



นโยบายและแนวทางการกำกับดูแลการใช้งานปัญญาประดิษฐ์ (AI Governance) การประปาส่วนภูมิภาค พ.ศ. ๒๕๖๙

๑. บทนำ

ปัญญาประดิษฐ์ (Artificial Intelligence หรือ AI) รวมถึง Generative AI มีความสามารถในการสร้างสรรค์เนื้อหาได้หลากหลายรูปแบบ เช่น ข้อความ ภาพ วิดีโอ เสียง ชุดคำสั่งหรือโปรแกรม (Source Code) เป็นต้น โดยการสั่งการผ่านข้อความ หรือคำสั่ง (Prompt) ที่ผู้ใช้งานเป็นผู้กำหนด อีกทั้งยังสามารถสร้างเนื้อหาได้อย่างสมจริงและโต้ตอบกับผู้ใช้งานได้คล้ายคลึงกับมนุษย์ กำลังเปลี่ยนแปลงวิธีการทำงานอย่างรวดเร็ว การประปาส่วนภูมิภาค (กปภ.) ในฐานะหน่วยงานสาธารณูปโภคขั้นพื้นฐานและเป็นโครงสร้างพื้นฐานสารสนเทศที่สำคัญของชาติ (Critical Information Infrastructure - CII) ตระหนักถึงศักยภาพของ AI ในการยกระดับประสิทธิภาพการดำเนินงาน ปรับปรุงการให้บริการ และบริหารจัดการทรัพยากรน้ำอย่างยั่งยืน

อย่างไรก็ตาม เมื่อมีการนำเทคโนโลยี Generative AI มาใช้ในการสร้างสรรค์ และเกิดข้อมูลเท็จที่ไม่ได้ตั้งใจให้เกิดความเข้าใจผิด (Misinformation) และข้อมูลเท็จที่จงใจให้เกิดความเข้าใจผิด (Disinformation) กรณีเช่นนี้จะถูกจัดเป็นหนึ่งในความเสี่ยงระดับโลกที่รุนแรงที่สุด ตามรายงาน Global Risks Report 2023 - 2024 ของ World Economic Forum ที่เนื้อหาที่ได้จากการสร้างสรรค์จากเทคโนโลยี Generative AI ที่อาจทำให้ผู้ใช้งานเชื่อได้ว่าข้อมูลนั้นถูกต้อง มีความน่าเชื่อถือและสามารถนำไปใช้งานได้ทันที โดยไม่จำเป็นต้องพิจารณาถึงความเหมาะสมและข้อจำกัดของเทคโนโลยี โดยเฉพาะอย่างยิ่งข้อจำกัดทั้งด้านภาษาไทย ความเข้าใจในบริบทหรือวัฒนธรรมของประเทศไทย หรือบริบทของหน่วยงาน ซึ่งอาจก่อให้เกิดผลกระทบในเชิงลบต่อบุคคล องค์กร สังคมและประเทศชาติ เช่น

- ข้อมูลเท็จที่ไม่ได้ตั้งใจให้เกิดความเข้าใจผิด (Misinformation) และข้อมูลเท็จที่จงใจให้เกิดความเข้าใจผิด (Disinformation)
- เนื้อหาไม่ถูกต้องตามข้อเท็จจริง (Confabulation)
- เนื้อหาที่มีความอันตรายหรือรุนแรง (Dangerous or Violent Recommendations)
- เนื้อหาส่งผลกระทบต่อความเป็นส่วนตัวของข้อมูล (Data Privacy)
- เนื้อหาส่งผลกระทบต่อประเด็นที่มีความอ่อนไหว และมีเนื้อหาที่ขัดต่อหลักจริยธรรม (Sensitive and Ethics Context)
- เนื้อหาที่มีข้อมูลที่ไม่ถูกต้องครบถ้วนหรือถูกทำให้บิดเบือน (Information Integrity)
- เนื้อหาที่สร้างก่อให้เกิดอคติ แบ่งแยกและการเลือกปฏิบัติ (Bias and Discrimination)

• ข้อมูลส่วนบุคคล...

- ข้อมูลส่วนบุคคลรั่วไหล (Personal Data Leakage)
 - การละเมิดทรัพย์สินทางปัญญา (Intellectual Property Infringement)
 - ความมั่นคงปลอดภัยของข้อมูล (Information Security)
 - ผลกระทบที่อาจเกิดขึ้นจากเทคโนโลยี Generative AI หรือเทคโนโลยีอื่นที่มีลักษณะเดียวกัน
- ดังนั้น จึงสมควรกำหนดนโยบายการกำกับดูแลการใช้งานปัญญาประดิษฐ์ (AI Governance)

เพื่อให้ใช้งานเทคโนโลยี Generative AI ที่ยอมรับได้ (Acceptable Use Policy: Generative AI) และนำมา
รังสรรค์งานของการประปาส่วนภูมิภาคขึ้น

๒. วัตถุประสงค์

๒.๑ เพื่อกำหนดโครงสร้างการกำกับดูแล กลยุทธ์ และทิศทางการประยุกต์ใช้ AI ใน กปภ.

๒.๒ เพื่อให้ ผู้บริหาร พนักงาน ผู้รับจ้าง และ ลูกจ้าง กปภ. มีแนวทางปฏิบัติ (Acceptable Use) ในการใช้งาน AI อย่างถูกต้อง เหมาะสม และมีประสิทธิภาพ

๒.๓ เพื่อให้การใช้เทคโนโลยี Generative AI เป็นไปตามกฎหมาย กฎ ระเบียบ ประกาศ
ข้อบังคับ หรืออื่น ๆ ที่เกี่ยวข้อง ตลอดจนป้องกันความเสี่ยงที่เกิดจากการใช้งานผิดวัตถุประสงค์ ที่ก่อให้เกิด
ผลกระทบที่เกิดขึ้นกับบุคคล สำนักงาน สังคมและประเทศชาติ

๒.๔ เพื่อป้องกันความเสี่ยงและผลกระทบเชิงลบต่อบุคคล องค์กร สังคม และให้สอดคล้องกับ
กฎหมาย ระเบียบ และมาตรฐานสากล (เช่น ISO/IEC 42001, PDPA, Cybersecurity Act)

๒.๕ เพื่อเสริมสร้างความตระหนักรู้และความเข้าใจเกี่ยวกับการใช้เทคโนโลยี Generative AI
อย่างมีความรับผิดชอบ รอบคอบ สามารถใช้งานได้อย่างถูกต้องเหมาะสมและเป็นไปตามแนวทางที่ กปภ. กำหนดไว้

๓. ขอบเขต

ขอบเขตของนโยบายฉบับนี้มีเนื้อหาครอบคลุมสำหรับการนำเทคโนโลยี Generative AI มาใช้
เพื่อการรังสรรค์งานตามภารกิจของ กปภ. ที่กำหนดให้ ผู้บริหาร พนักงาน ผู้รับจ้าง และ ลูกจ้าง กปภ. ที่ได้รับ
มอบหมายให้ดำเนินงานตามภารกิจของ กปภ. (ต่อไปนี้จะเรียกว่า "ผู้ใช้งาน") ที่ใช้งานเทคโนโลยี Generative AI ใน
การดำเนินการตามภารกิจที่กฎหมายกำหนดไว้

๔. โครงสร้างการกำกับดูแลและกลยุทธ์ (AI Governance Structure & Strategy)

เพื่อให้การใช้งาน AI มีทิศทางที่ชัดเจนและสอดคล้องกับเป้าหมายองค์กร กปภ. กำหนด
โครงสร้างและกลยุทธ์ ดังนี้

- คณะกรรมการกำกับดูแลการประยุกต์ใช้ AI (AI Governance Council) ทำหน้าที่โดย คณะกรรมการ
กำกับดูแลเทคโนโลยีดิจิทัลของ กปภ. (DTSC) มีหน้าที่กำหนดเป้าหมาย ขอบเขต นโยบาย ทิศทางการ
ดำเนินงาน ความเสี่ยงระดับองค์กร และงบประมาณดำเนินงาน
- คณะทำงานกำกับดูแลการใช้งานปัญญาประดิษฐ์ ทำหน้าที่พิจารณากันกรอง ติดตามการใช้งาน
ประเมินภัยคุกคาม และบริหารจัดการความเสี่ยงให้อยู่ในระดับที่ยอมรับได้

๕. คำนิยาม

“สำนักงาน/หน่วยงาน” หมายถึง การประปาส่วนภูมิภาค หรือ กปภ.

“ผู้บริหารสูงสุด” หมายถึง ผู้ว่าการการประปาส่วนภูมิภาค

“ผู้บังคับบัญชา” หมายถึง หัวหน้าหน่วยงานและผู้บริหารของ กปภ.

“ผู้ปฏิบัติงาน” หมายถึง พนักงาน กปภ.

“ผู้รับจ้าง” หมายถึง บุคคลหรือนิติบุคคลที่ลงนามเป็นคู่สัญญากับสำนักงาน รวมทั้ง ตัวแทนหรือลูกจ้าง หรือผู้รับจ้างช่วง ที่อยู่ในความรับผิดชอบของผู้รับจ้างตามสัญญานั้น ๆ

“นโยบาย” หมายถึง นโยบายและแนวทางการกำกับดูแลการใช้งานปัญญาประดิษฐ์ (AI Governance) การประปาส่วนภูมิภาค

“ปัญญาประดิษฐ์ (Artificial Intelligence : AI)” หมายถึง เทคโนโลยีที่ถูกพัฒนาขึ้นเพื่อให้ระบบประมวลผลของคอมพิวเตอร์ หุ่นยนต์ เครื่องจักร หรืออุปกรณ์อิเล็กทรอนิกส์ต่าง ๆ มีคุณสมบัติหรือพฤติกรรมใกล้เคียงมนุษย์ตามวัตถุประสงค์ที่มนุษย์กำหนด เช่น การเรียนรู้ การรับรู้และตอบสนองต่อสภาพแวดล้อม การให้เหตุผล และการแก้ไขปัญหา เป็นต้น

“เทคโนโลยี Generative AI” หมายถึง เทคโนโลยี AI ประเภทหนึ่งที่มีความสามารถในการสร้างเนื้อหาใหม่ในหลากหลายรูปแบบตามข้อความหรือคำสั่ง (Prompt) ที่มนุษย์เป็นผู้กำหนด เช่น ข้อความ ภาพ วิดีโอ ชุดคำสั่งหรือโปรแกรม (Source Code) หรือรูปแบบอื่น เป็นต้น

“ข้อมูลส่วนบุคคล” หมายถึง ข้อมูลเกี่ยวกับบุคคลธรรมดาซึ่งทำให้สามารถระบุตัวบุคคลนั้นได้ ไม่ว่าทางตรงหรือทางอ้อม แต่ไม่รวมถึงข้อมูลของผู้ถึงแก่กรรมโดยเฉพาะ

“ข้อมูลส่วนบุคคลอ่อนไหว” หมายถึง ข้อมูลส่วนบุคคลตามที่ถูกบัญญัติไว้ในมาตรา ๒๖ แห่งพระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. ๒๕๖๒ ซึ่งได้แก่ ข้อมูลเชื้อชาติ เผ่าพันธุ์ ความคิดเห็นทางการเมือง ความเชื่อในลัทธิ ศาสนาหรือปรัชญา พฤติกรรมทางเพศ ประวัติอาชญากรรม ข้อมูลสุขภาพ ความพิการ ข้อมูลสหภาพแรงงาน ข้อมูลพันธุกรรม ข้อมูลชีวภาพ หรือข้อมูลอื่นใดซึ่งกระทบต่อเจ้าของข้อมูลส่วนบุคคล ในทำนองเดียวกันตามที่คณะกรรมการคุ้มครองข้อมูลส่วนบุคคลประกาศกำหนด

๖. แนวทางการใช้เทคโนโลยี AI ที่ยอมรับได้

๖.๑ การใช้ที่ยอมรับได้ (Acceptable Use) การใช้งาน AI ต้องเป็นไปเพื่อสนับสนุนภารกิจและเป็นประโยชน์ต่อ กปภ. เท่านั้น เช่น

- การวิเคราะห์ข้อมูลและสร้างรายงาน
- การเขียนชุดคำสั่งหรือโปรแกรม (Source Code)
- การร่างเอกสาร นโยบาย สื่อประชาสัมพันธ์ และการสื่อสารภายในองค์กร

๖.๒ ข้อห้ามและจริยธรรมปัญญาประดิษฐ์ (Prohibited Use & AI Ethics)

- ห้ามใช้ AI ตัดสินใจในเรื่องที่มีความเสี่ยงสูง (การแพทย์ กฎหมาย การเงิน สิทธิของบุคคล) แทนมนุษย์
- ห้ามใช้สร้างเนื้อหาที่ผิดกฎหมาย บิดเบือนความจริง (Fake News) หรือทำให้เกิดความเสื่อมเสีย
- ห้ามใช้เพื่อก่อให้เกิดการหยุดชะงัก ขัดขวาง หรือรบกวนระบบสารสนเทศและบริการของ กปภ.

• ห้ามสร้างเนื้อหา...

- ห้ามสร้างเนื้อหาที่ส่งเสริมการเหยียดเชื้อชาติ ศาสนา เพศ วัย หรือขัดต่อหลักสิทธิมนุษยชน

๖.๓ การคุ้มครองข้อมูลส่วนบุคคลและการรักษาความลับ (Data Privacy & Confidentiality)

- ห้ามนำข้อมูลภายใน ข้อมูลความลับ (เช่น รหัสผ่าน สัญญา) ข้อมูลส่วนบุคคล และข้อมูลอ่อนไหว ไปใช้กับบริการ AI ทั่วไปโดยเด็ดขาด

๖.๔ การรักษาความมั่นคงปลอดภัย (Information Security)

- ห้ามป้อนข้อมูลที่กระทบต่อระบบ (เช่น API Key, System Configuration) ลงใน AI
- ชุดคำสั่งหรือโปรแกรม (Source Code) ที่สร้างโดย AI ต้องได้รับการตรวจสอบความถูกต้องและค้นหาช่องโหว่ (Vulnerability Scan) ก่อนนำมาใช้งานเสมอ
- หากพบเหตุละเมิดความมั่นคงปลอดภัยจาก AI ต้องรายงานผู้บังคับบัญชาและปฏิบัติตามแผนเผชิญเหตุ (Incident Response) ทันที

๖.๕ การลดอคติและการตรวจสอบ (Bias Reduction & Human-in-the-Loop)

- ผู้ใช้งานมีหน้าที่พิจารณา "ตรวจสอบความถูกต้องและอคติ" ของเนื้อหาที่สร้างโดย AI ก่อนนำไปใช้งานหรือเผยแพร่ทุกครั้ง
- หากนำเนื้อหาจาก AI ไปเผยแพร่ ควรมีการระบุว่า “เนื้อหานี้ได้รับความช่วยเหลือจากเทคโนโลยี AI” เพื่อความโปร่งใส

๖.๖ ทรัพย์สินทางปัญญา (Intellectual Property)

- ต้องระมัดระวังไม่ให้เกิดการใช้ AI ไปละเมิดลิขสิทธิ์ เครื่องหมายการค้า หรือทรัพย์สินทางปัญญาของผู้อื่น

๗. การพัฒนาบุคลากรและการสื่อสาร

- กปภ. จัดให้มีการสื่อสารและฝึกอบรมเพื่อสร้างความตระหนักรู้ (Awareness) ด้านการใช้ AI ความเสี่ยง และข้อจำกัด ให้กับพนักงานทุกระดับอย่างต่อเนื่อง
- สนับสนุนการพัฒนาองค์ความรู้เฉพาะทางให้เหมาะสมกับหน้าที่ความรับผิดชอบ

๘. การติดตาม การรายงาน และการทบทวน

- ให้คณะทำงานฯ ติดตามและบริหารจัดการปัญหาด้าน AI อย่างเป็นธรรมต่อผู้มีส่วนได้ส่วนเสีย
- รายงานผลการดำเนินงานให้คณะกรรมการกำกับดูแลเทคโนโลยีดิจิทัลของ กปภ. (DTSC) ทราบอย่างน้อยปีละ ๑ ครั้ง
- ให้มีการทบทวนนโยบายฉบับนี้ อย่างน้อยปีละ ๑ ครั้ง หรือเมื่อมีการเปลี่ยนแปลงทางเทคโนโลยีอย่างมีนัยสำคัญ

๙. โทษของการไม่ปฏิบัติตาม

ผู้ใดฝ่าฝืนหรือไม่ปฏิบัติตามนโยบายฉบับนี้ อาจถูกพิจารณาดำเนินการทางวินัยตามข้อบังคับของ กปภ. หรือตามข้อตกลงในสัญญาจ้าง และอาจได้รับโทษตามกฎหมายที่เกี่ยวข้อง (เช่น พ.ร.บ. ว่าด้วยการกระทำความผิดเกี่ยวกับคอมพิวเตอร์ หรือ PDPA เป็นต้น)

๑๐. ข้อควรปฏิบัติและไม่ควรปฏิบัติ (Do's & Don'ts) ของการใช้งานเทคโนโลยี Generative AI

๑๐.๑ พฤติกรรมที่ควรทำ (Do's)

- ตรวจสอบ ความถูกต้องของข้อมูลจาก AI ก่อนนำไปใช้
- ไม่นำข้อมูลที่เป็นความลับ ของสำนักงานไปใช้งานร่วมกับเทคโนโลยี Generative AI
- เผยแพร่ข้อมูล ที่ถูกต้องและไม่บิดเบือน
- อ้างอิง แหล่งที่มาของข้อมูล และโปร่งใสเมื่อใช้ AI ช่วยสร้างสรรค์งาน
- ตรวจสอบช่องโหว่ (Vulnerability Scan) ซุดคำสั่งหรือโปรแกรม (Source Code) ที่สร้างโดย AI ก่อนนำไปพัฒนาต่อ
- ปฏิบัติตาม นโยบายและแนวทางของสำนักงาน นโยบายความมั่นคงปลอดภัย และ PDPA อย่างเคร่งครัด
- หลีกเลี่ยง การสร้างเนื้อหาที่มีอคติ หรือเลือกปฏิบัติ
- หลีกเลี่ยง การใช้เนื้อหาที่อาจละเมิดลิขสิทธิ์ หรือทรัพย์สินทางปัญญา
- หลีกเลี่ยง การใช้งานเทคโนโลยี Generative AI ในทางที่ผิดกฎหมายหรือขัดต่อหลักธรรมาภิบาลคุณธรรม และจริยธรรม

๑๐.๒ พฤติกรรมที่ไม่ควรทำ (Don'ts)

- นำข้อมูล ไปใช้ทันทีโดยไม่ได้มีการตรวจสอบความถูกต้อง
- นำข้อมูล ที่เป็นความลับของสำนักงานไปใช้งานร่วมกับเทคโนโลยี Generative AI
- เผยแพร่ข้อมูล ที่ไม่ถูกต้อง (Misinformation) หรือบิดเบือน (Disinformation)
- ไม่มีการอ้างอิง แหล่งที่มาของข้อมูล
- สร้างเนื้อหาที่มีอคติ หรือเลือกปฏิบัติ
- ใช้เนื้อหาที่อาจละเมิดลิขสิทธิ์ หรือทรัพย์สินทางปัญญา
- ไม่ปฏิบัติตามนโยบายและแนวทางของสำนักงาน
- ใช้งานเทคโนโลยี Generative AI ในทางที่ผิดกฎหมาย หรือขัดต่อหลักธรรมาภิบาล คุณธรรม และจริยธรรม
- ห้าม นำข้อมูลขึ้นความลับ ข้อมูลส่วนบุคคล ไปป้อนใส่ AI สาธารณะ
- ห้าม นำผลลัพธ์จาก AI ไปใช้งานทันทีโดยไม่ตรวจสอบ (Copy & Paste)
- ห้าม ใช้ AI สร้างข้อมูลเท็จ บิดเบือน หรือสร้างความเกลียดชัง
- ห้าม นำ Code จาก AI ไปใช้ในระบบโดยไม่ผ่านการตรวจสอบช่องโหว่ความปลอดภัย

๑๑. แอปพลิเคชันหรือบริการ Generative AI

ประเภทบริการ	ตัวอย่างแอปพลิเคชัน (สำหรับการใช้งานทั่วไป)
เครื่องมือสร้างข้อความ	ChatGPT, Claude, Gemini
เครื่องมือสร้างภาพ	DALL-E, Midjourney, Stable Diffusion
เครื่องมือสร้างและพัฒนาโค้ด	Claude, Google Antigravity, GitHub Copilot, Codeium
เครื่องมือสร้างเสียง/วิดีโอ	Google VEO, YouTube Create, ElevenLabs, RunwayML
เครื่องมือสำหรับงานวิจัย	Perplexity, NotebookLM

๑๒. การพัฒนา...

