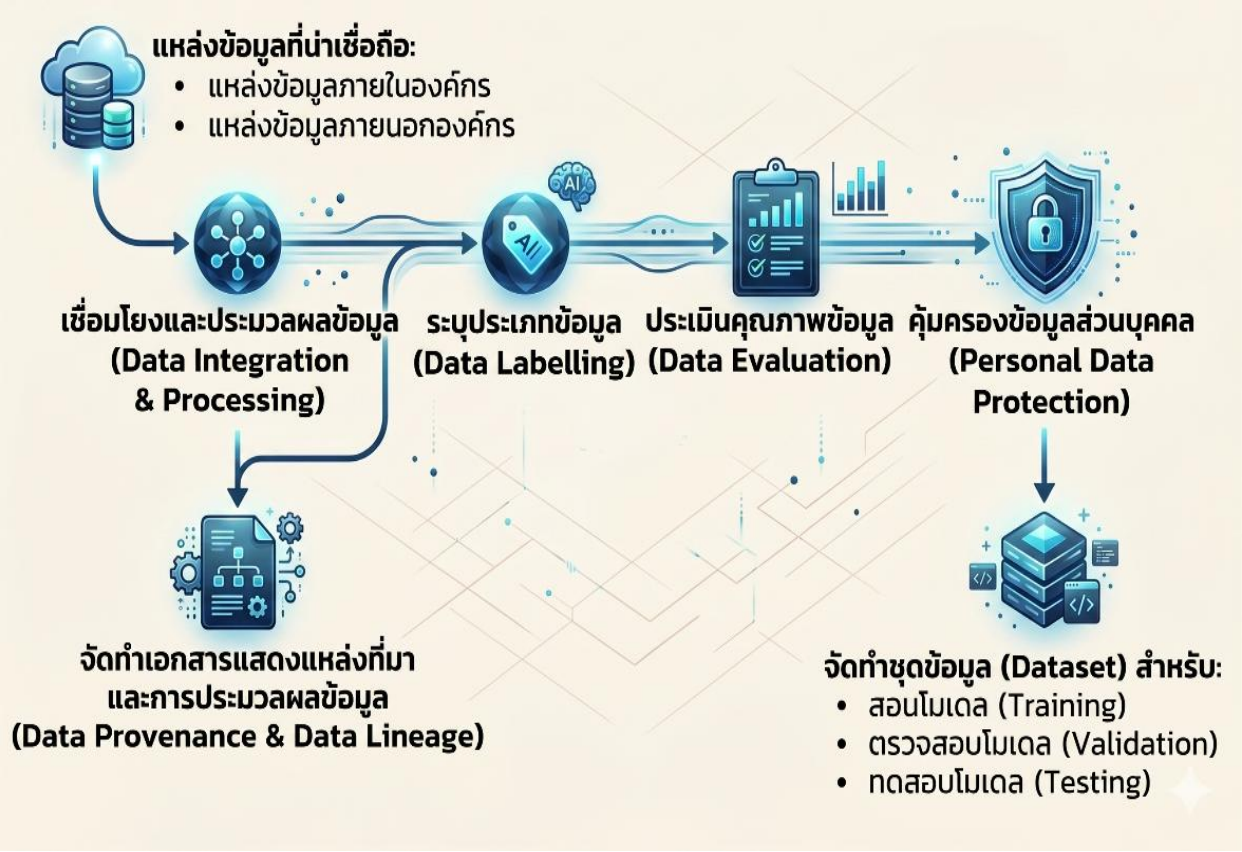
o กำหนดเป้าหมายในการประยุกต์ใช้ AI ตามลำดับความสำคัญ โดยพิจารณาจากประโยชน์ที่จะได้รับ ความพร้อมขององค์กร หลักการจริยธรรมปัญญาประดิษฐ์ กฎหมายและข้อกำหนดที่ต้องดำเนินการให้สอดคล้อง รวมถึง ความซับซ้อนและเวลาที่จำเป็นต้องใช้ในการดำเนินการ

o กำหนดกลยุทธ์ในการบริหารจัดการข้อมูล

o กำหนดแผนปฏิบัติงานในการประยุกต์ใช้AI (AI Roadmap)

o วิเคราะห์ความเสี่ยงและผลกระทบที่อาจเกิดขึ้นจากการประยุกต์ใช้ AI รวมถึงการกำหนดระดับการมีส่วนร่วมของมนุษย์ในการทำงานของ AI และมาตรการในการควบคุมความเสี่ยงที่เหมาะสม เพื่อควบคุมความเสี่ยงให้อยู่ในขอบเขตที่ยอมรับได้

