Krungthai
กรุงไทย

นโยบายการกำกับดูแลความเสี่ยงการใช้งานปัญญาประดิษฐ์

(Artificial Intelligence : AI Risk Governance Policy)

เวอร์ชั่น 1.13

17 กุมภาพันธ์ 2569

Krungthai
กรุงไทย

สารบัญ

Krungthai
กรุงไทย

นโยบายการกำกับดูแลความเสี่ยงการใช้งานปัญญาประดิษฐ์

ส่วนที่ 1. บทนำ

1. เหตุผลในการออกนโยบาย

ปัจจุบัน ผู้ให้บริการทางการเงินได้นำระบบปัญญาประดิษฐ์ (Artificial Intelligence: AI) มาใช้งานเพิ่มมากขึ้น เพื่อช่วยเพิ่มประสิทธิภาพในการดำเนินธุรกิจ ซึ่งรวมถึงการตัดสินใจทางธุรกิจ การสนับสนุนการปฏิบัติงานหรือระบบงานสำคัญ และการให้บริการแก่ลูกค้า อย่างไรก็ตาม ด้วยเหตุที่ระบบ AI มีกลไกที่สามารถเรียนรู้และต่อยอดจากข้อมูลเพื่อสร้างผลลัพธ์ที่ผู้ให้บริการทางการเงินสามารถนำไปใช้ประกอบการตัดสินใจในเรื่องต่าง ๆ หรือให้ระบบ AI ทำงานแทนมนุษย์ การนำระบบ AI มาใช้งานจึงอาจเป็นการขยายความเสี่ยงในการดำเนินธุรกิจ หรือทำให้ความเสี่ยงมีรูปแบบเปลี่ยนแปลงไป โดยเฉพาะอย่างยิ่งหากระบบ AI ไม่สามารถให้ผลลัพธ์ที่ถูกต้อง น่าเชื่อถือ โปร่งใส และอธิบายได้

บมจ. ธนาคารกรุงไทย (“ธนาคาร”) สนับสนุนการนำระบบ AI มาใช้งาน เพื่อช่วยเพิ่มประสิทธิภาพในการดำเนินธุรกิจ รวมถึงเพื่อพัฒนานวัตกรรมที่ทำให้การให้บริการสามารถตอบสนองความต้องการของลูกค้าได้ดียิ่งขึ้น โดยในขณะเดียวกัน ต้องมีการควบคุมและบริหารจัดการความเสี่ยงอย่างรัดกุม เพื่อไม่ให้การใช้งานระบบ AI ส่งผลกระทบต่อการดำเนินธุรกิจ รวมทั้งเพื่อให้ลูกค้าได้รับการดูแลและคุ้มครองในการใช้บริการอย่างเหมาะสม

ธนาคารจึงออกนโยบายการกำกับดูแลความเสี่ยงการใช้งานปัญญาประดิษฐ์ เพื่อเป็นแนวทางในการบริหารจัดการความเสี่ยงของการใช้งานระบบ AI ให้เป็นไปอย่างรัดกุมและเหมาะสมตามลักษณะการใช้งาน โดยสอดคล้องกับหลักการในการประยุกต์ใช้ AI อย่างมีความรับผิดชอบ (Responsible AI) นโยบายการกำกับดูแลความเสี่ยงการใช้งานปัญญาประดิษฐ์ นี้เป็นไปตามกฎหมายที่เกี่ยวข้อง โดยสอดคล้องกับข้อกำหนดของหน่วยงานกำกับฯ และ ได้อ้างอิงจากมาตรฐานสากลสำหรับการบริหารจัดการความเสี่ยงของการใช้งานระบบ AI ซึ่งได้แก่ ISO/IEC 23894:2023, ISO/IEC 42001:2023, NIST Artificial Intelligence Risk Management Framework เป็นต้น

2. วัตถุประสงค์ของนโยบาย

นโยบายฉบับนี้มีวัตถุประสงค์หลักในการกำหนดหลักการและแนวทางในการบริหารจัดการความเสี่ยงจากการใช้งานระบบ AI โดยสามารถนำไปประยุกต์ใช้ตามลักษณะการใช้งาน ไม่ว่าจะเป็นระบบ AI ที่พัฒนาขึ้นเองหรือระบบ AI ของบุคคลภายนอกที่นำมาใช้ เพื่อให้การใช้งานระบบ AI มีความเหมาะสมและสอดคล้องกับหลักการที่ดีที่ได้รับการยอมรับในระดับสากล รวมถึงการปฏิบัติตามกฎหมายที่เกี่ยวข้องอย่างเคร่งครัด เช่น กฎหมายคุ้มครองข้อมูลส่วนบุคคลและกฎหมายทรัพย์สินทางปัญญา

3. ขอบเขตการใช้นโยบาย

นโยบายการกำกับดูแลการใช้งาน AI ฉบับนี้ครอบคลุมถึง

3.1. พนักงานทั้งหมด รวมถึงพนักงานปฏิบัติงานไม่เต็มเวลา (Part time) พนักงานชั่วคราว (Temporary) และพนักงานฝึกหัด (Trainee)

3.2. บริษัทย่อย

Commented [A1]: BOT แนวนโยบายธนาคารแห่งประเทศไทย เรื่อง การบริหารจัดการความเสี่ยงของการใช้งานระบบปัญญาประดิษฐ์
1. เหตุผลในการออกแนวนโยบาย
ปัจจุบันผู้ให้บริการทางการเงินนำระบบปัญญาประดิษฐ์ (Artificial Intelligence: AI) มาใช้งานมากขึ้น เพื่อช่วยเพิ่มประสิทธิภาพในการดำเนินธุรกิจ ซึ่งรวมถึงการตัดสินใจทางธุรกิจ การสนับสนุนการปฏิบัติงานหรือระบบงานสำคัญ และการให้บริการแก่ลูกค้า อย่างไรก็ตาม เนื่องจากระบบ AI มีความสามารถเรียนรู้และพัฒนาต่อยอดจากข้อมูลเพื่อสร้างผลลัพธ์ที่ผู้ให้บริการทางการเงินสามารถ ...

Commented [A2]: BOT แนวนโยบายธนาคารแห่งประเทศไทย เรื่อง การบริหารจัดการความเสี่ยงของการใช้งานระบบปัญญาประดิษฐ์
1. เหตุผลในการออกแนวนโยบาย
ธนาคารแห่งประเทศไทย (ธปท.) สนับสนุนให้ผู้ให้บริการทางการเงินสามารถนำระบบ AI มาใช้งานเพื่อช่วยเพิ่มประสิทธิภาพในกา ...

Commented [A3]: BOT แนวนโยบายธนาคารแห่งประเทศไทย เรื่อง การบริหารจัดการความเสี่ยงของการใช้งานระบบปัญญาประดิษฐ์
1. เหตุผลในการออกแนวนโยบาย
ธปท. จึงออกแนวนโยบายเกี่ยวกับการบริหารจัดการความเสี่ยงจากการใช้งานระบบ AI เพื่อให้ผู้ให้บริการทางการเงินนำไปใช้อ้างอิง ...

Commented [A4]: BOT แนวนโยบายธนาคารแห่งประเทศไทย เรื่อง การบริหารจัดการความเสี่ยงของการใช้งานระบบปัญญาประดิษฐ์
2. วัตถุประสงค์ของแนวนโยบาย
เพื่อให้ผู้ให้บริการทางการเงินมีแนวทางที่สามารถนำไปประยุกต์ใช้สำหรับการบริหารจัดการความเสี่ยงของการใช้งานระบบ AI ได้อย่ ...

Commented [A5]: BOT แนวนโยบายธนาคารแห่งประเทศไทย เรื่อง การบริหารจัดการความเสี่ยงของการใช้งานระบบปัญญาประดิษฐ์
คำถาม – คำตอบแนบท้าย
5. หากผู้ให้บริการทางการเงินไม่ได้เป็นผู้พัฒนาระบบ AI ขึ้นเอง โดยนำระบบ AI แบบสำเร็จรูปมาใช้งานเท่านั้น เช่น การนำ ...

Commented [A6]: BOT แนวนโยบายธนาคารแห่งประเทศไทย เรื่อง การบริหารจัดการความเสี่ยงของการใช้งานระบบปัญญาประดิษฐ์
คำถาม – คำตอบแนบท้าย ...