๑๒. การพัฒนาหรือการใช้งานปัญญาประดิษฐ์ (AI) ผู้บริหาร พนักงาน ผู้รับจ้าง และ ลูกจ้าง กปภ. ควรปฏิบัติตามคู่มือและแนวทางการกำกับดูแลการใช้งานปัญญาประดิษฐ์ (AI) ของการประปาส่วนภูมิภาค

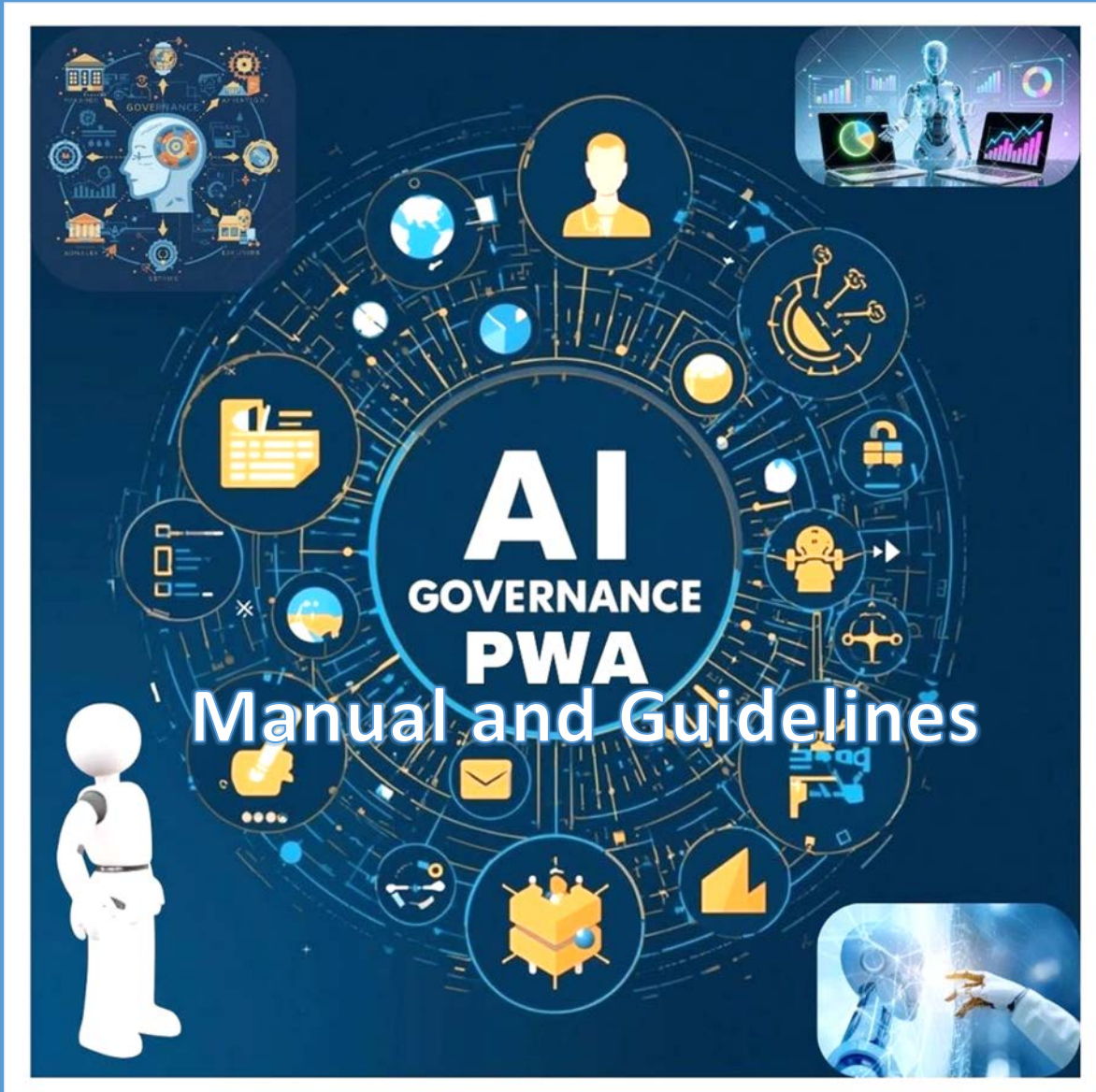
ทั้งนี้ นโยบายมีผลบังคับใช้กับ ผู้บริหาร พนักงาน ผู้รับจ้าง และลูกจ้างทุกคน

ประกาศ ณ วันที่ ๒๔ เดือน กรกฎาคม พ.ศ. ๒๕๖๙



(นายจักรพงศ์ คำจันทร์)
ผู้ว่าการการประปาส่วนภูมิภาค

คู่มือและแนวทางการกำกับดูแลการใช้งานปัญญาประดิษฐ์ (AI)
ของการประปาส่วนภูมิภาค



จัดทำโดยคณะกรรมการกำกับดูแลการใช้งานปัญญาประดิษฐ์
การประปาส่วนภูมิภาค (PWA AI Governance Working Group)
มิถุนายน ๒๕๖๙

คู่มือและแนวทางการกำกับดูแลการใช้งานปัญญาประดิษฐ์ (AI)
ของการประปาส่วนภูมิภาค

การอนุมัติเอกสาร

การอนุมัติ	ตำแหน่ง	วันที่
ผู้อนุมัติ	ผู้ช่วยผู้ว่าการ (ดิจิทัลและสารสนเทศ)	๒๑ กรกฎาคม ๒๕๖๙
ผู้ทบทวน	ผู้อำนวยการกองบริหารระบบข้อมูลสารสนเทศ	๑๘ มิถุนายน ๒๕๖๙
ผู้จัดทำ	หัวหน้างานวิเคราะห์ข้อมูลสารสนเทศ	๒๘ พฤศจิกายน ๒๕๖๘

บันทึกประวัติการจัดทำ/แก้ไขเอกสาร

ประวัติปรับปรุงเอกสาร			
ครั้งที่	วันที่	รายละเอียดการจัดทำ/แก้ไข	ผู้จัดทำ/แก้ไข
	๒๘ พ.ย. ๖๘	จัดทำคู่มือและแนวทางการกำกับดูแลการใช้งานปัญญาประดิษฐ์ (AI) ของการประสานส่วนภูมิภาคฉบับใหม่	งานวิเคราะห์ข้อมูลสารสนเทศ กบข.
๑	๑๒ มิ.ย. ๖๙	เสนอคณะทำงานกำกับดูแลการใช้งานปัญญาประดิษฐ์ (AI Governance Working Group)	เลขานุการคณะทำงานฯ
๒	๑๘ มิ.ย. ๖๙	เสนอคณะกรรมการกำกับดูแลการใช้งานปัญญาประดิษฐ์ (AI Governance Council) หรือคณะกรรมการกำกับดูแลเทคโนโลยีดิจิทัลของ กปภ. (DTSC)	ผู้อำนวยการกองบริหารระบบข้อมูลสารสนเทศ

บทนำ

คู่มือและแนวทางการกำกับดูแลการใช้งานปัญญาประดิษฐ์ของการประปาส่วน จัดทำขึ้นเพื่อกำหนดกรอบธรรมาภิบาลปัญญาประดิษฐ์ (AI Governance Framework) รวมถึงหลักเกณฑ์และแนวปฏิบัติสำหรับการพัฒนา การจัดหา การประยุกต์ใช้ และการบริหารจัดการระบบปัญญาประดิษฐ์ (AI) และ Generative AI ภายในองค์กร ให้เป็นไปอย่างมีประสิทธิภาพ โปร่งใส ตรวจสอบได้ มีความรับผิดชอบ และสอดคล้องกับกฎหมายและมาตรฐานที่เกี่ยวข้อง

การนำ AI มาใช้เป็นกลไกสำคัญในการขับเคลื่อน กปภ. สู่การเป็นองค์กรดิจิทัลอัจฉริยะ ช่วยเพิ่มประสิทธิภาพการดำเนินงาน การบริหารจัดการทรัพยากรน้ำ การลดน้ำสูญเสีย และการยกระดับการให้บริการประชาชน อย่างไรก็ตาม ในฐานะหน่วยงานผู้ให้บริการสาธารณูปโภคขั้นพื้นฐานและมีบทบาทเกี่ยวข้องกับโครงสร้างพื้นฐานสารสนเทศที่สำคัญ (Critical Information Infrastructure: CII) การใช้งาน AI จำเป็นต้องอยู่ภายใต้การกำกับดูแลที่เหมาะสม เพื่อบริหารจัดการความเสี่ยงด้านความมั่นคงปลอดภัยไซเบอร์ การคุ้มครองข้อมูลส่วนบุคคล ความถูกต้องน่าเชื่อถือของผลลัพธ์ และประเด็นด้านจริยธรรม

คู่มือฉบับนี้พัฒนาขึ้นโดยอ้างอิงแนวปฏิบัติและมาตรฐานที่ได้รับการยอมรับทั้งในระดับสากลและระดับประเทศ อาทิ NIST AI Risk Management Framework (AI RMF), ISO/IEC ๔๒๐๐๑, แนวทาง Generative AI ของ ETDA, แนวปฏิบัติด้านจริยธรรม AI ของ สดช. และ สวทช. รวมถึงแนวทางด้านความมั่นคงปลอดภัยไซเบอร์ของ สกมช. เพื่อนำมาประยุกต์ใช้ให้เหมาะสมกับบริบทและภารกิจของ กปภ.

กรอบการกำกับดูแล AI ของ กปภ. ยึดหลักสำคัญ ได้แก่ ความโปร่งใสและสามารถอธิบายได้ (Transparency and Explain ability) ความรับผิดชอบและตรวจสอบได้ (Accountability) ความมั่นคงปลอดภัย และการคุ้มครองข้อมูล (Security and Privacy) ความเป็นธรรม (Fairness) และการกำกับดูแลโดยมนุษย์ (Human Oversight) ตลอดวงจรชีวิตของระบบ AI (AI Life Cycle)

ทั้งนี้ การดำเนินงานด้าน AI ของ กปภ. มุ่งสนับสนุนยุทธศาสตร์ด้านที่ ๖ “สร้างคุณค่าด้วยเทคโนโลยีดิจิทัลและนวัตกรรม” ภายใต้วิสัยทัศน์ “SMART Water” เพื่อยกระดับประสิทธิภาพองค์กร สร้างคุณค่าแก่ผู้มีส่วนได้ส่วนเสีย และขับเคลื่อนองค์กรสู่ความมั่นคงและความยั่งยืนในระยะยาว ผ่านการใช้เทคโนโลยี AI ที่มีความน่าเชื่อถือ ปลอดภัย และเกิดประโยชน์สูงสุดต่อประชาชนและสังคม

บทสรุปผู้บริหาร (Executive Summary)

๑. ภาพรวมและบริบทเชิงกลยุทธ์ของการกำกับดูแลปัญญาประดิษฐ์

การประสานส่วนภูมิภาค (กปภ.) ตระหนักถึงบทบาทสำคัญของเทคโนโลยีปัญญาประดิษฐ์ (Artificial Intelligence: AI) ในการขับเคลื่อนการเปลี่ยนผ่านสู่การเป็นองค์กรดิจิทัลอัจฉริยะ และการยกระดับประสิทธิภาพการดำเนินงานในทุกมิติ จึงได้กำหนดนโยบายและแนวปฏิบัติการกำกับดูแลปัญญาประดิษฐ์ เพื่อเป็นกรอบในการบริหารจัดการ การพัฒนา และการประยุกต์ใช้ AI อย่างมีประสิทธิภาพ โปร่งใส มีจริยธรรม และรับผิดชอบต่อสังคม

นโยบายและแนวปฏิบัติฉบับนี้จัดทำขึ้นภายใต้กรอบการกำกับดูแลปัญญาประดิษฐ์ (AI Governance Framework) ที่สอดคล้องกับแผนยุทธศาสตร์ AI แห่งชาติ (พ.ศ. ๒๕๖๕-๒๕๗๐) นโยบายรัฐบาลดิจิทัล แผนยุทธศาสตร์องค์กร แผนปฏิบัติการดิจิทัล และสถาปัตยกรรมองค์กร (Enterprise Architecture: EA) ของ กปภ. เพื่อสนับสนุนการนำ AI มาใช้ในการสร้างคุณค่า เพิ่มประสิทธิภาพการดำเนินงาน และยกระดับการให้บริการประชาชนอย่างยั่งยืน

ในฐานะหน่วยงานผู้ให้บริการสาธารณะขั้นพื้นฐานด้านน้ำประปา และเป็นหน่วยงานที่มีบทบาทเกี่ยวข้องกับโครงสร้างพื้นฐานสำคัญทางสารสนเทศ (Critical Information Infrastructure: CII) กปภ. จำเป็นต้องกำกับดูแลการใช้ AI อย่างเหมาะสม เพื่อให้มั่นใจว่าระบบ AI มีความมั่นคงปลอดภัย เชื่อถือได้ โปร่งใส และสามารถบริหารจัดการความเสี่ยงที่อาจเกิดขึ้นได้อย่างมีประสิทธิภาพ ทั้งในด้านการดำเนินงาน ความมั่นคงปลอดภัยไซเบอร์ การคุ้มครองข้อมูลส่วนบุคคล ตลอดจนด้านกฎหมายและจริยธรรม

การกำกับดูแลปัญญาประดิษฐ์จึงเป็นกลไกเชิงยุทธศาสตร์ที่สำคัญในการส่งเสริมให้ กปภ. สามารถนำ AI มาใช้เพื่อสนับสนุนภารกิจขององค์กร อาทิ การบริหารจัดการทรัพยากรน้ำ การลดน้ำสูญเสีย การเพิ่มประสิทธิภาพการผลิตและจ่ายน้ำ การยกระดับคุณภาพการให้บริการลูกค้า และการวิเคราะห์เชิงคาดการณ์ เพื่อการบริหารจัดการโครงสร้างพื้นฐานได้อย่างมีประสิทธิภาพ

กรอบการกำกับดูแลปัญญาประดิษฐ์ของ กปภ. จึงมุ่งเน้นการสร้างที่น่าเชื่อถือ (Trustworthiness) ความโปร่งใส (Transparency) ความรับผิดชอบ (Accountability) ความมั่นคงปลอดภัย (Security) และการกำกับดูแลโดยมนุษย์ (Human Oversight) ตลอดวงจรชีวิตของระบบปัญญาประดิษฐ์ (AI Life Cycle) เพื่อให้การนำ AI มาใช้เกิดประโยชน์สูงสุดต่อองค์กร ผู้มีส่วนได้ส่วนเสีย และประชาชนผู้รับบริการ ภายใต้หลักธรรมาภิบาลและการพัฒนาที่ยั่งยืน

๒. หลักการและเป้าหมายของกรอบการกำกับดูแล (AI Governance Framework)

กรอบการกำกับดูแล AI ของ กปภ. ถูกสร้างขึ้นบนหลักการพื้นฐานที่สำคัญ ๖ ประการ ซึ่งเป็นหัวใจสำคัญของการสร้างความไว้วางใจในการใช้เทคโนโลยี

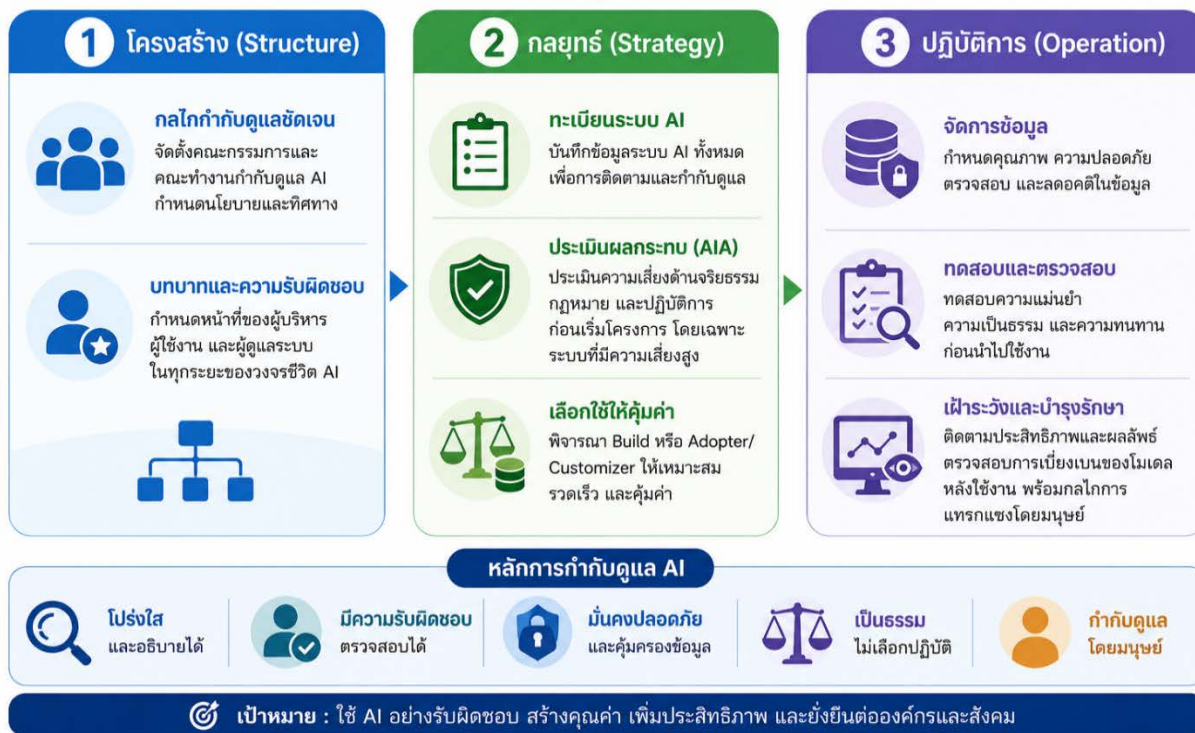
หลักการสำคัญ	คำอธิบายความสำคัญสำหรับ กปภ.
๑.ความสามารถในการแข่งขันและความยั่งยืน (Competitiveness and Sustainable Development)	ใช้ AI เพื่อสร้างคุณค่า เพิ่มประสิทธิภาพ และพัฒนาองค์กรอย่างยั่งยืน
๒.กฎหมาย จริยธรรม และการกำกับโดยมนุษย์ (Laws, Ethics, and International Standards)	ใช้ AI อย่างถูกต้องตามกฎหมาย มีจริยธรรม และมนุษย์สามารถควบคุมได้ ระบบ AI ที่มีความเสี่ยงสูงต้องมีกลไกให้มนุษย์เข้ามาตรวจสอบ แทรกแซง หรือตัดสินใจขั้นสุดท้ายได้เสมอ
๓.ความโปร่งใสและความรับผิดชอบ (Transparency and Accountability)	AI ต้องอธิบายได้ ตรวจสอบได้ และมีผู้รับผิดชอบที่ชัดเจนสามารถระบุผู้รับผิดชอบได้อย่างชัดเจนในการออกแบบ พัฒนา และตัดสินใจของระบบ AI โดยเฉพาะในกรณีที่เกี่ยวข้องกับการให้บริการสาธารณะ
๔.ความมั่นคงปลอดภัยและความเป็นส่วนตัว (Security and Privacy)	AI ต้องได้รับการปกป้องจากภัยคุกคามไซเบอร์ และทำงานได้อย่างแม่นยำและทนทาน ไม่เกิดความผิดพลาดที่ส่งผลต่อการให้บริการปกป้องข้อมูลและระบบให้ปลอดภัย ตามกฎหมายคุ้มครองข้อมูลส่วนบุคคล (PDPA) อย่างเคร่งครัด
๕.ความเป็นธรรมและไม่เลือกปฏิบัติ (Fairness)	ลดอคติของ AI และคำนึงถึงความเท่าเทียมของผู้ใช้งานทุกกลุ่ม ไม่สร้างหรือขยายอคติ (Bias) ในการใช้งานหรือให้บริการ
๖.ความน่าเชื่อถือ (Reliability)	AI ต้องมีความถูกต้อง แม่นยำ และสามารถปรับปรุงได้อย่างต่อเนื่อง

