ประกาศสำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ

เรื่อง แนวปฏิบัติจริยธรรมด้านปัญญาประดิษฐ์

เพื่อให้การกำกับดูแลงานด้านปัญญาประดิษฐ์ (Artificial Intelligence : AI) ที่ดำเนินการโดยบุคลากรของสำนักงาน ผู้ที่ร่วมวิจัยหรือรับการสนับสนุนการวิจัยจากสำนักงาน ภาคเอกชนที่ใช้พื้นที่ภายในอุทยานวิทยาศาสตร์ประเทศไทย และอาคารอื่น ๆ ของสำนักงาน และผู้รับจ้างช่วงที่เกี่ยวข้อง เป็นไปด้วยความเรียบร้อยและมีประสิทธิภาพ อันจะนำไปสู่ความเชื่อมั่นในงานด้านปัญญาประดิษฐ์ให้มีการดำเนินงานอย่างเหมาะสม

ดังนั้น เพื่อให้สอดคล้องกับนโยบายจริยธรรมด้านปัญญาประดิษฐ์ของสำนักงาน คณะกรรมการจริยธรรมปัญญาประดิษฐ์ของสำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ จึงได้มีมติเห็นชอบให้มีแนวปฏิบัติจริยธรรมด้านปัญญาประดิษฐ์ของสำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ สำนักงานจึงออกประกาศแนวปฏิบัติจริยธรรมด้านปัญญาประดิษฐ์ รายละเอียดปรากฏตามเอกสารแนบ

ทั้งนี้ ตั้งแต่บัดนี้เป็นต้นไป

ประกาศนี้ใช้ลายมือชื่ออิเล็กทรอนิกส์ซึ่งมีผลใช้บังคับได้

ประกาศ ณ วันที่ ๓๑ มีนาคม พ.ศ. ๒๕๖๕

(นายณรงค์ ศิริเลิศวรกุล)

ผู้อำนวยการ

สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ

แนวปฏิบัติจริยธรรมด้านปัญญาประดิษฐ์
ของสำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ

บทที่ 1 บทนำ

ปัญญาประดิษฐ์ (artificial intelligence) เป็นเทคโนโลยีที่ถูกใช้อย่างแพร่หลาย และกำลังได้รับความนิยมอย่างมาก เนื่องด้วยศักยภาพของปัญญาประดิษฐ์ที่สามารถประยุกต์ใช้กับหลากหลายวงการ ขณะเดียวกันหากผู้ที่เกี่ยวข้อง วิจัย ออกแบบ พัฒนา หรือใช้ในทางที่ไม่ถูกต้อง ละเมิดกฎหมายหรือจริยธรรมพื้นฐาน ทั้งที่ตั้งใจหรือรู้เท่าไม่ถึงการณ์ก็ตาม พลังอำนาจของปัญญาประดิษฐ์ รวมถึงวิทยาศาสตร์ข้อมูล (data science) ที่ใช้ปัญญาประดิษฐ์ เช่น การเรียนรู้ด้วยเครื่อง (machine learning) การเรียนรู้เชิงลึก (deep learning) หรือใช้อัลกอริทึมที่ขับเคลื่อนด้วยข้อมูล (data-driven algorithm) ที่พัฒนาอย่างรวดเร็วนี้ ก็อาจนำมาซึ่งภัยร้ายแรงต่อมนุษย์ สังคม และสิ่งแวดล้อมได้เช่นกัน หากไม่มีการกำกับดูแลอย่างเหมาะสม

แนวปฏิบัติจริยธรรมด้านปัญญาประดิษฐ์ ของสำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ (สวทช.) ฉบับนี้ จัดทำขึ้นโดยอ้างอิงจากแนวทางจริยธรรมสำหรับปัญญาประดิษฐ์ที่น่าเชื่อถือ (Ethics guidelines for trustworthy AI) โดยกลุ่มผู้เชี่ยวชาญระดับสูงด้านปัญญาประดิษฐ์ของคณะกรรมาธิการยุโรป (High-Level Expert Group on Artificial Intelligence) และเอกสารแนวปฏิบัติจริยธรรมปัญญาประดิษฐ์ (Thailand AI Ethics Guideline) ของสำนักงานคณะกรรมการดิจิทัลเพื่อเศรษฐกิจและสังคมแห่งชาติ กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม เพื่อใช้เป็นแนวปฏิบัติในการวิจัย ออกแบบ พัฒนา ประยุกต์ใช้ และถ่ายทอดเทคโนโลยีด้านปัญญาประดิษฐ์ และวิทยาศาสตร์ข้อมูลที่ใช้ปัญญาประดิษฐ์หรืออัลกอริทึมที่ขับเคลื่อนด้วยข้อมูล ให้อยู่ภายใต้หลักการด้านจริยธรรมปัญญาประดิษฐ์ที่เป็นสากล คำนึงถึงบริบททางสังคมและกฎหมายที่เกี่ยวข้อง รวมถึงป้องกันและลดผลกระทบหรือความเสียหายที่อาจเกิดขึ้น โดยเป็นหลักปฏิบัติสำหรับบุคลากรของ สวทช. ผู้ที่ร่วมวิจัยหรือรับทุนวิจัยของ สวทช. ภาคเอกชนที่ใช้พื้นที่ภายในอุทยานวิทยาศาสตร์ประเทศไทย และเขตนวัตกรรมระเบียงเศรษฐกิจพิเศษภาคตะวันออก (EECi) รวมถึงผู้รับจ้างช่วงที่เกี่ยวข้อง

ทั้งนี้ เพื่อส่งเสริมให้เกิดการพัฒนาที่รวดเร็ว และไม่เป็นอุปสรรคหรือทำให้เกิดความล่าช้า ข้อปฏิบัติในแนวปฏิบัติฉบับนี้ จึงถูกแบ่งออกเป็นกลุ่ม ๆ ซึ่งมีระดับการบังคับที่แตกต่างกัน เพื่อให้สอดคล้องกับกลุ่มความเสี่ยงของระบบปัญญาประดิษฐ์ และระยะของโครงการ โดยหากเป็นข้อปฏิบัติที่ต้องอาศัยการพิจารณาเลือกแนวทางการดำเนินงาน แนวปฏิบัติฉบับนี้ ได้กำหนดให้นักวิจัย ผู้ช่วยวิจัย หรือผู้ปฏิบัติงานขอรับคำปรึกษาและคำแนะนำจากผู้บังคับบัญชา หรือคณะกรรมการประจำองค์กรที่เกี่ยวข้องก่อน เพื่อเลือกแนวทางการดำเนินงานที่เหมาะสมสำหรับการปฏิบัติงานต่อไป

บทที่ 2 นิยาม

ระบบปัญญาประดิษฐ์ (artificial intelligence system, AI system) หมายถึง ซอฟต์แวร์ที่ถูกพัฒนาขึ้นด้วยเทคนิคในทางปัญญาประดิษฐ์ที่สามารถให้ผลลัพธ์ต่าง ๆ เช่น การคาดการณ์ การให้คำแนะนำ หรือการตัดสินใจ ซึ่งมีอิทธิพลต่อสิ่งแวดล้อมที่เกี่ยวข้อง ทั้งนี้ ต้องเป็นการทำงานเพื่อวัตถุประสงค์ที่มนุษย์ได้กำหนดไว้โดยเทคนิคทางปัญญาประดิษฐ์ ได้แก่ [1]

- การเรียนรู้ด้วยเครื่อง (machine learning) ประกอบด้วย การเรียนรู้แบบมีผู้สอน (supervised learning), การเรียนรู้แบบไม่มีผู้สอน (unsupervised learning) และการเรียนรู้แบบเสริมกำลัง (reinforcement learning) โดยมีการใช้เทคนิคที่หลากหลาย รวมถึงการเรียนรู้เชิงลึก (deep learning)
- เทคนิคที่อาศัยการใช้ตรรกะและองค์ความรู้ ประกอบด้วย การแทนความรู้ (knowledge representation), การโปรแกรมตรรกะเชิงอุปนัย (inductive (logic) programming), ฐานความรู้ (knowledge base), เครื่องอนุมานและการให้เหตุผลเชิงนิรนัย (inference and deductive engine), การให้เหตุผลในเชิงสัญลักษณ์ (symbolic reasoning) และระบบผู้เชี่ยวชาญ (expert system)
- เทคนิคทางสถิติ ประกอบด้วย การประมาณค่าพารามิเตอร์แบบเบส์ (Bayesian estimation) และการค้นหาและการหาค่าที่เหมาะที่สุด (search and optimization methods)
- ศาสตร์หุ่นยนต์ (robotics) ที่จัดเป็นหุ่นยนต์อัจฉริยะ ซึ่งมีการรับรู้, มีตัวตรวจจับ, มีความสามารถในการตัดสินใจ, มีตัวขับเคลื่อนของตัวเองที่ใช้ปัญญาประดิษฐ์เป็นตัวควบคุม หรือมีการประยุกต์เทคนิคอื่น ๆ ร่วมกับระบบไซเบอร์-กายภาพ (cyber-physical system) [2]

ทั้งนี้ ระบบปัญญาประดิษฐ์สามารถแบ่งได้เป็น 3 ประเภท ตามระดับความสามารถ ได้แก่ [3]

1. Artificial narrow intelligence (ANI) หรือ weak AI คือ ปัญญาประดิษฐ์ที่มีระดับความสามารถในการทำงานอยู่ในวงจำกัด ยังไม่มีความสามารถที่ใกล้เคียงกับมนุษย์ โดยปัญญาประดิษฐ์ที่ใช้งานในปัจจุบันทั้งหมดจัดอยู่ในประเภท ANI เช่น ระบบการจดจำใบหน้า (facial recognition), ระบบจดจำเสียงพูด (speech recognition) ระบบการสั่งงานด้วยเสียง (voice assistant) และรถยนต์ไร้คนขับ เป็นต้น

---

[1] DLA Piper. (2021). *The Future Regulation of Technology: EU AI Regulation Handbook*. https://www.dlapiper.com/~/media/files/insights/publications/2021/05/ai-regs-handbook.pdf

[2] High-Level Expert Group on Artificial Intelligence. (2019). *A definition of Artificial Intelligence: main capabilities and scientific disciplines*.

[3] O'Carroll, B. (2017). *What are the 3 types of AI? A guide to narrow, general, and super artificial intelligence*. Retrieved 22 March from https://codebots.com/artificial-intelligence/the-3-types-of-ai-is-the-third-even-possible

2. Artificial general intelligence (AGI) หรือ strong AI คือ ปัญญาประดิษฐ์ที่มีระดับความสามารถเทียบเท่ากับสติปัญญาของมนุษย์ มีความสามารถในการเรียนรู้ แก้ปัญหา คิดวิเคราะห์ เข้าใจ และทำสิ่งใด ๆ ได้เหมือนกับที่มนุษย์ทำได้ ซึ่งในปัจจุบัน นักวิจัยยังไม่สามารถพัฒนา AGI ขึ้นมาได้

3. Artificial super intelligence (ASI) คือ ปัญญาประดิษฐ์ที่มีระดับความสามารถเหนือสติปัญญาของมนุษย์ ในทางทฤษฎี ASI สามารถทำทุกอย่างได้ดีกว่ามนุษย์ เช่น มีความสามารถในการจดจำและวิเคราะห์ข้อมูลได้ดีและรวดเร็วกว่า มีความสามารถในการตัดสินใจ และแก้ปัญหาได้เหนือกว่าที่มนุษย์ทำได้ รวมถึงมีความคิดสร้างสรรค์ทางวิทยาศาสตร์ ภูมิปัญญา และทักษะทางสังคม [4]

ความเสี่ยงของระบบปัญญาประดิษฐ์ หมายถึง ผลกระทบหรือเหตุการณ์ที่ไม่พึงประสงค์ซึ่งอาจเกิดขึ้นจากการพัฒนาและ/หรือการนำปัญญาประดิษฐ์ไปใช้งาน โดยสามารถแบ่งได้เป็น 4 กลุ่ม ได้แก่ [5]

1. กลุ่มที่มีความเสี่ยงที่ไม่สามารถยอมรับได้ (unacceptable risk) เช่น ระบบปัญญาประดิษฐ์ที่เป็นอันตรายคุกคามต่อความปลอดภัย ความเป็นอยู่ หรือ ละเมิดสิทธิขั้นพื้นฐานของประชาชน

2. กลุ่มที่มีความเสี่ยงสูง (high risk) ดังต่อไปนี้ ซึ่งจะต้องปฏิบัติตามข้อปฏิบัติ และหลักการจริยธรรมที่เกี่ยวข้องอย่างเคร่งครัด จึงจะสามารถวางขายในตลาดได้

2.1. เทคโนโลยีที่ใช้ในการระบุตัวตน

2.2. โครงสร้างพื้นฐานที่สำคัญ ซึ่งอาจก่อให้เกิดความเสี่ยงต่อชีวิตและสุขภาพของประชาชน เช่น การคมนาคม

2.3. การศึกษาและการอบรมวิชาชีพที่อาจเป็นตัวกำหนดโอกาสการเข้าถึงการศึกษา เช่น การให้คะแนนสอบ

2.4. การจ้างงาน และการบริหารจัดการทรัพยากรบุคคล เช่น ซอฟต์แวร์ที่ใช้ในการคัดเลือกเอกสารประวัติการทำงาน (CV)