3.3. สาขาต่างประเทศ

3.4. บุคคลภายนอก บุคคลหรือนิติบุคคลภายนอกซึ่งเป็นผู้ให้บริการด้านเทคโนโลยีสารสนเทศแทนธนาคารฯ หรือเป็นผู้ที่มีการเชื่อมต่อกับระบบเทคโนโลยีสารสนเทศของธนาคารฯ หรือเป็นผู้ที่สามารถเข้าถึงข้อมูลในระดับชั้นข้อมูลสารสนเทศที่ใช้เฉพาะภายในองค์กร (Internal Use Only) ขึ้นไป หรือข้อมูลลูกค้าของธนาคาร ในรูปแบบอิเล็กทรอนิกส์ได้ โดยกรณีของสาขาต่างประเทศให้รวมถึงสำนักงานใหญ่หรือสาขาอื่นในต่างประเทศที่เป็นนิติบุคคลเดียวกันด้วย ทั้งนี้ บุคคลภายนอกไม่ครอบคลุมถึงลูกค้าที่ใช้ผลิตภัณฑ์และบริการของธนาคารฯ

นโยบายฉบับนี้ครอบคลุมถึงการกำกับดูแล การใช้และพัฒนาระบบ AI ทั้งหมดตามที่กำหนดนิยามไว้ในหัวข้อ "นิยามและคำจำกัดความ" และรวมถึง แต่ไม่จำกัดเพียงการกำกับดูแล การใช้ผลลัพธ์ ระบบ และผลิตภัณฑ์ที่สร้างโดย AI ซึ่งได้แก่ เนื้อหาหรือข้อมูลในกระบวนการทางธุรกิจ ผลิตภัณฑ์ หรือบริการใด ๆ และหากมีข้อสงสัยเกี่ยวกับประเภทของเทคโนโลยีที่เข้าข่ายนิยามของ AI ให้ระมัดระวังและปฏิบัติตามคำแนะนำของนโยบายฉบับนี้

4. คำประกาศความเสี่ยงที่รับได้ (Risk Appetite Statement)

ธนาคารมีการกำหนดคำประกาศความเสี่ยงที่รับได้ (Risk Appetite Statement) ให้สอดคล้องกับทิศทางการดำเนินธุรกิจของธนาคาร

"กำกับดูแลการบริหารความเสี่ยงด้าน AI ให้เป็นไปตามหลักการประยุกต์ใช้ AI อย่างมีความรับผิดชอบ (Responsible AI) โดยบริหารจัดการให้อยู่ในระดับความเสี่ยงต่ำ"

5. หลักการบริหารจัดการความเสี่ยงของการใช้งานระบบ AI

ธนาคารบริหารจัดการความเสี่ยงจากการใช้งานระบบ AI ให้เหมาะสมกับลักษณะการใช้งาน โดยสอดคล้องกับหลักการ ดังนี้

5.1. มีความรับผิดชอบต่อการทำงานของระบบ AI รวมถึงการตัดสินใจที่อ้างอิงจากผลลัพธ์ของระบบ AI

5.2. มีการกำกับดูแลการใช้งานระบบ AI ให้เป็นไปตามหลักการใช้งานระบบ AI อย่างรับผิดชอบ (responsible AI) ที่ได้รับการยอมรับโดยทั่วไป เช่น หลักการใช้งานระบบ AI ที่มีความเป็นธรรม (fairness) มีจริยธรรม (ethics) มีความรับผิดชอบ (accountability) และมีความโปร่งใส (transparency) (หลักการ FEAT)

5.3. มีการบริหารจัดการความเสี่ยงของระบบ AI ครอบคลุมกระบวนการพัฒนาและการใช้งานระบบ AI (AI lifecycle) อย่างรัดกุมและเหมาะสม โดยระบบ AI มีความมั่นคงปลอดภัย ถูกต้อง และเชื่อถือได้ รวมถึงสามารถควบคุมความเสี่ยงและอธิบายที่มาของผลลัพธ์จากระบบ AI ได้

5.4. มีแนวทางการดูแลและคุ้มครองลูกค้าที่โปร่งใสและเป็นธรรม

6. หลักการประยุกต์ใช้ AI อย่างมีความรับผิดชอบ (Responsible AI)

หลักการ FEAT และแนวทางของสำนักงานคณะกรรมการดิจิทัลเพื่อเศรษฐกิจและสังคมแห่งชาติ (สดช.)

(1) **ความเป็นธรรม** (Fairness) การพัฒนา การบริหารจัดการ และการใช้งาน AI ต้องยึดหลักการดูแลและคุ้มครองลูกค้าอย่างเป็นธรรม เท่าเทียม หลากหลาย และครอบคลุม หลีกเลี่ยงการใช้ข้อมูลที่มีอคติ (Bias) หรือการ

Commented [A7]: BOT แนวนโยบายธนาคารแห่งประเทศไทย เรื่อง การบริหารจัดการความเสี่ยงของการใช้งานระบบปัญญาประดิษฐ์

4.4 แนวทางบริหารจัดการความเสี่ยงของการใช้งานระบบ AI

ส่วนที่ 1 ธรรมาภิบาลของการนำระบบ AI มาใช้งาน (governance)

(3) การบริหารจัดการความเสี่ยงจากการใช้งานระบบ AI โดยครอบคลุมการดำเนินการ ดังนี้

(3.1) ระบุความเสี่ยงจากการใช้งานระบบ AI และกำหนดเป้าหมายของ risk appetite ขององค์กรอย่างชัดเจน รวมทั้งจัดให้มีการประเมินและติดตามความเสี่ยงของการใช้งานระบบ AI อย่างต่อเนื่อง เพื่อให้มีการควบคุมความเสี่ยงเหมาะสมกับลักษณะการใช้งานภายใต้กรอบนโยบายการบริหารจัดการความเสี่ยงที่ยอมรับได้ขององค์กร

Commented [A8]: **BOT แนวนโยบายธนาคารแห่งประเทศไทย เรื่อง การบริหารจัดการความเสี่ยงของการใช้งานระบบปัญญาประดิษฐ์**

4.3 หลักการบริหารจัดการความเสี่ยงของการใช้งานระบบ AI

ผู้ให้บริการทางการเงินควรบริหารจัดการความเสี่ยงของการใช้งานระบบ AI ให้เหมาะสมกับลักษณะการใช้งาน โดยสอดคล้องกับหลักการ ดังนี้

(1) ผู้ให้บริการทางการเงินมีความรับผิดชอบต่อการทำงานของระบบ AI รวมถึงการตัดสินใจที่อ้างอิงจากผลลัพธ์ของระบบ AI

(2) ผู้ให้บริการทางการเงินมีการกำกับดูแลการใช้งานระบบ AI ให้เป็นไปตามหลักการใช้งานระบบ AI อย่างรับผิดชอบ (responsible AI) ที่ได้รับการยอมรับโดยทั่วไป เช่น หลักการใช้งานระบบ AI ที่มีความเป็นธรรม (fairness) มีจริยธรรม (ethics) มีความรับผิดชอ...

Commented [A9]: Monetary Authority of Singapore (MAS) หลักการ FEAT - Principles to Promote Fairness, Ethics, Accountability and Transparency (FEAT) in the Use of Artificial Intelligence and Data Analytics in Singapore's Financial Sector

4 Summary of Principles
Fairness
Justifiability
1. Individuals or groups of individuals are not systematically disadvantaged through AIDA-driven decisions unless these decisions can be justified.
2. Use of personal attributes as input factors for AIDA-driven decisions is justified.

Accuracy and Bias
3. Data and models used for AIDA-driven decisions are regularly reviewed and validated for accuracy and relevance, and to minimize unintentional bias.

ออกแบบโมเดลที่อาจนำไปสู่การเลือกปฏิบัติอย่างไม่เหมาะสม (Discrimination) รวมถึงสามารถพิสูจน์ถึงความเป็นธรรมสำหรับทุกฝ่ายที่เกี่ยวข้อง

(2) **มีจริยธรรม** (Ethics) การพัฒนา การบริหารจัดการ และการใช้งาน AI ต้องสอดคล้องกับจริยธรรม กฎหมาย และข้อกำหนดที่เกี่ยวข้อง รวมถึงจรรยาบรรณธุรกิจธนาคาร (Code of Conduct) โดยผลลัพธ์หรือการตัดสินใจที่เกิดจาก AI ต้องมีมาตรฐานทางจริยธรรมไม่น้อยกว่าการตัดสินใจโดยมนุษย์ (Human-driven decisions) โดยเคารพต่อความเป็นส่วนตัว เกียรติ สิทธิเสรีภาพ และสิทธิมนุษยชน

(3) **ความรับผิดชอบ** (Accountability) การพัฒนา การบริหารจัดการ และการใช้งาน AI ต้องมีความรับผิดชอบต่อการทำงานของระบบ AI และการตัดสินใจที่อ้างอิงจากผลลัพธ์ของระบบ AI มีการเฝ้าติดตามความผิดปกติ มีความสามารถในการสืบย้อนกลับ (Traceability) และสามารถวินิจฉัยปัญหาและความผิดพลาดได้ (Diagnosability) นอกจากนี้ต้องมีการกำหนดหน้าที่และความรับผิดชอบ (Role and Responsibility) รวมถึงความรับผิดชอบต่อผลของการกระทำ (Accountability) ต่อผลกระทบที่เกิดขึ้นจาก AI ตามหน้าที่ที่รับผิดชอบได้

(4) **ความโปร่งใส** (Transparency) การพัฒนา การบริหารจัดการ และการใช้งาน AI ต้องเปิดเผยการใช้ AI ต่อผู้ใช้งานหรือลูกค้าอย่างเพียงพอและเหมาะสมกับความเสี่ยงของกิจกรรมที่มีการใช้งาน สามารถอธิบายการทำงาน กระบวนการพัฒนา การตัดสินใจของระบบได้อย่างชัดเจน (Explainability) รวมถึงความเสี่ยงและผลกระทบที่อาจเกิดขึ้นจากผลลัพธ์หรือการตัดสินใจดังกล่าว เพื่อสามารถปฏิบัติต่อผู้ใช้งานหรือลูกค้าได้อย่างเป็นธรรม

(5) **ความน่าเชื่อถือ** (Reliability) การพัฒนา การบริหารจัดการ และการใช้งาน AI ต้องมีความน่าเชื่อถือและความมั่นใจในการใช้งานต่อสาธารณะ โดยต้องมีการพัฒนาด้วยความยึดมั่นในความถูกต้อง (Accuracy) ความน่าเชื่อถือ สามารถทนทานต่อเหตุการณ์ที่อาจเกิดความผิดพลาด (Robustness) และสามารถสร้างผลลัพธ์ได้เหมือนเดิม (Reproducibility) นอกจากนี้ต้องมีการควบคุมคุณภาพของข้อมูล (Quality of data) รวมถึงกำหนดกระบวนการและช่องทางรับความคิดเห็น (Feedback) จากผู้ใช้งาน ให้ข้อมูลเพิ่มเติม รับเรื่องร้องเรียนแจ้งปัญหาที่พบ และมีการตอบสนอง หรือดำเนินการแก้ไขปัญหาที่พบได้ทันท่วงที

