

## ข้อเสนอเชิงนโยบายในการจัดทำกฎหมายปัญญาประดิษฐ์ของสหภาพยุโรป: หลักเกณฑ์และแนวทางปฏิบัติเกี่ยวกับการพัฒนาเทคโนโลยี AI ที่มีความเสี่ยงสูง (High-risk AI applications)

คณะกรรมการการยุโรปได้เผยแพร่เอกสาร “White Paper on Artificial Intelligence – A European Approach to Excellence and Trust” โดยมีสาระสำคัญเป็นการเสนอนโยบายการดำเนินการด้านปัญญาประดิษฐ์ของสหภาพยุโรปและที่เกี่ยวข้องกับการพัฒนาปัญญาประดิษฐ์ที่มีความเสี่ยงสูง (High-risk AI applications) ซึ่งได้เสนอแนวทางการกำหนดหลักเกณฑ์การพิจารณาการเป็นเทคโนโลยีปัญญาประดิษฐ์ที่มีความเสี่ยงสูงไว้สองประการ ได้แก่

**1** ปัญญาประดิษฐ์นั้นจะนำไปใช้ในภาคอุตสาหกรรมที่มีความอ่อนไหวหรือโดยลักษณะของการปฏิบัติงานอาจมีความเสี่ยงที่จะเกิดผลกระทบในวงกว้าง เช่น การดำเนินงานด้านการแพทย์และสาธารณสุข การขนส่ง พลังงาน การจ้างงาน ชายแดน การอพยพ การศาล และกิจการสาธารณะอื่น ๆ ของรัฐ

**2** ระดับความรุนแรงของผลกระทบที่อาจเกิดขึ้นจากการปฏิบัติงานของปัญญาประดิษฐ์ เช่น มีความเสี่ยงต่อชีวิตในการบาดเจ็บหรือเสียชีวิต ความเสียหายต่อทรัพย์สิน หรือที่กระทบต่อสิทธิของบุคคลหรือบุคคล เป็นต้น

### ข้อเสนอหลักเกณฑ์และแนวปฏิบัติในการจัดทำข้อกำหนดการพัฒนาเทคโนโลยีปัญญาประดิษฐ์ที่มีความเสี่ยงสูง (High-risk AI applications)

ชุดข้อมูลที่ใช้ทดสอบระบบ AI (Training data)

ต้องอยู่บนพื้นฐานของการเคารพศักดิ์ศรีความเป็นมนุษย์ (Human Dignity) และสิทธิขั้นพื้นฐานของประชาชน (Fundamental Rights) ห้ามการเลือกปฏิบัติทางเพศหรือเชื้อชาติ และอยู่ภายใต้ขอบเขตกฎหมายของสหภาพยุโรป โดยเฉพาะที่เกี่ยวข้องกับการรักษาความเป็นส่วนตัวและการคุ้มครองข้อมูลส่วนบุคคล (Privacy Protection) เช่น General Data Protection Regulation (GDPR)

การเก็บรักษาข้อมูล (Data and record-keeping)

เพื่อประโยชน์ในการตรวจสอบย้อนกลับกรณีการทำงานหรือการตัดสินใจของระบบ AI มีปัญหาควรเก็บรักษาข้อมูล ดังต่อไปนี้

- ชุดข้อมูลที่ถูกต้องที่ใช้ในการฝึกฝนและทดสอบระบบ AI พร้อมคำอธิบายลักษณะที่สำคัญ และวิธีการเลือกใช้ชุดข้อมูลดังกล่าว
- ข้อมูลที่ระบบกำหนดขึ้นเอง
- ข้อมูลประกอบที่เกี่ยวข้องกับการเขียนโปรแกรม วิธีการทดสอบ กระบวนการและเทคนิคที่ใช้ในการสร้าง ทดสอบ และตรวจสอบความถูกต้องของระบบ AI รวมถึงข้อมูลเรื่องความปลอดภัยและข้อห้ามเรื่องการเลือกปฏิบัติ

ข้อกำหนดนี้จะช่วยให้เกิดความโปร่งใส ส่งเสริมกลไกการตรวจสอบและความรับผิดชอบระหว่างผู้มีส่วนเกี่ยวข้องตั้งแต่ขั้นตอนการสร้างไปจนถึงการนำมาใช้ ซึ่งจะเป็นการกำกับดูแลตลอดวงจรของการใช้ปัญญาประดิษฐ์

การให้ข้อมูลแก่สาธารณะ (Information to be provided)