เป้าหมายหลัก คือการส่งเสริมการใช้ AI เพื่อยกระดับประสิทธิภาพและสร้างความเชื่อมั่นให้กับผู้มีส่วนได้ส่วนเสียทุกฝ่าย พร้อมทั้งบริหารจัดการความเสี่ยงให้สอดคล้องกับสถานะ CII ขององค์กร

๓. องค์ประกอบหลักของกรอบการกำกับดูแล AI ของ กปภ. (AI Governance Framework ๓ เสาหลัก)

คู่มือนี้ได้กำหนดแนวทางการกำกับดูแลที่ครอบคลุม ๓ องค์ประกอบหลัก เพื่อให้เกิดการปฏิบัติที่สอดคล้องและเป็นระบบทั่วทั้งองค์กร

กรอบการกำกับดูแล AI (3 เสาหลัก)



๓.๑ เสาหลักที่ ๑: โครงสร้าง (Structure) เน้นการจัดตั้งกลไกการบริหารจัดการและกำหนดบทบาท ความรับผิดชอบที่ชัดเจน จัดตั้งคณะกรรมการกำกับดูแลการประยุกต์ใช้งานปัญญาประดิษฐ์ (AI Governance Council) เพื่อกำหนดนโยบาย ทิศทางเชิงกลยุทธ์ และอนุมัติโครงการ AI ที่มีความเสี่ยงสูง จัดตั้งคณะทำงานกำกับดูแลการใช้งานปัญญาประดิษฐ์ (AI Governance Working Group) ทำหน้าที่เป็นหน่วยงานกลางในการประสานงาน ติดตามการปฏิบัติ และให้คำแนะนำด้านเทคนิคและจริยธรรมการกำหนดบทบาทความรับผิดชอบ (Roles and Responsibilities) กำหนดหน้าที่ความรับผิดชอบที่ชัดเจนสำหรับผู้บริหาร ผู้ใช้งาน และผู้ดูแลระบบ ในทุกระยะของวงจรชีวิต AI (AI Life Cycle)

๓.๒ เสาหลักที่ ๒: กลยุทธ์ (Strategy) เน้นการวางแผนและการจัดการความเสี่ยงเชิงรุก การจัดทำทะเบียนระบบ AI (AI Register) บันทึกข้อมูลสำคัญของระบบ AI ทั้งหมดในองค์กร เพื่อให้สามารถติดตามความเสี่ยงและสถานะการกำกับดูแลได้ มีการประเมินผลกระทบ AI (AI Impact Assessment - AIA) เป็นขั้นตอนบังคับสำหรับโครงการ AI ใหม่ทุกโครงการ โดยเฉพาะ AI ที่มีความเสี่ยงสูง (High-Risk AI) เช่น การตัดสินใจเกี่ยวกับการบริหารจัดการน้ำโดยตรง เพื่อประเมินความเสี่ยงด้านจริยธรรม กฎหมาย และปฏิบัติการล่วงหน้าการเลือกรูปแบบการประยุกต์ใช้ที่คุ้มค่า เน้นการประเมินศักยภาพและต้นทุนในการพัฒนา (Build) หรือการปรับใช้ (Adopter/Customizer) ระบบ AI ที่มีอยู่แล้ว เพื่อให้เกิดความรวดเร็วและคุ้มค่า

๓.๓ เสาหลักที่ ๓: ปฏิบัติการ (Operation) เน้นการกำกับดูแลการทำงานจริงตลอดวงจรชีวิตของระบบ AI (AI Life Cycle)

- การจัดการข้อมูล (Data Governance) กำหนดมาตรฐานคุณภาพ ความสมบูรณ์ และความปลอดภัยของชุดข้อมูลที่ใช้ในการฝึก AI รวมถึงการตรวจสอบและลดอคติ (Bias Mitigation) ในข้อมูล
- การทดสอบและการตรวจสอบ (Testing and Validation) กำหนดเกณฑ์การทดสอบที่เข้มงวดสำหรับการดำเนินงาน ความแม่นยำ ความเป็นธรรม และความทนทานของระบบ AI ก่อนนำไปใช้งานจริง (Pre-Deployment)
- การเฝ้าระวังและการบำรุงรักษา (Monitoring and Maintenance) การติดตามประสิทธิภาพ ผลลัพธ์ และการตรวจสอบการเบี่ยงเบนของโมเดล (Model Drift) อย่างต่อเนื่อง หลังการใช้งานจริง (Post-Deployment) พร้อมกลไกการแทรกแซงโดยมนุษย์

๔. การนำไปปฏิบัติและผลกระทบที่คาดหวัง

การนำกรอบการกำกับดูแลนี้ไปปฏิบัติอย่างจริงจังและครอบคลุมทั่วทั้งองค์กรจะก่อให้เกิดผลกระทบที่สำคัญในหลายด้าน เช่น

- **การปฏิบัติตามกฎหมาย** กปภ. จะสามารถมั่นใจได้ว่าการใช้เทคโนโลยี AI เป็นไปตามข้อกำหนดทางกฎหมายที่เกี่ยวข้องทั้งหมด เช่น พ.ร.บ. คุ้มครองข้อมูลส่วนบุคคล และ พ.ร.บ. ความมั่นคงปลอดภัยไซเบอร์
- **การยกระดับความเชื่อมั่น** การดำเนินงานอย่างโปร่งใสและรับผิดชอบจะเสริมสร้างความไว้วางใจจากผู้ใช้น้ำ และผู้มีส่วนได้ส่วนเสีย
- **การเพิ่มประสิทธิภาพอย่างยั่งยืน** การบริหารจัดการความเสี่ยงอย่างเป็นระบบจะช่วยลดความผิดพลาดและค่าใช้จ่ายที่อาจเกิดขึ้นจากการใช้ AI ที่ไม่มีการกำกับดูแล ส่งผลให้การใช้ AI สามารถสร้างมูลค่าเพิ่มได้อย่างยั่งยืนและปลอดภัย
- **การเป็นผู้บุกเบิก** กปภ. จะสามารถเป็นต้นแบบและผู้บุกเบิกการนำ AI ที่มีความรับผิดชอบต่อมาใช้ในภาคสาธารณสุขของประเทศไทย ซึ่งเป็นส่วนสำคัญในการขับเคลื่อนการเปลี่ยนผ่านสู่ดิจิทัลของประเทศโดยรวม

การนำแผนงานเชิงกลยุทธ์และแนวทางปฏิบัตินี้ไปใช้อย่างจริงจัง จะช่วยให้ กปภ. ไม่เพียงแต่ปฏิบัติตามข้อกำหนดทางกฎหมายที่เกี่ยวข้องเท่านั้น แต่ยังสามารถเป็นผู้บุกเบิกการนำ AI ที่มีความรับผิดชอบต่อใช้ในภาคสาธารณสุขของประเทศไทย ซึ่งจะส่งผลสำคัญต่อเป้าหมายการเปลี่ยนผ่านสู่ดิจิทัลของประเทศโดยรวม และยกระดับความเชื่อมั่นให้กับผู้มีส่วนได้ส่วนเสียทุกฝ่าย ดังนั้นการลงทุนใน AI Governance คือการลงทุนในอนาคตที่ยั่งยืนของ กปภ. ซึ่งจะช่วยให้องค์กรสามารถควบคุมเทคโนโลยีให้เป็นไปในทิศทางที่สร้างสรรค์คุณค่าและปกป้องผลประโยชน์สาธารณะได้อย่างแท้จริง

สารบัญ

เรื่อง	หน้า
บทนำ	๓
บทสรุปผู้บริหาร (Executive Summary)	๔
บทที่ ๑ บทนำและหลักการพื้นฐาน	๙
๑.๑ ความเป็นมาและวัตถุประสงค์	๙
๑.๒ ขอบเขตการบังคับใช้	๙
๑.๓ หลักการธรรมาภิบาลปัญญาประดิษฐ์ (AI Governance Principles)	๑๐
๑.๔ Generative AI Governance	๑๑
บทที่ ๒ โครงสร้างการกำกับดูแลและการบริหารความเสี่ยง	๑๒
๒.๑ โครงสร้างการกำกับดูแล (AI Governance Structure)	๑๓
๒.๒ การบริหารความเสี่ยงปัญญาประดิษฐ์ (AI Risk Management)	๑๖
๒.๓ การปฏิบัติตามกฎหมายและข้อบังคับของไทย	๒๓
บทที่ ๓ แนวปฏิบัติเชิงปฏิบัติการด้านการใช้งานและความมั่นคงปลอดภัย	๒๖
๓.๑ การกำกับดูแลตลอดวงจรชีวิตระบบ AI (AI System Life Cycle Governance)	๒๖
๓.๒ แนวปฏิบัติการใช้งานที่ยอมรับได้ (Acceptable Use Policy - AUP)	๒๗
๓.๓ การจัดการข้อมูลเพื่อรองรับ AI	๓๐
บทที่ ๔ การตรวจสอบและการปรับปรุงอย่างต่อเนื่อง	๓๑
๔.๑ การตอบสนองต่อเหตุการณ์ผิดปกติ (Incident Response)	๓๑
๔.๒ การตรวจสอบและการรับรอง (Audit and Certification)	๓๑
๔.๓ การฝึกอบรมและการสร้างวัฒนธรรม AI	๓๑
บรรณานุกรม (References)	๓๒
ภาคผนวก	๓๔
ผนวก ก นิยามศัพท์ที่เกี่ยวข้อง	๓๔
ผนวก ข แบบฟอร์มรายงานการใช้งานปัญญาประดิษฐ์ (AI Usage Report)	๓๗
ผนวก ค รายการตรวจสอบความมั่นคงปลอดภัยสำหรับระบบ AI (AI Security Checklist)	๓๘
ผนวก ง รูปแบบรายงานเหตุการณ์ฉุกเฉินเฉพาะ AI (AI-Specific Incident Response Template)	๓๙

บทที่ ๑ บทนำ และหลักการพื้นฐาน

๑.๑ ความเป็นมาและวัตถุประสงค์

ปัญญาประดิษฐ์ (AI) รวมถึง Generative AI กำลังเปลี่ยนแปลงการทำงานขององค์กรอย่างรวดเร็ว การประสานส่วนภูมิภาค กปภ. ในฐานะหน่วยงานที่เป็นโครงสร้างพื้นฐานสารสนเทศที่สำคัญ (CII) ตระหนักถึงศักยภาพของ AI ในการยกระดับประสิทธิภาพการดำเนินงานและการให้บริการ อย่างไรก็ตาม การนำ AI มาใช้นั้นต้องเผชิญกับความเสี่ยงด้านจริยธรรม ความเป็นส่วนตัว และความมั่นคงปลอดภัย

คู่มือฉบับนี้จัดทำขึ้นเพื่อเป็นกรอบมาตรฐานในการพัฒนา จัดซื้อ ใช้งาน และกำกับดูแลระบบปัญญาประดิษฐ์ (AI) เพื่อให้การนำเทคโนโลยีปัญญาประดิษฐ์ (AI) มาใช้ในการบริหารงานและการให้บริการประชาชนของ กปภ. เป็นไปอย่างปลอดภัย มีประสิทธิภาพ ควบคู่ไปกับการคุ้มครองสิทธิเสรีภาพของประชาชน สอดคล้องกับพระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. ๒๕๖๒ (PDPA) และพระราชบัญญัติการรักษาความมั่นคงปลอดภัยไซเบอร์ พ.ศ. ๒๕๖๒ โดยสอดคล้องกับกฎหมายไทยและมาตรฐานสากล เช่น OECD AI Principles, UNESCO AI Ethics รวมถึง พ.ร.บ. คุ้มครองข้อมูลส่วนบุคคล (PDPA) จึงจำเป็นต้องมีนโยบายกำกับดูแลที่ชัดเจน

วัตถุประสงค์หลักของคู่มือ

๑. เพื่อเป็นกรอบแนวทางและกลไกในการกำกับดูแลการประยุกต์ใช้ปัญญาประดิษฐ์ (AI) ให้เป็นไปอย่างมีธรรมาภิบาลและมั่นคงปลอดภัย

๒. เพื่อกำหนดกรอบการกำกับดูแลและการใช้เทคโนโลยีปัญญาประดิษฐ์ (AI) ในองค์กรให้เกิดความโปร่งใส มีความรับผิดชอบปลอดภัย และสอดคล้องกับกฎหมายและจริยธรรม พร้อมสนับสนุนการสร้างนวัตกรรมที่ยั่งยืน

๓. เพื่อกำหนดบทบาท หน้าที่ และความรับผิดชอบสำหรับการใช้งานระบบปัญญาประดิษฐ์ (AI)

๔. เพื่อส่งเสริมให้เกิดการพัฒนาและการประยุกต์ใช้ปัญญาประดิษฐ์ (AI) อย่างปลอดภัย และน่าเชื่อถือ สอดคล้องกับยุทธศาสตร์ขององค์กร

๑.๒ ขอบเขตการบังคับใช้

คู่มือนี้ครอบคลุมการใช้งานปัญญาประดิษฐ์ (AI) ทุกประเภท (รวมถึง Generative AI/LLM) ที่ กปภ. จัดหา พัฒนา หรืออนุญาตให้ใช้ในการสนับสนุนการดำเนินงานตามภารกิจของ กปภ. มีผลบังคับใช้กับผู้บริหาร พนักงาน ลูกจ้าง ผู้ปฏิบัติงาน และผู้รับจ้างของ กปภ.

๑.๓ หลักการธรรมาภิบาลปัญญาประดิษฐ์ (AI Governance Principles)

การประยุกต์ใช้ AI ของ กปภ. ต้องยึดมั่นในหลักการสำคัญ (AI Ethics Principles) ซึ่งรวมถึงความสอดคล้องกับภารกิจของ กปภ. ตามแนวทาง ETDA และมาตรฐานสากล ดังนี้

๑. ความสามารถในการแข่งขันและการพัฒนาอย่างยั่งยืน (Competitiveness and Sustainable Development) AI ควรถูกสร้างและใช้งานเพื่อสร้างประโยชน์และคุณค่าให้แก่มนุษย์ สังคม เศรษฐกิจ และสิ่งแวดล้อมอย่างยั่งยืน พร้อมทั้งช่วยเพิ่มขีดความสามารถในการแข่งขันของประเทศ

๒. ความสอดคล้องกับกฎหมาย จริยธรรม และมาตรฐานสากล (Laws, Ethics, and International Standards) การพัฒนาและใช้งาน AI ต้องเคารพสิทธิมนุษยชน สิทธิเสรีภาพ ความเป็นส่วนตัว และสอดคล้องกับบรรทัดฐานทางกฎหมายและจริยธรรม โดยต้องยึดหลักการออกแบบที่มีมนุษย์เป็นศูนย์กลาง (Human-Centered Design) ควรมีกลไกให้มนุษย์ต้องเป็นผู้ควบคุมและสามารถแทรกแซงหรือขัดขวางการปฏิบัติการของระบบ เพื่อควบคุมความเสี่ยงที่อาจมีผลกระทบกับมนุษย์ได้ และไม่ควรออกแบบ AI เพื่อใช้กำหนดชะตาชีวิตมนุษย์

๓. ความโปร่งใสและความรับผิดชอบต่อผลของการกระทำ (Transparency and Accountability) ระบบ AI ต้องมีความโปร่งใส สามารถอธิบายเหตุผลหรือตรรกะเบื้องหลังการตัดสินใจได้ (Explain ability) และมีความสามารถในการตรวจสอบย้อนหลังได้ (Traceability) รวมถึงต้องมีการกำหนดผู้รับผิดชอบที่ชัดเจนต่อผลกระทบที่เกิดขึ้นจากการทำงาน ระบบ AI ต้องสามารถอธิบายที่มาที่ไปได้ ตรวจสอบย้อนกลับได้ตั้งแต่แหล่งที่มาของชุดข้อมูล กระบวนการทำงาน และการตัดสินใจของ AI

- ต้องมีการบันทึกขั้นตอนวิธีการและโมเดล AI ที่ใช้เพื่อให้สามารถตรวจสอบย้อนหลังได้
- ให้ผู้ใช้งานและผู้มีส่วนได้เสียเข้าใจว่าระบบ AI ทำงานอย่างไรและตัดสินใจอย่างไร

ระบบผู้รับผิดชอบผลลัพธ์ได้อย่างชัดเจน ต้องสามารถระบุตัวบุคคลหรือหน่วยงานผู้รับผิดชอบต่อผลลัพธ์ที่เกิดจาก AI สำหรับผลของการตัดสินใจของระบบ AI

- กำหนดผู้รับผิดชอบที่ชัดเจนในทุกโครงการ AI
- มีช่องทางรับข้อร้องเรียนและตรวจสอบผลกระทบจากการใช้ AI

๔. ความมั่นคงปลอดภัยและความเป็นส่วนตัว (Security and Privacy) การนำ AI มาใช้ต้องให้ความสำคัญสูงสุดกับการป้องกันความเสี่ยงด้านความมั่นคงปลอดภัยทางไซเบอร์ ปฏิบัติตามกฎหมายคุ้มครองข้อมูลส่วนบุคคลอย่างเคร่งครัด และต้องมีกลไกให้มนุษย์สามารถแทรกแซงหรือควบคุมระบบได้ในกรณีที่เกิดเหตุฉุกเฉิน ป้องกันภัยคุกคามและคุ้มครองข้อมูลตาม PDPA ระบบ AI ต้องได้รับการออกแบบโดยใช้หลักการป้องกันความเสี่ยง เพื่อป้องกันการโจมตีจากภัยคุกคาม ภัยอันตรายที่อาจเกิดขึ้น สอดคล้องกับนโยบายความมั่นคงปลอดภัยด้านสารสนเทศของ กปภ. รวมถึงการคุ้มครองข้อมูลส่วนบุคคล ต้องมั่นใจว่าได้รับการคุ้มครองตามสิทธิและกฎหมายที่เกี่ยวข้อง (PDPA)