(6) **ความมั่นคงปลอดภัยและความเป็นส่วนตัว** (Security and Privacy) การพัฒนา การบริหารจัดการ และการใช้งาน AI ต้องดำเนินมาตรการด้านความมั่นคงปลอดภัยทางเทคโนโลยีสารสนเทศและความเป็นส่วนตัว ภายใต้กรอบหลักการรักษาความลับของระบบและข้อมูล (Confidentiality) ความถูกต้องเชื่อถือได้ของระบบและข้อมูล (Integrity) และความพร้อมใช้งานของเทคโนโลยีสารสนเทศ (Availability) และปฏิบัติตามกฎหมายคุ้มครองข้อมูลส่วนบุคคล รวมถึงมีกลไกที่ให้มนุษย์สามารถแทรกแซงเพื่อควบคุมการทำงานของ AI ในกรณีที่เกิดเหตุการณ์หรือความเสี่ยงที่มีผลกระทบกับมนุษย์ได้

7. การทบทวนนโยบาย

นโยบายการกำกับดูแลความเสี่ยงการใช้งานปัญญาประดิษฐ์ จะต้องได้รับการทบทวนและปรับปรุงให้เป็นปัจจุบัน โดยฝ่ายบริหารความเสี่ยงด้านเทคโนโลยีสารสนเทศ อย่างน้อยทุก 1 ปี หรือเมื่อมีการเปลี่ยนสำคัญเกิดขึ้น ทั้งด้านการจัดองค์กร ระเบียบวิธีปฏิบัติ หรือระบบงาน รวมทั้งอนุมัติโดยคณะกรรมการธนาคาร เพื่อให้มั่นใจได้ว่านโยบายฯ ที่ปรับเปลี่ยนนั้น สามารถป้องกันภัยคุกคาม และจุดอ่อนของธนาคารฯ รวมทั้งความเสี่ยงซึ่งอาจจะเพิ่มขึ้นเมื่อ เกิดการเปลี่ยนแปลงในการดำเนินธุรกิจได้

Commented [A10]: ETDA AI Governance Guideline for Executive
หลักการจริยธรรมปัญญาประดิษฐ์ของไทยตามแนวทางของ สดช.
05 ความเท่าเทียม หลากหลาย ครอบคลุม และเป็นธรรม (Fairness)
ในการส่งเสริมการประยุกต์ใช้ AI อย่างเป็นธรรมและลดความมีอคติ (Bias) ระบบควรถูกออกแบบและพัฒนาโดยคำนึงความหลากหลาย (Diversity) ลดการเอนเอียง แบ่งแยก และเลือกปฏิบัติ (Discrimination) ต่อบุคคลหรือกลุ่มคนที่มีคุณลักษณะที่ต่างกัน (อาทิ อายุ เพศ ลักษณะทางกายภาพ เชื้อชาติ) โดยเฉพาะกลุ่มผู้ด้อยโอกาสในสังคม รวมถึงสามารถพิสูจน์ถึงความเป็นธรรมสำหรับทุกฝ่ายที่เกี่ยวข้อง

Commented [A11R10]: SEC **3. กรอบการกำกับดูแลการใช้งาน AI/ML**
3.3 หลักการที่ควรคำนึงถึง

Commented [A12R10]: OIC AI Governance Guideline

Commented [A13]: Monetary Authority of Singapore (MAS)

Commented [A14]: ETDA AI Governance Guideline for Executive

Commented [A15R14]: SEC **3. กรอบการกำกับดูแลการใช้งาน AI/ML**

Commented [A16]: Monetary Authority of Singapore (MAS)

Commented [A17]: BOT BOT แนวนโยบายธนาคารแห่งประเทศไทย เรื่อง การบริหารจัดการความเสี่ยงของการใช้งานระบ

Commented [A18R17]: SEC **3. กรอบการกำกับดูแลการใช้งาน AI/ML**

Commented [A19R17]: **MDES สรุปร่างหลักการของกฎหม**

Commented [A20R17]: ETDA AI Governance Guideline for Executive

Commented [A21]: Monetary Authority of Singapore (MAS)

Commented [A22]: OIC AI Governance Guideline

Commented [A23R22]: SEC **3. กรอบการกำกับดูแลการใช้งาน AI/ML**

Commented [A24R22]: ETDA AI Governance Guideline for Executive

Commented [A25]: ETDA AI Governance Guideline for Executive

Commented [A26]: ETDA AI Governance Guideline for Executive

กระบวนการนี้ อาจจะมีผลกระทบให้ระดับการควบคุมของนโยบายเพิ่มขึ้นหรือลดลง ทั้งนี้ขึ้นอยู่กับประเภทของการเปลี่ยนแปลง ซึ่งการเปลี่ยนแปลงใดๆ ในนโยบายหรือแนวทางการปฏิบัติการการกำกับดูแลการใช้งานปัญญาประดิษฐ์ต้องได้รับการอนุมัติและประกาศให้ทราบกันโดยทั่วไป

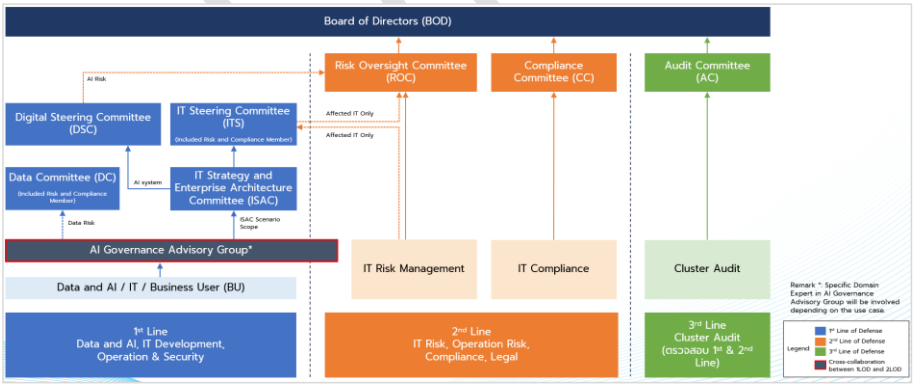
ส่วนที่ 2. ธรรมาภิบาลของการนำระบบ AI มาใช้งาน (Governance)

1. โครงสร้างการกำกับดูแล

ธนาคารยึดหลักการแบ่งแยกหน้าที่ความรับผิดชอบ 3 ระดับ (Three Lines of Defense) โดยมีการแบ่งแยกหน้าที่ความรับผิดชอบอย่างชัดเจนระหว่างการทำหน้าที่ พัฒนาและการใช้งานระบบ AI บริหารความเสี่ยงในการพัฒนาและการใช้งานระบบ AI และกำกับดูแลการปฏิบัติตามกฎเกณฑ์ที่เกี่ยวข้องกับการพัฒนาและการใช้งานระบบ AI รวมถึงการตรวจสอบด้าน AI เพื่อให้มีการถ่วงดุลอำนาจกันอย่างอิสระ และมี**คณะที่ปรึกษาการกำกับดูแล** AI (AI Governance Advisory Group) ทำหน้าที่ในการให้คำปรึกษาและแนะนำเกี่ยวกับการใช้เทคโนโลยี AI เพื่อบริหารจัดการความเสี่ยงที่เกี่ยวข้องกับ AI โดยเฉพาะ คณะที่ปรึกษาประกอบด้วยผู้แทนจากหน่วยงานหลัก ได้แก่

- ฝ่าย Data Protection
- ส่วนงานการกำกับดูแลข้อมูล (Data governance office)
- กลยุทธ์เทคโนโลยีสารสนเทศเพื่อขับเคลื่อนธุรกิจ (Enterprise Architecture)
- หน่วยงานบริหารความปลอดภัยข้อมูลสารสนเทศ (IT Security Department)
- ฝ่ายบริหารความเสี่ยงด้านเทคโนโลยีสารสนเทศ (IT Risk)
- ฝ่ายบริหารความเสี่ยงด้านปฏิบัติการ (Operational Risk)
- หน่วยงาน Digital Banking Regulations and Compliance Review (IT Compliance)
- ฝ่ายนิติการ (Legal)

รูปที่ 1 แผนภาพโครงสร้างการกำกับดูแล AI



2. การกำกับดูแลความเสี่ยงจากการใช้งานระบบ AI

2.1. ระบุความเสี่ยงจากการใช้งานระบบ AI และกำหนดเป้าหมายของ risk appetite ของธนาคารอย่างชัดเจน รวมทั้งจัดให้มีการประเมินและติดตามความเสี่ยงของการใช้งานระบบ AI อย่างต่อเนื่อง เพื่อให้สามารถควบคุมความเสี่ยงได้อย่างเหมาะสมกับลักษณะการใช้งาน ภายใต้กรอบนโยบายการบริหารจัดการความเสี่ยงที่ธนาคารยอมรับได้