๑๒.๔ AI Lifecycle



ขั้นตอนการจัดเตรียมข้อมูลสำหรับการประยุกต์ใช้ AI

๑) กำหนดคุณสมบัติและคุณภาพของข้อมูลที่เหมาะสม เพื่อให้ AI สามารถทำงานได้อย่างมีประสิทธิภาพตามเป้าหมายที่กำหนด

๒) สรรหาแหล่งข้อมูล ทำการเชื่อมโยงข้อมูลและประมวลผลข้อมูลเพื่อเตรียมข้อมูลสำหรับสอน ตรวจสอบ และทดสอบโมเดล

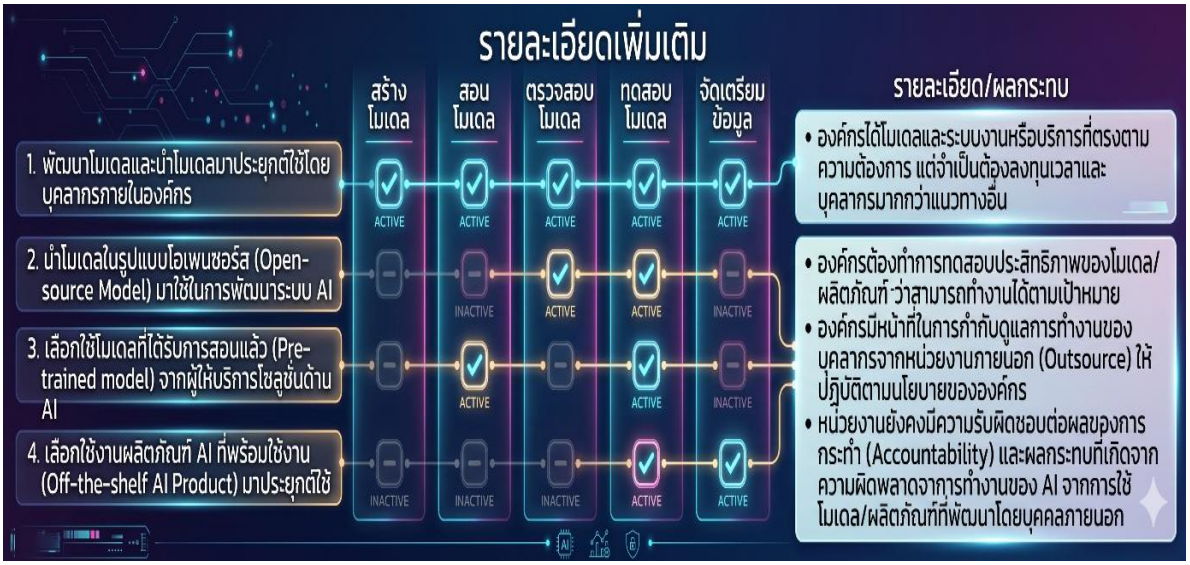
๓) จัดทำเอกสารแสดงแหล่งที่มาและการประมวลผลข้อมูล (Data Provenance & Data Lineage)

๔) ระบุข้อมูล (Labelling Data) อย่างถูกต้อง เพื่อให้ AI มีข้อมูลและรับรู้ความหมายของข้อมูลได้อย่างถูกต้อง

๕) ประเมินคุณภาพของข้อมูล (Data Evaluation) และปรับปรุงข้อมูลให้มีคุณภาพตามที่องค์กรกำหนด

๖) จัดให้มีการเก็บข้อมูลส่วนบุคคลอย่างมั่นคงปลอดภัยและดำเนินมาตรการเพื่อคุ้มครองข้อมูลส่วนบุคคล เช่น ป้องกับกับการเข้าถึง (Access Control) msเข้ารหัสลับข้อมูล (Encryption) การทำข้อมูลส่วนบุคคลให้เป็นข้อมูลนิรมาม (Anononymization) เป็นต้น

แนวทางในการสร้างโมเดลและนำโมเดลมาปรับใช้ ดังนี้



๑๒.๕ AI Service Provision

การสื่อสารนโยบายในการให้บริการ และข้อมูลที่เกี่ยวข้องกับการใช้งาน AI รวมถึงเปิดให้มีช่องทางการสื่อสารกับผู้ใช้งานเพื่อรับเสียงสะท้อนจากการใช้งานจริง จะช่วยสนับสนุนให้เกิดความโปร่งใสในการบริการ เช่น

๑. ประกาศนโยบายและข้อมูลทั่วไปเดียวกับการใช้งาน AI (Policy and General Disclosure) เช่น นโยบายในการใช้งาน AI (AI Usage Policy) แนวทางการปฏิบัติตามหลักการจริยธรรมปัญญาประดิษฐ์ (AI Ethics Principles) นโยบายด้านความมั่นคงปลอดภัย (Security Policy) นโยบายความเป็นส่วนตัว (Privacy Policy) เป็นต้น