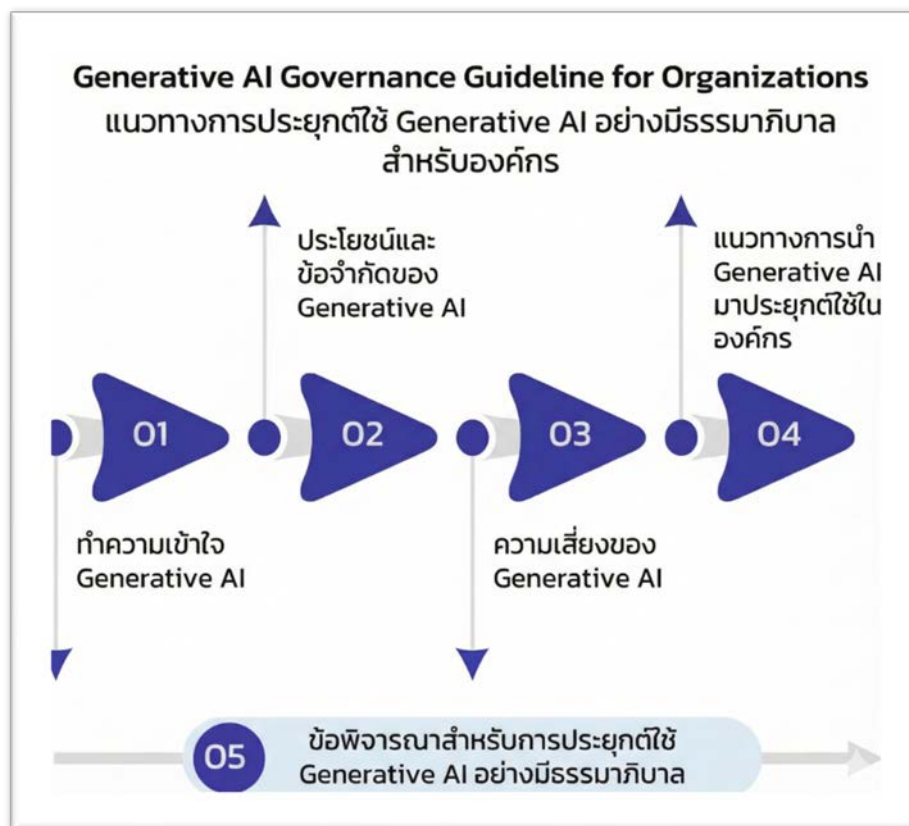
- ปฏิบัติตามกฎหมายคุ้มครองข้อมูลส่วนบุคคล (PDPA/GDPR)
- มีมาตรการป้องกันการรั่วไหลของข้อมูลและการโจมตีระบบ

๕. ความเท่าเทียม หลากหลาย ครอบคลุม และเป็นธรรม (Fairness) ระบบ AI ต้องถูกออกแบบมาเพื่อลดความเอนเอียง (Bias) และต้องไม่ก่อให้เกิดการเลือกปฏิบัติ (Discrimination) โดยต้องคำนึงถึงความหลากหลายเพื่อให้เกิดความเป็นธรรมต่อคนทุกกลุ่ม โดยเฉพาะอย่างยิ่งการปกป้องกลุ่มเปราะบางหรือผู้ด้อยโอกาสในสังคม ระบบ AI ต้องจัดอคติ (Bias) ต่อเชื้อชาติ ศาสนา เพศ หรือสถานะทางสังคม ที่อาจเกิดขึ้นจากข้อมูลหรืออัลกอริทึม

- ป้องกันการเกิดอคติ (Bias) ในโมเดล AI
- ตรวจสอบและประเมินผลกระทบต่อกลุ่มเปราะบาง

๖. ความน่าเชื่อถือ (Reliability) ระบบ AI จะต้องมีความถูกต้องแม่นยำ (Accuracy) มีความทนทานต่อสภาพแวดล้อมหรือข้อผิดพลาด (Robustness) และสามารถสร้างผลลัพธ์ที่สม่ำเสมอและทำซ้ำได้ (Reproducibility) ก่อนนำมาใช้งานจริง ระบบต้องทำงานได้แม่นยำและรับฟัง Feedback ระบบ AI ควรได้รับการสนับสนุนให้มีความน่าเชื่อถือและความมั่นใจในการใช้งาน มีกระบวนการและช่องทางรับผลสะท้อนกลับ (Feedback) จากผู้ใช้งาน

๑.๔ Generative AI Governance



ที่มา: สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์ (สพธอ.) และ กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม. (๒๕๖๔). เอกสารแนวปฏิบัติจริยธรรมปัญญาประดิษฐ์ (Thailand AI Ethics Guideline)

บทที่ ๒

โครงสร้างการกำกับดูแลปัญญาประดิษฐ์ และการบริหารความเสี่ยง

๒.๑ ธรรมาภิบาลในการประยุกต์ใช้ AI (AI Governance)

ธรรมาภิบาลในการประยุกต์ใช้ AI คือ หลักการกำกับดูแลและการปฏิบัติงานในทุกกระบวนการที่เกี่ยวข้องกับการนำปัญญาประดิษฐ์ (AI) มาใช้งาน โดยมีการจัดให้มีมาตรการในการกำกับดูแลผ่านการกำหนดนโยบาย ขั้นตอนปฏิบัติ และ เครื่องมือในการปฏิบัติงาน

ทั้งนี้ เพื่อให้มั่นใจได้ว่าการประยุกต์ใช้ AI นั้นสามารถบรรลุตามเป้าหมายขององค์กรได้อย่างมีความรับผิดชอบ โดยคำนึงถึง

- ความสอดคล้องตามหลักการ จริยธรรมปัญญาประดิษฐ์
 - ความสอดคล้องตาม กฎหมายและข้อกำหนดที่เกี่ยวข้อง
 - และการมีกลไก **ควบคุมความเสี่ยง** ที่อาจส่งผลกระทบต่อบุคคลที่เกี่ยวข้อง องค์กร และสังคมโดยกว้าง
- มีการจัดตั้ง คณะกรรมการกำกับดูแลการประยุกต์ใช้ AI ซึ่งมีหน้าที่หลักในการขับเคลื่อนธรรมาภิบาล AI ในองค์กร ดังนี้

๑. **กำหนดทิศทางการดำเนินการ (Direction):** โดยการกำหนดกลยุทธ์ในการประยุกต์ใช้ AI และนโยบายที่เกี่ยวข้อง
๒. **เฝ้าติดตาม (Monitoring):** ติดตามประสิทธิภาพ (Performance) ของการประยุกต์ใช้ AI รวมถึงการปฏิบัติตามนโยบายและข้อกำหนดต่าง ๆ (Conformance) ทั้งภายในและภายนอกองค์กร
๓. **ประเมินผล (Evaluation):** ประเมินผลการประยุกต์ใช้งาน AI ในปัจจุบันและอนาคต โดยพิจารณาจากปัจจัยที่เกี่ยวข้องทั้งภายในและภายนอก เช่น เป้าหมายในการประยุกต์ใช้ AI ภัยคุกคามและโอกาสจากการประยุกต์ใช้ AI ความเปลี่ยนแปลงด้านเทคโนโลยี ประสิทธิภาพและประสิทธิผลในการกำกับดูแล เป็นต้น



ที่มา: สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์ (สพธอ.) และ กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม. (2564). แนวทางการประยุกต์ใช้ปัญญาประดิษฐ์อย่างมีธรรมาภิบาล สำหรับผู้บริหารองค์กร (AI Governance Guideline for Executives)

๒.๑ โครงสร้างการกำกับดูแล (AI Governance Structure) ของ กปภ.

คณะกรรมการกำกับดูแลการประยุกต์ใช้ AI (AI Governance Council) (ซึ่งอาจใช้ คณะกรรมการกำกับดูแลเทคโนโลยีดิจิทัลของ กปภ. (DTSC) หรือคณะกรรมการธรรมาภิบาลข้อมูล เป็น ผู้รับผิดชอบหลัก) จะทำหน้าที่กำหนดทิศทางเชิงกลยุทธ์และมีความรับผิดชอบสูงสุดต่อความเสี่ยง

ตารางที่ ๑ บทบาท หน้าที่และความรับผิดชอบหลัก

บทบาท	หน้าที่ความรับผิดชอบ
<p>คณะกรรมการกำกับดูแลการประยุกต์ใช้ AI Governance (AI Governance Council)</p>	<p>คณะกรรมการกำกับดูแลเทคโนโลยีดิจิทัลของ กปภ. (DT Steering Committee: DTSC) / คณะกรรมการกำกับดูแลข้อมูลของ กปภ. (Data Governance Council) จะทำหน้าที่เป็น คณะกรรมการกำกับดูแลการประยุกต์ใช้ AI (AI Governance Council) ประกอบด้วยผู้บริหารระดับสูง ผู้ทรงคุณวุฒิ ทำหน้าที่กำหนดทิศทางเชิงกลยุทธ์และนโยบายเพื่อขับเคลื่อนการประยุกต์ใช้ AI ภายในองค์กร รวมถึงมีหน้าที่ในการเฝ้าติดตาม และประเมินผล การประยุกต์ใช้ AI อย่างต่อเนื่อง ได้แก่</p> <ol style="list-style-type: none"> ๑. กำหนดนโยบาย มาตรฐาน และทิศทางการใช้งาน AI (AI Roadmap) ๒. ติดตามประสิทธิภาพ (Performance) ของการประยุกต์ใช้ AI รวมถึง การปฏิบัติงานตามนโยบายและข้อกำหนดต่าง ๆ (Conformance) ทั้ง ภายในและภายนอกองค์กร ๓. ประเมินผลการประยุกต์ใช้งาน AI ในปัจจุบัน โดยพิจารณาจากปัจจัยที่เกี่ยวข้องทั้งภายในและภายนอก เช่น เป้าหมายในการประยุกต์ใช้ AI ภัยคุกคามและโอกาสจากการประยุกต์ใช้ AI ความเปลี่ยนแปลงด้าน เทคโนโลยี ประสิทธิภาพและประสิทธิผลในการกำกับดูแล ๔. กำหนดระดับความเสี่ยงด้านปัญญาประดิษฐ์ที่องค์กรยอมรับได้ ๕. อนุมัติ พิจารณา หรือตัดสินใจด้านต่าง ๆ ที่เกี่ยวข้องกับการประยุกต์ใช้ AI เช่น <ul style="list-style-type: none"> • อนุมัตินโยบายและการใช้งานปัญญาประดิษฐ์อย่างมีจริยธรรม • จัดสรรทรัพยากร (งบประมาณ บุคลากร เทคโนโลยี) ให้เพียงพอต่อการดำเนินงานด้านปัญญาประดิษฐ์และความมั่นคงปลอดภัย • ส่งเสริม วัฒนธรรมการใช้งานปัญญาประดิษฐ์อย่างมีความรับผิดชอบทั่วทั้งองค์กร
<p>๒. คณะทำงานกำกับดูแลการใช้งานปัญญาประดิษฐ์ (AI Governance Working Group)</p>	<p>ประกอบด้วยตัวแทนผู้เชี่ยวชาญหลากหลายสาขา (Cross-Functional) อาทิ ด้านจริยธรรม กฎหมาย เทคโนโลยีดิจิทัล ธรรมาภิบาลข้อมูล ทรัพยากรบุคคล ความเสี่ยงและควบคุมภายใน สื่อสารองค์กรและลูกค้าสัมพันธ์ และความมั่นคงปลอดภัยไซเบอร์ ทำหน้าที่</p> <ol style="list-style-type: none"> ๑. ทบทวนนโยบายการกำกับดูแลการใช้งานปัญญาประดิษฐ์ (AI Governance) ของ กปภ. ให้มีความเหมาะสมกับปัจจุบัน ๒. เสนอเป้าหมาย ทิศทางการดำเนินงานการใช้งาน AI (AI Roadmap) และรายงานผลการดำเนินงานด้าน AI ของ กปภ. ให้กับคณะกรรมการกำกับดูแลการประยุกต์ใช้ AI (AI Governance Council) ทราบอย่างน้อยปีละครั้ง

บทบาท	หน้าที่ความรับผิดชอบ
	<p>๓. กำหนดแนวทางการใช้งาน AI ของ กปภ. ตามแนวทางของสำนักงานพัฒนาธุรกรรมอิเล็กทรอนิกส์ กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม หรือหน่วยงานของภาครัฐที่เกี่ยวข้อง</p> <p>๔. ตรวจสอบและสอบทานกระบวนการต่าง ๆ ที่มีการใช้งาน AI พร้อมรายงานคณะกรรมการกำกับดูแลการประยุกต์ใช้ AI ทราบอย่างน้อยปีละครั้ง</p> <p>๕. ติดตาม ตรวจสอบ และประเมินผลการใช้ AI พร้อมเสนอแนวทางปรับปรุงพัฒนา</p> <p>๖. รายงานเหตุการณ์ละเมิดความมั่นคงปลอดภัยที่เกิดจากการใช้งาน AI ให้กับผู้ที่เกี่ยวข้องทราบโดยทันที</p> <p>๗. สื่อสารเสริมสร้างความตระหนัkd้านการใช้ AI ให้กับพนักงานทุกระดับอย่างต่อเนื่อง</p>
<p>๓. ทีมพัฒนาและดำเนินงานระบบปัญญาประดิษฐ์</p>	<p>ผู้ออกแบบ พัฒนา และฝึกสอนโมเดลปัญญาประดิษฐ์</p> <ol style="list-style-type: none"> ๑) นำข้อกำหนดด้านความปลอดภัยมาปฏิบัติ ๒) สร้างโมเดลที่ทนทานต่อการโจมตี ๓) ใช้เทคนิคการคุ้มครองข้อมูลส่วนบุคคล ๔) จัดทำเอกสารที่เกี่ยวข้องกับโมเดล เพื่อระบุข้อจำกัด โอกาสที่อาจเกิดขึ้น และประสิทธิภาพของโมเดล
<p>๔. ทีมงานเทคนิคด้านความมั่นคงปลอดภัย</p>	<p>เป็นผู้เชี่ยวชาญที่ให้คำปรึกษา ตรวจสอบ และเฝ้าระวัง (“ผู้แนะนำและผู้ทวนสอบ”)</p> <ol style="list-style-type: none"> ๑) นำกระบวนการการสร้างแบบจำลองภัยคุกคามสำหรับระบบปัญญาประดิษฐ์ในระยการออกแบบ ๒) ดำเนินการทดสอบเจาะระบบสำหรับปัญญาประดิษฐ์ เพื่อค้นหาช่องโหว่ในระยการทวนสอบ ๓) ออกแบบและติดตั้งระบบบันทึกข้อมูลและเฝ้าระวังสำหรับตรวจจับการโจมตีที่จำเพาะเจาะจงกับปัญญาประดิษฐ์ ๔) พัฒนาคู่มือการรับมือเหตุการณ์สำหรับสถานการณ์ที่เกี่ยวข้องกับปัญญาประดิษฐ์โดยเฉพาะ
<p>๕. ทีมงานกฎหมายและกำกับการปฏิบัติตาม</p>	<p>ดูแลให้การดำเนินงานของระบบปัญญาประดิษฐ์สอดคล้องกับกฎหมาย กฎระเบียบ และมาตรฐานที่เกี่ยวข้อง</p> <ol style="list-style-type: none"> ๑) ติความและให้คำแนะนำเกี่ยวกับข้อกฎหมายคุ้มครองข้อมูลส่วนบุคคล เช่น PDPA ๒) ทบทวนและอนุมัตินโยบายการจัดเก็บและทำลายข้อมูล ๓) ประเมินความรับผิดทางกฎหมายที่อาจเกิดขึ้นจากการตัดสินใจที่ผิดพลาดของปัญญาประดิษฐ์

บทบาท	หน้าที่ความรับผิดชอบ
๖. เจ้าหน้าที่คุ้มครองข้อมูลส่วนบุคคล	กำกับดูแลและให้คำปรึกษาด้านการคุ้มครองข้อมูลส่วนบุคคล <ol style="list-style-type: none"> ๑) ดำเนินการประเมินผลกระทบด้านการคุ้มครองข้อมูลส่วนบุคคลสำหรับโครงการปัญญาประดิษฐ์ ๒) ให้คำปรึกษาเกี่ยวกับความชอบด้วยกฎหมายในการประมวลผลข้อมูลเพื่อใช้ฝึกสอนปัญญาประดิษฐ์ ๓) เป็นจุดประสานงานสำหรับเจ้าของข้อมูลในการใช้สิทธิ์ของตนที่เกี่ยวข้องกับระบบปัญญาประดิษฐ์
๗. ผู้ใช้งานระบบปัญญาประดิษฐ์	บุคลากรที่ใช้งานหรือมีปฏิสัมพันธ์กับระบบปัญญาประดิษฐ์โดยตรง <ol style="list-style-type: none"> ๑) ตระหนักรู้ว่ากำลังใช้งานหรือได้รับผลจากการทำงานของระบบปัญญาประดิษฐ์ ๒) ปฏิบัติตามนโยบายการใช้งานที่ยอมรับได้ของระบบนั้น ๆ ๓) รายงานพฤติกรรมของระบบที่ผิดปกติ น่าสงสัย หรืออาจก่อให้เกิดความเสียหาย

๒.๒ การบริหารความเสี่ยงปัญญาประดิษฐ์ (AI Risk Management)

กปภ. ต้องบูรณาการการบริหารความเสี่ยง AI เข้ากับกรอบการบริหารความเสี่ยงองค์กร (ERM) โดยใช้หลักการตามมาตรฐานที่เกี่ยวข้อง

๑. การประเมินความเสี่ยง (AI Risk Assessment) ต้องวิเคราะห์ความเสี่ยงและผลกระทบที่อาจเกิดขึ้นจากการประยุกต์ใช้ AI โดยต้องมีการประเมินความเสี่ยงเบื้องต้น (Initial Risk Assessment) สำหรับระบบ AI ที่จะนำมาใช้งาน

๒. การปฏิบัติตามกฎหมาย (Compliance Check) ต้องตรวจสอบความสอดคล้องกับกฎหมายและข้อบังคับของไทย เช่น พ.ร.บ. ไซเบอร์ พ.ศ. ๒๕๖๒ และ พ.ร.บ. คุ้มครองข้อมูลส่วนบุคคล พ.ศ. ๒๕๖๒

๓. การประเมินผลกระทบด้านข้อมูลส่วนบุคคล (DPIA) ต้องมีการดำเนินการประเมินผลกระทบด้านการคุ้มครองข้อมูลส่วนบุคคล (DPIA) หากระบบ AI เกี่ยวข้องกับข้อมูลส่วนบุคคล

๔. พันธกรณีสำหรับระบบความเสี่ยงสูง: หากระบบ AI ถูกจัดอยู่ในกลุ่มความเสี่ยงสูง จะต้องมีการจัดทำเอกสารข้อมูลทางเทคนิค (Technical Documentation) และคำแนะนำการใช้งานระบบในรูปแบบดิจิทัลหรือรูปแบบอื่น

๒.๒.๑ การบูรณาการความเสี่ยงปัญญาประดิษฐ์เข้ากับกรอบการบริหารความเสี่ยงองค์กร (Integrating AI Risks into ERM)

ในการดำเนินการบูรณาการความเสี่ยงอย่างเป็นระบบ องค์กรควรใช้มาตรฐานและกรอบการทำงานที่ได้รับการยอมรับในระดับสากล ได้แก่