เพื่อส่งเสริมการสร้าง AI อย่างมีความรับผิดชอบ จำเป็นต้องจัดทำข้อกำหนดการแจ้งข้อมูลแก่สาธารณะ โดยเฉพาะ AI ที่มีความเสี่ยงสูงโดยอาจกำหนดให้ผู้พัฒนาต้องแจ้งข้อมูลเกี่ยวกับวัตถุประสงค์ในการสร้าง ความสามารถและข้อจำกัดของ AI และปฏิสัมพันธ์ที่อาจเกิดขึ้นระหว่าง AI และมนุษย์ โดยต้องเป็นข้อมูลที่มีความชัดเจน เหมาะสม กระชับ และเข้าใจได้ง่าย

ความเที่ยงตรงและแม่นยำของระบบปฏิบัติการ (Robustness and accuracy)

AI จะต้องมีความสามารถทางเทคนิคที่แม่นยำตามวัตถุประสงค์ในการสร้าง โดยอาจจัดทำข้อกำหนดเพื่อลดความเสี่ยงจากการปฏิบัติงานที่ผิดพลาดของ AI ดังนี้

- ข้อกำหนดเกี่ยวกับความเที่ยงตรงของระบบ AI เช่น การกำหนดให้มีการแสดงระดับความแม่นยำในวงจรของระบบปฏิบัติการ
- ข้อกำหนดให้ AI มีความสามารถแสดงผลซ้ำ
- ข้อกำหนดเกี่ยวกับการจัดการความผิดพลาดในวงจรของระบบ
- ข้อกำหนดเกี่ยวกับการป้องกันการโจมตีทั้งการโจมตีอย่างเปิดเผยและการโจมตีที่ซับซ้อน รวมถึงมาตรการบรรเทาผลกระทบจากการถูกโจมตีและมาตรการเพื่อความมั่นคงปลอดภัยกรณีเกิดเหตุฉุกเฉิน

การกำกับดูแลโดยมนุษย์ (Human Oversight)

การนำ AI มาใช้ในบางวัตถุประสงค์จำเป็นต้องมีการแทรกแซงโดยการตัดสินใจของมนุษย์ ซึ่งอาจมีรายละเอียดที่แตกต่างไปตามวัตถุประสงค์ ดังมีตัวอย่างการกำหนดมาตรการกำกับดูแลโดยมนุษย์ ดังนี้

ข้อกำหนดเฉพาะสำหรับการใช้เทคโนโลยีชีวมิติ (Biometrics)

White Paper ได้กำหนดให้ AI ที่เกี่ยวข้องกับการใช้เทคโนโลยีชีวมิติหรือระบบการจดจำใบหน้า (Biometric & Face Recognition) เป็นเทคโนโลยีในกลุ่มที่มีความเสี่ยงสูง ซึ่งขณะนี้คณะกรรมการการยุโรปอยู่ระหว่างจัดทำข้อเสนอหลักเกณฑ์การใช้เทคโนโลยีการจดจำใบหน้าเป็นการเฉพาะ

- กำหนดให้ AI สามารถแสดงผลได้ต่อเมื่อได้รับการตรวจสอบความถูกต้องของข้อมูลจากมนุษย์แล้วเท่านั้น เช่น การปฏิเสธการรับสวัสดิการประกันสังคม
- AI สามารถแสดงผลได้ในทันที โดยที่มนุษย์สามารถตรวจสอบข้อมูลได้ภายหลัง เช่น การปฏิเสธใบสมัครบัตรเครดิตที่อาจดำเนินการเบื้องต้นได้โดยระบบ AI แต่ต้องได้รับการตรวจสอบโดยเจ้าหน้าที่ในภายหลัง
- มนุษย์สามารถตรวจสอบ/ระงับการแสดงผลหรือการทำงานของ AI ได้ในทันที (real time) เช่น การจัดให้มีปุ่มหยุดและควบคุมรหัสสำหรับรถยนต์ไร้คนขับกรณีพบสิ่งผิดปกติ
- กำหนดความสามารถและข้อจำกัดของ AI ตั้งแต่ขั้นตอนการสร้างระบบ เช่น กำหนดให้รถยนต์ไร้คนขับสามารถหยุดการทำงานในสภาวะที่มีทัศนวิสัยไม่ปลอดภัยหรือการกำหนดขอบเขตการรักษาระยะห่างที่แน่นอนจากรถคันก่อนหน้า เป็นต้น