2.5. การเข้าถึงบริการภาคเอกชน และบริการสาธารณะที่จำเป็น เช่น เครื่องมือในการประเมินคะแนนเครดิตสำหรับพิจารณาการให้สินเชื่อ

2.6. การบังคับใช้กฎหมาย ซึ่งอาจละเมิดสิทธิขั้นพื้นฐานของประชาชน เช่น การประเมินความน่าเชื่อถือของหลักฐาน

2.7. การอพยพ ลี้ภัย การจัดการคนเข้าเมืองและการผ่านแดน เช่น การตรวจสอบเอกสารเดินทาง

2.8. กระบวนการยุติธรรม และประชาธิปไตย เช่น เครื่องมือที่ช่วยในการพิจารณาคดี

3. กลุ่มที่มีความเสี่ยงจำกัด (limited risk) เช่น ระบบปัญญาประดิษฐ์ที่ต้องมีความโปร่งใสมากเป็นพิเศษ ตัวอย่างเช่น แชตบอต (chatbot)

---

[4] Bostrom, N. (1998). How long before superintelligence? *International Journal of Futures Studies*, *2*. https://www.nickbostrom.com/superintelligence.html

[5] European Commission. (2021). *Europe fit for the Digital Age: Commission proposes new rules and actions for excellence and trust in Artificial Intelligence*. https://ec.europa.eu/commission/presscorner/detail/en/IP_21_1682

4. กลุ่มที่มีความเสี่ยงต่ำ (minimal risk) เช่น วิดีโอเกมที่ใช้ระบบปัญญาประดิษฐ์ หรือโปรแกรมกรองจดหมายขยะ (spam filter)

วิทยาศาสตร์ข้อมูล (data science) หมายถึง สหสาขาวิชาที่ใช้วิธีการในสาขาวิชาต่าง ๆ เช่น คณิตศาสตร์ สถิติ ปัญญาประดิษฐ์ อัลกอริทึม และวิธีทางวิทยาศาสตร์ ในการสกัดข้อมูลหรือองค์ความรู้สำคัญจากข้อมูลขนาดใหญ่ที่เก็บรวบรวมมา โดยวิทยาศาสตร์ข้อมูลครอบคลุมตั้งแต่การเตรียมข้อมูลสำหรับการวิเคราะห์และประมวลผล (เช่น การทำความสะอาดข้อมูล การรวบรวมและจัดการข้อมูล) การวิเคราะห์ข้อมูลขั้นสูง และการนำเสนอผลและรูปแบบองค์ความรู้ที่ซ่อนอยู่ในข้อมูลดังกล่าว เพื่อนำไปใช้ประโยชน์ต่อไป [6]

ปัญญาประดิษฐ์ที่ไว้วางใจได้ (trustworthy AI) หมายถึง ปัญญาประดิษฐ์ที่ประกอบไปด้วย 3 องค์ประกอบ ที่สามารถพบได้ตลอดวัฏจักรชีวิตของระบบปัญญาประดิษฐ์ [7] ได้แก่

1. ความชอบด้วยกฎหมาย
2. มาตรฐานจริยธรรม
3. ความทนทาน ทั้งจากมุมเทคนิคและสังคม แม้จะเกิดความผิดพลาดก็ไม่ก่อให้เกิดอันตราย หรือความเสียหายใด ๆ โดยไม่ตั้งใจ

ผู้วิจัยปัญญาประดิษฐ์ หมายถึง ผู้ค้นคว้าหาองค์ความรู้ปัญญาประดิษฐ์ใหม่ ๆ หรือเพื่อประยุกต์ใช้องค์ความรู้ปัญญาประดิษฐ์ที่มีอยู่ โดยมีกระบวนการคิดอย่างเป็นระบบ และใช้วิธีการหรือเทคนิคที่ได้รับการยอมรับจากทุกศาสตร์ที่เกี่ยวข้องว่ามีความน่าเชื่อถือ [8]

ผู้ออกแบบปัญญาประดิษฐ์ หมายถึง ผู้มีหน้าที่ศึกษาและวิเคราะห์ความต้องการของผู้ใช้งานปัญญาประดิษฐ์มาจัดทำแบบแผนในการสร้างปัญญาประดิษฐ์ เพื่อให้สามารถใช้งานได้จริงและตรงตามความต้องการของผู้ใช้งานปัญญาประดิษฐ์ [8]

ผู้พัฒนาปัญญาประดิษฐ์ หมายถึง ผู้มีหน้าที่พัฒนาปัญญาประดิษฐ์ตามแบบแผนที่ผู้ออกแบบปัญญาประดิษฐ์ออกแบบไว้ และทำการทดสอบการใช้งานปัญญาประดิษฐ์เพื่อตรวจสอบว่าปัญญาประดิษฐ์สามารถใช้งานได้จริงและตรงตามความต้องการ [8]

ผู้ใช้งานปัญญาประดิษฐ์ หมายถึง บุคคลหรือกลุ่มบุคคลผู้เป็นผู้รับบริการปัญญาประดิษฐ์ โดยมีหน้าที่ระบุปัญหาและความต้องการที่ต้องการให้ปัญญาประดิษฐ์ให้ความช่วยเหลือและแก้ไขปัญหาให้ โดยหมายความรวมถึงประชาชนทุกคนและกลุ่มคนส่วนน้อย เช่น ผู้ด้อยโอกาส ผู้พิการ และผู้ทุพพลภาพ [8]

---

[6] IBM Cloud Education. (2020). *Data Science*. https://www.ibm.com/cloud/learn/data-science-introduction

[7] Commission, E. (2019). *Ethics guidelines for trustworthy AI*. Retrieved 22 March from https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

[8] สุภาภรณ์ เกียรติสิน, มนัสศิริ จันสุทธิรางกูร, ปรัชญ์ สง่างาม, & ยุทธพงศ์ อุณหทวีทรัพย์. (2021). เอกสารแนวปฏิบัติจริยธรรมปัญญาประดิษฐ์ (*Thailand AI Ethics Guideline)*. สำนักงานคณะกรรมการดิจิทัลเพื่อเศรษฐกิจและสังคมแห่งชาติ กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม.

ความสามารถในการสืบย้อน (traceability) หมายถึง ความสามารถในการตรวจสอบย้อนกลับไปตั้งแต่แหล่งที่มาของชุดข้อมูล กระบวนการทำงานและการตัดสินใจของปัญญาประดิษฐ์ เพื่อใช้ในการเฝ้าระวังตรวจสอบความผิดปกติที่พบ และสามารถวินิจฉัยปัญหาที่ทำให้เกิดความผิดพลาดล้มเหลวได้ [8]

สินค้าที่ใช้ได้สองทาง (dual-use item) หมายถึง เทคโนโลยีที่สามารถนำมาใช้ได้ทั้งในทางที่ก่อให้เกิดประโยชน์ และก่อให้เกิดโทษหรืออันตราย เช่น ปัญญาประดิษฐ์ที่สามารถนำมาใช้เพื่อให้เกิดความสงบสุข ในขณะเดียวกันก็สามารถนำมาใช้เพื่อวัตถุประสงค์ด้านปฏิบัติการทางทหารได้เช่นกัน [9]

ระยะเริ่มต้นของการพัฒนาปัญญาประดิษฐ์ หมายถึง ระยะเริ่มต้นของการวิจัย มีการเก็บรวบรวมและประมวลผลข้อมูลเพื่อใช้สอนและพัฒนาปัญญาประดิษฐ์

ระยะทดสอบการใช้งานในวงแคบ หมายถึง ระยะที่เริ่มนำปัญญาประดิษฐ์ที่พัฒนาขึ้น มาทดสอบในสภาพแวดล้อมควบคุม ทั้งนี้ เพื่อประโยชน์ในการปรับปรุงระบบให้มีความถูกต้องมากขึ้นและหลีกเลี่ยงอคติ

ระยะที่นำไปใช้งานในวงกว้าง หมายถึง ระยะที่นำปัญญาประดิษฐ์ที่พัฒนาขึ้น ไปใช้งานจริงเป็นการทั่วไปนอกสภาพแวดล้อมควบคุม โดยอาจเป็นการนำไปใช้งานในเชิงธุรกิจ หรือการให้บริการสาธารณะ

บทที่ 3 หลักการที่เกี่ยวข้องกับจริยธรรมปัญญาประดิษฐ์ (AI Ethics Principles)

นอกจากการดำเนินงานวิจัย ออกแบบ พัฒนา ประยุกต์ใช้ และถ่ายทอดเทคโนโลยีปัญญาประดิษฐ์จะต้องสอดคล้องกับหลักกฎหมายที่เกี่ยวข้องแล้ว การดำเนินงานดังกล่าวยังต้องอยู่บนพื้นฐานของหลักการด้านจริยธรรมปัญญาประดิษฐ์ที่เป็นสากล รวมถึงคำนึงถึงบริบททางสังคมด้วย หลักการด้านจริยธรรมที่ควรคำนึงถึงมี 7 ข้อ ดังนี้

1. ความเป็นส่วนตัว (privacy)
2. ความมั่นคงและปลอดภัย (security and safety)
3. ความไว้วางใจ (reliability)
4. ความเป็นธรรม เท่าเทียม และไม่แบ่งแยก (fairness and non-discrimination)
5. ความโปร่งใสและอธิบายได้ (transparency and explainability)
6. ภาระความรับผิด (accountability)
7. มนุษย์เป็นผู้ควบคุมปัญญาประดิษฐ์ เพื่อความยั่งยืนของมนุษยชาติ (human oversight and human agency)

1. ความเป็นส่วนตัว (privacy)

ปัญญาประดิษฐ์ควรถูกออกแบบ ให้สามารถปกป้องความเป็นส่วนตัว และเคารพต่อสิทธิเสรีภาพของบุคคล การนำข้อมูลส่วนบุคคลของผู้ใดไปใช้งาน รวมถึงการเผยแพร่และใช้ประโยชน์ผลลัพธ์

---

[9] Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., Dafoe, A., Scharre, P., Zeitzoff, T., Filar, B., Anderson, H., Roff, H., Allen, G., Steinhardt, J., Flynn, C., hÉigeartaigh, S., Beard, S., Belfield, H., Farquhar, S., & Amodei, D. (2018). The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation.

จากการประมวลผลและการตัดสินใจของปัญญาประดิษฐ์ ต้องแจ้งให้ผู้ใช้งานปัญญาประดิษฐ์ทราบล่วงหน้าถึงข้อมูลที่จะถูกเก็บรวบรวมและลักษณะการนำข้อมูลดังกล่าวไปใช้ ซึ่งต้องได้รับการยินยอมจากเจ้าของข้อมูลก่อน

การออกแบบและพัฒนาปัญญาประดิษฐ์ควรอยู่บนพื้นฐานของการปกป้องสิทธิเสรีภาพและศักดิ์ศรีของมนุษย์ให้คงอยู่ และต้องเคารพต่อกฎ ระเบียบ หรือกฎหมายภายในประเทศที่เกี่ยวข้องกับการคุ้มครองข้อมูลส่วนบุคคล เช่น พระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562 เป็นต้น นอกจากนี้ ยังรวมถึงการเคารพต่อกฎหมายระหว่างประเทศ เช่น General Data Protection Regulation (GDPR) ด้วย

2. ความมั่นคงและปลอดภัย (security and safety)

ปัญญาประดิษฐ์ควรถูกสร้างให้มีความมั่นคงและปลอดภัย รวมถึงป้องกันภัยอันตรายที่อาจเกิดขึ้นต่อมนุษย์ สิ่งแวดล้อม สังคม เศรษฐกิจ และประเทศ จากการใช้งานปัญญาประดิษฐ์ที่มากจนเกินไป (overused) และที่ไม่ถูกต้องเหมาะสม (misused) โดยการใช้งานและการตัดสินใจของปัญญาประดิษฐ์ควรเกิดจากความตั้งใจของมนุษย์ หรือมีกลไกให้มนุษย์สามารถแทรกแซงการดำเนินการต่าง ๆ เพื่อลดความเสี่ยงจากอันตรายที่อาจเกิดขึ้น เช่น การตัดสินใจที่ผิดพลาด การคุกคามจากผู้ไม่ประสงค์ดี และการนำไปใช้ในทางที่ผิด ซึ่งรวมถึงการใช้เสมือนเป็นอาวุธ (weaponization) และการทำให้เข้าใจผิด หรือให้ข้อมูลที่นำไปสู่การเข้าใจผิด (misinformation)

เนื่องจากในการวิจัย ออกแบบ และพัฒนาปัญญาประดิษฐ์ อาจใช้ชุดข้อมูลที่อ่อนไหว ชุดข้อมูลที่เป็นความลับ หรือชุดข้อมูลที่เป็นข้อมูลส่วนบุคคล ซึ่งหากถูกเข้าถึงโดยผู้ไม่ประสงค์ดีแล้วก็อาจนำมาซึ่งความเสียหายต่อผู้มีส่วนเกี่ยวข้องได้ จึงควรคำนึงถึงหลักการรักษาความมั่นคงปลอดภัยและการคุ้มครองข้อมูลส่วนบุคคลด้วยเช่นกัน

3. ความไว้วางใจ (reliability)