Commented [A27]: BOT แนวนโยบายธนาคารแห่งประเทศไทย เรื่อง การบริหารจัดการความเสี่ยงของการใช้งานระบบปัญญาประดิษฐ์
4.4 แนวทางบริหารจัดการความเสี่ยงของการใช้งานระบบ AI
ส่วนที่ 1 ธรรมาภิบาลของการนำระบบ AI มาใช้งาน (governance)
(1.2) ดูแลให้มีการกำหนดบทบาทหน้าที่ในการกำกับดูแลและการบริหารความเสี่ยงในการพัฒนาและการใช้งานระบบ AI อย่างชัดเจน ครอบคลุมหน่วยงานที่มีหน้าที่ความรับผิดชอบทั้ง 3 ระดับตามหลักการ three lines of defence

Commented [A28R27]: NCSA AI Security Guideline
๔.๑ การกำหนดบทบาทหน้าที่และความรับผิดชอบ
รากฐานที่สำคัญที่สุดของการกำกับดูแลปัญญาประดิษฐ์ที่มีประสิทธิภาพ คือ การกำหนดบทบาท ความรับผิดชอบ และอำนาจหน้าที่ที่เกี่ยวข้องกับความมั่นคงปลอดภัยของระบบปัญญาประดิษฐ์อย่างชัดเจนและเป็นลายลักษณ์อักษร การขาดความชัดเจนในเรื่องนี้มักนำไปสู่ช่องว่างด้านความรับผิดชอบ ซึ่งเป็นบ่อเกิดของความเสี่ยงที่ร้ายแรง
องค์กรควรจัดตั้งโครงสร้างบุคลากรที่ครอบคลุมตั้งแต่ระดับยุทธศาสตร์ไปจนถึงระดับปฏิบัติการ โดยแต่ละบทบาทมีหน้าที่ที่จำเพาะเจาะจงต่อวงจรชีวิตของระบบปัญญาประดิษฐ์

Commented [A29]: BOT แนวนโยบายธนาคารแห่งประเทศไทย เรื่อง การบริหารจัดการความเสี่ยงของการใช้งานระบบปัญญาประดิษฐ์
4.4 แนวทางบริหารจัดการความเสี่ยงของการใช้งานระบบ AI
ส่วนที่ 1 ธรรมาภิบาลของการนำระบบ AI มาใช้งาน (governance)
(3) การบริหารจัดการความเสี่ยงจากการใช้งานระบบ AI โดยครอบคลุมการดำเนินการ ดังนี้
(3.1) ระบุความเสี่ยงจากการใช้งานระบบ AI และกำหนดเป้าหมายของ risk appetite ขององค์กรอย่างชัดเจน รวมทั้งจัดให้มีการประเมินและติดตามความเสี่ยงของการใช้งานระบบ AI อย่างต่อเนื่อง เพื่อให้มีการควบคุมความเสี่ยงเหมาะสมกับลักษณะการใช้งาน ภายใต้กรอบนโยบายการบริหารจัดการความเสี่ยงที่ยอมรับได้ขององค์กร

2.2. จำแนกประเภทความเสี่ยงสำหรับระบบ AI และระบุสิ่งที่ต้องดำเนินการ โดยอ้างอิงหัวข้อ 2.3 หมวดหมู่ความเสี่ยงของปัญญาประดิษฐ์ แนวปฏิบัติการใช้ปัญญาประดิษฐ์อย่างมั่นคงปลอดภัย (AI Security Guidelines) จากสำนักงานคณะกรรมการการรักษาความมั่นคงปลอดภัยไซเบอร์แห่งชาติ (สกมช.) ได้แก่

2.2.1. **ความเสี่ยงที่ยอมรับไม่ได้** (Unacceptable) ระบบ AI ที่มีคุณลักษณะการทำงานที่ก่อให้เกิดภัยคุกคามต่อความปลอดภัย สิทธิขั้นพื้นฐาน และคุณค่าที่เป็นที่ยอมรับในสังคมสากล และไม่เป็นไปตามหลักเกณฑ์ว่าด้วยการบริหารจัดการด้านการให้บริการแก่ลูกค้าอย่างเป็นธรรม (Market conduct) กฎหมายและกฎเกณฑ์ที่เกี่ยวข้อง โดยระบบ AI ประเภทนี้ ห้ามมีการริเริ่ม พัฒนา จัดหา หรือนำมาใช้งานภายในธนาคารโดยสิ้นเชิง

2.2.2. **ความเสี่ยงสูง** (High) ระบบ AI ที่อาจส่งผลกระทบเชิงลบอย่างมีนัยสำคัญต่อสุขภาพ ความปลอดภัย หรือสิทธิขั้นพื้นฐานของบุคคลหรือลูกค้า ระบบ AI ประเภทนี้ จำเป็นต้องอยู่ภายใต้การกำกับดูแลที่เข้มงวดและต้องปฏิบัติตามข้อกำหนดดังต่อไปนี้ ตลอดวงจรชีวิตของระบบ AI

1) การบริหารจัดการความเสี่ยง (risk management)

2) การกำกับดูแลข้อมูล (data governance)

3) การจัดทำเอกสารทางเทคนิค

4) การเก็บบันทึกการทำงาน (record-keeping)

5) การให้ข้อมูลและความโปร่งใส

6) การกำกับดูแลโดยมนุษย์ (human oversight)

7) การรับประกันความแม่นยำ (accuracy) ความทนทาน (robustness) และความมั่นคงปลอดภัยทางไซเบอร์

8) มีระบบบริหารคุณภาพ (quality management system) เพื่อให้แน่ใจว่าปฏิบัติตามข้อกำหนดตลอดวงจรชีวิตของระบบ AI

2.2.3. **ความเสี่ยงจำกัด** (Limited) ระบบ AI ที่ความเสี่ยงหลักเกิดจากการขาดความโปร่งใส ซึ่งอาจทำให้ผู้ใช้ไม่ตระหนักว่ากำลังมีปฏิสัมพันธ์กับระบบ AI เช่น ระบบสนทนาอัตโนมัติ (Chatbots) ระบบที่สร้างเนื้อหาประเภท Deepfake ระบบที่สร้างเนื้อหาอื่น ๆ จากปัญญาประดิษฐ์ (AI-generated Content) โดยระบบ AI ประเภทนี้ต้องเปิดเผยข้อมูลเกี่ยวกับการให้บริการด้วยระบบ AI ให้ผู้ใช้งานรับทราบอย่างชัดเจน และปฏิบัติตามเป็นไปตามหลักการประยุกต์ใช้ AI อย่างมีความรับผิดชอบ (Responsible AI)

2.2.4. **ความเสี่ยงน้อยที่สุด** (Minimal) ระบบ AI ประเภทอื่น ๆ ที่ไม่เข้าข่ายทั้งสามระดับข้างต้น และ มีความเสี่ยงต่ำ เช่น ระบบกรองอีเมลขยะ ระบบ AI ประเภทนี้ ธนาคารส่งเสริมให้การพัฒนาและการใช้งานเป็นไปตามหลักการประยุกต์ใช้ AI อย่างมีความรับผิดชอบ (Responsible AI)

2.3. พิจารณาความเสี่ยงและผลกระทบจากการใช้งานระบบ AI ที่มีต่อการดำเนินธุรกิจและการให้บริการแก่ลูกค้า โดยเฉพาะกรณีที่มีการนำระบบ AI มาใช้กับงานหลักที่เกี่ยวข้องกับการตัดสินใจเชิงกลยุทธ์ หรืองานที่ส่งผลกระทบโดยตรงต่อลูกค้า โดยในการนำระบบ AI มาใช้ในลักษณะดังกล่าว พิจารณาให้มนุษย์มีส่วนร่วมในการตัดสินใจ (human in the loop) หรือกำกับดูแล (human over the loop) ตามความเหมาะสมในแต่ละกรณี โดยพิจารณาระดับการมีส่วนร่วมหรือการกำกับดูแลตามความเสี่ยงและผลกระทบของการใช้งาน

2.3.1. ระดับความเสี่ยงที่อาจส่งผลกระทบเชิงลบ

Commented [A30]: NCSA AI Security Guidelines ๒.๓ หมวดหมู่ความเสี่ยงของปัญญาประดิษฐ์

Commented [A31]: NCSA AI Security Guidelines ๒.๓ หมวดหมู่ความเสี่ยงของปัญญาประดิษฐ์
ระดับที่ ๑ ความเสี่ยงที่ยอมรับไม่ได้
คำนิยาม ระบบปัญญาประดิษฐ์ที่ถูกจัดประเภทว่ามีคุณลักษณะการทำงานที่ก่อให้เกิดภัยคุกคามต่อความปลอดภัย สิทธิขั้นพื้นฐาน และคุณค่าที่เป็นที่ยอมรับในสังคมสากล
ตัวอย่าง
๑) ระบบการให้คะแนนทางสังคม การใช้ระบบปัญญาประดิษฐ์ โดยหน่วยงานภาครัฐ ...

Commented [A32]: NCSA AI Security Guidelines ๒.๓ หมวดหมู่ความเสี่ยงของปัญญาประดิษฐ์
ระดับที่ ๒ ความเสี่ยงสูง ...

Commented [A33R32]: MDES สรุปร่างหลักการของกฎหมายว่าด้วยปัญญาประดิษฐ์
▪ หลักการของหน้าที่สำหรับผู้ประยุกต์ใช้ปัญญาประดิษฐ์ความเสี่ยงสูง ...

Commented [A34]: EU AI Act
Article 17: Quality Management System
1. Providers of high-risk AI systems shall put a quality management system in place that ensures compliance with this Regulation. That system shall be documented in a systematic and orderly manner in the form of written ...

Commented [A35]: NCSA AI Security Guidelines ๒.๓ หมวดหมู่ความเสี่ยงของปัญญาประดิษฐ์
ระดับที่ ๓ ความเสี่ยงจำกัด ...