๑. COSO ERM Framework (Integrating with Strategy and Performance) เป็นกรอบการทำงานที่ช่วยเชื่อมโยงการบริหารความเสี่ยงเข้ากับกลยุทธ์และผลการดำเนินงาน
๒. ISO ๓๑๐๐๐:๒๐๑๘ (Risk Management - Guidelines) เป็นมาตรฐานสากลที่ให้แนวทางและหลักการพื้นฐานสำหรับกระบวนการบริหารความเสี่ยงทั้งหมด
๓. ISO/IEC ๔๒๐๐๑:๒๐๒๓ Information technology - Artificial intelligence - Management system เป็นมาตรฐานระบบบริหารจัดการแรกของโลกที่มุ่งเน้นด้าน AI โดยเฉพาะ โดยใช้หลักการของระบบบริหารจัดการ (Management System Standard - MSS) เช่นเดียวกับ ISO ๙๐๐๑ (คุณภาพ) หรือ ISO/IEC ๒๗๐๐๑ (ความปลอดภัยข้อมูล)
๔. ISO/IEC ๒๓๘๔๔:๒๐๒๓ Information technology - Artificial intelligence - Guidance on risk management ซึ่งเป็นส่วนขยาย (Extension) จากมาตรฐาน ISO ๓๑๐๐๐:๒๐๑๘ Risk management - Guidelines ที่ อธิบายรายละเอียดเพิ่มเติมในส่วนของการบริหารจัดการความเสี่ยงจากการประยุกต์ใช้ AI
๕. Artificial Intelligence Risk Management Framework (AI RMF ๑.๐) โดย NIST เป็นแนวทางปฏิบัติที่ยืดหยุ่นและเป็นไปตามความสมัครใจ (Voluntary) ที่ถูกออกแบบมาเพื่อช่วยให้องค์กรบริหารจัดการความเสี่ยงที่อาจเกิดขึ้นตลอดวงจรชีวิตของระบบ AI ได้อย่างมีประสิทธิภาพ

๒.๒.๒ กระบวนการบูรณาการความเสี่ยงปัญญาประดิษฐ์เข้ากับประเภทความเสี่ยงใน ERM

๑. **ความเสี่ยงด้านกลยุทธ์:** ความเสี่ยงที่ส่งผลกระทบต่อความสามารถในการบรรลุเป้าหมายทางธุรกิจในระยะยาว ตัวอย่างเช่น การเสื่อมถอยของโมเดล (Model Degradation / Concept Drift) ที่ทำให้การตัดสินใจทางกลยุทธ์ผิดพลาด
๒. **ความเสี่ยงด้านปฏิบัติการ:** ความเสี่ยงที่จะเกิดความเสียหายจากความล้มเหลวของกระบวนการภายใน บุคลากร หรือระบบ ตัวอย่างเช่น การโจมตีห่วงโซ่อุปทาน (ใช้โมเดลที่ถูกฝัง Backdoor) หรือ Prompt Injection ที่ส่งผลให้ระบบอัตโนมัติของฝ่ายปฏิบัติการหยุดทำงาน
๓. **ความเสี่ยงด้านการเงิน:** ความเสี่ยงที่จะเกิดความสูญเสียทางการเงินทั้งทางตรงและทางอ้อม ตัวอย่างเช่น การรั่วไหลของทรัพย์สินทางปัญญา (คู่แข่งใช้ Model Extraction สร้างโมเดลลอกเลียนแบบ) ทำให้สูญเสียความได้เปรียบทางการแข่งขัน
๔. **ความเสี่ยงด้านการปฏิบัติตามกฎหมายและกฎระเบียบ:** ความเสี่ยงที่จะถูกลงโทษจากการไม่ปฏิบัติตามกฎหมาย ตัวอย่างเช่น การละเมิดกฎหมายคุ้มครองข้อมูลส่วนบุคคล (โมเดลเปิดเผยข้อมูลส่วนบุคคลที่ละเอียดอ่อนของลูกค้า)
๕. **ความเสี่ยงด้านชื่อเสียง:** ความเสี่ยงที่จะเกิดความเสียหายต่อภาพลักษณ์และความไว้วางใจขององค์กร ตัวอย่างเช่น การสร้างผลลัพธ์ที่มีอคติและไม่เป็นธรรม (ระบบคัดกรองงานมีอคติ) หรือ การถูกใช้เป็นเครื่องมือสร้าง Deep fake ของผู้บริหาร

๒.๒.๓ กระบวนการบริหารความเสี่ยงปัญญาประดิษฐ์ตามแนวทาง ISO/IEC ๒๓๘๙๔

การจัดการความเสี่ยงปัญญาประดิษฐ์ในทางปฏิบัติควรดำเนินการตามกระบวนการที่เป็นวงจรและต่อเนื่อง ตามมาตรฐาน ISO/IEC ๒๓๘๙๔:๒๐๒๓ ซึ่งประกอบด้วย ๔ ขั้นตอนสำคัญ

๑. **การระบุความเสี่ยง (Risk Identification)** ค้นหาและระบุภัยคุกคามและช่องโหว่ที่อาจเกิดขึ้นตลอดวงจรชีวิตของระบบปัญญาประดิษฐ์ ครอบคลุมความเสี่ยงทางเทคนิค (เช่น Data Poisoning, Prompt Injection) และความเสี่ยงเชิงองค์ประกอบ (ข้อมูล โมเดล แอปพลิเคชัน โครงสร้างพื้นฐาน)
๒. **การวิเคราะห์และประเมินความเสี่ยง (Risk Analysis and Assessment)** วิเคราะห์โอกาสที่จะเกิดและผลกระทบของความเสี่ยงนั้น ๆ (มักใช้ตารางเมทริกซ์ Likelihood × Impact Matrix) เพื่อจัดลำดับความสำคัญ (สูง กลาง ต่ำ)
๓. **การตอบสนองต่อความเสี่ยง (Risk Response)** กำหนดแนวทางการจัดการความเสี่ยงที่ประเมินไว้ โดยเลือกจาก ๔ แนวทางหลัก ได้แก่
 - **หลีกเลี่ยง (Avoid)** ยกเลิกกิจกรรมที่ก่อให้เกิดความเสี่ยง
 - **บรรเทา (Mitigate)** กำหนดมาตรการควบคุมเพื่อลดโอกาสเกิดหรือผลกระทบ (เช่น การควบคุมการเข้าถึง การเข้ารหัสลับ)
 - **ถ่ายโอน (Transfer)** โอนย้ายความรับผิดชอบไปยังบุคคลที่สาม (เช่น การทำประกันภัย)
 - **ยอมรับ (Accept)** ยอมรับความเสี่ยงหากอยู่ในระดับที่องค์กรรับได้
๔. **การติดตามและทบทวน (Monitoring and Review)** ตรวจสอบประสิทธิภาพของมาตรการควบคุมที่กำหนดไว้และทบทวนสถานะความเสี่ยงอย่างสม่ำเสมอ

๒.๒.๔ ทะเบียนความเสี่ยงด้านความมั่นคงปลอดภัยปัญญาประดิษฐ์

ทะเบียนความเสี่ยง (Risk Register) คือเอกสารกลางที่ใช้รวบรวม ติดตาม และจัดการความเสี่ยงทั้งหมดของระบบปัญญาประดิษฐ์อย่างต่อเนื่อง

ข้อมูลในทะเบียนความเสี่ยงจะมาจากผลลัพธ์ของกระบวนการบริหารความเสี่ยงในแต่ละขั้นตอน

- **ขั้นตอนที่ ๑ การระบุความเสี่ยง** → สร้างรายการความเสี่ยง (รหัสความเสี่ยง ชื่อความเสี่ยง รายละเอียด หมวดหมู่ความเสี่ยง)
- **ขั้นตอนที่ ๒ การวิเคราะห์และประเมิน** → ประเมินระดับความรุนแรง (โอกาสที่จะเกิด ผลกระทบ ระดับความเสี่ยง)
- **ขั้นตอนที่ ๓ การตอบสนอง** → กำหนดแผนจัดการ (แนวทางการตอบสนอง มาตรการควบคุม ผู้รับผิดชอบ กำหนดแล้วเสร็จ)
- **ขั้นตอนที่ ๔ การติดตามและทบทวน** → ปรับปรุงสถานะ (สถานะ วันที่ทบทวนล่าสุด วันที่ทบทวนครั้งถัดไป)

๒.๒.๕ การประเมินความเสี่ยง จะถูกแบ่งออกเป็น ๒ มิติหลัก คือ โอกาสที่จะเกิดขึ้น (Likelihood) และ ผลกระทบเชิงลบ (Adverse Impact) โดยมีการกำหนดค่านิยามของระดับคะแนน ๑-๕ ไว้อย่างชัดเจน ดังนี้

๑. ระดับโอกาสที่จะเกิดขึ้น (Likelihood) หมายถึง ความน่าจะเป็นของการเกิดเหตุการณ์ความเสี่ยงนั้น ๆ ทั้งจากสถิติในอดีตและโอกาสที่จะเกิดขึ้นในอนาคต แบ่งเป็น ๕ ระดับ ได้แก่

ระดับ ๕ สูงมาก โอกาสเกิดมากกว่า ๙๙% (เกิดบ่อยมาก หรืออาจเกิดขึ้นได้ภายในรอบวันถึงสัปดาห์)

ระดับ ๔ สูง โอกาสเกิดมากกว่า ๕๐% (เกิดขึ้นได้ง่าย หรืออาจเกิดขึ้นได้ภายในรอบสัปดาห์ถึงรอบเดือน)

ระดับ ๓ ปานกลาง โอกาสเกิดมากกว่า ๑๐% (เคยเกิดขึ้นแล้ว หรืออาจเกิดขึ้นได้ภายในรอบปี)

ระดับ ๒ น้อย โอกาสเกิดมากกว่า ๑% (เป็นไปได้แต่ยังไม่เคยเกิดขึ้น หรืออาจเกิดขึ้นได้ภายในรอบหลายปี)

ระดับ ๑ น้อยมาก โอกาสเกิดน้อยกว่า ๑% (เป็นไปได้เฉพาะกรณีฉุกเฉิน หรือเกิดขึ้นได้ยากมากแม้ในระยะยาว)

๒. ระดับผลกระทบ (Impact) หมายถึง ความรุนแรงของความเสียหายที่ส่งผลกระทบต่อองค์กรในมิติต่าง ๆ เช่น ประสิทธิภาพการดำเนินงาน ระยะเวลาที่ระบบ IT หยุดชะงัก (Downtime) ความเสียหายทางการเงิน ภาพลักษณ์ และความไม่พึงพอใจของผู้ใช้บริการ แบ่งเป็น ๕ ระดับ ได้แก่

ระดับ ๕ ผลกระทบสูงมาก

การดำเนินงาน พลาดเป้าหมาย > ๔๐%

ระบบสารสนเทศ หยุดชะงัก (Downtime) > ๒๔ ชั่วโมง

การเงิน เสียหาย > ๕% ของรายได้ปีที่ผ่านมา

ภาพลักษณ์/บริการ กระทบองค์กรในภาพรวม หรือผู้ให้บริการไม่พึงพอใจ > ๒๐%

ความปลอดภัย เกิดอันตรายถึงชีวิต หรืออันตรายต่อระบบแกนหลักขององค์กร

ระดับ ๔ ผลกระทบสูง

การดำเนินงาน พลาดเป้าหมาย > ๓๐% - ๔๐%

ระบบสารสนเทศ หยุดชะงัก > ๑๒ - ๒๔ ชั่วโมง

การเงิน เสียหาย > ๓% - ๕% ของรายได้ปีที่ผ่านมา

ภาพลักษณ์/บริการ กระทบหลายส่วนงาน หรือผู้ให้บริการไม่พึงพอใจ > ๑๖% - ๒๐%

ความปลอดภัย บาดเจ็บสาหัส หรือใช้เวลานานมากในการกู้ระบบแกนหลัก

ระดับ ๓ ผลกระทบปานกลาง

การดำเนินงาน พลาดเป้าหมาย > ๒๐% - ๓๐%

ระบบสารสนเทศ หยุดชะงัก > ๓ - ๑๒ ชั่วโมง

การเงิน เสียหาย > ๑% - ๓% ของรายได้ปีที่ผ่านมา

ภาพลักษณ์/บริการ กระทบส่วนงานภายใน หรือผู้ให้บริการไม่พึงพอใจ > ๑๑% - ๑๕%

ความปลอดภัย บาดเจ็บต้องรักษาตัว หรือใช้เวลาไม่นานในการกู้ระบบแกนหลัก

ระดับ ๒ ผลกระทบน้อย

การดำเนินงานพลาดเป้าหมาย > ๑๐% - ๒๐%

ระบบสารสนเทศ หยุดชะงัก > ๑ - ๓ ชั่วโมง

การเงิน เสียหาย > ๐.๕% - ๑% ของรายได้ปีที่ผ่านมา

ภาพลักษณ์/บริการ กระทบบางส่วนของงานภายใน หรือผู้ใช้บริการไม่พึงพอใจ > ๖% - ๑๐%

ความปลอดภัย บาดเจ็บเล็กน้อย หรือมีผลกระทบน้อยมากต่อระบบแกนหลัก

ระดับ ๑ ผลกระทบน้อยมาก

การดำเนินงานพลาดเป้าหมาย <= ๑๐%

ระบบสารสนเทศ หยุดชะงัก <= ๑ ชั่วโมง

การเงิน เสียหาย <= ๐.๕% ของรายได้ปีที่ผ่านมา

ภาพลักษณ์/บริการ กระทบเพียงบางทีม/บางบุคคล หรือผู้ใช้บริการไม่พึงพอใจ <= ๕%

ความปลอดภัย แค่ทำให้รำคาญใจ หรือไม่มีผลต่อระบบแกนหลักเลย

การกำหนดระดับความเสี่ยง (Risk Level) ของระบบปัญญาประดิษฐ์และเทคโนโลยีสารสนเทศ คำนวณจากการนำระดับคะแนน โอกาสที่จะเกิดขึ้น (Likelihood: ๑-๕) คูณด้วย ระดับผลกระทบ (Impact: ๑-๕) ซึ่งจะได้ผลลัพธ์เป็นค่าคะแนนความเสี่ยงรวมตั้งแต่ ๑ ถึง ๒๕ จากการบริหารความเสี่ยงระดับองค์กร (ERM) ของ กปภ. และแนวทางการจัดทำทะเบียนความเสี่ยงตามแนวปฏิบัติของ สกมช. สามารถอธิบายและจัดกลุ่มช่วงคะแนน ๑-๒๕ ออกเป็น ๔ ระดับหลัก พร้อมแนวทางการจัดการ ดังนี้

๑. คะแนน ๑ - ๔ : ระดับความเสี่ยงต่ำ (Low Risk)

ลักษณะ เหตุการณ์ที่มีโอกาสเกิดขึ้นน้อยมากถึงน้อย และหากเกิดขึ้นก็มีผลกระทบต่อองค์กรน้อยมาก

ตัวอย่างจากเอกสาร ความเสี่ยง AI-SEC-๐๐๖ (ปัญหาโมเดลเสื่อมสภาพในระบบงานที่ไม่สำคัญ) ประเมินโอกาสเกิดระดับ ๓ คูณ ผลกระทบระดับ ๑ ได้คะแนนรวม ๓ ถือเป็นความเสี่ยงระดับ "ต่ำ"

การจัดการ ยอมรับความเสี่ยง (Accept) ไม่จำเป็นต้องเพิ่มมาตรการควบคุมหรือลงทุนแก้ปัญหาเพิ่มเติมในขณะนี้ แต่ให้เฝ้าติดตามผลเป็นระยะว่าความเสี่ยงมีแนวโน้มรุนแรงขึ้นหรือไม่

๒. คะแนน ๕ - ๙ : ระดับความเสี่ยงปานกลาง (Medium Risk)

ลักษณะ เหตุการณ์ที่อาจเกิดขึ้นได้บ้าง และมีผลกระทบในระดับที่องค์กรพอจะบริหารจัดการได้ หรือมีผลกระทบน้อยแต่เกิดบ่อย

ตัวอย่างจากเอกสาร ความเสี่ยงระดับองค์กร R๖ ของ กปภ. (มาตรการป้องกันภัยคุกคามทางไซเบอร์ไม่เพียงพอ) ประเมินโอกาสเกิดระดับ ๓ คูณ ผลกระทบระดับ ๓ ได้คะแนนรวม ๙ จัดอยู่ในระดับ "ปานกลาง"

การจัดการ ควบคุมด้วยมาตรการที่มีอยู่ หรือพิจารณากำหนดมาตรการบรรเทาความเสี่ยงเพิ่มเติมหากมีความคุ้มค่า เพื่อป้องกันไม่ให้ความเสี่ยงยกระดับสูงขึ้น

๓. คะแนน ๑๐ - ๑๕ : ระดับความเสี่ยงสูง (High Risk)

ลักษณะ เหตุการณ์ที่อาจเกิดขึ้นได้ง่ายและมีผลกระทบรุนแรง หรือเป็นเหตุการณ์ที่โอกาสเกิดน้อยแต่หากเกิดขึ้นแล้วจะสร้างความเสียหายอย่างหนัก

ตัวอย่างจากเอกสาร

ความเสี่ยง AI-SEC-๐๐๔ (การถูกโจมตีทางไซเบอร์แบบ DoS บน API ของ AI) โอกาสเกิดระดับ ๒ คุณ ผลกระทบระดับ ๕ ได้คะแนนรวม ๑๐ (ระดับสูง)

ความเสี่ยง AI-SEC-๐๐๑ (การโจมตีด้วยข้อมูลปนเปื้อน Data Poisoning) โอกาสเกิดระดับ ๓ คุณ ผลกระทบระดับ ๔ ได้คะแนนรวม ๑๒ (ระดับสูง)

ความเสี่ยงระดับองค์กรของ กปภ. R๔/R๕ โอกาสเกิดระดับ ๕ คุณ ผลกระทบระดับ ๓ ได้คะแนนรวม ๑๕ (ระดับสูง)

(หมายเหตุ ในเกณฑ์ของ สกมช. บางกรณีประเมินคะแนน ๑๕ ไว้ในกลุ่มความเสี่ยงสูงมาก)

การจัดการ บรรเทา (Mitigate) หรือ ถ่ายโอน (Transfer) ต้องจัดทำแผนจัดการความเสี่ยงและกำหนดมาตรการควบคุมอย่างชัดเจน เช่น การใช้เครื่องมือคัดกรองข้อมูล หรือการถ่ายโอนความเสี่ยงโดยใช้บริการระบบคลาวด์ที่มีระบบป้องกันภัยคุกคามในตัว

๔. คะแนน ๑๖ - ๒๕ : ระดับความเสี่ยงสูงมาก / วิกฤต (Very High / Critical Risk)

ลักษณะ เหตุการณ์ที่มีโอกาสเกิดขึ้นสูงถึงสูงมาก และส่งผลกระทบรุนแรงมากถึงขั้นทำให้ระบบหลักหยุดชะงักหรือเกิดความเสียหายทางการเงินและชื่อเสียงอย่างมหาศาล

ตัวอย่างจากเอกสาร

ความเสี่ยงระดับองค์กรของ กปภ. R๒ และ R๓ โอกาสเกิดระดับ ๔ คุณ ผลกระทบระดับ ๔ ได้คะแนนรวม ๑๖ (สูงมาก)

ความเสี่ยงระดับองค์กรของ กปภ. R๑ โอกาสเกิดระดับ ๕ คุณ ผลกระทบระดับ ๔ ได้คะแนนรวม ๒๐ (สูงมาก)

ความเสี่ยง AI-SEC-๐๐๕ (การนำโมเดล AI จากแหล่งที่ไม่น่าเชื่อถือมาใช้งาน) โอกาสเกิดระดับ ๔ คุณ ผลกระทบระดับ ๕ ได้คะแนนรวม ๒๐ จัดอยู่ในระดับ "วิกฤต"

การจัดการ หลีกเลี่ยง (Avoid) หรือต้องมีมาตรการเข้าจัดการอย่างเร่งด่วนและเข้มงวดที่สุด เป็นระดับความเสี่ยงที่องค์กร "ยอมรับไม่ได้" ผู้บริหารระดับสูงต้องเข้ามามีส่วนร่วมตัดสินใจ หากเป็นระบบ AI อาจต้องไม่อนุมัติให้จัดหาหรือระงับการใช้งานระบบทันทีจนกว่าจะแก้ไขความเสี่ยงได้