ผู้วิจัย ออกแบบ หรือพัฒนาปัญญาประดิษฐ์จะต้องสามารถสร้างความไว้ใจและความเชื่อมั่นในการใช้งานปัญญาประดิษฐ์ ซึ่งยังไม่ทราบผลกระทบจากการตัดสินใจของระบบที่จะเกิดขึ้นได้แน่ชัดให้แก่สาธารณชนได้ โดยการวิจัยและพัฒนาระบบการวิเคราะห์ ประมวลผล และตัดสินใจของปัญญาประดิษฐ์ให้แม่นยำถูกต้อง สร้างผลลัพธ์ที่เชื่อถือได้ และสร้างผลลัพธ์แบบเดียวกันใหม่ได้ (reproducible) รวมถึงมีการควบคุมคุณภาพของข้อมูลที่นำมาใช้งาน เพื่อป้องกันไม่ให้เกิดการตัดสินใจที่ผิดพลาด ซึ่งอาจส่งผลต่อความน่าเชื่อถือของระบบปัญญาประดิษฐ์ได้

4. ความเป็นธรรม เท่าเทียม และไม่แบ่งแยก (fairness and non-discrimination)

ปัญญาประดิษฐ์ควรถูกออกแบบ และนำไปใช้งานเพื่อส่งเสริมความเป็นธรรม ความเท่าเทียม ความหลากหลาย ความสามัคคี และความยุติธรรมของคนทุกกลุ่มในสังคม โดยที่ไม่เกิดความอคติ หรือความเอนเอียงใด ๆ รวมถึงการให้โอกาสประชาชนทุกคนในสังคม เช่น กลุ่มคนด้อยโอกาส ผู้พิการและผู้ทุพพลภาพ ให้ได้รับประโยชน์จากปัญญาประดิษฐ์ได้อย่างทั่วถึงและเท่าเทียม โดยไม่แบ่งแยกเชื้อชาติ สีผิว และไม่ก่อให้เกิดความเหลื่อมล้ำทางสังคม

5. ความโปร่งใสและอธิบายได้ (transparency and explainability)

ปัญญาประดิษฐ์ควรได้รับการออกแบบและนำไปใช้ โดยให้มนุษย์สามารถรู้ได้ว่ากำลังใช้ปัญญาประดิษฐ์อยู่ เข้าใจได้ว่าข้อมูลถูกนำไปใช้อย่างไร เข้าใจได้ถึงกระบวนการการตัดสินใจ การคาดการณ์การกระทำต่าง ๆ ได้ และกำกับดูแลและตรวจสอบปัญญาประดิษฐ์ได้ โดยควรทำให้มีการแปลผลการดำเนินการของระบบให้เป็นข้อมูลที่สามารถอธิบายและเข้าใจได้โดยมนุษย์ สามารถตรวจสอบย้อนกลับไปยังแหล่งที่มาของชุดข้อมูลที่ได้รับ กระบวนการทำงาน และการตัดสินใจของปัญญาประดิษฐ์ รวมถึงสถานที่ เวลา และวิธีการนำปัญญาประดิษฐ์ไปใช้ เพื่อใช้เฝ้าระวัง ตรวจสอบความผิดปกติ วินิจฉัยปัญหา และหาผู้รับผิดชอบได้อย่างมีประสิทธิภาพ อีกทั้งเพื่อช่วยสร้างความเชื่อมั่นให้แก่ผู้ใช้งานปัญญาประดิษฐ์

การสื่อสารคำอธิบายผลการดำเนินการ ความสามารถ และข้อจำกัดของปัญญาประดิษฐ์ ควรทำอย่างทันการณ์และเหมาะสมกับระดับความเชี่ยวชาญของผู้ที่ต้องการข้อมูลดังกล่าว

6. ภาระความรับผิด (accountability)

การวิจัย ออกแบบ หรือพัฒนาปัญญาประดิษฐ์ ต้องมีกลไกที่ทำให้เกิดความมั่นใจถึงการรับผิดชอบต่อผลกระทบที่เกิดจากปัญญาประดิษฐ์ของผู้ที่มีส่วนร่วมในการวิจัย ออกแบบ พัฒนา และนำไปใช้งาน ซึ่งต้องสามารถตรวจสอบย้อนกลับถึงผู้รับผิดชอบได้โดยชัดเจน รวมถึงมีกลไกแก้ไขปัญหา หรือรับผิดชอบต่อความเสียหายที่อาจเกิดขึ้น ตามภาระหน้าที่ของตนได้อย่างเพียงพอ

อีกทั้งควรตระหนักถึงบทบาทสำคัญของผู้มีส่วนเกี่ยวข้องทุกคนที่มีส่วนร่วมในการออกแบบ พัฒนา และนำไปใช้งาน ซึ่งผู้ที่มีส่วนได้ส่วนเสียเหล่านั้น จะต้องมีการปรึกษาร่วมกันเกี่ยวกับการทำงานของปัญญาประดิษฐ์อย่างเหมาะสม รวมถึงมีการวางแผนถึงการบริหารจัดการความเสี่ยง หรือผลกระทบในระยะยาวที่อาจเกิดขึ้น

7. มนุษย์เป็นผู้ควบคุมปัญญาประดิษฐ์ เพื่อความยั่งยืนของมนุษยชาติ (human oversight and human agency)

ในการออกแบบและการนำปัญญาประดิษฐ์ไปใช้งานนั้น ควรกำหนดให้คำนึงถึงมนุษย์เป็นหลัก (human centric) และคงไว้ซึ่งความสามารถในการควบคุมปัญญาประดิษฐ์และสิทธิให้มนุษย์เป็นผู้ตัดสินใจ ในขั้นตอนการตัดสินใจที่เป็นกระบวนการสำคัญ

นอกจากนี้ ระบบปัญญาประดิษฐ์จะต้องถูกออกแบบ และนำมาใช้เพื่อให้เกิดประโยชน์ต่อมนุษยชาติ ไม่ก่อให้เกิดอันตรายต่อมนุษย์ คำนึงถึงความเป็นอยู่ที่ดี และส่งเสริมคุณค่าของมนุษย์ รวมถึงเป็นการนำปัญญาประดิษฐ์มาใช้เพื่อให้เกิดการพัฒนาที่ยั่งยืน ต่อทั้งสังคมและสิ่งแวดล้อม ตลอดจนทำให้เกิดการพัฒนาในด้านอื่น ๆ ต่อไป

บทที่ 4 แนวปฏิบัติจริยธรรมด้านปัญญาประดิษฐ์

4.1 แนวปฏิบัติสำหรับผู้มีส่วนเกี่ยวข้องในกลุ่มวิชาการ

แนวปฏิบัติสำหรับผู้มีส่วนเกี่ยวข้องในกลุ่มวิชาการ แบ่งออกเป็น 3 ระดับ โดยใช้ความรุนแรงของผลกระทบต่อผู้มีส่วนเกี่ยวข้อง สังคม และสิ่งแวดล้อม เป็นเกณฑ์ โดยแบ่งได้ดังนี้

- **ระดับที่ 1 เป็นแนวปฏิบัติที่ต้องปฏิบัติตาม** ซึ่งเป็นแนวปฏิบัติพื้นฐานที่ไม่สามารถละเว้นได้ เนื่องจากหากไม่ปฏิบัติตามอาจก่อให้เกิดผลกระทบ อันตราย และความเสียหายต่อผู้มีส่วนเกี่ยวข้อง สังคม และสิ่งแวดล้อม ในวงกว้าง และอยู่ในระดับที่ร้ายแรง โดยมีสัญลักษณ์แสดงระดับการปฏิบัติตาม ดังนี้

  ✓✓ หมายถึง กำหนดให้ปฏิบัติตามโดยไม่มีข้อยกเว้น

  ✓ หมายถึง กำหนดให้ปฏิบัติเพื่อเตรียมพร้อมให้กับการวิจัยในขั้นต่อไป

- **ระดับที่ 2 เป็นแนวปฏิบัติที่ควรปฏิบัติตาม** ซึ่งสามารถพิจารณาได้ตามความเหมาะสม ความจำเป็น หรือข้อจำกัดที่มี โดยหากไม่ปฏิบัติตามอาจก่อให้เกิดผลกระทบ อันตราย และความเสียหายต่อผู้มีส่วนเกี่ยวข้อง สังคม และสิ่งแวดล้อมบางส่วน และอยู่ในระดับที่ไม่ร้ายแรง โดยมีสัญลักษณ์แสดงระดับการปฏิบัติตาม ดังนี้

  ✓✓ หมายถึง ควรปฏิบัติเพื่อลดโอกาสการเกิดปัญหา หรือผลกระทบต่าง ๆ ต่อสังคม

  ✓ หมายถึง ข้อแนะนำให้พิจารณาปฏิบัติตาม (highly recommended)

- **ระดับที่ 3 เป็นแนวปฏิบัติที่ดี (best practices)** ที่จะนำไปสู่ความเป็นเลิศตามเป้าหมาย ทำให้เป็นที่ยอมรับในวงการปัญญาประดิษฐ์ โดยมีสัญลักษณ์แสดงระดับการปฏิบัติตาม ดังนี้

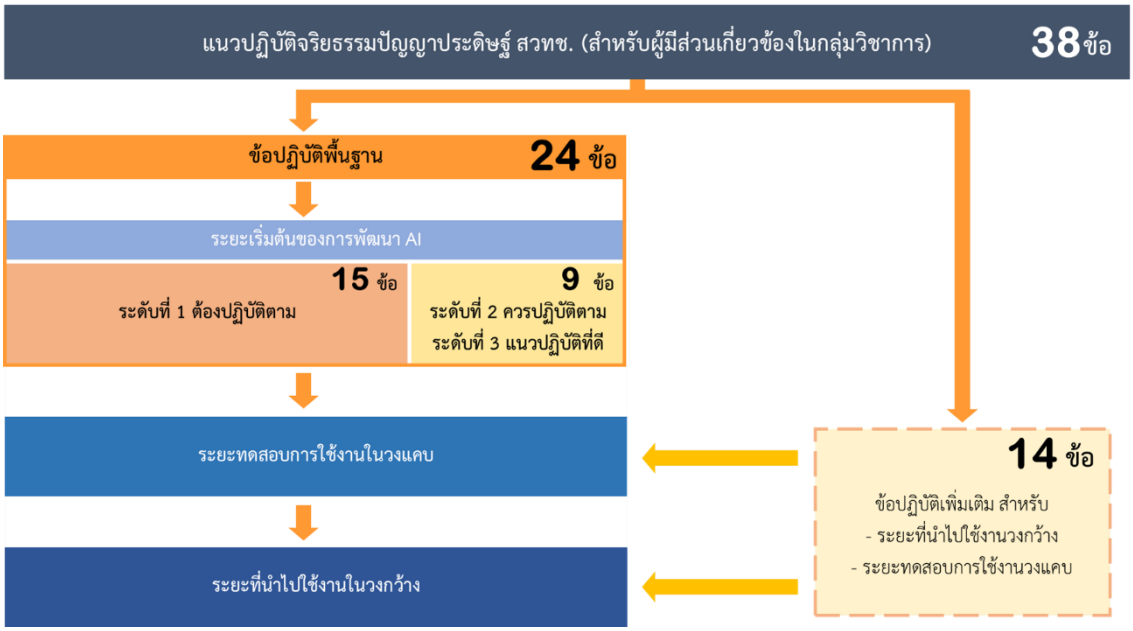
  ✓✓ หมายถึง การปฏิบัติเพื่อสนับสนุนให้มีการดำเนินงานที่สอดคล้องตามมาตรฐานสากล และนำไปสู่การเป็นที่ยอมรับในวงการปัญญาประดิษฐ์

  ✓ หมายถึง การปฏิบัติเพื่อลดโอกาสการถูกตั้งคำถามด้านจริยธรรมปัญญาประดิษฐ์

ตารางที่ 1 แสดงแนวปฏิบัติสำหรับผู้มีส่วนเกี่ยวข้องในกลุ่มวิชาการแบ่งตามหลักการจริยธรรม (ethics domain) และระดับความสำคัญ (priority)

	ระดับความสำคัญ (priority)		
หลักการจริยธรรม (ethics domain)	**ระดับที่ 1 ต้องปฏิบัติตาม**	**ระดับที่ 2 ควรปฏิบัติตาม**	**ระดับที่ 3 แนวปฏิบัติที่ดี**
A. คุณสมบัติผู้วิจัยปัญญาประดิษฐ์ และผู้ร่วมโครงการ	ข้อ 1	ข้อ 2	ข้อ 3
B. ความเป็นส่วนตัว (privacy)	ข้อ 4 ข้อ 5 ข้อ 6	-	-
C. ความมั่นคงและปลอดภัย (security and safety)	ข้อ 7 ข้อ 8 ข้อ 9	ข้อ 10	-
D. ความไว้วางใจ (reliability)	ข้อ 11*	ข้อ 12*	-
E. ความเป็นธรรม เท่าเทียม และไม่แบ่งแยก (fairness and non-discrimination)	ข้อ 13	ข้อ 14 ข้อ 15	ข้อ 16* ข้อ 17*
F. ความโปร่งใสและอธิบายได้ (transparency and explainability)	ข้อ 18 ข้อ 19	ข้อ 20 ข้อ 21*	ข้อ 22*
G. ภาระความรับผิด (accountability)	ข้อ 23 ข้อ 24	ข้อ 25*	ข้อ 26*
H. มนุษย์เป็นผู้ควบคุมปัญญาประดิษฐ์เพื่อความยั่งยืนของมนุษยชาติ (human oversight and human agency)	ข้อ 27 ข้อ 28	ข้อ 29*	ข้อ 30* ข้อ 31
I. การบริหารจัดการเพื่อลดความเสี่ยงและผลกระทบจากการพัฒนาและการนำปัญญาประดิษฐ์ไปใช้	ข้อ 32 ข้อ 33* ข้อ 34*	ข้อ 35 ข้อ 36	ข้อ 37* ข้อ 38*