Commented [A36]: NCSA AI Security Guidelines ๒.๓ หมวดหมู่ความเสี่ยงของปัญญาประดิษฐ์
ระดับที่ ๔ ความเสี่ยงน้อยที่สุด ...

Commented [A37]: MDES สรุปร่างหลักการของกฎหมายว่าด้วยปัญญาประดิษฐ์
▪ หลักการของหน้าที่สำหรับผู้ประยุกต์ใช้ปัญญาประดิษฐ์ความเสี่ยงสูง ...

Commented [A38]: BOT แนวนโยบายธนาคารแห่งประเทศไทย เรื่อง การบริหารจัดการความเสี่ยงของการใช้งานระบบปัญญาประดิษฐ์
4.4 แนวทางบริหารจัดการความเสี่ยงของการใช้งานระบบ AI ...

2.3.2. ความรุนแรงของผลกระทบเชิงลบที่อาจเกิดขึ้น

2.3.3. ผลกระทบที่เกิดขึ้นนั้นสามารถย้อนกลับไปแก้ไขได้หรือไม่

2.3.4. ความเป็นไปได้หรือความเหมาะสมในการเข้าไปแทรกแซงการทำงานโดยมนุษย์

2.4. ในการนำระบบ AI มาใช้สนทนา หรือสื่อสารกับลูกค้าแทนมนุษย์ พิจารณาแนวทางการแจ้งให้ลูกค้าทราบเกี่ยวกับการให้บริการด้วยระบบ AI รวมทั้งพิจารณาจัดให้มีทางเลือกสำหรับลูกค้าในการติดต่อกับเจ้าหน้าที่ในกรณีที่ลูกค้าไม่ประสงค์จะสนทนาหรือสื่อสารกับระบบ AI

2.5. กรณีที่เป็นการใช้งาน AI ที่มีความเสี่ยงสูง หรือส่งผลกระทบโดยตรงต่อผู้ใช้หรือลูกค้า ต้องมีการทดสอบสิทธิภาพของระบบ AI ให้สอดคล้องกับระดับความเสี่ยงและผลกระทบที่อาจเกิดขึ้น และแจ้งผลการทดสอบไปยังผู้พัฒนาระบบและผู้เกี่ยวข้องเพื่อดำเนินการแก้ไข

2.6. การกำกับดูแลและบริหารความเสี่ยงในการใช้บริการ AI จากบุคคลภายนอกด้านเทคโนโลยีสารสนเทศ

2.6.1. จัดให้มีมาตรการควบคุมและติดตามการใช้งาน AI จากผู้ให้บริการภายนอก ตามนโยบายการบริหารจัดการความเสี่ยงจากบุคคลภายนอก (Third Party Risk Management Policy)

- กรณีเลือกใช้บริการจากบุคคลภายนอกด้านเทคโนโลยีสารสนเทศ เช่น การใช้ Generative AI จัดให้มีมาตรการควบคุมตามความเหมาะสม กำหนดขอบเขตการใช้งานภายในธนาคาร และข้อกำหนดเกี่ยวกับสิ่งที่ควรและไม่ควรปฏิบัติในการใช้งาน Generative AI ของธนาคาร
- กรณีจ้างผู้ให้บริการภายนอกพัฒนาระบบ AI ต้องมีการประเมินความเสี่ยงและข้อจำกัดจากการใช้งานระบบ AI และข้อมูลของบุคคลภายนอก ตลอดจนการพิจารณาข้อกำหนดในสัญญาผู้ให้บริการ เช่น การพัฒนาระบบ AI บนหลักการประยุกต์ใช้ AI อย่างมีความรับผิดชอบ (Responsible AI) ความโปร่งใสของโมเดล แหล่งข้อมูลที่ใช้ฝึกสอนโมเดล สิทธิในการตรวจสอบ ข้อกำหนดการให้บริการ และกระบวนการในการปรับปรุงโมเดล เป็นต้น

2.6.2. ประเมินระดับการพึ่งพาผู้ให้บริการ AI จากผู้ให้บริการภายนอกอย่างสม่ำเสมอ โดยให้คำนึงถึงความเสี่ยงด้านปฏิบัติการ (Operational Risk) และความเสี่ยงด้านการกระจุกตัว (Concentration Risk) ของผู้ให้บริการ AI ซึ่งอาจส่งผลกระทบให้ธุรกิจเกิดการหยุดชะงัก

2.6.3. กำหนดมาตรการควบคุมความเสี่ยงเพิ่มเติม ตามลักษณะการใช้งานและความเสี่ยงของระบบ เช่น การพิจารณาความจำเป็นของระบบสำรอง (System Redundancy) การกำกับดูแลการใช้ข้อมูลของระบบ AI การติดตั้งกลไก คำสั่ง หรือกระบวนที่สามารถปิดหรือหยุดการทำงานของระบบ AI ได้ทันที (Kill Switch) เพื่อหยุดการทำงานของระบบภายใต้เงื่อนไขที่กำหนด การจัดทำและทดสอบแผนฉุกเฉิน (Contingency Plan) รวมถึงการเปลี่ยนไปใช้ระบบอื่นตามความจำเป็น และการจัดทำแผนการสิ้นสุดการใช้บริการ (Exit Plan)

2.7. การกำกับดูแลและบริหารความเสี่ยงด้านความต่อเนื่องทางธุรกิจ (Business & IT Continuity) ของระบบ AI โดยระบบ AI ที่ถูกนำมาใช้ในกระบวนการสำคัญของธนาคาร (Critical Business Process) กำหนดให้มีแนวทางด้านความต่อเนื่องทางธุรกิจและระบบงาน (Business & IT Continuity) ที่ครอบคลุมทั้งการเตรียมความพร้อมและการทดสอบ ดังนี้

2.7.1. การเตรียมความพร้อม

Commented [A39]: ETDA AI Governance Guideline for Executive
6.2 การบริหารจัดการความเสี่ยงจากการประยุกต์ใช้ AI (AI Risk Management)
ระดับการมีส่วนร่วมของมนุษย์ (Level of Human Involvement)
จากตัวอย่างของการมีส่วนร่วมของมนุษย์ทั้ง 3 ระดับ จะพบว่าปัจจัยในการพิจารณาเลือกระดับการมีส่วนร่วมของมนุษย์นั้นมีความหลากหลายแตกต่างกันไปตามบริบทของการประยุกต์ใช้ AI ซึ่งองค์กรจำเป็นต้องมีการพิจารณาเลือกระดับการมีส่วนร่วมของมนุษย์อย่างเหมาะสม โดยอาจพิจารณาจากปัจจัย ดังต่อไปนี้
1.ระดับความเสี่ยงที่อาจส่งผลกระทบเชิงลบ
2.ความรุนแรงของผลกระทบเชิงลบที่อาจเกิดขึ้น

Commented [A40]: BOT แนวนโยบายธนาคารแห่งประเทศไทย เรื่อง การบริหารจัดการความเสี่ยงของการใช้งานระบบปัญญาประดิษฐ์
4.4 แนวทางบริหารจัดการความเสี่ยงของการใช้งานระบบ AI
ส่วนที่ 1 ธรรมาภิบาลของการนำระบบ AI มาใช้งาน (governance)
(3) การบริหารจัดการความเสี่ยงจากการใช้งานระบบ AI

Commented [A41]: OIC **AI Governance Guideline**
ส่วนที่ 2 : ความทนทานและการรักษาความมั่นคงปลอดภัยระบบ AI (Robustness and AI Security)
2.1 การประเมินประสิทธิภาพการทำงานของระบบ AI
บริษัทควรประเมินประสิทธิภาพในการทำงานของระบบ AI อย่างสม่ำเสมอ โดยคำนึงถึงวัตถุประสงค์การใช้งาน วิเคราะห์ความถูกต้องของโมเดลและข้อมูล เช่น การวิเคราะห์ความไวของโมเดลต่อการเปลี่ยนแปลง ของข้อมูล (Sensitivity Analysis) การทดสอบ

Commented [A42]: SEC 7.2.2 การพัฒนาโมเดล
สำหรับ AI/ML ที่ถูกนำมาใช้งานในกระบวนการทางธุรกิจที่มีความเสี่ยงสูง ผู้ประกอบธุรกิจควรหลีกเลี่ยงการใช้งาน AI/ML หรือโมเดลที่มนุษย์ไม่สามารถควบคุมหรือระงับการทำงานใด ๆ ได้ (Human-out-of-the-Loop) และควรจัดให้ที่มีช่องทางให้มนุษย์เข้ามามีส่วนร่วมในการควบคุมการทำงานหรือการตัดสินใจของ AI/ML (Human-in-the-Loop หรือ Human-over-the-Loop3) หรือจัดให้

Commented [A43]: OIC AI Governance Guideline
ส่วนที่ 1 : การกำกับดูแลการใช้งาน AI (AI Governance)
1.4 การกำกับดูแลการใช้บริการ AI จากบุคคลภายนอกด้านเทคโนโลยีสารสนเทศ
ครอบคลุมในเรื่องดังต่อไปนี้
1.4.1 บริษัทควรมีมาตรการควบคุมและติดตามการใช้งาน AI จากการใช้บริการจากบุคคลภายนอกด้านเทคโนโลยีสารสนเทศ ตามก

- ระบบ AI ต้องมีแผนความต่อเนื่อง (Contingency/Fallback Plan) และขั้นตอน manual override เพื่อรองรับกรณีระบบขัดข้องหรือไม่สามารถให้บริการได้
- แผนดังกล่าวต้องบูรณาการเข้ากับแผนความต่อเนื่องของกระบวนการหลักที่ระบบ AI สนับสนุน