ตารางที่ ๒ ตัวอย่างตารางทะเบียนความเสี่ยง

รหัส	ชื่อความเสี่ยง	รายละเอียด	หมวดหมู่	โอกาส (๑-๕)	ผลกระทบ (๑-๕)	ระดับ ความ เสี่ยง	แนว ทางการ ตอบสนอง	มาตรการควบคุม	ผู้รับผิดชอบ	กำหนด เสร็จ	สถานะ
AI- SEC- ๐๐๑	Data Poisoning Attack	ผู้ไม่หวังดีแฝงข้อมูลที่เป็นพิษเข้ามาในชุดข้อมูลฝึกสอน ทำให้โมเดลเรียนรู้และตัดสินใจผิดพลาด	ปฏิบัติการ	๓ (กลาง)	๔ (สูง)	๑๒ (สูง)	บรรเทา	๑. ตรวจสอบความถูกต้องของแหล่งข้อมูล ๒. ใช้ Anomaly Detection ตรวจสอบจับข้อมูลผิดปกติ ๓. จำกัดสิทธิ์การเข้าถึงข้อมูลฝึกสอน	Head of Data Science	๓๑/ ๑๒/ ๒๕๖๘	กำลัง ดำเนินการ
AI- SEC- ๐๐๒	ผลลัพธ์ที่มีอคติ (Biased Outcome)	โมเดลคัดกรองผู้สมัครงานมีอคติต่อเพศ/เชื้อชาติเนื่องจากข้อมูลที่ใช้สอนมีอคติแฝงอยู่	ชื่อเสียง / ปฏิบัติ ตาม กฎหมาย	๔ (สูง)	๔ (สูง)	๑๖ (สูง มาก)	หลีกเลี่ยง (Avoid)	๑. ใช้เครื่องมือตรวจสอบอคติ ๒. จัดตั้งทีมตรวจสอบด้านจริยธรรม ปัญญาประดิษฐ์ ๓. ปรับปรุงชุดข้อมูลให้มีความหลากหลาย	Chief People Officer	๓๐/ ๐๙/ ๒๕๖๘	กำลัง ดำเนินการ

๒.๒.๖ ความสัมพันธ์ระหว่างทะเบียนความเสี่ยงและตัวชี้วัดความเสี่ยงหลัก (KRI Mapping)

หลังจากที่มีทะเบียนความเสี่ยงแล้ว ขั้นตอนต่อไปคือ การกำหนด ตัวชี้วัดความเสี่ยงหลัก (Key Risk Indicators - KRIs) เพื่อใช้เป็นระบบแจ้งเตือนล่วงหน้า เพื่อติดตาม "ความเสี่ยง" ที่วินิจฉัยไว้ในทะเบียนความเสี่ยง หากค่า KRI ผิดปกติ ก็เป็นสัญญาณเตือนว่าความเสี่ยงกำลังจะเกิดขึ้นจริง ดังนั้น KRIs จึงเป็นตัวชี้วัดที่ออกแบบมาเพื่อติดตามความเสี่ยงและวัดประสิทธิผลของมาตรการควบคุม

ตารางที่ ๓ ตัวอย่างการเชื่อมโยงความเสี่ยงด้านปัญญาประดิษฐ์ตัวชี้วัดความเสี่ยงหลัก (KRI Mapping)

รหัสความเสี่ยง (จากทะเบียน ความเสี่ยง)	ชื่อความเสี่ยง	มาตรการควบคุม (ตัวอย่าง)	ตัวชี้วัดความเสี่ยงหลัก ที่เกี่ยวข้อง	เกณฑ์แจ้งเตือน
AI-SEC-๐๐๑	Data Poisoning Attack	ใช้ Anomaly Detection ตรวจสอบ ข้อมูลผิดปกติ	* จำนวนครั้งที่ระบบ ตรวจจับข้อมูลผิดปกติ ต่อสัปดาห์	> ๑๐ ครั้ง/สัปดาห์ (Yellow) > ๒๐ ครั้ง/สัปดาห์ (Red)
AI-SEC-๐๐๒	ผลลัพธ์ที่มี อคติ	ใช้เครื่องมือ ตรวจสอบ อคติ	* ค่าคะแนนความเท่า เทียม ของโมเดล เช่น Disparate Impact Ratio	< ๐.๙ (Yellow) < ๐.๘ (Red)

๒.๓ การปฏิบัติตามกฎหมายและข้อบังคับของไทย

ระบบปัญญาประดิษฐ์ต้องอยู่ภายใต้กรอบของกฎหมายและกฎระเบียบที่มีผลบังคับใช้ การละเลยการปฏิบัติตามข้อกำหนดเหล่านี้อาจนำไปสู่ความรับผิดทางกฎหมาย โทษปรับ และความเสียหายต่อชื่อเสียงขององค์กรอย่างรุนแรง กฎหมายไทยสองฉบับที่มีผลกระทบโดยตรงต่อการพัฒนาและใช้งานระบบปัญญาประดิษฐ์คือ พระราชบัญญัติการรักษาความมั่นคงปลอดภัยไซเบอร์ พ.ศ. ๒๕๖๒ และพระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. ๒๕๖๒

๑) พระราชบัญญัติการรักษาความมั่นคงปลอดภัยไซเบอร์ พ.ศ. ๒๕๖๒ (พ.ร.บ. ไซเบอร์)

วัตถุประสงค์ จัดตั้งกลไกการป้องกัน รับมือ และลดความเสี่ยงจากภัยคุกคามทางไซเบอร์ที่อาจส่งผลกระทบต่อเสถียรภาพและความมั่นคงของประเทศ

การประยุกต์ใช้กับ AI หากระบบ AI ถูกใช้เป็นส่วนหนึ่งหรือใช้ในการควบคุมจัดการระบบที่จัดเป็นโครงสร้างพื้นฐานสำคัญทางสารสนเทศ (CII) ระบบ AI นั้นจะตกอยู่ภายใต้ข้อบังคับของ พ.ร.บ. ไซเบอร์โดยอัตโนมัติ

ข้อกำหนด องค์กรที่เป็นเจ้าของระบบ (หน่วยงานควบคุมหรือกำกับดูแล) ต้องปฏิบัติตามมาตรฐานและแนวปฏิบัติขั้นต่ำที่คณะกรรมการการรักษาความมั่นคงปลอดภัยไซเบอร์แห่งชาติ (สกมช.) กำหนด

๒) พระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. ๒๕๖๒ (PDPA)

วัตถุประสงค์ ให้สิทธิแก่บุคคลธรรมดา (เจ้าของข้อมูล) ในการควบคุมการเก็บ ใช้ และเปิดเผยข้อมูลส่วนบุคคลของตน

การประยุกต์ใช้กับ AI ตลอดวงจรชีวิตของระบบ AI เกี่ยวข้องกับ PDPA อย่างแนบแน่น ทั้งการรวบรวมข้อมูลเพื่อ "ฝึกสอน" และการนำข้อมูลมา "ประมวลผล" เพื่อให้ได้ผลลัพธ์

คำแนะนำ ทีมกฎหมายและ DPO ควรเข้ามามีส่วนร่วมตั้งแต่ระยะขั้นตอนแนวคิดของโครงการ

๓) พระราชบัญญัติว่าด้วยการกระทำความผิดเกี่ยวกับคอมพิวเตอร์ พ.ศ. ๒๕๕๐ (พ.ร.บ. คอมฯ) และฉบับแก้ไขเพิ่มเติม

การประยุกต์ใช้กับ AI ข้อมูลนำเข้า (Prompt) และข้อมูลผลลัพธ์ (ข้อความ ภาพ วิดีโอ) ถือเป็นข้อมูลคอมพิวเตอร์ภายใต้ พ.ร.บ. นี้

ความรับผิดของผู้ให้บริการ (มาตรา ๑๕) ผู้ให้บริการหรือเจ้าของแพลตฟอร์มอาจต้องร่วมรับผิด หากให้ความร่วมมือ ยินยอม หรือรู้เห็นเป็นใจให้มีการกระทำความผิดเกิดขึ้นในระบบของตน เช่น การขาดมาตรการป้องกันทางเทคนิคที่สมเหตุสมผลสำหรับ Generative AI

ความรับผิดของผู้ใช้งาน ผู้ใช้งานต้องรับผิดชอบโดยตรงต่อเนื้อหาที่ตนสั่งให้ระบบ AI สร้างขึ้น

- o **การนำเข้าข้อมูลเท็จ** (มาตรา ๑๔ (๑) และ (๒)) การสร้าง Prompt เพื่อให้ AI สร้างเนื้อหาอันเป็นเท็จที่น่าจะเกิดความเสียหายต่อความมั่นคงทางเศรษฐกิจของประเทศ หรือก่อให้เกิดความตื่นตระหนก แล้วนำไปเผยแพร่ ถือเป็นความผิดโดยตรง
- o **การตัดต่อภาพผู้อื่น** (มาตรา ๑๖) การใช้ Generative AI ตัดต่อหรือดัดแปลงภาพของบุคคลอื่นในลักษณะที่น่าจะทำให้เสียชื่อเสียง ถูกดูหมิ่น หรือถูกเกลียดชัง แล้วนำไปเผยแพร่ เป็นความผิด
- o **ความผิดเกี่ยวกับความมั่นคงแห่งราชอาณาจักร** (มาตรา ๑๔ (๓)) การใช้ Generative AI สร้างดัดแปลง หรือตัดต่อภาพของพระบรมวงศานุวงศ์ในลักษณะที่เป็นการหมิ่นประมาท ดูหมิ่น หรือแสดงความอาฆาตมาดร้าย ถือเป็นความผิดร้ายแรง

มาตรการควบคุมที่จำเป็น: องค์กรที่อนุญาตให้ใช้งาน Generative AI ควรกำหนดให้เนื้อหาลักษณะดังกล่าวอยู่ใน **ระดับความเสี่ยงที่ยอมรับไม่ได้** และต้องมีการออกแบบกลไกทางเทคนิคเพื่อป้องกันผลลัพธ์ที่เป็นอันตราย (Output Filtering) และบังคับใช้นโยบายการใช้เทคโนโลยีที่ยอมรับได้ (AUP)

๔) ข้อพิจารณาเพิ่มเติม การรักษาอธิปไตยทางข้อมูลในระบบปัญญาประดิษฐ์

หมายถึง อำนาจในการควบคุมและบังคับใช้กฎหมายของประเทศไทยเหนือข้อมูลที่ถูกสร้าง จัดเก็บ และประมวลผลภายในประเทศ

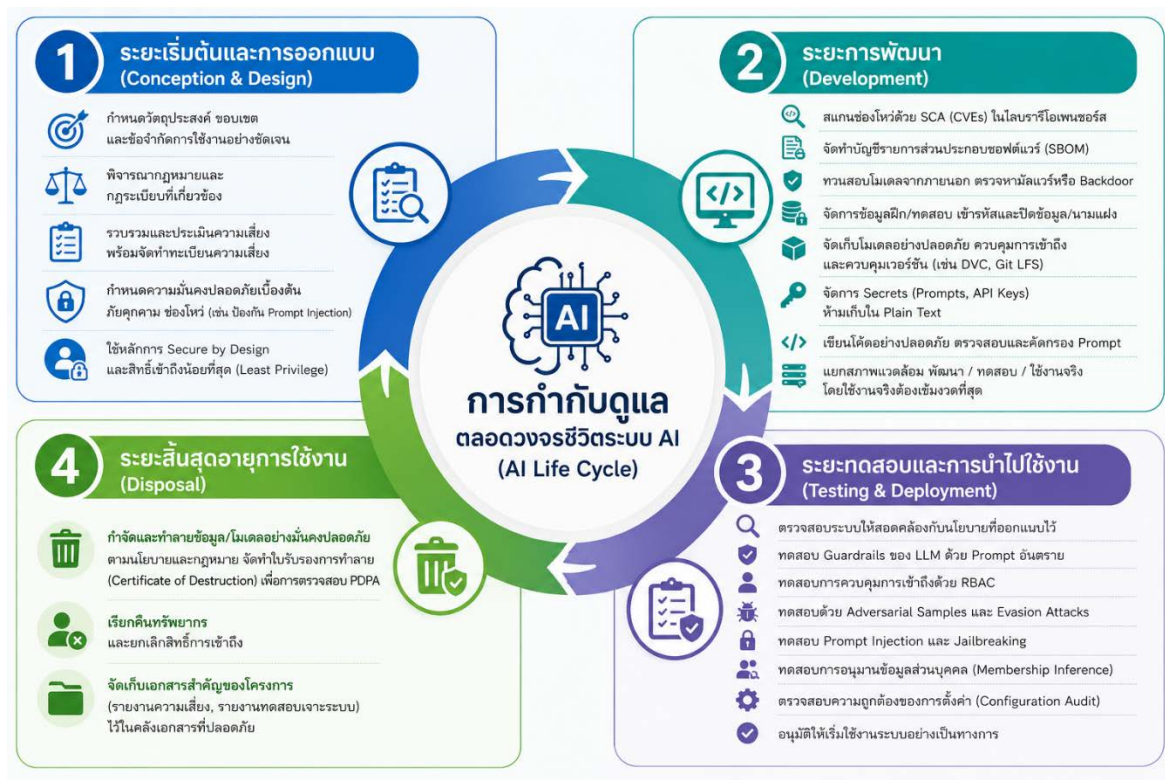
- O หลักการด้านที่ตั้งของข้อมูล (Data Location Principle) ข้อมูลที่สำคัญของชาติ จะต้องถูกจัดเก็บและประมวลผลอยู่ในอาณาเขตของประเทศไทย ตามประกาศ สกมช. เรื่อง มาตรฐานด้านความมั่นคงปลอดภัยไซเบอร์ระบบคลาวด์ พ.ศ. ๒๕๖๗ ชุด ข้อมูลที่ใช้ฝึกสอนและข้อมูลที่เกิดจากการทำงานของระบบ AI ที่มีความสำคัญต่อองค์กร จะต้องถูกจัดเก็บและประมวลผลบนโครงสร้างพื้นฐานที่อยู่ภายใต้เขตอำนาจกฎหมาย ไทย
- O หลักการจำแนกข้อมูลตามระดับผลกระทบ (Data Classification Principle) การ ควบคุมต้องจำแนกตามความสำคัญและผลกระทบที่อาจเกิดขึ้น (ระดับต่ำ กลาง สูง) ตามประกาศ สกมช. เรื่อง มาตรฐานการกำหนดคุณลักษณะความมั่นคงปลอดภัย ไซเบอร์ ข้อมูลผลกระทบระดับสูง (เช่น ข้อมูลสาธารณะสุข) ต้องใช้มาตรการด้านที่ตั้ง และการควบคุมที่เข้มงวดสูงสุด ข้อมูลส่วนบุคคลต้องถูกจัดระดับผลกระทบด้านการ รักษาความลับ ไม่น้อยกว่าระดับกลาง
- O หลักการความรับผิดชอบร่วมกันในบริการคลาวด์ (Shared Responsibility in Cloud Services) หากใช้บริการ AI บนคลาวด์สาธารณะ การรักษาอธิปไตยทางข้อมูล จะเป็นความรับผิดชอบร่วมกันระหว่างผู้ให้บริการ (CSC) และผู้ให้บริการ (CSP) องค์กร ต้องตรวจสอบสัญญาอย่างละเอียดเพื่อยืนยันว่า CSP ปฏิบัติตามกฎหมายไทยและ ยืนยัน ที่ตั้งทางภูมิศาสตร์ ของการจัดเก็บและประมวลผลข้อมูล

บทที่ ๓

แนวปฏิบัติเชิงปฏิบัติการด้านการใช้งานและความมั่นคงปลอดภัย

๓.๑ การกำกับดูแลตลอดวงจรชีวิตระบบ AI (AI Life Cycle Governance)

การกำกับดูแลต้องครอบคลุมทุกระยะของวงจรชีวิตของระบบ AI (AI Life Cycle)



๑. ระยะเริ่มต้นและการออกแบบ (Conception & Design)

- ต้องกำหนดวัตถุประสงค์ ขอบเขต และข้อจำกัดการใช้งานของระบบ AI อย่างชัดเจน
- ต้องพิจารณากฎหมายและกฎระเบียบที่เกี่ยวข้อง
- รวบรวมและประเมินความเสี่ยงพร้อมจัดทำทะเบียนความเสี่ยง
- ข้อกำหนดด้านความมั่นคงปลอดภัยเบื้องต้น ภัยคุกคาม ช่องโหว่ (เช่น กลไกป้องกัน Prompt Injection)
- ต้องใช้หลักการ Secure by Design และการกำหนดสิทธิ์เข้าถึงที่น้อยที่สุด (Principle of Least Privilege)

๒. ระยะการพัฒนา (Development)

- ใช้เครื่องมือ Software Composition Analysis (SCA) เพื่อสแกนหาช่องโหว่ (CVEs) ในไลบรารีโอเพนซอร์ส
- จัดทำบัญชีรายการส่วนประกอบซอฟต์แวร์ (SBOM) เพื่อติดตามส่วนประกอบทั้งหมด
- ทวนสอบโมเดลที่ฝึกไว้ล่วงหน้าจากภายนอกเพื่อตรวจหลุมลับหรือ Backdoor

- ชุดข้อมูลฝึกสอน/ทดสอบ บังคับใช้นโยบายการ เข้ารหัสลับข้อมูลขณะจัดเก็บ และใช้เทคนิคการปิด ข้อมูล/นามแฝง

- โมเดลปัญญาประดิษฐ์ จัดเก็บในที่ปลอดภัย, ควบคุมการเข้าถึง, และใช้ระบบ ควบคุมเวอร์ชัน ของ โมเดล (เช่น DVC หรือ Git LFS)

- ข้อมูลลับ (Secrets) จัดการ Prompts และ API Keys เสมือนเป็นข้อมูลลับ และห้ามเก็บเป็นข้อความ ธรรมดา (Plain Text) ในซอร์สโค้ด

- การเขียนโค้ดอย่างปลอดภัย โค้ดที่รับ Prompt ต้องผ่านการตรวจสอบและคัดกรองอย่างเข้มงวด (Input Validation หรือ Prompt Sanitization)

- การแบ่งแยกสภาพแวดล้อม แยกสภาพแวดล้อม การพัฒนา การทดสอบ และ การใช้งานจริง ออกจาก กัน โดยสภาพแวดล้อมใช้งานจริงต้องมีมาตรการควบคุมที่เข้มงวดที่สุด

๓. ระยะทดสอบและการนำไปใช้งาน (Testing & Deployment)

- ตรวจสอบระบบที่สร้างขึ้นสอดคล้องกับนโยบายที่กำหนดไว้ในระยะออกแบบ (เช่น ทวนสอบการกำกับ ดูแลข้อมูล, การทวนสอบสถาปัตยกรรม)

- ทดสอบ Guardrails ของ LLM โดยการสร้าง Prompt ที่มีเนื้อหาอันตรายเพื่อทวนสอบว่าระบบ สามารถปฏิเสธการตอบสนองได้

- ทดสอบการควบคุมการเข้าถึงด้วยสิทธิ์ผู้ใช้งานทั่วไป (RBAC)

- ทดสอบด้วยตัวอย่างที่เป็นปฏิปักษ์ (Adversarial Samples) วัดความทนทานต่อ Evasion Attacks

- ทดสอบ Prompt Injection และ Jailbreaking สำหรับ LLM

- ทดสอบการอนุมานข้อมูลส่วนบุคคล เช่น Membership Inference

- ตรวจสอบความถูกต้องของการตั้งค่า (Configuration Audit) การติดตั้งและเปิดใช้งานมาตรการ ควบคุมความปลอดภัยทั้งหมด