*ระดับความสำคัญของแนวปฏิบัติมีความแตกต่างกันตามระยะของโครงการ โดยข้อปฏิบัติดังกล่าวไม่บังคับสำหรับโครงการที่อยู่ในระยะเริ่มต้นของการพัฒนาปัญญาประดิษฐ์



ภาพที่ 1 สรุปจำนวนข้อปฏิบัติสำหรับผู้มีส่วนเกี่ยวข้องในกลุ่มวิชาการ ที่ดำเนินโครงการอยู่ในระยะเริ่มต้นของการพัฒนาปัญญาประดิษฐ์[10] ระยะทดสอบการใช้งานในวงแคบ[11] และระยะที่นำไปใช้งานในวงกว้าง[12]

[10] ระยะเริ่มต้นของการพัฒนาปัญญาประดิษฐ์ หมายถึง ระยะเริ่มต้นของการวิจัย มีการเก็บรวบรวมและประมวลผลข้อมูลเพื่อใช้สอนและพัฒนาปัญญาประดิษฐ์

[11] ระยะทดสอบการใช้งานในวงแคบ หมายถึง ระยะที่เริ่มนำปัญญาประดิษฐ์ที่พัฒนาขึ้น มาทดสอบในสภาพแวดล้อมควบคุม ทั้งนี้เพื่อประโยชน์ในการปรับปรุงระบบให้มีความถูกต้องมากขึ้นและหลีกเลี่ยงอคติ

[12] ระยะที่นำไปใช้งานในวงกว้าง หมายถึง ระยะที่นำปัญญาประดิษฐ์ที่พัฒนาขึ้น ไปใช้งานจริงเป็นการทั่วไป นอกสภาพแวดล้อมควบคุม โดยอาจเป็นการนำไปใช้งานในเชิงธุรกิจ หรือการให้บริการสาธารณะ

ตารางที่ 2 แสดงแนวทางการจำแนกโครงการ/กิจกรรมที่เกี่ยวข้องกับปัญญาประดิษฐ์ ตามแนวปฏิบัติจริยธรรมปัญญาประดิษฐ์ สวทช.

เกณฑ์ที่ใช้ในการจำแนก	กิจกรรม
1. โครงการ/กิจกรรมของท่านมีความเกี่ยวข้องกับปัญญาประดิษฐ์หรือเทคโนโลยี ดังต่อไปนี้ หรือไม่ *(พิจารณานิยามของปัญญาประดิษฐ์ และ วิทยาศาสตร์ข้อมูล ในบทที่ 2 นิยาม)*	
☐ ใช่ โปรดระบุ ☐ ปัญญาประดิษฐ์ (artificial intelligence) โปรดระบุ ☐ การเรียนรู้ด้วยเครื่อง (machine learning): supervised learning, unsupervised learning, reinforcement learning และ deep learning ☐ เทคนิคที่อาศัยการใช้ตรรกะและองค์ความรู้: knowledge representation, inductive (logic) programming, knowledge base, inference and deductive engine, symbolic reasoning และ expert system ☐ เทคนิคทางสถิติ: Bayesian estimation และ search and optimization method ☐ หุ่นยนต์ (robotics) ที่จัดเป็นหุ่นยนต์อัจฉริยะ ☐ วิทยาศาสตร์ข้อมูล (data science)	พิจารณาข้อ 2
☐ ไม่ใช่	ไม่เข้าข่ายต้องปฏิบัติตามแนวปฏิบัติฯ ฉบับนี้
2. วัตถุประสงค์ของการวิจัย ออกแบบ และพัฒนาระบบปัญญาประดิษฐ์ หรือเทคโนโลยี ในโครงการ/กิจกรรมของท่าน	
โปรดระบุ _____	พิจารณาข้อ 3
3. ระบบปัญญาประดิษฐ์ของท่าน มีความเสี่ยงอยู่ในระดับใด *(พิจารณากลุ่มความเสี่ยงของระบบปัญญาประดิษฐ์ ในบทที่ 2 นิยาม ความเสี่ยงของระบบปัญญาประดิษฐ์)*	
☐ กลุ่มที่มีความเสี่ยงที่ไม่สามารถยอมรับได้ (unacceptable risk)	ไม่สามารถดำเนินการได้
☐ กลุ่มที่มีความเสี่ยงสูง (high risk) ได้แก่ เทคโนโลยีที่เกี่ยวข้องกับประเด็นดังต่อไปนี้ ☐ การระบุตัวตน ☐ โครงสร้างพื้นฐานที่สำคัญ ซึ่งมีอิทธิพลต่อชีวิตและสุขภาพของประชาชน ☐ การศึกษาและการอบรมวิชาชีพที่อาจกำหนดโอกาสการเข้าถึงการศึกษา ☐ การจ้างงาน และการบริหารจัดการทรัพยากรบุคคล ☐ การเข้าถึงบริการภาคเอกชน และบริการสาธารณะที่จำเป็น ☐ การบังคับใช้กฎหมาย ซึ่งอาจละเมิดสิทธิขั้นพื้นฐานของประชาชน	พิจารณาข้อ 4

เกณฑ์ที่ใช้ในการจำแนก	กิจกรรม
☐ การอพยพ ลี้ภัย การจัดการคนเข้าเมืองและการผ่านแดน ☐ กระบวนการยุติธรรม และประชาธิปไตย	
☐ กลุ่มที่มีความเสี่ยงจำกัด (limited risk) หรือ กลุ่มที่มีความเสี่ยงต่ำ (minimal risk)	พิจารณาข้อ 4
4. ระบบปัญญาประดิษฐ์หรือเทคโนโลยีของท่าน มีการเก็บข้อมูล ในหมวดหมู่ตามธรรมาภิบาลข้อมูลภาครัฐ ดังต่อไปนี้	
☐ ข้อมูลสาธารณะ	พิจารณาข้อ 6
☐ ข้อมูลส่วนบุคคล	พิจารณาข้อ 5
☐ ข้อมูลความมั่นคง	พิจารณาข้อ 6
☐ ข้อมูลความลับทางราชการ	พิจารณาข้อ 6
5. ระบบปัญญาประดิษฐ์หรือเทคโนโลยีของท่าน มีการเก็บข้อมูลส่วนบุคคล ซึ่งเป็นข้อมูลเกี่ยวกับบุคคล ซึ่งทำให้สามารถระบุตัวบุคคลนั้นได้ ไม่ว่าทางตรงหรือทางอ้อม แต่ไม่รวมถึงข้อมูลของผู้ถึงแก่กรรมโดยเฉพาะ[13] ได้แก่ ข้อมูลส่วนบุคคลเกี่ยวกับประเด็นดังต่อไปนี้	
☐ ข้อมูลส่วนบุคคล (personal data) โดยไม่รวมถึงข้อมูลของผู้ถึงแก่กรรมและข้อมูลของนิติบุคคลที่ไม่ใช่ข้อมูลส่วนบุคคลตาม พ.ร.บ. คุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562 ☐ ชื่อ-นามสกุล ☐ เลขประจำตัวประชาชน ☐ ที่อยู่ และ/หรือ เบอร์โทรศัพท์ ☐ วันเกิด และ/หรือ เพศ ☐ การศึกษา และ/หรือ อาชีพ ☐ รูปถ่าย ☐ ข้อมูลทางการเงิน ☐ ข้อมูลส่วนบุคคลที่มีความละเอียดอ่อน (sensitive personal data) ☐ เชื้อชาติ เผ่าพันธุ์ ☐ ความคิดเห็นทางการเมือง ☐ ความเชื่อในลัทธิ ศาสนาหรือปรัชญา ☐ พฤติกรรมทางเพศ ☐ ประวัติอาชญากรรม ☐ ข้อมูลสุขภาพ ความพิการ ☐ ข้อมูลสหภาพแรงงาน	พิจารณาข้อ 6

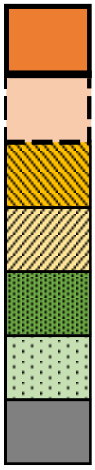
[13] พระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562

เกณฑ์ที่ใช้ในการจำแนก	กิจกรรม
☐ ข้อมูลพันธุกรรม ข้อมูลชีวภาพ ☐ ข้อมูลชีวมาตรอื่น ๆ เช่น ใบหน้า 2 มิติ ลายพิมพ์นิ้วมือ ม่านตา (iris) จอประสาทตา (retina) เป็นต้น ☐ ข้อมูลอื่นใดตามที่คณะกรรมการคุ้มครองข้อมูลส่วนบุคคลประกาศกำหนด	
6. ท่านมีการกำหนดระดับการเปิดเผยข้อมูลที่ใช้ในระบบปัญญาประดิษฐ์ หรือเทคโนโลยีของท่าน อย่างไร	
☐ ข้อมูลสาธารณะ ☐ ข้อมูลเปิดเผยภายใน สวทช. ☐ ข้อมูลที่ต้องได้รับอนุญาตจากศูนย์ ☐ ข้อมูลที่ต้องได้รับอนุญาตจากเจ้าของข้อมูล (กรณีหน่วยงานภายนอก) ☐ ข้อมูลปกปิด	พิจารณาข้อ 7
7. ในปัจจุบัน ระบบปัญญาประดิษฐ์หรือเทคโนโลยีของท่านมี ระดับความพร้อมของเทคโนโลยี (technology readiness level: TRL) อยู่ในระดับใด และท่านมีเป้าหมายอยู่ที่ระดับใด	
ระดับความพร้อมของเทคโนโลยี ในปัจจุบัน คือ _____ ระดับความพร้อมของเทคโนโลยี ตามเป้าหมาย คือ _____	พิจารณาข้อ 8
8. โครงการ/กิจกรรมที่เกี่ยวข้องกับระบบปัญญาประดิษฐ์หรือเทคโนโลยีของท่าน อยู่ในระยะการวิจัยใด	
☐ ระยะเริ่มต้นของการพัฒนาปัญญาประดิษฐ์ : ระยะเริ่มต้นของการวิจัย มีการเก็บรวบรวมและประมวลผลข้อมูลเพื่อใช้สอนและพัฒนาปัญญาประดิษฐ์	ประเมินจริยธรรมปัญญาประดิษฐ์ 24 ข้อ
☐ ระยะทดสอบการใช้งานในวงแคบ: ระยะที่เริ่มนำปัญญาประดิษฐ์ที่พัฒนาขึ้น มาทดสอบในสภาพแวดล้อมควบคุม ทั้งนี้ เพื่อประโยชน์ในการปรับปรุงระบบให้มีความถูกต้องมากขึ้นและหลีกเลี่ยงอคติ	ประเมินจริยธรรมปัญญาประดิษฐ์ 37 ข้อ (กลุ่ม a)
☐ ระยะที่นำไปใช้งานในวงกว้าง: ระยะที่นำปัญญาประดิษฐ์ที่พัฒนาขึ้น ไปใช้งานจริงเป็นการทั่วไปนอกสภาพแวดล้อมควบคุม โดยอาจเป็นการนำไปใช้งานในเชิงธุรกิจ หรือการให้บริการสาธารณะ	ประเมินจริยธรรมปัญญาประดิษฐ์ 37 ข้อ (กลุ่ม b)

ตารางที่ 3 แสดงข้อปฏิบัติสำหรับผู้มีส่วนเกี่ยวข้องในกลุ่มวิชาการ ที่ดำเนินโครงการอยู่ในระยะเริ่มต้นของการพัฒนาปัญญาประดิษฐ์ ระยะทดสอบการใช้งานในวงแคบ และระยะที่นำไปใช้งานในวงกว้าง

		ระยะเริ่มต้นของการพัฒนา AI (24 ข้อ)				ระยะทดสอบการใช้งานในวงแคบ (37 ข้อ กลุ่ม a)				ระยะที่นำไปใช้งานในวงกว้าง (37 ข้อ กลุ่ม b)			
ส่วนที่ 1 คุณสมบัติผู้วิจัย และผู้ร่วมโครงการ	A	A1	1			A1	1			A1	1		
		A2	2			A2	2			A2	2		
		A3	3			A3	3			A3	3		
ส่วนที่ 2 การดำเนินการตามหลักการจริยธรรมปัญญาประดิษฐ์	B	B1	4	5	6	B1	4	5	6	B1	4	5	6
		B2				B2				B2			
		B3				B3				B3			
	C	C1	7	8	9	C1	7	8	9	C1	7	8	9
		C2	10			C2	10			C2	10		
		C3				C3				C3			
	D	D1	11			D1	11			D1	11		
		D2	12			D2	12			D2	12		
		D3				D3				D3			
	E	E1	13			E1	13			E1	13		
		E2	14	15		E2	14	15		E2	14	15	
		E3	16	17		E3	16	17		E3	16	17	
	F	F1	18	19		F1	18	19		F1	18	19	
		F2	20	21		F2	20	21		F2	20	21	
		F3	22			F3	22			F3	22		
	G	G1	23	24		G1	23	24		G1	23	24	
		G2	25			G2	25			G2	25		
		G3	26			G3	26			G3	26		
	H	H1	27	28		H1	27	28		H1	27	28	
		H2	29			H2	29			H2	29		
		H3	30	31		H3	30	31		H3	30	31	
ส่วนที่ 3 การบริหารจัดการเพื่อลดความเสี่ยงและผลกระทบ	I	I1	32	33	34	I1	32	33	34	I1	32	33	34
		I2	35	36		I2	35	36		I2	35	36	
		I3	37	38		I3	37	38		I3	37	38	