2.7.2. การทดสอบความต่อเนื่อง (Testing)

- ระบบ AI ต้องถูกรวมอยู่ในขอบเขตการทดสอบ BCP และ ITCP ของหน่วยงานที่เกี่ยวข้อง
- พิจารณาการทดสอบสถานการณ์จำลองเฉพาะของ AI เช่น การกู้คืนระบบจากความล้มเหลวของโมเดล หรือการโจมตีที่ทำให้ผลลัพธ์ของ AI เบี่ยงเบน (AI Attack Scenario)
- ซักซ้อมแผนอย่างสม่ำเสมอ เพื่อทดสอบประสิทธิภาพของกระบวนการ

2.7.3. การรายงานและติดตามผล

- ผลการทดสอบและการบริหารความต่อเนื่องของระบบ AI ควรถูกรายงานต่อคณะกรรมการที่เกี่ยวข้อง เพื่อประกอบการประเมินความพร้อมด้าน Operational Resilience ของธนาคาร

Commented [A44]: Content added as per feedback from Op Risk team

ส่วนที่ 3. การบริหารจัดการและควบคุมความเสี่ยงในการพัฒนาและการใช้งานระบบ AI

1. ลักษณะความเสี่ยงของการใช้งานระบบ AI

การใช้งานระบบ AI อาจเกิดความเสี่ยงได้ในกระบวนการเรียนรู้ข้อมูลเพื่อพัฒนาระบบ AI ซึ่งครอบคลุมตั้งแต่การเตรียมข้อมูล การพัฒนาโมเดล และการทดสอบวัดผลโมเดล ตลอดจนกระบวนการนำระบบ AI มาใช้งาน โดยสามารถจำแนกความเสี่ยงออกเป็น 3 ด้าน ดังนี้

1.1. ความเสี่ยงด้านข้อมูล

เนื่องจากในกระบวนการเรียนรู้ข้อมูล โมเดลอาจใช้ข้อมูลที่ไม่มีคุณภาพ เช่น ข้อมูลที่ไม่เป็นปัจจุบัน ไม่ถูกต้อง หรือไม่หลากหลายเพียงพอ ส่งผลให้โมเดลอาจสร้างผลลัพธ์ที่ไม่น่าเชื่อถือ เชิงลบ มีความโน้มเอียง หรือเป็นเท็จ หรืออาจใช้ข้อมูลมีการละเมิดทรัพย์สินทางปัญญานำไปสู่การละเมิดทางกฎหมายได้ นอกจากนี้ ในกระบวนการเรียนรู้ข้อมูลและกระบวนการนำระบบ AI มาใช้งาน หากการรักษาความมั่นคงปลอดภัยของข้อมูลไม่รัดกุมเพียงพอ หรือมีช่องโหว่ของระบบที่อาจทำให้เป็นเป้าหมายของการถูกโจมตี อาจส่งผลให้ข้อมูลในระดับชั้นข้อมูลสารสนเทศที่ใช้เฉพาะภายในองค์กร (Internal Use Only) ขึ้นไป ข้อมูลลูกค้าของธนาคาร ข้อมูลส่วนบุคคล หรือข้อมูลอ่อนไหว ถูกเปิดเผยแก่ผู้ที่ไม่เกี่ยวข้องหรือรั่วไหลออกไปภายนอกได้

1.2. ความเสี่ยงด้านการพัฒนาโมเดล

เนื่องจากโมเดลอาจเรียนรู้และสร้างเงื่อนไขที่ไม่สามารถอธิบายที่มาของผลลัพธ์ได้ (explainability) หรืออาจสร้างผลลัพธ์ที่ไม่น่าเชื่อถือ เชิงลบ โน้มเอียง หรือเป็นเท็จ อันส่งผลให้เกิดความเสี่ยงต่อการดำเนินธุรกิจ ไม่ว่าจะเป็นการตัดสินใจทางธุรกิจ การสนับสนุนการปฏิบัติงานหรือระบบงานสำคัญ และการให้บริการแก่ลูกค้า เป็นต้น

1.3. ความเสี่ยงด้านภัยคุกคามทางไซเบอร์

เนื่องจากอาจมีการโจมตีทางไซเบอร์ในรูปแบบต่าง ๆ ซึ่งรวมถึงการโจมตีในรูปแบบใหม่ ๆ ที่เจาะจงกับระบบ AI เช่น

- Prompt injection ซึ่งเป็นการสร้างคำถามที่แอบแฝงเป้าหมายหลักเพื่อหลบเลี่ยงการตรวจจับ หรือหลอกให้ระบบ AI ทำงานผิดไปจากที่ได้มีการพัฒนาออกแบบไว้
- Model inversion ซึ่งเป็นการเข้าถึงข้อมูลโครงสร้างของโมเดลหรือข้อมูลที่ใช้ในการเรียนรู้ของโมเดลโดยมิชอบ
- Data poisoning ซึ่งเป็นการปรับแต่งชุดข้อมูลที่ใช้ในการเรียนรู้ของโมเดลด้วยข้อมูลที่ไม่ถูกต้อง เพื่อทำให้โมเดลสร้างผลลัพธ์ที่ผิดพลาด หรือทำให้เกิดช่องโหว่ที่สามารถถูกโจมตีได้ อันจะทำให้ประสิทธิภาพของระบบ AI ลดลง
- Adversarial attack ซึ่งเป็นการป้อนข้อมูลที่มีลักษณะคล้ายคลึงกับข้อมูลปกติ แต่ส่งผลให้โมเดลทำงานผิดพลาดระหว่างการใช้งาน

ทั้งนี้ควรพิจารณาความเสี่ยงด้านอื่น ๆ จากการนำระบบ AI มาใช้งาน ซึ่งขึ้นอยู่กับลักษณะการนำไปใช้งาน (use case) เช่น ความเสี่ยงด้านกลยุทธ์ ความเสี่ยงด้านการเงิน ความเสี่ยงด้านชื่อเสียง ความเสี่ยงด้านปฏิบัติการและกฎหมาย ความเสี่ยงด้านเทคโนโลยีสารสนเทศ

Commented [A45]: OIC **AI Governance Guideline**
ส่วนที่ 1 : การกำกับดูแลการใช้งาน AI (AI Governance)
1.3 การบริหารจัดการความเสี่ยงการใช้งาน AI
1.3.1 การประเมินความเสี่ยงตามลักษณะการใช้งาน AI และผลกระทบเชิงลบที่อาจเกิดขึ้น โดยเฉพาะที่มีความเสี่ยงทางด้านพฤติกรรมทางการตลาด (Market Conduct) และความเสี่ยงด้านความมั่นคง ทางการเงิน (Prudential Risk) สอดคล้องตามการกำกับดูแลตามระดับความเสี่ยง (Risk-Based Proportionality) ทั้งนี้ บริษัทควรพิจารณาความเสี่ยงที่มีโอกาสเกิดขึ้นจากการใช้งาน AI เช่น
- การละเมิดทรัพย์สินทางปัญญา (Intellectual Property Infringement) ข้อมูลที่ระบบ AI นำมาใช้ อาจมีการละเมิดทรัพย์สินทางปัญญานำไปสู่การละเมิดทางกฎหมายได้

Commented [A46]: KTB ERM

Commented [A47]: BOT แนวนโยบายธนาคารแห่งประเทศไทย เรื่อง การบริหารจัดการความเสี่ยงของการใช้งานระบบปัญญาประดิษฐ์
4.2 ลักษณะความเสี่ยงของการใช้งานระบบ AI
การใช้งานระบบ AI อาจเกิดความเสี่ยงได้ในกระบวนการเรียนรู้ข้อมูลเพื่อพัฒนาระบบ AI ซึ่งครอบคลุมตั้งแต่การเตรียมข้อมูล การพัฒนาโมเดล และการทดสอบวัดผลโมเดล ตลอดจนกระบวนการนำระบบ AI มาใช้งาน โดยสามารถจำแนกความเสี่ยงออกเป็น 3 ด้าน ดังนี้
(1) ความเสี่ยงด้านข้อมูล
จากเหตุที่ในกระบวนการเรียนรู้ข้อมูล โมเดลอาจใช้ข้อมูลที่ไม่มีคุณภาพ เช่น ไม่เป็นปัจจุบัน ไม่ถูกต้อง หรือไม่หลากหลายเพียงพอ ส่งผลให้โมเดลอาจสร้างผลลัพธ์ที่ไม่น่าเชื่อถือ เชิงลบ โน้มเอียง หรือเป็นเท็จ นอกจากนี้ ในกระบวนการเรียนรู้ข้อมูลและกระบวนการนำระบบ AI มาใช้งาน หากการรักษาความมั่นคงปลอดภัยของข้อมูลไม่รัดกุมเพียงพอ หรือมีช่องโหว่ของระบบที่อาจทำให้เป็นเป้าหมายของการถูกโจมตี อาจส่งผลให้ข้อมูลสำคัญ ข้อมูลส่วนบุคคล หรือข้อมูลอ่อนไหว ถูกเปิดเผยแก่ผู้ที่ไม่เกี่ยวข้องหรือรั่วไหลออกไปภายนอกได้
(2) ความเสี่ยงด้านการพัฒนาโมเดล
จากเหตุที่โมเดลอาจเรียนรู้และสร้างเงื่อนไขที่ไม่สามารถอธิบายที่มาของผลลัพธ์ได้ (explainability) หรืออาจสร้างผลลัพธ์ที่ไม่น่าเชื่อถือ เชิงลบ โน้มเอียง หรือเป็นเท็จ อันส่งผลให้เกิดความเสี่ยงต่อการดำเนินธุรกิจ ไม่ว่าจะในการตัดสินใจทางธุรกิจ การสนับสนุนการปฏิบัติงานหรือระบบงานสำคัญ และการให้บริการแก่ลูกค้า เป็นต้...