- การอนุมัติให้เริ่มใช้งานระบบอย่างเป็นทางการ

๔. ระยะสิ้นสุดอายุการใช้งาน (Disposal)

- การกำจัดและทำลายข้อมูลและโมเดลอย่างมั่นคงปลอดภัย ต้องเป็นไปตามนโยบายขององค์กรและ กฎหมาย จัดทำ ใบรับรองการทำลาย (Certificate of Destruction) เพื่อเป็นหลักฐานสำคัญสำหรับการตรวจสอบ ด้าน PDPA

- การเรียกคืนทรัพยากรและยกเลิกสิทธิ์การเข้าถึง

- จัดเก็บเอกสารสำคัญทั้งหมดของโครงการ (รายงานความเสี่ยง, รายงานทดสอบเจาะระบบ) ไว้ในคลัง เอกสารที่ปลอดภัยตามนโยบายองค์กร

๓.๒ แนวปฏิบัติการใช้งานที่ยอมรับได้ (Acceptable Use Policy - AUP)

การปฏิบัติตามแนวทางที่กำหนด สำหรับพนักงาน ลูกจ้าง ผู้ปฏิบัติงาน และผู้รับจ้างของ กปภ. ตาม เอกสารแนวปฏิบัติการใช้งานที่ยอมรับได้ (Acceptable Use Policy - AUP)

๓.๒.๑ การใช้งานที่ยอมรับได้ (Acceptable Use) การใช้งาน AI ต้องเป็นไปเพื่อสนับสนุนภารกิจและเป็นประโยชน์ต่อ กปภ. เท่านั้น เช่น

- การวิเคราะห์ข้อมูลและสร้างรายงาน
- การเขียนชุดคำสั่งหรือโปรแกรม (Source Code)
- การร่างเอกสาร นโยบาย สื่อประชาสัมพันธ์ และการสื่อสารภายในองค์กร

๓.๒.๒ ข้อห้ามและจริยธรรมปัญญาประดิษฐ์ (Prohibited Use & AI Ethics)

- ห้ามใช้ AI ตัดสินใจในเรื่องที่มีความเสี่ยงสูง (การแพทย์ กฎหมาย การเงิน สิทธิของบุคคล) แทนมนุษย์
- ห้ามใช้สร้างเนื้อหาที่ผิดกฎหมาย บิดเบือนความจริง (Fake News) หรือทำให้เกิดความเสียหาย
- ห้ามใช้เพื่อก่อให้เกิดการหยุดชะงัก ขัดขวาง หรือรบกวนระบบสารสนเทศและบริการของ กปภ.
- ห้ามสร้างเนื้อหาที่ส่งเสริมการเหยียดเชื้อชาติ ศาสนา เพศ วัย หรือขัดต่อหลักสิทธิมนุษยชน

๓.๒.๓ การคุ้มครองข้อมูลส่วนบุคคลและการรักษาความลับ (Data Privacy & Confidentiality)

- ห้ามนำข้อมูลภายใน ข้อมูลความลับ (เช่น รหัสผ่าน สัญญา) ข้อมูลส่วนบุคคล และข้อมูลอ่อนไหว ไปใช้กับบริการ AI ทั่วไปโดยเด็ดขาด

๓.๒.๔ การรักษาความมั่นคงปลอดภัย (Information Security)

- ห้ามป้อนข้อมูลที่กระทบต่อระบบ (เช่น API Key, System Configuration) ลงใน AI
- ซอร์สโค้ด (Source Code) ที่สร้างโดย AI ต้องได้รับการตรวจสอบความถูกต้องและค้นหาช่องโหว่ (Vulnerability Scan) ก่อนนำมาใช้งานเสมอ
- หากพบเหตุละเมิดความมั่นคงปลอดภัยจาก AI ต้องรายงานผู้บังคับบัญชาและปฏิบัติตามแผนเผชิญเหตุ (Incident Response) ทันที

๓.๒.๕ การลดอคติและการตรวจสอบ (Bias Reduction & Human-in-the-Loop)

- ผู้ใช้งานมีหน้าที่พิจารณา "ตรวจสอบความถูกต้องและอคติ" ของเนื้อหาที่สร้างโดย AI ก่อนนำไปใช้งานหรือเผยแพร่ทุกครั้ง
- หากนำเนื้อหาจาก AI ไปเผยแพร่ ควรมีการระบุไว้ “เนื้อหานี้ได้รับความช่วยเหลือจากเทคโนโลยี AI” เพื่อความโปร่งใส

๓.๒.๖ ทรัพย์สินทางปัญญา (Intellectual Property)

- ต้องระมัดระวังไม่ให้เกิดการใช้ AI ไปละเมิดลิขสิทธิ์ เครื่องหมายการค้า หรือทรัพย์สินทางปัญญาของผู้อื่น

๓.๒.๗ ข้อควรปฏิบัติและไม่ควรปฏิบัติ (Do's & Don'ts) ของการใช้งานเทคโนโลยี Generative AI

ผู้บริหารและพนักงานทุกคนต้องใช้ AI อย่างระมัดระวังและมีความรับผิดชอบ

พฤติกรรมที่ควรทำ (Should Do)	พฤติกรรมที่ไม่ควรทำ (Should Not Do)
✓ ตรวจสอบความถูกต้องและความสมเหตุสมผลของผลลัพธ์จากปัญญาประดิษฐ์ทุกครั้ง	X เชื่อผลการวิเคราะห์หรือการตัดสินใจของระบบปัญญาประดิษฐ์ทันทีโดยไม่ใช้วิจารณญาณ
✓ ใช้ปัญญาประดิษฐ์เป็นเครื่องมือสนับสนุนโดยมนุษย์เป็นผู้ตัดสินใจขั้นสุดท้าย	X ใช้ระบบปัญญาประดิษฐ์ในการตัดสินใจที่สำคัญโดยไม่มี การกำกับดูแลจากมนุษย์
✓ ใช้ข้อมูลที่สะอาดและได้รับอนุญาตป้อนเข้าสู่ระบบปัญญาประดิษฐ์	X ป้อนข้อมูลที่เป็นความลับ ข้อมูลส่วนบุคคล หรือข้อมูล ที่ไม่น่าเชื่อถือเข้าสู่ระบบ
✓ อ้างอิงและให้เครดิตว่ามีการใช้ปัญญาประดิษฐ์ช่วยในการทำงานตามความเหมาะสม	X นำผลงานจากระบบปัญญาประดิษฐ์ไปใช้โดยอ้างว่าเป็น ผลงานของตนเองทั้งหมด
✓ ปฏิบัติตามนโยบายและแนวทางขององค์กรอย่างเคร่งครัด	X ใช้จากระบบปัญญาประดิษฐ์ในทางที่ผิดกฎหมาย หรือ ขัดต่อหลักธรรมาภิบาล คุณธรรม และจริยธรรม
✓ ตรวจสอบช่องโหว่ (Vulnerability Scan) ซอร์สโค้ดที่สร้างโดย AI ก่อนนำไปพัฒนาต่อ	X เผยแพร่ข้อมูล ที่ไม่ถูกต้อง (Misinformation) หรือ บิดเบือน (Disinformation) ไม่มีการอ้างอิง แหล่งที่มาของ ข้อมูล
✓ หลีกเลี่ยง <ul style="list-style-type: none"> • การสร้างเนื้อหาที่มีอคติ หรือเลือกปฏิบัติ • การใช้เนื้อหาที่อาจละเมิดลิขสิทธิ์ หรือ ทรัพย์สินทางปัญญา • การใช้งานเทคโนโลยี Generative AI ในทางที่ผิดกฎหมายหรือขัดต่อหลักธรรมาภิบาล คุณธรรมและจริยธรรม 	X ห้าม <ul style="list-style-type: none"> • นำข้อมูลชั้นความลับ ข้อมูลส่วนบุคคล ไปป้อนใส่ AI สาธารณะ • นำผลลัพธ์จาก AI ไปใช้งานทันทีโดยไม่ตรวจสอบ (Copy & Paste) • ใช้ AI สร้างข้อมูลเท็จ บิดเบือน หรือสร้างความเกลียดชัง • นำ Code จาก AI ไปใช้ในระบบโดยไม่ผ่านการตรวจสอบช่องโหว่ความปลอดภัย • สร้างเนื้อหาที่มีอคติ หรือเลือกปฏิบัติ • ใช้เนื้อหาที่อาจละเมิดลิขสิทธิ์ หรือทรัพย์สินทางปัญญา • นำข้อมูลที่เป็นความลับของสำนักงานไปใช้งานร่วมกับเทคโนโลยี Generative AI

๓.๓ การจัดการข้อมูลเพื่อรองรับ AI

การประมวลผลข้อมูลสำหรับ AI ต้องเป็นไปตาม "นโยบายธรรมาภิบาลข้อมูลของ กปภ."

๑. การคุ้มครองข้อมูลส่วนบุคคล ห้ามนำข้อมูลที่มีความสำคัญและมีผลกระทบสูงมากต่อความมั่นคงปลอดภัยของ กปภ. ออกไปใช้งานภายนอก ต้องมีการจัดทำเอกสารข้อตกลงเรื่องการรักษาความปลอดภัยของข้อมูลสารสนเทศ (NDA) สำหรับพนักงานและบุคคลภายนอกที่เกี่ยวข้อง

๒. การจำแนกข้อมูล ต้องกำหนดชั้นความลับและความอ่อนไหวของข้อมูลอย่างชัดเจน ชุดข้อมูลที่มีระดับความเสี่ยงสูง เช่น ข้อมูลส่วนบุคคล (ระดับ ๓ หรือ ๔) ต้องได้รับการคุ้มครองอย่างเคร่งครัด

๓. การบูรณาการข้อมูล ต้องกำหนดมาตรฐานทางเทคนิคสำหรับการเชื่อมโยงและแลกเปลี่ยนข้อมูล โดยแนะนำให้ใช้รูปแบบที่เป็นมาตรฐานเปิด (Open Format) เช่น XML และ JSON ข้อมูลต้องได้รับการตรวจสอบคุณภาพ กำจัดข้อมูลซ้ำ และแก้ไขข้อมูลที่มีค่าว่าง

บทที่ ๔

การตรวจสอบและการปรับปรุงอย่างต่อเนื่อง

๔.๑ การตอบสนองต่อเหตุการณ์ผิดปกติ (Incident Response)

๑. การเฝ้าระวัง ต้องมีการจัดเก็บข้อมูลจราจรทางคอมพิวเตอร์ (Logging) ของระบบ AI ที่ถูกต้อง ครบถ้วน เพื่อใช้ติดตามตรวจสอบร่องรอยการเข้าใช้งานระบบ

๒. การรายงานเหตุการณ์ เมื่อพบเหตุการณ์ผิดปกติ (IT Incident) ต้องมีการบันทึกเหตุการณ์โดยใช้แบบฟอร์มที่กำหนด. การแก้ไขปัญหาความผิดปกติของระบบ AI ควรมีกระบวนการหา สาเหตุที่แท้จริง (Root Cause) และกำหนดแผนการแก้ไข (Remediation Plan) เพื่อป้องกันไม่ให้เกิดเหตุการณ์ซ้ำ

๓. ความต่อเนื่องทางธุรกิจ กปภ. ต้องมีแผนฉุกเฉินด้านเทคโนโลยีสารสนเทศ (IT contingency plan) เพื่อรองรับเหตุการณ์ที่ระบบเกิดหยุดชะงัก และต้องมีการซักซ้อมแผนรับมือภัยคุกคามทางไซเบอร์ (Cyber incident response plan)

๔.๒ การตรวจสอบและการรับรอง (Audit and Certification)

๑. การตรวจสอบ ต้องมีการตรวจสอบการดำเนินงานตามนโยบาย AI Governance อย่างน้อยปีละ ๑ ครั้ง วัตถุประสงค์หลักของการตรวจประเมินคือการยืนยันความสอดคล้องกับหลักเกณฑ์ที่กำหนดไว้ เช่น ข้อกำหนดของมาตรฐาน นโยบายองค์กร หรือข้อกำหนด

๒. รายงานสถานะและความเสี่ยงของโครงการ AI ต่อผู้บริหารอย่างน้อยปีละ ๑ ครั้ง

๓. การรับรองตามมาตรฐานสากล กปภ. ควรพิจารณาการได้รับการรับรองตามมาตรฐาน ISO/IEC ๒๗๐๐๑ และมาตรฐานด้านการบริหารจัดการ AI เช่น ISO/IEC ๔๒๐๐๑:๒๐๒๓

๔.๓ การฝึกอบรมและการสร้างวัฒนธรรม AI

๑. การฝึกอบรม กปภ. ต้องให้มีการพัฒนาองค์ความรู้บุคลากรที่เกี่ยวข้องกับการใช้งาน การกำกับดูแล และการประยุกต์ใช้ AI ให้เหมาะสมกับหน้าที่และความรับผิดชอบ

๒. การเสริมสร้างวัฒนธรรม ให้อาจารย์มีการสื่อสารเสริมสร้างความตระหนักรู้ด้านการใช้ AI ให้กับพนักงานทุกระดับอย่างต่อเนื่อง

บรรณานุกรม (References)

หลักการสากล มาตรฐาน และกรอบการกำกับดูแล

ENISA. (๒๐๒๓). Cybersecurity of AI and standardization. European Union Agency for Cybersecurity.

International Organization for Standardization. (๒๐๒๒). ISO/IEC ๒๗๐๐๑:๒๐๒๒ Information security, cybersecurity and privacy protection — Information security management systems — Requirements. ISO.

International Organization for Standardization. (๒๐๒๓). ISO/IEC ๔๒๐๐๑:๒๐๒๓ Artificial intelligence - Management system. [อ้างอิงจากแหล่งที่มาอื่นที่กล่าวถึง ISO/IEC ๔๒๐๐๑:๒๐๒๓]

International Organization for Standardization. (๒๐๒๓). ISO/IEC ๒๓๘๙๔:๒๐๒๓ Artificial intelligence - Guidance on risk management. [อ้างอิงจากแหล่งที่มาอื่นที่กล่าวถึง ISO/IEC ๒๓๘๙๔:๒๐๒๓]

OECD. (๒๐๒๒). OECD Framework for the Classification of AI systems. Retrieved from https://alcassapappert.adv.br/๒๐๒๑/wp-content/uploads/๒๐๒๒/๐๒/OECD_Framework_๒๐๒๒_๑๖๔๕๕๓๒๔๖๘.pdf

U.S. Department of Homeland Security. (๒๐๒๔). Mitigating artificial intelligence (AI) risk: Safety and security guidelines for critical infrastructure owners and operators. DHS.

UNESCO. (๒๐๒๑). Recommendation on the Ethics of Artificial Intelligence. [อ้างอิงจากแหล่งที่มาอื่นที่กล่าวถึง UNESCO Recommendation]

กฎหมายและข้อบังคับของไทย

พระราชบัญญัติการรักษาความมั่นคงปลอดภัยไซเบอร์ พ.ศ. ๒๕๖๒. (๒๕๖๒, ๒๗ พฤษภาคม). ราชกิจจานุเบกษา, ๑๓๖(๖๙ ก).

พระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. ๒๕๖๒. (๒๕๖๒, ๒๗ พฤษภาคม). ราชกิจจานุเบกษา, ๑๓๖(๖๙ ก). (หน้า ๑๕)

แนวปฏิบัติและเอกสารทางวิชาการของไทย

คณะนิติศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, ศูนย์วิจัยกฎหมายและการพัฒนา. (๒๕๖๕). Thailand Artificial Intelligence Guidelines ๑.๐: แนวปฏิบัติเกี่ยวกับมาตรฐานการใช้ปัญญาประดิษฐ์.

สำนักงานคณะกรรมการการรักษาความมั่นคงปลอดภัยไซเบอร์แห่งชาติ (สกมช.). (ม.ป.ป.). แนวปฏิบัติการใช้ปัญญาประดิษฐ์อย่างมั่นคงปลอดภัย (AI Security Guidelines).

สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์ (สพธอ.). (ม.ป.ป.). แนวทางการประยุกต์ใช้ Generative AI อย่างมีธรรมาภิบาลสำหรับองค์กร.

สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์ (สพธอ.) และ กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม. (๒๕๖๔). เอกสารแนวปฏิบัติจริยธรรมปัญญาประดิษฐ์ (Thailand AI Ethics Guideline).

สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ (สวทช.). (ม.ป.ป.). แนวปฏิบัติจริยธรรมด้านปัญญาประดิษฐ์ (AI Ethics Guidelines).

สุภาภรณ์ เกียรติสิน, มนัสศิริ จันสุทธินางกูร, ปรัชญ์ สง่างาม, & ยุทธพงศ์ อุณหทวีทรัพย์. (๒๐๒๑). เอกสารแนวปฏิบัติจริยธรรมปัญญาประดิษฐ์ (Thailand AI Ethics Guideline). สำนักงานคณะกรรมการดิจิทัลเพื่อเศรษฐกิจและสังคมแห่งชาติ กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม.

เอกสารภายในของการประปาส่วนภูมิภาค (กปภ.)

การประปาส่วนภูมิภาค. (๒๕๖๘). นโยบาย AI Governance การประปาส่วนภูมิภาค ปี ๒๕๖๘.

การประปาส่วนภูมิภาค. (๒๕๖๘). แผนปฏิบัติการดิจิทัลของ กปภ. ปีงบประมาณ ๒๕๖๙-๒๕๗๓.

การประปาส่วนภูมิภาค. (๒๕๖๘). นโยบายความมั่นคงปลอดภัยด้านสารสนเทศของ กปภ. (พ.ศ. ๒๕๖๘).

การประปาส่วนภูมิภาค. (๒๕๖๘). นโยบายธรรมาภิบาลข้อมูลและแนวปฏิบัติของ กปภ. ทบพนปี พ.ศ. ๒๕๖๘.

การประปาส่วนภูมิภาค. (๒๕๖๘). สถาปัตยกรรมองค์กรของ กปภ. ปี ๒๕๖๙-๒๕๗๓ ทบพนปี ๒๕๖๘.