๒. ให้ข้อมูลเกี่ยวกับการใช้งาน AI แก่ผู้ใช้งาน เช่น

o แจ้งผู้ใช้งานทราบว่ากำลังใช้งาน รับบริการ หรือทำงานร่วมกับ AI ผ่านการแจ้งเตือนบนแอพพลิเคชั่น

o แจ้งวิธีการใช้งาน ข้อห้ามในการใช้งาน ความสามารถ ข้อจำกัด ผลลัพธ์จากการตัดสินใจของ AI รวมถึงวิธีการและเหตุผลเบื้องหลังการทำงานของ AI ผ่านทางคู่มือการใช้งาน คำถามที่มีที่พบบ่อย (FAQ) และข้อตกลงการใช้บริการ (Terms and Conditions) เป็นต้น

o กรณีที่ระบบเปิดโวทาสให้ผู้ใช้งานปิดการทำงานของ AI ได้ด้วยตนเอง เช่น รถยนต์ที่มีระบบการขับเคลื่อนอัตโนมัติด้วยตนเอง เป็นต้น องค์กรควรมีการสื่อสารเพื่อให้ผู้ใช้งานทราบถึงขั้นตอนการปิดการทำงานดังกล่าว

๓. เปิดช่องทางการติดต่อสื่อสารเพื่อเปิดรับความคิดเห็น (Feedbock) ประเด็นปัญหา (Issue) และความผิดพลาด (Error) เพื่อให้องค์กรสามารถแก้ไขปัญหาได้ทันท่วงที อีกทั้งนำมาปรับปรุงและป้องกันปัญหาที่อาจเกิดขึ้นในอนาคต

หลักการจริยธรรมปัญญาประดิษฐ์ ตามแนวของ สดช. (ONDE's AI Ethics Principles)	องค์ประกอบในการสนับสนุนธรรมาภิบาลในการประยุกต์ใช้ AI		
	การกำหนดโครงสร้างการกำกับดูแล (AI Governance Structure)	การกำหนดกลยุทธ์ในการประยุกต์ใช้ AI (AI Strategy)	การกำกับดูแลการปฏิบัติงานที่เกี่ยวข้องกับ AI (AI Operation)
1. ความสามารถในการแข่งขันและการพัฒนาอย่างยั่งยืน (Competitiveness and Sustainable Development)	✓	✓	
2. ความสอดคล้องกับกฎหมาย จริยธรรม และมาตรฐานสากล (Laws, Ethics, and International Standards)	✓	✓	✓
3. ความโปร่งใสและความรับผิดชอบต่อผลของการกระทำ (Transparency and Accountability)	✓		✓
4. ความมั่นคงปลอดภัยและความเป็นส่วนตัว (Security and Privacy)			✓
5. ความเท่าเทียม หลากหลาย ครอบคลุม และเป็นธรรม (Fairness)			✓
6. ความน่าเชื่อถือ (Reliability)			✓

ตารางแสดงความสัมพันธ์ระหว่างหลักการจริยธรรมปัญญาประดิษฐ์และองค์ประกอบหลักในการสนับสนุนธรรมาภิบาลในการประยุกต์ใช้ AI

เอกสารอ้างอิง

กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม. *แนวทางจริยธรรมปัญญาประดิษฐ์ (AI Ethics Guideline)*. กรุงเทพฯ: กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม.

ศักดิ์ เสกขุนทด. *หลักการจริยธรรมปัญญาประดิษฐ์ การใช้ AI อย่างรับผิดชอบ ธรรมาภิบาลในการประยุกต์ใช้ AI* [เอกสารประกอบการบรรยาย]. สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์.

ศูนย์ธรรมาภิบาลปัญญาประดิษฐ์ (AIGC). *แนวทางการประยุกต์ใช้ Generative AI อย่างมีธรรมาภิบาลสำหรับองค์กร (Generative AI Governance Guideline for Organizations)*. สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์ (ETDA).

สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์. *แนวทางการประยุกต์ใช้ปัญญาประดิษฐ์อย่างมีธรรมาภิบาล สำหรับผู้บริหารองค์กร*. สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์ (ETDA)

กรมการแพทย์แผนไทย และการแพทย์ทางเลือก

กลุ่มงานบริหารจัดการข้อมูลขนาดใหญ่และ
ธรรมาภิบาลข้อมูล

2569

0 2149 5697 (1412)

dme.dtam@gmail.com