2. การบูรณาการความเสี่ยงของระบบ AI เข้ากับกรอบการบริหารความเสี่ยงของธนาคาร

ความเสี่ยงที่เกิดจากระบบ AI ไม่ใช่เป็นเพียงความเสี่ยงทางเทคโนโลยีที่จำกัดอยู่แค่ในฝ่ายเทคโนโลยีสารสนเทศเท่านั้น แต่เป็นความเสี่ยงทางธุรกิจที่สามารถส่งผลกระทบในระดับองค์กรได้ ดังนั้น การบริหารจัดการความเสี่ยง AI ในลักษณะแยกส่วน จึงไม่มีประสิทธิภาพเพียงพอ จำเป็นต้องบูรณาการความเสี่ยงที่จำเพาะเจาะจงกับ AI เข้าไปในกรอบการบริหารความเสี่ยงระดับองค์กร (Enterprise Risk Management - ERM) ของธนาคาร

ทั้งนี้ การดำเนินการในส่วนต่าง ๆ ให้เป็นไปตามนโยบายและคู่มือของแต่ละประเภทความเสี่ยงที่ธนาคารกำหนดไว้

3. การควบคุมความเสี่ยง

กำหนดแนวทางควบคุมความเสี่ยงของการใช้งานระบบ AI ครอบคลุมด้านข้อมูลที่ระบบ AI ใช้ในการเรียนรู้ ด้านการพัฒนาโมเดล และด้านภัยคุกคามทางไซเบอร์เพื่อให้การใช้งานระบบ AI อยู่ภายใต้กรอบความถูกต้อง เชื่อถือได้ และโปร่งใส ตลอดจนมีความมั่นคงปลอดภัยจากภัยคุกคามไซเบอร์รูปแบบใหม่ที่อาจเกิดขึ้น

Commented [A48]: NCSA AI Security Guidelines
๔.๒ การบูรณาการความเสี่ยงปัญญาประดิษฐ์เข้ากับกรอบการบริหารความเสี่ยงองค์กร (Integrating AI Risks into ERM)
ความเสี่ยงที่เกิดจากระบบปัญญาประดิษฐ์ไม่ใช่เป็นเพียงความเสี่ยงทางเทคโนโลยีที่จำกัดอยู่แค่ในฝ่ายเทคโนโลยีสารสนเทศเท่านั้น แต่เป็นความเสี่ยงทางธุรกิจที่สามารถส่งผลกระทบในระดับองค์กรได้ ดังนั้น การบริหารจัดการความเสี่ยงปัญญาประดิษฐ์ในลักษณะแยกส่วน จึงไม่มีประสิทธิภาพเพียงพอ องค์กรจึงจำเป็นต้องบูรณาการความเสี่ยงที่จำเพาะเจาะจงกับปัญญาประดิษฐ์ เข้าไปในกรอบการบริหารความเสี่ยงระดับองค์กร (Enterprise Risk Management - ERM) ที่มีอยู่เดิม
กระบวนการนี้ทำให้มั่นใจได้ว่าผู้บริหารระดับสูงและคณะกรรมการบริษัทจะสามารถมองเห็น ประเมิน และตัดสินใจเกี่ยวกับความเสี่ยงปัญญาประดิษฐ์ได้ในบริบทเดียวกับความเสี่ยงอื่น ๆ ขององค์กร เช่น ความเสี่ยงทางการเงิน การตลาด หรือการปฏิบัติการ ซึ่งนำไปสู่การจัดสรรทรัพยากรและการกำหนดกลยุทธ์ที่สอดคล้องกันทั่วทั้งองค์กร

Commented [A49]: BOT แนวนโยบายธนาคารแห่งประเทศไทย เรื่อง การบริหารจัดการความเสี่ยงของการใช้งานระบบปัญญาประดิษฐ์
4.4 แนวทางบริหารจัดการความเสี่ยงของการใช้งานระบบ AI
ส่วนที่ 2 การพัฒนาและการรักษาความมั่นคงปลอดภัยของการใช้งานระบบ AI (development and security)
ผลลัพธ์ที่คาดหวัง : ผู้ให้บริการทางการเงินมีแนวทางควบคุมความเสี่ยงของการใช้งานระบบ AI ครอบคลุมด้านข้อมูลที่ระบบ AI ใช้ในการเรียนรู้ ด้านการพัฒนาโมเดล และด้านภัยคุกคามทางไซเบอร์ เพื่อให้การใช้งานระบบ AI อยู่ภายใต้กรอบความถูกต้อง เชื่อถือได้ และโปร่งใส ตลอดจนมีความมั่นคงปลอดภัยจากภัยคุกคามไซเบอร์รูปแบบใหม่ที่อาจเกิดขึ้น โดยผู้ให้บริการทางการเงินควรดำเนินการดังนี้

ส่วนที่ 4. ภาคผนวก

1. นิยามและคำจำกัดความ

1.1. “ระบบปัญญาประดิษฐ์ (Artificial Intelligence: AI)” หรือ “ระบบ AI” หมายถึงระบบที่เลียนแบบปัญญาของมนุษย์ โดยสามารถทำงานต่าง ๆ เช่น เรียนรู้ จดจำ ตัดสินใจ ปฏิบัติงาน หรือสร้างสรรค์เนื้อหาใหม่ ผ่านกลไกการเรียนรู้จากข้อมูลและการสร้างเงื่อนไขอัตโนมัติ ทั้งนี้ ไม่รวมถึงระบบอัตโนมัติที่มนุษย์เป็นผู้กำหนดขั้นตอนการทำงานของระบบ

1.2. “Machine Learning” AI ประเภทหนึ่งที่มีความสามารถในการเรียนรู้ หรือปรับปรุงประสิทธิภาพการทำงานของตน โดยพัฒนาและใช้อัลกอริทึมในการวิเคราะห์ และเรียนรู้จากข้อมูลที่ได้รับจากการฝึกสอน หรือสภาพแวดล้อม

1.3. “Generative AI” เป็นระบบ AI ที่มีความสามารถในการสร้างข้อมูลใหม่ ที่ไม่เคยมีมาก่อนด้วยตัวเอง เช่น ข้อความ รูปภาพ เสียง วีดีโอ หรือรูปแบบอื่น

1.4. “การตัดสินใจเชิงกลยุทธ์” หมายถึงงานหลักที่เกี่ยวข้องกับการตัดสินใจเชิงกลยุทธ์ อ้างอิงตามประกาศธนาคารแห่งประเทศไทยว่าด้วยหลักเกณฑ์การใช้บริการจากพันธมิตรทางธุรกิจ (business partner) ของสถาบันการเงิน และประกาศธนาคารแห่งประเทศไทยว่าด้วยหลักเกณฑ์การกำกับดูแลการใช้บริการจากผู้ให้บริการสนับสนุนการประกอบธุรกิจ (business facilitator) ของสถาบันการเงินเฉพาะกิจ เช่น งานอนุมัติสินเชื่อ งานอนุมัติเปิดบัญชี งานอนุมัติการฝาก ถอน หรือโอนเงิน

1.5. “Human in the Loop” หมายถึงการใช้งานระบบ AI ที่มนุษย์มีส่วนร่วมในการควบคุมการทำงานหรือตัดสินใจทั้งหมด โดยมีระบบ AI ทำหน้าที่ให้คำแนะนำหรือข้อมูล แต่ไม่สามารถทำงานหรือตัดสินใจในการดำเนินการใด ๆ ได้โดยปราศจากมนุษย์

1.6. “Human over the Loop” หมายถึงการใช้งานระบบ AI ที่ระบบ AI สามารถทำงานหรือตัดสินใจได้โดยอัตโนมัติ แต่ยังคงจำเป็นต้องมีมนุษย์ในการกำกับดูแล และมนุษย์สามารถเข้าควบคุมหรือระงับการทำงานได้เมื่อพบความผิดพลาดหรืออาจก่อให้เกิดผลกระทบเชิงลบ

1.7. “Risk Appetite” หมายถึงระดับหรือประเภท และเกณฑ์ของความเสี่ยงที่ธนาคารจะยอมรับได้เพื่อช่วยให้บรรลุเป้าหมาย ซึ่งระดับความเสี่ยงที่ยอมรับได้ที่ธนาคารจะกำหนดอาจจะเป็นค่าเดียวหรือเป็นช่วง ขึ้นอยู่กับความเหมาะสมของปัจจัยนั้น ๆ

1.8. “หลักการจริยธรรมปัญญาประดิษฐ์ของไทยตามแนวทางของสำนักงานคณะกรรมการดิจิทัลเพื่อเศรษฐกิจและสังคมแห่งชาติ (สดช.)” ประกอบด้วย ความสามารถในการแข่งขันและการพัฒนาอย่างยั่งยืน (Competitiveness and Sustainable Development), ความสอดคล้องกับกฎหมาย จริยธรรม และมาตรฐานสากล (Laws, Ethics, and International Standards), ความโปร่งใสและความรับผิดชอบต่อผลของการกระทำ (Transparency and Accountability), ความมั่นคงปลอดภัยและความเป็นส่วนตัว (Security and Privacy), ความเท่าเทียม หลากหลาย ครอบคลุม และเป็นธรรม (Fairness), ความน่าเชื่อถือ (Reliability)