ภาคผนวก

ผนวก ก

นิยามศัพท์ที่เกี่ยวข้อง

คำศัพท์	ความหมาย
ปัญญาประดิษฐ์ (Artificial Intelligence: AI)	ปัญญาประดิษฐ์เป็นเทคโนโลยีที่ถูกพัฒนาขึ้นเพื่อให้คอมพิวเตอร์มีคุณสมบัติหรือพฤติกรรมใกล้เคียงมนุษย์ เช่น การเรียนรู้ การรับรู้ และตอบสนองต่อสภาพแวดล้อม การให้เหตุผล และการแก้ไขปัญหา เป็นต้น ตามวัตถุประสงค์ที่มนุษย์กำหนด
ระบบปัญญาประดิษฐ์ (AI System)	ระบบที่สามารถทำการคาดการณ์ ให้คำแนะนำ หรือตัดสินใจที่มีอิทธิพลต่อสภาพแวดล้อมทางกายภาพหรือเสมือน
ปัญญาประดิษฐ์แบบรู้สร้าง (Generative AI)	เทคโนโลยี AI ประเภทหนึ่งที่มีความสามารถในการสร้างเนื้อหาใหม่ ในหลากหลายรูปแบบ เช่น ข้อความ ภาพ หรือ ซอร์สโค้ด ตามคำสั่ง (Prompt) ที่มนุษย์เป็นผู้กำหนด เพื่อสร้างเนื้อหาใหม่ที่คล้ายคลึงกับข้อมูลที่ได้รับฝึกฝน
Large Language Model (LLM)	โมเดลภาษาขนาดใหญ่ที่รับข้อความหรือคำสั่ง (Input) ในรูปแบบภาษา และนำไปสร้างผลลัพธ์ (Output) ที่มีความสามารถในด้านภาษาที่หลากหลาย เช่น การสร้างข้อความใหม่ การแปลภาษา การสรุปความ การวิเคราะห์ข้อความ เป็นต้น
Prompt Engineering	การสร้างและปรับแต่งข้อความหรือคำสั่งเพื่อให้ Generative AI สร้างผลลัพธ์ (Output) ที่ดีที่สุดและตรงตามความต้องการ
Artificial Neural Network (ANN)	โครงข่ายของเซลล์ประสาทเทียม (Artificial Neuron) ที่คล้ายกับการเชื่อมต่อเซลล์ประสาท (Neuron) ในสมองมนุษย์ โดยในแต่ละเซลล์ประสาทเทียมนั้น มีหน้าที่ในการรับข้อมูลและนำไปประมวลผลเพื่อสร้างเป็นผลลัพธ์ จากนั้นจึงส่งต่อผลลัพธ์ไปยังเซลล์ประสาทเทียมในชั้น (Layer) ถัดไปเพื่อประมวลผลต่อ

Machine Learning (ML):	เทคโนโลยี AI ประเภทหนึ่งที่ทำางานหรือสร้างผลลัพธ์บนพื้นฐานของข้อมูลที่ได้รับจากการฝึกฝนหรือจากสภาพแวดล้อม
Deep Learning (DL):	Machine Learning ประเภทหนึ่งที่ประมวลผลผ่านโครงข่ายประสาทเทียม (Artificial Neural Network: ANN) จำนวนหลายชั้น (Layer) ที่ถูกสร้างขึ้นจากข้อมูลที่ได้รับการฝึกฝนเพื่อให้สามารถทำงานหรือสร้างผลลัพธ์ที่มีประสิทธิภาพดียิ่งขึ้น
Artificial Neural Network (ANN):	โครงข่ายของเซลล์ประสาทเทียม (Artificial Neuron) ที่คล้ายกับการเชื่อมต่อเซลล์ประสาท (Neuron) ในสมองมนุษย์ โดยในแต่ละเซลล์ประสาทเทียมนั้น มีหน้าที่ในการรับข้อมูลและนำไปประมวลผลเพื่อสร้างเป็นผลลัพธ์ จากนั้นจึงส่งต่อผลลัพธ์ไปยังเซลล์ประสาทเทียมในชั้น (Layer) ถัดไปเพื่อประมวลผลต่อ
AI Foundation Model	โมเดล AI ประเภท Generative AI ที่ได้รับการฝึกฝนด้วยข้อมูลขนาดใหญ่ โดยมีวัตถุประสงค์เพื่อให้สามารถสร้างเนื้อหาใหม่ที่คล้ายคลึงกับข้อมูลที่ได้รับการฝึกฝน
ธรรมาภิบาลปัญญาประดิษฐ์ (AI Governance)	ชุดของหลักเกณฑ์ นโยบาย และแนวทางปฏิบัติที่กำกับดูแลผลกระทบทางจริยธรรม กฎหมาย และสังคมของการใช้ AI
ความรับผิดชอบ AI (AI Accountability)	บทบาทของการรับผิดชอบต่อการทำงานและการตัดสินใจของ AI รวมถึงผลกระทบที่ระบบ AI มีต่อบุคคลและสังคม
ความสามารถในการสืบย้อน AI (AI Traceability)	ความสามารถในการตรวจสอบย้อนกลับไปตั้งแต่แหล่งที่มาของชุดข้อมูล กระบวนการทำงาน และการตัดสินใจของ AI เพื่อใช้ในการเฝ้าระวังและวินิจฉัยปัญหา
การมีส่วนร่วมการกำกับดูแล AI โดยมนุษย์ (AI Human in the loop)	การมีส่วนร่วมของมนุษย์ในการควบคุมการทำงานหรือการตัดสินใจของระบบ AI โดยเฉพาะในสถานการณ์ที่มีความเสี่ยงสูง โดยมี AI ทำหน้าที่ในการให้คำแนะนำ หรือข้อมูล แต่ไม่สามารถทำงานหรือตัดสินใจใน การดำเนินการใด ๆ ได้ โดยปราศจากมนุษย์

การมีส่วนร่วมการกำกับดูแล AI โดยไม่มีมนุษย์ Human-over-the-loop	เป็นระดับการมีส่วนร่วมที่ AI สามารถทำงานหรือตัดสินใจได้โดยอัตโนมัติ (ไม่ต้องรอการตัดสินใจจากมนุษย์) แต่ยังคงจำเป็นต้องมีมนุษย์ในการกำกับดูแล อีกทั้งมนุษย์ยังสามารถเข้าควบคุมหรือระงับ การทำงานได้ หากพบความผิดพลาดหรือเกิดเหตุการณ์ที่อาจก่อให้เกิดผลกระทบเชิงลบ
AI ทำงานโดยไม่มีมนุษย์ Human-out-of-the-loop	เป็นระดับการมีส่วนร่วมที่ AI ควบคุมการทำงานหรือตัดสินใจทั้งหมดได้ โดยมนุษย์ไม่สามารถเข้าควบคุม แทรกแซง หรือระงับการทำงานได้
AI ความเสี่ยงสูง (AI High Risk)	ระบบ AI ที่มีศักยภาพในการส่งผลกระทบเชิงลบอย่างมีนัยสำคัญต่อสุขภาพ ความปลอดภัย หรือสิทธิขั้นพื้นฐานของบุคคล และต้องอยู่ภายใต้การกำกับดูแลที่เข้มงวด
คุณภาพข้อมูล (Data Quality)	การกำหนดเกณฑ์คุณภาพของชุดข้อมูลเพื่อให้มีความถูกต้องครบถ้วน เป็นปัจจุบัน และตรงตามความต้องการของผู้ใช้
ข้อมูลส่วนบุคคล (Personal Data)	ข้อมูลเกี่ยวกับบุคคลซึ่งทำให้สามารถระบุตัวบุคคลนั้นได้ไม่ว่าทางตรงหรือทางอ้อม

ผนวก ข

แบบฟอร์มรายงานการใช้งานปัญญาประดิษฐ์ (AI Usage Reporting Form)

แบบฟอร์มนี้ใช้สำหรับบันทึกและรายงานรายละเอียดของระบบ AI ที่ใช้งาน/พัฒนา เพื่อให้สามารถติดตามและสร้างความรับผิดชอบ

หัวข้อ	รายละเอียดที่ต้องรายงาน
ข้อมูลทั่วไปของระบบ	ชื่อระบบ/โมเดล เจ้าของผลิตภัณฑ์ (Product Owner) หน่วยงานผู้ใช้งานหลัก วัตถุประสงค์การใช้งานที่ชัดเจน
การจำแนกความเสี่ยง	ระดับความเสี่ยง (สูง/จำกัด/น้อยที่สุด) ผลกระทบหลัก (เช่น ด้านการเงิน/ปฏิบัติการ/ชื่อเสียง) มาตรการบรรเทาความเสี่ยงเบื้องต้น
ข้อมูลที่ใช้และธรรมาภิบาล	แหล่งที่มาของชุดข้อมูล (Data Provenance) มีข้อมูลส่วนบุคคล/ข้อมูลอ่อนไหวหรือไม่ การใช้เทคนิคปกป้องข้อมูล (เช่น Encryption, Anonymization) การประเมินคุณภาพข้อมูล
Human Oversight	ระดับการกำกับดูแลโดยมนุษย์ (เช่น Human-over-the-Loop) จุดที่มนุษย์มีอำนาจในการแทรกแซง/ตัดสินใจ

ผนวก ค

รายการตรวจสอบความมั่นคงปลอดภัยสำหรับระบบ AI (AI Security Checklist)

รายการตรวจสอบนี้จัดทำขึ้นตามแนวทางของ สกมช. และหลักการธรรมาภิบาลข้อมูลของ กปภ..

รหัส	รายการตรวจสอบ	สถานะ
☐ CPT-๐๑	ได้กำหนดวัตถุประสงค์ ขอบเขต และข้อจำกัดการใช้งานของระบบ AI อย่างชัดเจน	
☐ CPT-๐๓	ได้ดำเนินการประเมินความเสี่ยงเบื้องต้นสำหรับระบบปัญญาประดิษฐ์ ตามกรอบการทำงาน เช่น ISO ๒๓๘๔๔:๒๐๒๓	
☐ CPT-๐๔	ได้ดำเนินการประเมินผลกระทบด้านการคุ้มครองข้อมูลส่วนบุคคล (DPIA) (กรณีที่เกี่ยวข้องกับข้อมูลส่วนบุคคล)	
☐ DCM-๐๑	ต้องมีมนุษย์ตรวจสอบ/อนุมัติผลลัพธ์ของ AI เสมอ ก่อนนำไปใช้ในการตัดสินใจสำคัญ	
☐ DSP-๐๓	มีการจัดเก็บรักษาข้อมูลบันทึกเหตุการณ์/กิจกรรม Logs อัตโนมัติไว้อย่างเพียงพอตามที่กฎหมายกำหนด	
☐ SEC-๐๑	ได้ยืนยันว่าไม่มีการป้อนข้อมูลที่เป็นความลับ หรือข้อมูลส่วนบุคคลเข้าระบบ AI สาธารณะที่ไม่ได้รับอนุมัติ	

ผนวก ง

รูปแบบรายงานเหตุการณ์ฉุกเฉินเฉพาะ AI (AI-Specific Incident Response Template)

แม่แบบนี้ใช้สำหรับการบันทึกและตอบสนองต่อเหตุการณ์ผิดปกติด้านความมั่นคงปลอดภัยหรือจริยธรรมของระบบ AI

ลำดับ	รายการ	รายละเอียด
๑.	ข้อมูลเหตุการณ์	วันที่/เวลาที่เกิด ระบบ AI ที่ได้รับผลกระทบ ผู้รับผิดชอบ
๒.	ประเภทภัยคุกคามเฉพาะ AI	เช่น Data Poisoning, Model Evasion, Prompt Injection, ผลลัพธ์มีอคติ การรั่วไหลของโมเดล
๓.	ผลกระทบ	ผลกระทบต่อลูกค้า การดำเนินงาน ข้อมูลส่วนบุคคล (ต้องพิจารณาความเสียหายอย่างรุนแรง)
๔.	มาตรการตอบสนองที่ดำเนินการ	การระงับการทำงานโดยมนุษย์ (Human Intervention), การจำกัดผลลัพธ์
๕.	สาเหตุที่แท้จริง (Root Cause)	ระบุข้อบกพร่องทางเทคนิค ข้อบกพร่องของกระบวนการ หรือปัจจัยมนุษย์
๖.	บทเรียนและแผนการแก้ไข (Remediation Plan)	ระบุมาตรการที่จะดำเนินการเพื่อป้องกันไม่ให้เกิดเหตุการณ์ซ้ำ

1. แบบฟอร์มรายงานเหตุการณ์ความเสี่ยงและความผิดพลาดของระบบปัญญาประดิษฐ์ (AI Incident Report)

ส่วนที่	รายการ	ข้อมูลพื้นฐาน
1	ข้อมูลพื้นฐานของเหตุการณ์ (Incident Details)	
1.1	วันที่/เวลาที่เกิดเหตุ/ตรวจพบ	วันที่: _____ เวลา: _____
1.2	ระบบ AI ที่ได้รับผลกระทบ	_____
1.3	ผู้รับผิดชอบ/ผู้รายงาน	ชื่อ: _____ ตำแหน่ง/ฝ่าย: _____
2	ประเภทภัยคุกคามเฉพาะ AI (AI-Specific Threat Classification)	
2.1	ระดับความเสี่ยงเริ่มต้น	() สูง () ปานกลาง () ต่ำ
2.2	ประเภทภัยคุกคาม (เลือกได้มากกว่า 1)	() Data Poisoning () Model Evasion () Prompt Injection () ผลลัพธ์มีอคติ (Bias) () การรั่วไหลของโมเดล () อื่น ๆ: _____
2.3	รายละเอียดเหตุการณ์โดยสรุป	_____
3	ผลกระทบและความเสียหาย (Impact Assessment)	
3.1	ผลกระทบต่อลูกค้า/การดำเนินงาน	() บริการหยุดชะงัก () การตัดสินใจที่ไม่เป็นธรรม () ความเสียหายต่อชื่อเสียง () การสูญเสียทางการเงินโดยประมาณ: _____ บาท
3.2	ผลกระทบต่อข้อมูลส่วนบุคคล	() ข้อมูลรั่วไหล/ถูกนำไปใช้ผิดวัตถุประสงค์ () ไม่พบผลกระทบที่ต้องแจ้งตาม PDPA
3.3	การพิจารณาความเสียหายรุนแรง	() ใช่ (ต้องแจ้งผู้บริหารระดับสูงทันที) () ไม่ใช่
4	มาตรการตอบสนองที่ดำเนินการ (Response Measures)	
4.1	มาตรการตอบสนองหลัก	() ระบุการทำงานโดยมนุษย์ (Human Intervention) () จำกัดผลลัพธ์/ถอนผลลัพธ์ที่ผิดพลาด () กักกันและวิเคราะห์หลักฐาน (Log/Data)
4.2	รายละเอียดมาตรการที่ทำ	_____
4.3	สถานะปัจจุบัน/แก้ไขเสร็จสิ้น	() แก้ไขแล้ว วันที่/เวลา: _____ () อยู่ระหว่างดำเนินการแก้ไข
5	สาเหตุที่แท้จริง (Root Cause Analysis)	
5.1	สาเหตุหลัก	() ข้อบกพร่องทางเทคนิค () ข้อบกพร่องของกระบวนการ/กำกับดูแล () ปัจจัยมนุษย์
5.2	คำอธิบายสาเหตุที่แท้จริง	_____
6	บทเรียนและแผนการแก้ไข (Remediation Plan)	
6.1	มาตรการป้องกันการเกิดซ้ำ	_____
6.2	ผู้รับผิดชอบแผนแก้ไข	ชื่อ: _____ ฝ่าย: _____
6.3	กำหนดเวลาแล้วเสร็จ	_____

2. แบบทดสอบระบบ AI ก่อนใช้งานจริง และแบบการตรวจสอบผลลัพธ์อย่างสม่ำเสมอ (AI Testing and Monitoring Form)

แบบฟอร์มนี้ใช้เพื่อยืนยันว่าระบบทำงานตามที่คาดหวัง มีความทนทาน และประสิทธิภาพคงที่

ขั้นตอน/หัวข้อ	เกณฑ์การตรวจสอบ/ผลการวัด (ตัวอย่าง)	ผลการตรวจสอบ (ผ่าน/ไม่ผ่าน)
ก่อนใช้งานจริง (Pre-Deployment Testing)		
ความแม่นยำ/ประสิทธิภาพ (Accuracy)	โมเดลมีความแม่นยำ (Accuracy Rate) ตามที่กำหนดไว้ (เช่น >95%)	[]
ความทนทาน/ความมั่นคง (Robustness)	ผ่านการทดสอบ Adversarial Testing/Prompt Injection (ไม่มีผลลัพธ์ที่ไม่พึงประสงค์)	[]
การจัดการอคติ (Bias/Fairness Check)	ผลลัพธ์ไม่แสดงอคติ/ความลำเอียงต่อกลุ่มประชากรใดกลุ่มหนึ่งอย่างมีนัยสำคัญ	[]
การตรวจสอบสม่ำเสมอ (Continuous Monitoring)		
การเบี่ยงเบนของโมเดล (Model Drift)	ค่าตัวชี้วัดประสิทธิภาพของโมเดล (เช่น F1-Score, Override Rate) อยู่ในเกณฑ์ที่กำหนด	[]
คุณภาพข้อมูลขาเข้า (Input Data Quality)	ข้อมูลขาเข้ามีความถูกต้อง/เป็นปัจจุบัน (Timeliness) และไม่พบการปนเปื้อนของข้อมูล (Data Poisoning)	[]
การเก็บบันทึก (Logging)	บันทึกกิจกรรมการทำงาน/การตัดสินใจของ AI (Audit Trail/Logs) ถูกจัดเก็บอย่างครบถ้วนและปลอดภัย	[]

3. แบบประเมินความเสี่ยง AI (AI Risk Assessment Form)

แบบฟอร์มนี้ใช้เพื่อวิเคราะห์โอกาสและผลกระทบของความเสี่ยงที่เกี่ยวข้องกับ AI (อ้างอิงแนวทาง ISO/IEC 23894:2023)

ความเสี่ยง	คำอธิบายความเสี่ยง	โอกาส (Likelihood: 1-5)	ผลกระทบ (Impact: 1-5)	ระดับ ความเสี่ยง (Risk Score)	แนวทางการ จัดการ/บรรเทา (Mitigation Plan)
Data Poisoning	ผู้ไม่หวังดีแทรกข้อมูลที่เป็นพิษในชุดข้อมูลฝึกสอน ทำให้โมเดลตัดสินใจผิดพลาด				
Model Bias	โมเดลแสดงผลลัพธ์ที่มีอคติ/เลือกปฏิบัติต่อกลุ่มใดกลุ่มหนึ่ง				
Prompt Injection	ผู้โจมตีสั่งงานนอกขอบเขตการทำงานของ AI ผ่านคำสั่งที่เป็นอันตราย				
Data Privacy Breach	ระบบ AI เปิดเผยข้อมูลส่วนบุคคลที่ละเอียดอ่อนของลูกค้า				

หมายเหตุ: ระดับความเสี่ยง = โอกาส x ผลกระทบ (เช่น 1-4 = ต่ำ, 5-12 = กลาง/สูง, 13-25 = สูงมาก/วิกฤต)

4. แบบประเมินความมั่นคงปลอดภัย AI (AI Security Assessment Form)

แบบฟอร์มนี้ใช้เพื่อตรวจสอบมาตรการควบคุมด้านความมั่นคงปลอดภัยไซเบอร์เฉพาะสำหรับระบบ AI (อ้างอิงแนวทาง Secure by Design)

มิติ/รายการตรวจสอบ	รายละเอียดการตรวจสอบ	สถานะ (ผ่าน/ไม่ผ่าน)
การออกแบบ (Secure Design)		
การสร้างแบบจำลองภัยคุกคาม	ได้จัดทำ Threat Model ที่ระบุภัยคุกคามที่จำเพาะเจาะจงกับ AI (เช่น Data Poisoning, Model Evasion)	[]
การกำหนดค่าเริ่มต้นปลอดภัย	มีการกำหนดค่าเริ่มต้น (Secure Defaults) สำหรับ API ทั้งหมด เช่น บังคับใช้การพิสูจน์ตัวตน (Authentication) และ จำกัดอัตราการร้องขอ (Rate Limiting)	[]
การพัฒนา (Secure Development)		
ความมั่นคงปลอดภัยห่วงโซ่อุปทาน	ไลบรารี/โมเดลของบุคคลที่สามที่นำมาใช้ (Open Source) ได้ผ่านการสแกนหาช่องโหว่แล้ว	[]
การคุ้มครองทรัพย์สิน	ไฟล์โมเดลและข้อมูลฝึกสอนถูกเข้ารหัสลับขณะจัดเก็บ (Encryption at Rest)	[]
การปฏิบัติการ (Secure Operation)		
การจำกัดสิทธิ์ใน Production	สิทธิ์การเข้าถึงในสภาพแวดล้อมใช้งานจริง (Production) ถูกกำหนดตามหลักการสิทธิ์น้อยที่สุด (Least Privilege)	[]
การเฝ้าระวังภัยคุกคาม	มีระบบเฝ้าระวังเพื่อตรวจจับความพยายามทำ Prompt Injection หรือ Jailbreaking	[]