วิธีการอ่านตาราง



หมายถึง ระดับที่ 1 เป็นแนวปฏิบัติที่ต้องปฏิบัติตาม โดยไม่มีข้อยกเว้น

หมายถึง ระดับที่ 1 เป็นแนวปฏิบัติที่ต้องปฏิบัติตาม เพื่อเตรียมพร้อมให้กับการวิจัยในขั้นต่อไป

หมายถึง ระดับที่ 2 เป็นแนวปฏิบัติที่ควรปฏิบัติตาม เพื่อลดโอกาสการเกิดปัญหา หรือผลกระทบต่างๆ ต่อสังคม

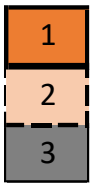
หมายถึง ระดับที่ 2 เป็นแนวปฏิบัติที่ควรปฏิบัติตาม เป็นข้อแนะนำที่ควรคำนึงถึงอย่างมาก (highly recommended)

หมายถึง ระดับที่ 3 เป็นแนวปฏิบัติที่ดี (best practices) เพื่อสนับสนุนให้มีการดำเนินงานที่สอดคล้องตามมาตรฐานสากล และนำไปสู่การเป็นที่ยอมรับในวงการปัญญาประดิษฐ์

หมายถึง ระดับที่ 3 เป็นแนวปฏิบัติที่ดี (best practices) เพื่อลดโอกาสการถูกตั้งคำถามด้านจริยธรรมปัญญาประดิษฐ์

หมายถึง ข้อปฏิบัติที่ไม่บังคับสำหรับโครงการที่อยู่ในระยะดังกล่าว

ตัวอย่างเช่น

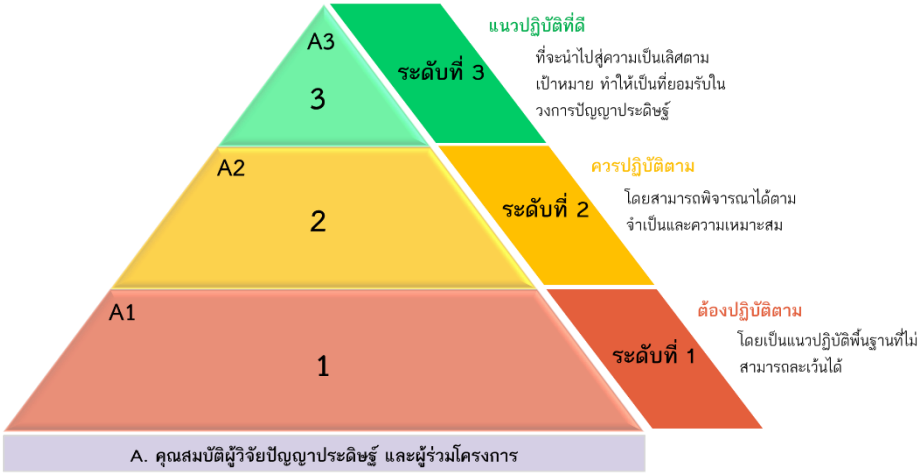


หมายถึง ข้อ 1 อยู่ในระดับที่ 1 เป็นแนวปฏิบัติที่ต้องปฏิบัติตาม โดยไม่มีข้อยกเว้น

หมายถึง ข้อ 2 อยู่ในระดับที่ 1 เป็นแนวปฏิบัติที่ต้องปฏิบัติตาม เพื่อเตรียมพร้อมให้กับการวิจัยในขั้นต่อไป

หมายถึง ข้อ 3 เป็นข้อปฏิบัติที่ไม่บังคับสำหรับโครงการที่อยู่ในระยะดังกล่าว

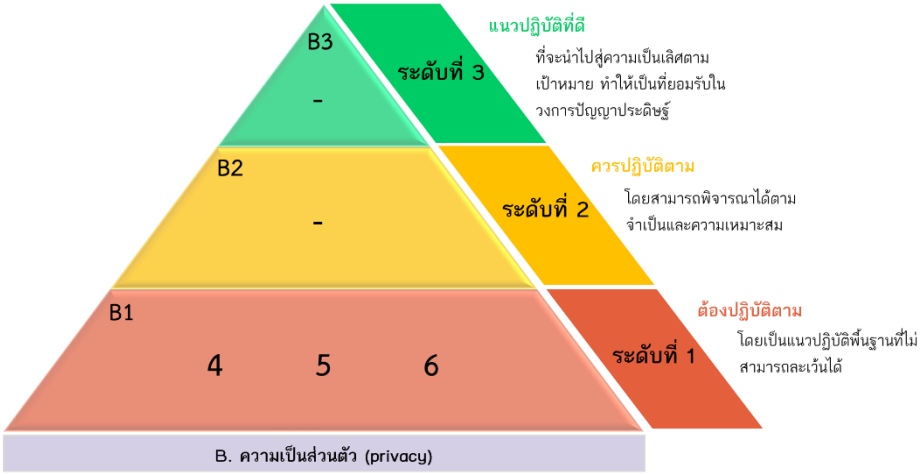
ส่วนที่ 1 คุณสมบัติผู้วิจัยปัญญาประดิษฐ์ และผู้ร่วมโครงการ



กลุ่ม A1	ระยะการวิจัยที่จะนำไปกำกับดูแล	
ข้อ 1 ผู้วิจัย ออกแบบ และพัฒนาปัญญาประดิษฐ์ และผู้ร่วมโครงการต้องศึกษาและทำความเข้าใจ นโยบายจริยธรรมด้านปัญญาประดิษฐ์ของ สวทช. รวมถึง กฎระเบียบ ข้อบังคับอื่น ๆ และกฎหมายที่เกี่ยวข้องกับหลักการจริยธรรมปัญญาประดิษฐ์ พร้อมทั้งปฏิบัติตามอย่างเคร่งครัด	✓✓ ✓ ✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง
กลุ่ม A2	ระยะการวิจัยที่จะนำไปกำกับดูแล	
ข้อ 2 ผู้วิจัย ออกแบบ และพัฒนาปัญญาประดิษฐ์ และผู้ร่วมโครงการควรมีความรู้ ทักษะ และความสามารถในด้านปัญญาประดิษฐ์เป็นอย่างดี มีประสบการณ์ที่เพียงพอ และสามารถปฏิบัติงานให้สอดคล้องกับหลักการจริยธรรมปัญญาประดิษฐ์ได้	✓✓ ✓ ✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง
กลุ่ม A3	ระยะการวิจัยที่จะนำไปกำกับดูแล	
ข้อ 3 ผู้วิจัย ออกแบบ และพัฒนาปัญญาประดิษฐ์ และผู้ร่วมโครงการสามารถขอรับคำปรึกษาจากผู้เชี่ยวชาญ หรือคณะกรรมการจริยธรรมปัญญาประดิษฐ์ สวทช. เพื่อให้ระบบปัญญาประดิษฐ์ที่พัฒนาขึ้นมีความถูกต้องสมบูรณ์ สอดคล้องตามหลักการจริยธรรมปัญญาประดิษฐ์ที่ดี	✓✓ ✓✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง

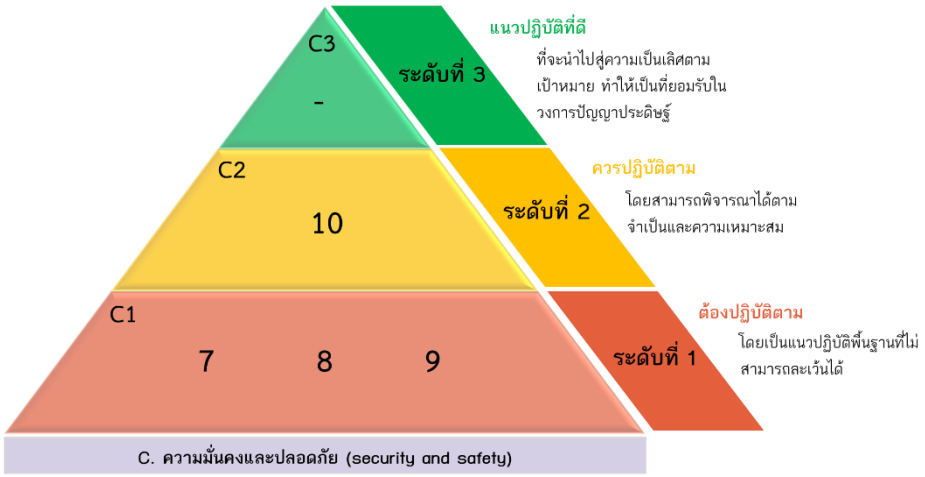
ส่วนที่ 2 การดำเนินการตามหลักการจริยธรรมปัญญาประดิษฐ์

<u>B. ความเป็นส่วนตัว (privacy)</u>



<u>กลุ่ม B1</u>	ระยะการวิจัยที่จะนำไปกำกับดูแล	
ข้อ 4 ต้องมีการประเมินประเภทและขอบเขตของข้อมูลที่ใช้ในระบบปัญญาประดิษฐ์ ว่ามีข้อมูลชนิดใดบ้าง มีข้อมูลส่วนบุคคลหรือไม่ และพิจารณาวิธีการพัฒนาหรือสอนระบบปัญญาประดิษฐ์ โดยใช้ข้อมูลที่อ่อนไหว หรือข้อมูลส่วนบุคคลให้น้อยที่สุด	✓ ✓✓ ✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง
ข้อ 5 ต้องปฏิบัติตาม พ.ร.บ. คุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562 โดยมีกระบวนการสำหรับตรวจสอบ และควบคุมข้อมูลส่วนบุคคล เช่น การขอความยินยอม (consent) จากเจ้าของข้อมูล และกลไกการขอระงับ การจำกัดการลบการใช้ข้อมูล และการเพิกถอนข้อมูลส่วนบุคคลในภายหลัง เพื่อป้องกันการละเมิดข้อมูลส่วนบุคคลโดยตั้งใจ และไม่ตั้งใจ	✓ ✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง
ข้อ 6 ต้องมีระบบการบริหารจัดการและกำกับดูแลข้อมูลที่สอดคล้องกับกฎหมายและมาตรฐานสากลต่าง ๆ ที่เกี่ยวข้อง เช่น มาตรฐาน ISO หรือ IEEE เป็นต้น มีการควบคุมคุณภาพของข้อมูล และจัดทำธรรมาภิบาลข้อมูล (data governance) ของโครงการ ที่ครอบคลุมตลอดวัฏจักรชีวิตของข้อมูล ตั้งแต่การเก็บรวบรวมข้อมูล การใช้ข้อมูล การถ่ายโอนข้อมูล การคุ้มครองข้อมูล โดยมีกระบวนการเพิ่มการรักษาความเป็นส่วนตัวของข้อมูล เช่น การเข้ารหัสข้อมูล (encryption) การทำให้ไม่สามารถบ่งบอกถึงตัวตนเจ้าของข้อมูลได้ (anonymization) หรือ การผสมข้อมูล (aggregation) เป็นต้น จนถึงขั้นตอนการทำลายข้อมูล รวมถึงมีการตรวจสอบความปลอดภัยของข้อมูลว่าไม่ถูกเปลี่ยนแปลงโดยไม่ได้รับอนุญาต	✓ ✓✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง

C. ความมั่นคงและปลอดภัย (security and safety)

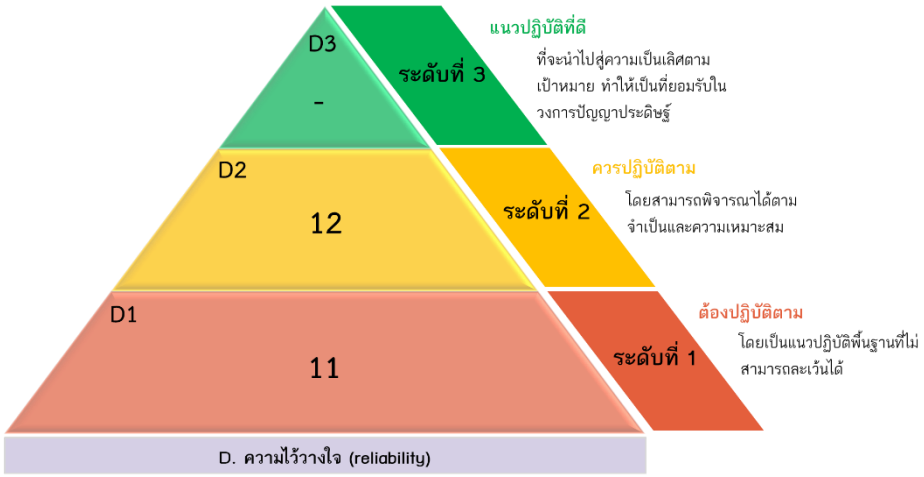


กลุ่ม C1	ระยะการวิจัยที่จะนำไปกำกับดูแล	
ข้อ 7 ต้องปฏิบัติตามนโยบายและแนวปฏิบัติในการรักษาความมั่นคงปลอดภัยด้านสารสนเทศตามที่ สวทช. กำหนด รวมถึงปฏิบัติตามมาตรการการรักษาความปลอดภัยของข้อมูล ตามประกาศกระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม โดยมีกลไกการรักษาความมั่นคงปลอดภัยของระบบปัญญาประดิษฐ์เพื่อป้องกันความเสี่ยง ผลกระทบ และอันตรายที่อาจเกิดขึ้นจากภัยคุกคามหรือการใช้งานปัญญาประดิษฐ์ในทางที่ผิด ซึ่งอาจส่งผลกระทบต่อองค์กร ข้อมูลส่วนบุคคล จริยธรรม มนุษย์ สังคมและสิ่งแวดล้อม	✓✓ ✓✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง
ข้อ 8 ต้องมีกลไกการตรวจสอบ เฝ้าระวัง แจ้งเตือนเหตุการณ์ภัยคุกคามด้านความมั่นคงปลอดภัย และการละเมิดความเป็นส่วนตัวในโครงสร้างพื้นฐานของระบบปัญญาประดิษฐ์ รวมถึงมีกลไกการติดตามแก้ไขปัญหาเมื่อระบบถูกโจมตี โดยต้องมีแผนสำรอง (fallback plan) ที่เพียงพอและเหมาะสม เช่น การปรับเปลี่ยนขั้นตอนของระบบ การร้องขอให้มนุษย์เป็นผู้ควบคุมก่อนที่ระบบจะดำเนินการใด ๆ เป็นต้น รวมถึงระบบปัญญาประดิษฐ์จะต้องมีความสามารถกลับคืนสู่สภาวะปกติได้ภายหลังจากการถูกโจมตี (resilience to attack)	✓✓ ✓✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง
ข้อ 9 ต้องมีการตรวจสอบการทำงานของระบบปัญญาประดิษฐ์ ว่ามีการทำงานและให้ผลลัพธ์อย่างไร เมื่อเกิดเหตุการณ์หรืออยู่ในสภาพแวดล้อมที่ไม่คาดคิด มีการประเมินผลกระทบที่อาจเกิดขึ้น รวมถึงต้องมีวิธีการแก้ไขและบรรเทาปัญหาที่เหมาะสม สอดคล้องกับมาตรฐานสากลที่เกี่ยวข้อง	✓ ✓✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง

ตัวอย่างเช่น ในยานยนต์ไร้คนขับจะต้องมีระบบป้องกันภัยและนำแผนสำรองมาใช้ (fail-safe mechanism) เช่น การส่งผ่านการควบคุมให้กับผู้ขับ การหยุดในช่องทางเดินรถอย่างปลอดภัย และการเคลื่อนที่ออกจากช่องทางเดินรถและหยุดอย่างปลอดภัย เป็นต้น [14]		
<u>กลุ่ม C2</u>	ระยะการวิจัยที่จะนำไปกำกับดูแล	
ข้อ 10 ควรพิจารณาระดับความเสี่ยงของระบบปัญญาประดิษฐ์ที่สามารถเป็นสินค้าที่ใช้ได้สองทาง (dual-use item) พร้อมกำหนดแนวทางการป้องกันการนำปัญญาประดิษฐ์ไปใช้ในทางที่ไม่พึงประสงค์ เช่น การไม่เผยแพร่ผลงานวิจัย การหลีกเลี่ยงการใช้ปัญญาประดิษฐ์ดังกล่าว หรือการเร่งเปิดเผยความเสี่ยงที่อาจเกิดจากปัญญาประดิษฐ์บางชนิด	✓ ✓ ✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง

[14] U.S. Department of Transportation. (2018). *A Framework for Automated Driving System Testable Cases and Scenarios*. National Highway Traffic Safety Administration (NHTSA). https://www.nhtsa.gov/sites/nhtsa.gov/files/documents/13882-automateddrivingsystems_092618_v1a_tag.pdf

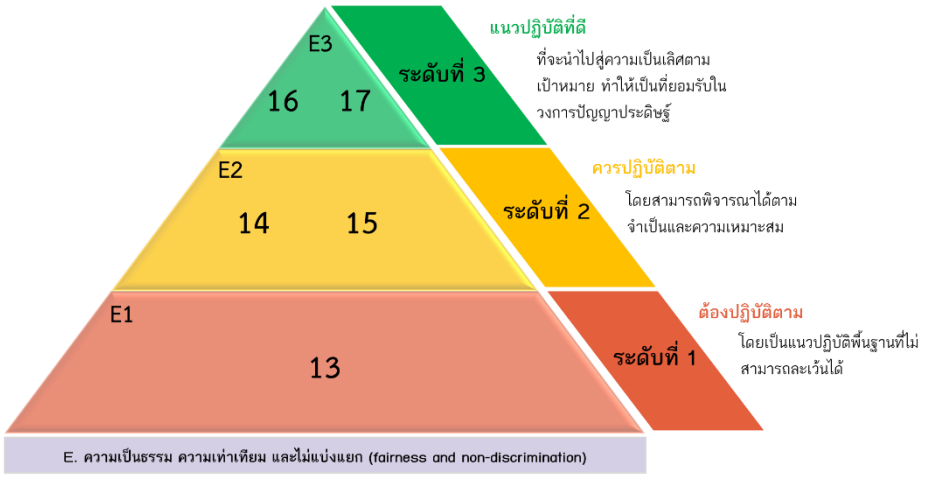
D. ความไว้วางใจ (reliability)



กลุ่ม D1	ระยะการวิจัยที่จะนำไปกำกับดูแล	
ข้อ 11 นักวิจัยจะต้องแจ้งให้ผู้ใช้งานทราบถึงระดับความสามารถและความแม่นยำของปัญญาประดิษฐ์ว่าอยู่ในระดับใด และมีข้อจำกัด หรือความเสี่ยงจากการใช้งานอย่างไร เพื่อให้ระบบปัญญาประดิษฐ์สามารถใช้งานได้ (usability) มีความสมบูรณ์ และแม่นยำสูงสุด ก่อนการเปิดใช้งาน รวมถึงช่วยให้ผู้ใช้งานสามารถเข้าใจการทำงานของระบบปัญญาประดิษฐ์ (understandability) และคาดการณ์ถึงผลกระทบที่อาจเกิดขึ้นได้ (predictability) สำหรับเกณฑ์การระบุความสามารถและความแม่นยำของปัญญาประดิษฐ์ ให้พิจารณาดังนี้ - กรณีที่อยู่ในระยะที่มีการทดสอบปัญญาประดิษฐ์กับคนในสังคม หรือมีการนำไปใช้งานในวงกว้าง นักวิจัยจะต้องอ้างอิงเกณฑ์การจำแนกระดับความสามารถและความแม่นยำที่มีความน่าเชื่อถือตามมาตรฐานสากลหรือมาตรฐานอุตสาหกรรมที่เกี่ยวข้องกับปัญญาประดิษฐ์นั้น ๆ เช่น กรณีของยานยนต์ไร้คนขับ (autonomous vehicle) สมาคมวิศวกรรมยานยนต์นานาชาติ (SAE International) ได้กำหนดระดับความสามารถของระบบขับเคลื่อนอัตโนมัติ (levels of driving automation) ไว้ 6 ระดับ ตั้งแต่ระดับ 0 คือผู้ขับต้องเป็นผู้ควบคุมรถทั้งหมด จนถึงระดับ 5 คือ ระบบขับเคลื่อนเป็นแบบอัตโนมัติโดยสมบูรณ์ ไม่ต้องการการควบคุมใด ๆ จากผู้ขับ เป็นต้น ทั้งนี้ นักวิจัยสามารถศึกษาตัวอย่างเกณฑ์มาตรฐานสากลที่ใช้ในการจำแนกระดับความสามารถและความแม่นยำของเทคโนโลยีปัญญาประดิษฐ์ต่าง ๆ ได้ ในภาคผนวกของแนวปฏิบัติฯ ฉบับนี้	- ✓✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง

- กรณีที่อยู่ในระยะทดสอบการใช้งานปัญญาประดิษฐ์ในวงแคบ นักวิจัยสามารถกำหนดเกณฑ์ที่ใช้ในการจำแนกและระบุระดับความสามารถและความแม่นยำของปัญญาประดิษฐ์ให้สอดคล้อง เหมาะสมกับงานวิจัยของตนเองได้		
<u>กลุ่ม D2</u>	ระยะการวิจัยที่จะนำไปกำกับดูแล	
ข้อ 12ควรมีวิธีการติดตาม และพิสูจน์ได้ว่า ระบบปัญญาประดิษฐ์สามารถบรรลุเป้าหมาย วัตถุประสงค์ และมีการนำไปใช้งานได้ตรงตามที่ผู้วิจัยและพัฒนาปัญญาประดิษฐ์ตั้งใจไว้ และมีหลักเกณฑ์ที่ใช้ในการตัดสินว่าระบบปัญญาประดิษฐ์สามารถให้ผลลัพธ์ที่ถูกต้องอย่างแท้จริง สามารถแสดงได้ว่ามีการใช้ข้อมูลที่ครอบคลุม และเป็นปัจจุบัน หรือ สามารถประเมินได้ว่าเมื่อใดที่ระบบต้องการข้อมูลเพิ่มเติม เพื่อให้ผลลัพธ์มีความถูกต้องมากขึ้น หรือทำให้อคติลดลง	- ✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง

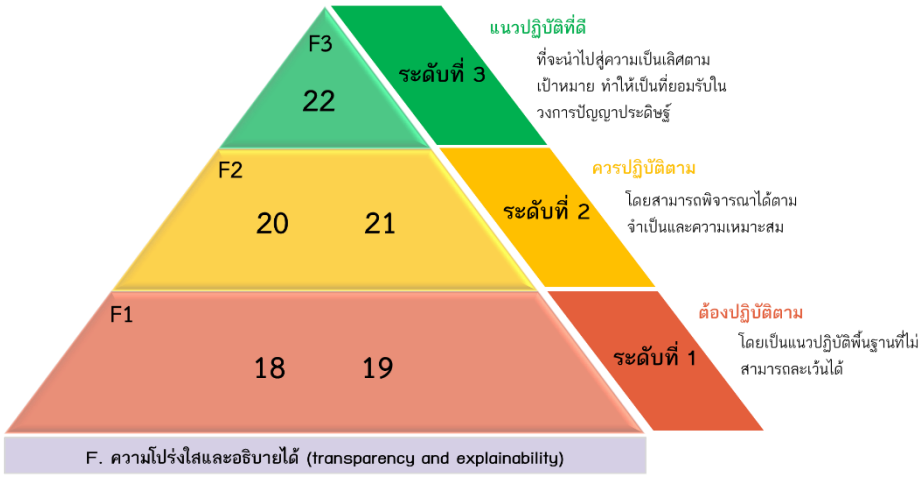
E. ความเป็นธรรม เท่าเทียม และไม่แบ่งแยก (fairness and non-discrimination)



กลุ่ม E1	ระยะการวิจัยที่จะนำไปกำกับดูแล	
ข้อ 13 ต้องมีวิธีหรือกลไกการหลีกเลี่ยงการสร้างความไม่เท่าเทียมและอคติในระบบปัญญาประดิษฐ์ โดยคำนึงถึงข้อมูลที่ถูกนำเข้าไปในระบบและการออกแบบอัลกอริทึม เช่น การใช้ชุดข้อมูลในการสอน และทดสอบที่มีความแตกต่างกัน เป็นข้อมูลที่มีคุณภาพ เชื่อถือได้ สามารถเป็นตัวแทนของประชากรที่ต้องการนำไปใช้ได้ เพื่อให้สามารถตรวจพบอคติของผลลัพธ์ที่อาจเกิดขึ้น และสามารถใช้ทดสอบกับข้อมูลใหม่ (unseen data) ได้	✓ ✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง
กลุ่ม E2	ระยะการวิจัยที่จะนำไปกำกับดูแล	
ข้อ 14 ควรมีความหลากหลายของข้อมูลที่เลือกนำมาใช้สอนอัลกอริทึม ซึ่งรวมถึง ความหลากหลายของประวัติภูมิหลังของผู้เข้าร่วมการทดสอบ ผู้ให้ข้อมูล และนักวิจัยในทีมพัฒนาระบบปัญญาประดิษฐ์ เพื่อช่วยลดความเสี่ยงในการเกิดความไม่เป็นธรรมขึ้นในระบบ และควรให้ผู้ใช้งานทุกกลุ่มได้ร่วมทดสอบระบบปัญญาประดิษฐ์ด้วย	✓ ✓✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง
ข้อ 15 ควรออกแบบและพัฒนาปัญญาประดิษฐ์ให้สามารถเข้าถึงได้ มีการออกแบบที่สามารถรองรับผู้ใช้งานได้ทุกกลุ่ม (universal design) และมีทางเลือกที่หลากหลายให้กับผู้ใช้งานในการดำเนินการเพื่อบรรลุถึงเป้าหมาย	✓ ✓✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง
กลุ่ม E3	ระยะการวิจัยที่จะนำไปกำกับดูแล	
ข้อ 16 ควรมีการจัดทำเอกสารเพื่อใช้แสดงข้อมูลการออกแบบ ขั้นตอนการทำงาน และแนวทางการนำระบบปัญญาประดิษฐ์ หรือองค์ความรู้ไปใช้ในสถานการณ์ที่แตกต่างกัน เพื่อรองรับผู้ใช้งานที่อาจมีพื้นฐานความรู้และความเข้าใจที่แตกต่างกัน	- ✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง

ข้อ 17 ควรพิจารณาถึงการมีส่วนร่วมของผู้มีส่วนเกี่ยวข้อง ในการพัฒนาและการใช้ระบบปัญญาประดิษฐ์	- ✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง

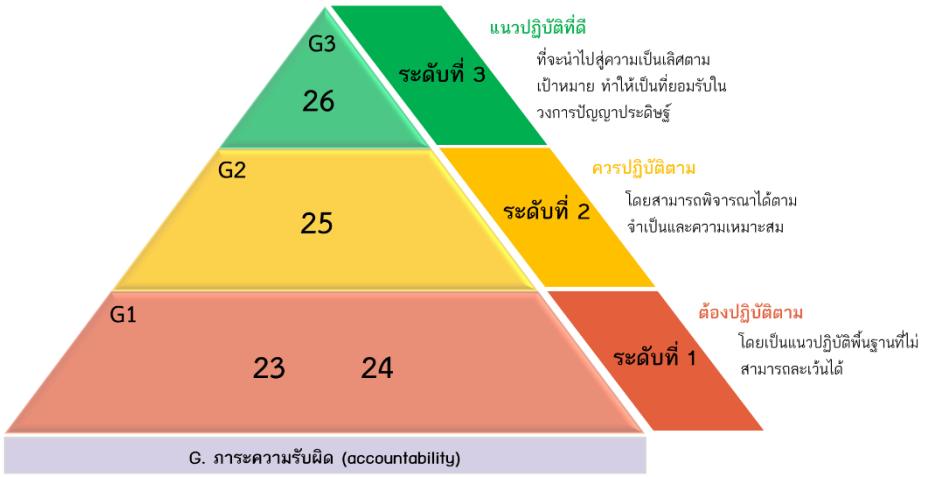
F. ความโปร่งใสและอธิบายได้ (transparency and explainability)



กลุ่ม F1	ระยะการวิจัยที่จะนำไปกำกับดูแล	
ข้อ 18 ต้องออกแบบให้ปัญญาประดิษฐ์มีความสามารถในการสืบย้อนกลับ (traceability) เพื่อใช้ในการเฝ้าระวัง ตรวจสอบความผิดปกติ และวินิจฉัยปัญหาที่เกิดขึ้น (diagnosability) ได้ โดยแนวทางการดำเนินงานเพื่อสร้างความสามารถในการสืบย้อนกลับ อ้างอิงจากแนวปฏิบัติจริยธรรมปัญญาประดิษฐ์ (Thailand AI Ethics Guideline) ประกอบด้วย • การบันทึกข้อมูลและกิจกรรมที่เกิดขึ้นระหว่างการวิจัย การพัฒนา การออกแบบ และการให้บริการ ตามลำดับเวลาของเหตุการณ์ (audit trail) • ข้อมูลชุดการสอนปัญญาประดิษฐ์ วิธีการในการเก็บรวบรวมและปรับแก้ การเคลื่อนย้ายข้อมูล ผลการตรวจวัดความแม่นยำของปัญญาประดิษฐ์ตลอดช่วงเวลาที่ดำเนินการ • โมเดลที่ใช้ออกแบบและอัลกอริทึมที่เลือกใช้ • การเปลี่ยนแปลงโปรแกรม (code) และผู้ที่ทำการเปลี่ยนแปลง • การบันทึกกระแสข้อมูลที่ไหลเข้าระบบทั้งหมดในช่วงเวลาที่มีการใช้งานระบบปัญญาประดิษฐ์ • จัดเก็บข้อมูลเพื่อการสืบย้อนในหน่วยจัดเก็บข้อมูลที่เหมาะสม เพื่อหลีกเลี่ยงการถูกลดทอนคุณภาพ หรือถูกเปลี่ยนแปลงแก้ไข และควรจัดเก็บไว้ตามระยะเวลาที่สอดคล้องกับกฎหมาย หรือข้อกำหนดในมาตรฐานอุตสาหกรรมที่นำระบบปัญญาประดิษฐ์ไปใช้งาน	✓ ✓✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง

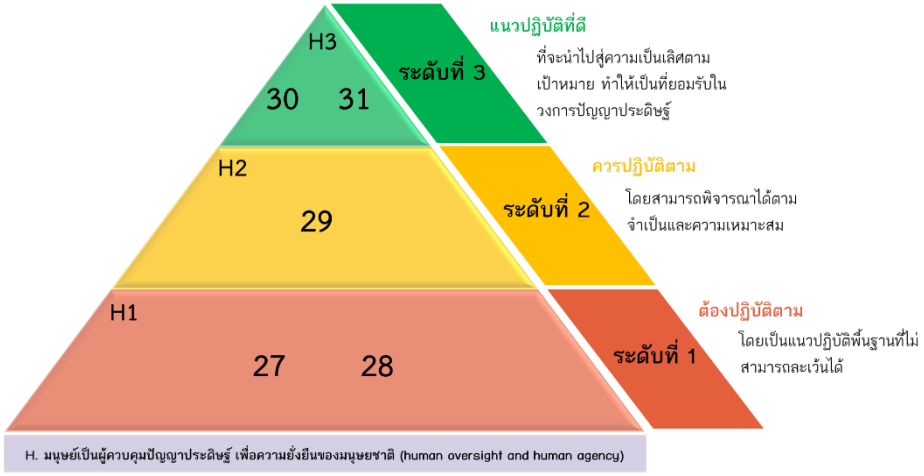
ข้อ 19 ต้องระบุวัตถุประสงค์ของระบบปัญญาประดิษฐ์ ผู้ที่จะได้รับประโยชน์จากปัญญาประดิษฐ์ คุณสมบัติ ข้อจำกัด รวมถึงข้อบกพร่องที่อาจเกิดขึ้น ของระบบปัญญาประดิษฐ์ให้ผู้ใช้งานทราบอย่างชัดเจน โดยใช้ภาษาและคำอธิบายที่สามารถสื่อสารให้ผู้ใช้งานทั่วไปเข้าใจได้	✓ ✓✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง
<u>กลุ่ม F2</u>	ระยะการวิจัยที่จะนำไปกำกับดูแล	
ข้อ 20 ควรเก็บรักษาหลักฐาน หรือเอกสารที่แสดงรายละเอียดการทบทวน วิเคราะห์ ดำเนินการแก้ไขข้อผิดพลาด และข้อยกเว้นต่าง ๆ ที่เกิดขึ้นในระบบปัญญาประดิษฐ์	✓ ✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง
ข้อ 21 ควรมีการกำหนดช่องทางการสื่อสารที่เข้าถึงได้ง่ายและรวดเร็ว สำหรับใช้รับผลสะท้อนกลับ (feedback) และเรื่องร้องขอให้ทบทวนการตัดสินใจที่เกิดจากระบบ (decision review) หรือความเสี่ยงต่าง ๆ ที่อาจเกิดขึ้น จากผู้ที่มีส่วนเกี่ยวข้องต่อระบบปัญญาประดิษฐ์ เพื่อให้มีการควบคุมปัญญาประดิษฐ์ทั้งระบบอย่างมีประสิทธิภาพ และมีกลไกการทบทวนการตัดสินใจที่เป็นธรรม และรวดเร็ว ในขณะเดียวกัน นักวิจัยก็ควรรายงานข้อผิดพลาดที่เกิดขึ้นในระบบปัญญาประดิษฐ์ ให้ผู้มีส่วนเกี่ยวข้องทราบถึงสถานการณ์ที่กำลังเกิดขึ้น ด้วยความโปร่งใส และตรงตามความเป็นจริง พร้อมหาสาเหตุเพื่อดำเนินการแก้ไขอย่างทันท่วงที	- ✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง
<u>กลุ่ม F3</u>	ระยะการวิจัยที่จะนำไปกำกับดูแล	
ข้อ 22 ควรมีกลไกสำหรับการให้ข้อมูลผู้ใช้งานปัญญาประดิษฐ์เกี่ยวกับสาเหตุและหลักเกณฑ์ที่อยู่เบื้องหลังการทำงานและผลลัพธ์ของระบบปัญญาประดิษฐ์ โดยเฉพาะเมื่อระบบปัญญาประดิษฐ์นั้น มีผลกระทบต่อมนุษย์อย่างมีนัยสำคัญ รวมถึงมีคำอธิบายที่สามารถทำให้ผู้ใช้งานปัญญาประดิษฐ์เข้าใจถึงผลลัพธ์ของระบบปัญญาประดิษฐ์ ซึ่งอาจเป็นการอธิบายที่ไม่ซับซ้อนที่ผู้ใช้งานสามารถเข้าถึงได้โดยง่าย และมีความน่าเชื่อถือได้ก็เพียงพอ ตัวอย่างเช่น ในกรณีที่เป็นการสร้างโมเดลของปัญญาประดิษฐ์ขึ้นมาใหม่ อาจใช้วิธีการให้ข้อมูลรายละเอียดของระบบ ที่คล้ายกับการเขียนบทความวิชาการ หรือ ในกรณีที่มีการปรับพารามิเตอร์บางอย่าง จากปัญญาประดิษฐ์ที่ผู้อื่นสร้างขึ้น ซึ่งส่งผลต่อประสิทธิภาพในระบบนิเวศของผู้ใช้งาน อาจใช้การอธิบายหน้าที่ของพารามิเตอร์นั้น ๆ เป็นต้น	- ✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง

G. ภาระความรับผิด (accountability)



กลุ่ม G1	ระยะการวิจัยที่จะนำไปกำกับดูแล	
ข้อ 23 ต้องมีการกำหนดบทบาท หน้าที่และความรับผิดชอบของผู้มีส่วนร่วมในโครงการทุกคนที่เกี่ยวข้องกับปัญญาประดิษฐ์ ซึ่งอยู่ในขอบเขตภาระหน้าที่ของตนเอง เพื่อสร้างความมั่นใจในการรับผิดชอบต่อผลกระทบที่อาจเกิดจากปัญญาประดิษฐ์ให้แก่ผู้ใช้งาน	✓✓ ✓✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง
ข้อ 24 ต้องมีการกำกับดูแลระบบปัญญาประดิษฐ์ให้สามารถตรวจสอบได้ (auditability) โดยนำผลหรือข้อมูลจากการสืบย้อนกลับของระบบปัญญาประดิษฐ์ (traceability) และกลไกการเก็บข้อมูลอื่น ๆ ที่ระบบไม่สามารถทำได้ เช่น การจัดทำเอกสาร การบันทึกข้อมูลการดำเนินงานและผลลัพธ์ของระบบโดยมนุษย์ มาใช้สนับสนุนกระบวนการการตรวจสอบดังกล่าว	✓ ✓✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง
กลุ่ม G2	ระยะการวิจัยที่จะนำไปกำกับดูแล	
ข้อ 25 ควรมีกระบวนการประเมินความเสี่ยงของระบบปัญญาประดิษฐ์ที่พิจารณาถึงผู้มีส่วนเกี่ยวข้องทั้งทางตรง และทางอ้อม เปิดโอกาสให้ผู้มีส่วนเกี่ยวข้องดังกล่าว สามารถรายงาน ข้อบกพร่อง ความเสี่ยง หรือ อคติที่อาจเกิดขึ้นในระบบปัญญาประดิษฐ์ได้	- ✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง
กลุ่ม G3	ระยะการวิจัยที่จะนำไปกำกับดูแล	
ข้อ 26 เมื่อผู้ใช้เกิดอันตราย หรือได้รับผลกระทบ ความเสียหายจากระบบปัญญาประดิษฐ์ เจ้าของระบบปัญญาประดิษฐ์ควรมีกลไกการรับผิดชอบต่อผลกระทบที่เกิดขึ้นกับผู้ใช้งาน ทั้งนี้ หน่วยงานต้นสังกัดควรให้ความสำคัญ และกำกับดูแลให้มีการรับผิดชอบที่เหมาะสมด้วย	- ✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง

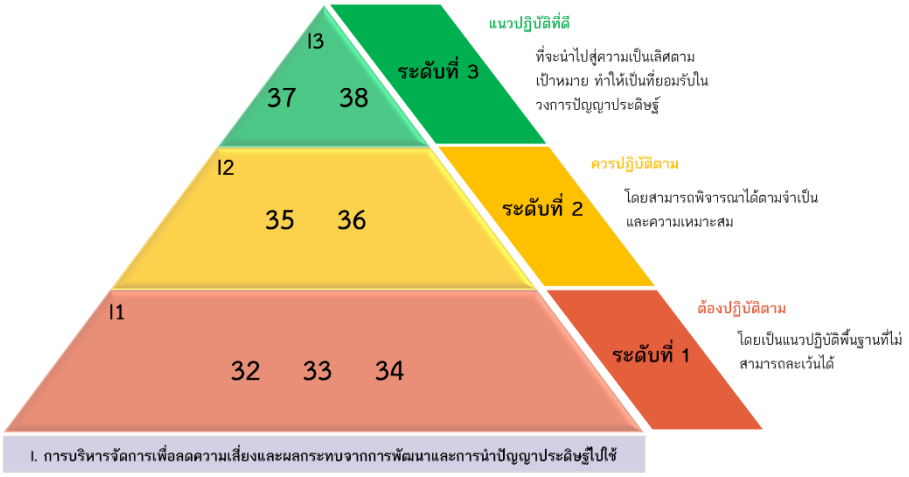
H. มนุษย์เป็นผู้ควบคุมปัญญาประดิษฐ์ เพื่อความยั่งยืนของมนุษยชาติ (human oversight and human agency)