2. นโยบายและเอกสารที่เกี่ยวข้อง

2.1. จรรยาบรรณธุรกิจธนาคาร (Code of Conduct)

Commented [A50]: BOT def

Commented [A51]: NCSA AI Security Guidelines ๒.๒ ประเภทของปัญญาประดิษฐ์ตามการใช้งาน

Commented [A52]: BOT def

Commented [A53]: BOT def

Commented [A54]: BOT def

Commented [A55]: ETDA

2.2. นโยบายการบริหารความเสี่ยงด้านเทคโนโลยีสารสนเทศ (IT Risk Management Policy)

2.3. นโยบายการบริหารจัดการความเสี่ยงจากบุคคลภายนอก (Third Party Risk Management Policy)

2.4. นโยบายการกำกับดูแลข้อมูล (Data Governance Policy)

2.5. นโยบายและกระบวนการรักษาความปลอดภัยสารสนเทศ (IT Security Policy)

3. มาตรฐานสากลและแนวปฏิบัติที่เกี่ยวข้อง

3.1. NIST Artificial Intelligence Risk Management Framework (AI RMF 1.0) เป็นชุดมาตรฐานและแนวปฏิบัติที่ช่วยให้องค์กรและบุคคลสามารถจัดการความเสี่ยงที่เกี่ยวข้องกับระบบ AI โดยกำหนดการดำเนินการและผลลัพธ์สำหรับการจัดการความเสี่ยงเกี่ยวกับ AI ที่แบ่งออกเป็นสี่ฟังก์ชัน

3.1.1. Govern: เน้นการกำหนดกรอบการบริหารและขั้นตอนการกำกับดูแล เพื่อควบคุมการจัดการความเสี่ยงที่เกี่ยวข้องกับ AI

3.1.2. Map: เน้นการเชื่อมโยงและตรวจสอบระบบ AI และความเสี่ยงที่เกี่ยวข้อง

3.1.3. Measure: เน้นความสำคัญของการวัดและประเมินความเสี่ยงที่เกี่ยวข้องกับ AI

3.1.4. Manage: จัดการมาตรการบริหารความเสี่ยงเกี่ยวกับ AI ให้มีการบริหารจัดการและการเพิ่มประสิทธิภาพของมาตรการอย่างต่อเนื่องและสม่ำเสมอ

3.2. ISO/IEC 42001:2023 (Information technology — Artificial intelligence — Management system) เป็นมาตรฐานสากลฉบับแรกของโลกสำหรับระบบการจัดการ AI โดยกำหนดกรอบการดำเนินงานเพื่อให้องค์กรสามารถพัฒนาและใช้งาน AI อย่างมีความรับผิดชอบครอบคลุมทั้งด้านความเสี่ยง จริยธรรม และความโปร่งใส

3.3. ISO/IEC 23894:2023 (Information technology — Artificial intelligence — Guidance on risk management) ให้คำแนะนำเกี่ยวกับการบริหารจัดการความเสี่ยงที่เกี่ยวข้องกับ AI โดยเฉพาะ ซึ่งสามารถนำไปปรับใช้ร่วมกับกรอบการบริหารความเสี่ยงขององค์กรได้

3.4. ISO/IEC 5338:2023 (Information technology — Artificial intelligence — AI system life cycle processes) กระบวนการวงจรชีวิตของระบบ AI ระบุกระบวนการต่าง ๆ ที่สนับสนุนการนิยาม การควบคุม การจัดการ การดำเนินการ และการปรับปรุงระบบ AI ในแต่ละขั้นตอนของวงจรชีวิตของระบบ AI กระบวนการเหล่านี้สามารถนำไปใช้ภายในองค์กรหรือโครงการในระหว่างการพัฒนาหรือการจัดหาระบบ AI

4. กฎหมาย กฎระเบียบ ข้อกำหนด ประกาศ และแนวปฏิบัติจากหน่วยงานภายนอกที่เกี่ยวข้อง

4.1. พระราชบัญญัติการรักษาความมั่นคงปลอดภัยไซเบอร์ พ.ศ. 2562

4.2. พระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562 (Personal Data Protection Act: PDPA)

4.3. แนวนโยบายธนาคารแห่งประเทศไทย เรื่อง การบริหารจัดการความเสี่ยงของการใช้งานระบบปัญญาประดิษฐ์ (ฉบับกันยายน พ.ศ. 2568)

4.4. แนวปฏิบัติการใช้ปัญญาประดิษฐ์อย่างมั่นคงปลอดภัย (AI Security Guidelines) จากสำนักงานคณะกรรมการการรักษาความมั่นคงปลอดภัยไซเบอร์แห่งชาติ (สกมช.) (ฉบับกันยายน พ.ศ. 2568)

Commented [A56]: ETDA สรุปร่างหลักการของกฎหมายว่าด้วยปัญญาประดิษฐ์
การบริหารจัดการความเสี่ยง (Risk Management)

Commented [A57]: ETDA สรุปร่างหลักการของกฎหมายว่าด้วยปัญญาประดิษฐ์
การบริหารจัดการความเสี่ยง (Risk Management)

Commented [A58]: NCSA AI Security Guidelines ๒.๓
มาตรฐานและกรอบการดำเนินงานสากล

Commented [A59]: NCSA AI Security Guidelines ๒.๔
กฎหมาย กฎระเบียบ และแนวปฏิบัติที่เกี่ยวข้องในประเทศไทย

4.5. แนวปฏิบัติเรื่องการกำกับดูแลการใช้งานปัญญาประดิษฐ์ สำหรับบริษัทประกันภัย (AI Governance Guideline) จากสำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย (คปภ.) (ฉบับพฤศจิกายน พ.ศ. 2568)

4.6. (ร่าง) หลักการของกฎหมายว่าด้วยปัญญาประดิษฐ์ จากกระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม (ฉบับพฤษภาคม พ.ศ. 2568)

4.7. AI Governance Guideline for Executive แนวทางการประยุกต์ใช้ปัญญาประดิษฐ์อย่างมีธรรมาภิบาลสำหรับผู้บริหารองค์กร จากสำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์ (สพธอ.) (ฉบับสิงหาคม พ.ศ. 2566)

4.8. กรอบการกำกับดูแลการใช้งานปัญญาประดิษฐ์ (Artificial Intelligence) และการเรียนรู้ของเครื่อง (Machine Learning) ในตลาดทุน จากสำนักงานคณะกรรมการกำกับหลักทรัพย์และตลาดหลักทรัพย์ (ฉบับธันวาคม พ.ศ. 2566)

5. ประวัติการทบทวน

เวอร์ชั่น	วันที่	ดำเนินการแก้ไขโดย	คำอธิบาย
V1.0	1 กันยายน 2568	PwC	ฉบับร่างเบื้องต้น
V1.1	3 กันยายน 2568	PwC	ปรับปรุงตามความคิดเห็นของ KTB
V1.2	29 กันยายน 2568	PwC	ปรับปรุงโครงสร้างเอกสารและเนื้อหาตามแนวนโยบายธนาคารแห่งประเทศไทย เรื่อง การบริหารจัดการความเสี่ยงของการใช้งานระบบปัญญาประดิษฐ์ (ฉบับกันยายน พ.ศ. 2568) และตามความคิดเห็นของ KTB
V1.3	14 ตุลาคม 2568	PwC	ปรับปรุงเนื้อหาตามแนวปฏิบัติการใช้ปัญญาประดิษฐ์อย่างมั่นคงปลอดภัย (AI Security Guidelines) (ฉบับกันยายน พ.ศ. 2568)
V1.4	27 ตุลาคม 2568	PwC	ปรับปรุงโครงสร้างการกำกับดูแลและเนื้อหาตามความคิดเห็นของ KTB
V1.5	31 ตุลาคม 2568	PwC	ปรับปรุงโครงสร้างการกำกับดูแลตามความคิดเห็นของ CISO
V1.6	7 พฤศจิกายน 2568	PwC	ปรับปรุงโครงสร้างการกำกับดูแลตามข้อมูลที่ได้รับจากการ Hearing กับผู้ที่เกี่ยวข้อง
V1.7	7 พฤศจิกายน 2568	PwC	ปรับปรุงโครงสร้างการกำกับดูแลตามข้อมูลที่ได้รับจากการ Hearing กับผู้ที่เกี่ยวข้อง
V1.8	16 พฤศจิกายน 2568	PwC	ปรับปรุงโครงสร้างเอกสารและเนื้อหาตามความคิดเห็นของ KTB

เวอร์ชั่น	วันที่	ดำเนินการแก้ไขโดย	คำอธิบาย
			ปรับปรุงชื่อนโยบาย ย้ายเนื้อหารายละเอียดไปยังเอกสารคู่มือการบริหารความเสี่ยงด้านปัญญาประดิษฐ์
V1.9	9 มกราคม 2569	PwC	ปรับปรุงโครงสร้างการกำกับดูแลตามความคิดเห็นของ CISO
V1.10	15 มกราคม 2569	PwC	ปรับปรุงโครงสร้างการกำกับดูแลตามความคิดเห็นของ KTB
V1.11	28 มกราคม 2569	PwC	ปรับปรุงภาคผนวก
V1.12	12 กุมภาพันธ์ 2569	PwC	ปรับปรุงเนื้อหาเอกสารตามความคิดเห็นของ KTB
V1.13	17 กุมภาพันธ์ 2569	PwC	ปรับปรุงเนื้อหาเอกสารและเพิ่มเนื้อหาอ้างอิงในคอมเมนต์