กลุ่ม H1	ระยะการวิจัยที่จะนำไปกำกับดูแล	
ข้อ 27 กรณีที่ยังอยู่ในช่วงวิจัยปัญญาประดิษฐ์ เช่น อยู่ในระยะเริ่มต้นของการพัฒนา หรือระยะทดสอบการใช้งานปัญญาประดิษฐ์ในวงแคบ ให้ใช้หลักการการออกแบบและการนำปัญญาประดิษฐ์ไปใช้งาน โดยให้คำนึงถึงมนุษย์เป็นหลัก (human centric) คงไว้ซึ่งสิทธิการเป็นผู้เลือกตัดสินใจโดยมนุษย์ หรือสามารถส่งต่ออำนาจการควบคุมและตัดสินใจไปยังมนุษย์ได้ รวมถึงคำนึงถึงประโยชน์ที่จะเกิดขึ้นกับมนุษย์ และสิ่งแวดล้อมอย่างยั่งยืน	✓✓ ✓✓ -	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง
ข้อ 28 ในกรณีที่เป็นระบบปัญญาประดิษฐ์แบบที่เรียนรู้ด้วยตัวเองหรือทำงานด้วยตัวเอง ต้องมีวิธีการตรวจสอบและการตอบสนองที่สามารถระบุสิ่งผิดปกติที่เกิดขึ้นได้ รวมทั้งควรมีขั้นตอนการหยุดการทำงานของระบบปัญญาประดิษฐ์อย่างปลอดภัย และสามารถระบุได้ว่า เป็นการหยุดการทำงานทั้งกระบวนการ หรือเป็นการส่งต่ออำนาจการควบคุมให้กับมนุษย์	✓✓ ✓✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง
กลุ่ม H2	ระยะการวิจัยที่จะนำไปกำกับดูแล	
ข้อ 29 ควรออกแบบปัญญาประดิษฐ์ให้มีกลไกในการอนุญาตให้มนุษย์สามารถเข้าแทรกแซงกระบวนการตัดสินใจของปัญญาประดิษฐ์ได้ รวมถึงการมีทางเลือกที่ให้มนุษย์ร่วมตัดสินใจในผลการตัดสินใจที่ไม่สามารถย้อนกลับได้ หรือยากที่จะย้อนกลับ หรือมีผลการทำลายล้าง หรือผลที่ไม่พึงประสงค์ หรือเกี่ยวข้องกับความเป็นความตาย	- ✓✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง

กลุ่ม H3	ระยะการวิจัยที่จะนำไปกำกับดูแล	
ข้อ 30 ระบบปัญญาประดิษฐ์ควรเป็นมิตรต่อสิ่งแวดล้อม และสร้างความยั่งยืนให้สังคม โดยลดผลกระทบทางสิ่งแวดล้อมที่อาจเกิดขึ้นจากวัฏจักรชีวิตของระบบปัญญาประดิษฐ์	- ✓✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง
ข้อ 31 ควรคำนึงถึงผลกระทบทางสังคมที่อาจเกิดขึ้นจากระบบปัญญาประดิษฐ์ เช่น ความเสี่ยงที่จะทำให้คนตกงาน หรือ ปัญหาการลดทักษะของแรงงาน	✓ ✓✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง

ส่วนที่ 3 การบริหารจัดการเพื่อลดความเสี่ยงและผลกระทบจากการพัฒนาและการนำปัญญาประดิษฐ์ไปใช้



กลุ่ม I1	ระยะการวิจัยที่จะนำไปกำกับดูแล	
ข้อ 32 ต้องมีการวิเคราะห์ ประเมิน และบริหารจัดการความเสี่ยงของผลกระทบหรือความเสียหายที่อาจเกิดขึ้นจากปัญหาด้านระบบเครือข่าย การประมวลผล การตัดสินใจที่ผิดพลาดของปัญญาประดิษฐ์ หรือพฤติกรรมที่ไม่ตั้งใจของระบบปัญญาประดิษฐ์ เช่น การกำหนดแนวทางในการจัดการความเสี่ยงที่ไม่สามารถยอมรับได้ การสื่อสารผลการวิเคราะห์และประเมินความเสี่ยงให้ผู้มีส่วนเกี่ยวข้องได้รับทราบ การวิเคราะห์เหตุการณ์ความเสี่ยงด้านจริยธรรมปัญญาประดิษฐ์ที่เกิดขึ้นเพื่อค้นหาสาเหตุของปัญหา	✓ ✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง
ข้อ 33 ต้องมีการเฝ้าระวัง บันทึก ตรวจสอบการทำงานของปัญญาประดิษฐ์ และดำเนินการแก้ไขเหตุการณ์ละเมิดด้านจริยธรรมปัญญาประดิษฐ์ รวมถึงกู้คืนระบบปัญญาประดิษฐ์ และจัดทำเอกสารกระบวนการกู้คืนและนำระบบกลับสู่สถานะเดิม ทั้งนี้ นักวิจัยจะต้องมีการทำความเข้าใจผลลัพธ์และความสามารถของปัญญาประดิษฐ์ เพื่อให้เกิดการพัฒนาปรับปรุงปัญญาประดิษฐ์ให้มีความน่าเชื่อถือได้อย่างต่อเนื่องสอดคล้องกับหลักจริยธรรมและทันต่อการเปลี่ยนแปลงที่รวดเร็วของเทคโนโลยีอยู่เสมอ	- ✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง
ข้อ 34 ในกรณีที่มีการโอนความรับผิดชอบการให้บริการระบบปัญญาประดิษฐ์ไปยังผู้ให้บริการรายอื่น จะต้องมั่นใจได้ว่า ผู้ให้บริการรายใหม่นั้นจะดูแลผู้ใช้งานระบบปัญญาประดิษฐ์ได้อย่างสมบูรณ์ และสอดคล้องตามแนวปฏิบัตินี้	- ✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง

กลุ่ม I2	ระยะการวิจัยที่จะนำไปกำกับดูแล	
ข้อ 35 ควรพิจารณาจัดทำแผนการทดสอบเพื่อตรวจสอบความมั่นคงและปลอดภัย ความเท่าเทียม ความหลากหลาย ความครอบคลุม ความเป็นธรรม และความน่าเชื่อถือของระบบ รวมถึงความสามารถในการสร้างผลลัพธ์แบบเดียวกัน (reproducibility) ของปัญญาประดิษฐ์ โดยการสร้างสภาพแวดล้อมและวิธีการที่ใช้ทดสอบปัญญาประดิษฐ์ให้มีความใกล้เคียงกับสภาพแวดล้อมจริง พร้อมบันทึกผลการทดสอบเก็บไว้เป็นข้อมูลเพื่อใช้สื่อสารกับผู้มีส่วนเกี่ยวข้องได้	✓ ✓ ✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง
ข้อ 36 ควรมีกระบวนการติดตามการนำปัญญาประดิษฐ์ไปใช้งาน และประเมินว่าปัญญาประดิษฐ์สามารถทำงานได้อย่างสอดคล้องกับจริยธรรม ปัญญาประดิษฐ์ภายใต้เงื่อนไขการปฏิบัติงานจริง รวมถึงควรมีการทบทวนความเสี่ยงด้านจริยธรรมในการนำปัญญาประดิษฐ์ไปใช้งานอยู่เสมอ	✓ ✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง
กลุ่ม I3	ระยะการวิจัยที่จะนำไปกำกับดูแล	
ข้อ 37 ควรระบุข้อกำหนดที่เกี่ยวข้องกับการนำปัญญาประดิษฐ์ไปใช้งานอย่างมีจริยธรรมที่ดี เช่น ความมั่นคงและปลอดภัย ความเป็นส่วนตัว ความเท่าเทียม ความหลากหลาย ความครอบคลุม ความเป็นธรรม และความน่าเชื่อถือในการประมวลผลข้อมูลของปัญญาประดิษฐ์ ในสัญญาจ้างผู้ให้บริการภายนอก (outsource) หรือ เมื่อผู้วิจัย ออกแบบ และพัฒนาปัญญาประดิษฐ์นำผลงานไปใช้กับภาคธุรกิจหรือภาคปกครอง	- - ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง
ข้อ 38 ควรสนับสนุนให้ผู้ใช้งานสามารถใช้งานระบบปัญญาประดิษฐ์ได้อย่างปลอดภัย สอดคล้องกับกฎหมาย และหลักการจริยธรรมปัญญาประดิษฐ์ที่เกี่ยวข้อง ดังนี้ • ส่งเสริมให้ผู้ใช้งานปัญญาประดิษฐ์ทราบถึงประโยชน์ ผลกระทบ และความเสี่ยงจากการใช้งานปัญญาประดิษฐ์ ตลอดจนวิธีการใช้งานและการทำงานร่วมกับปัญญาประดิษฐ์ เพื่อป้องกันการนำปัญญาประดิษฐ์ไปใช้ในทางที่ไม่พึงประสงค์ (สอดคล้องตามข้อปฏิบัติที่ 11, 19 และ 32) • ส่งเสริมผู้ใช้งานปัญญาประดิษฐ์ศึกษาและทราบถึงภาระความรับผิดชอบของตนเองที่จะเกิดขึ้น เมื่อใช้งานปัญญาประดิษฐ์ให้เข้าใจเพื่อให้ผู้ใช้งานสามารถวางแผนแนวทางการจำกัดขอบเขตของผลกระทบจากการใช้งานปัญญาประดิษฐ์ที่อาจเกิดขึ้นได้ (สอดคล้องตามข้อปฏิบัติที่ 19) • ส่งเสริมให้ผู้ใช้งานปัญญาประดิษฐ์ศึกษาทำความเข้าใจหลักการทำงานของปัญญาประดิษฐ์ พร้อมทั้งมีแนวทางตรวจสอบ	- ✓ ✓✓	ระยะเริ่มต้นของการพัฒนา AI ระยะทดสอบการใช้งานในวงแคบ ระยะที่นำไปใช้งานในวงกว้าง

ประเมินความน่าเชื่อถือของการตัดสินใจ/การวิเคราะห์ผลของปัญญาประดิษฐ์ (สอดคล้องตามข้อปฏิบัติที่ 11 และ 22) • ส่งเสริมให้ผู้ใช้งานปัญญาประดิษฐ์สามารถแจ้งปัญหาและส่งผลสะท้อนกลับ (feedback) ให้ผู้ควบคุม และผู้พัฒนาระบบปัญญาประดิษฐ์ทราบ เมื่อพบว่าปัญญาประดิษฐ์มีการทำงานที่ไม่ถูกต้อง หรือไม่สอดคล้องตามหลักจริยธรรมปัญญาประดิษฐ์ เพื่อนำไปสู่การปรับปรุงแก้ไขปัญญาประดิษฐ์ให้มีความถูกต้องและเหมาะสมต่อไป (สอดคล้องตามข้อปฏิบัติที่ 21 และ 25) • ส่งเสริมให้ผู้ใช้งานปัญญาประดิษฐ์ศึกษา ทำความเข้าใจ และปฏิบัติตามกฎหมายที่เกี่ยวข้องกับหลักการจริยธรรมปัญญาประดิษฐ์ รวมถึงนโยบาย กฎ ระเบียบ และข้อบังคับอื่น ๆ ที่เกี่ยวข้องขององค์กร เช่น นโยบายจริยธรรมด้านปัญญาประดิษฐ์ ของ สวทช. (สอดคล้องตามข้อปฏิบัติที่ 1, 5 และ 6)		

4.2 การกำหนดความรับผิดชอบในการปฏิบัติตามแนวปฏิบัติจริยธรรมด้านปัญญาประดิษฐ์

การกำหนดความรับผิดชอบในการปฏิบัติตามแนวปฏิบัติจริยธรรมด้านปัญญาประดิษฐ์ เพื่อใช้เป็นแนวทางในการนำแนวปฏิบัติฯ ไปใช้ดำเนินการ ซึ่งมีองค์ประกอบของผู้รับผิดชอบ และบทบาทหน้าที่ และความรับผิดชอบ ดังแสดงในตารางที่ 4

ตารางที่ 4 แสดงผู้รับผิดชอบและบทบาทหน้าที่ของผู้รับผิดชอบตามแนวปฏิบัติจริยธรรมด้านปัญญาประดิษฐ์

ลำดับ	ผู้รับผิดชอบ	บทบาท หน้าที่ และความรับผิดชอบ
1	คณะกรรมการจริยธรรมปัญญาประดิษฐ์ สวทช.	1. ให้คำปรึกษา คำแนะนำ และข้อเสนอแนะแก่นักวิจัย ผู้ช่วยวิจัย และผู้ที่มีส่วนเกี่ยวข้อง เกี่ยวกับการดำเนินการที่เกี่ยวข้องกับจริยธรรมด้านปัญญาประดิษฐ์ ที่ดำเนินการภายใน สวทช. และภายในพื้นที่ที่ สวทช. เป็นผู้ดูแลรับผิดชอบ ทั้งภายในอุทยานวิทยาศาสตร์ประเทศไทย เขตอุตสาหกรรมซอฟต์แวร์ประเทศไทย (Software Park Thailand) และเขตนวัตกรรมระเบียงเศรษฐกิจพิเศษภาคตะวันออก (EECi) ฯลฯ 2. บริหารความเสี่ยงด้านการละเมิดด้านจริยธรรมของโครงการที่เกี่ยวข้องกับปัญญาประดิษฐ์ 3. ประสานงานกับคณะกรรมการพัฒนา ส่งเสริม และสนับสนุนจริยธรรมการวิจัยในมนุษย์ สวทช. ในกรณีที่มีความจำเป็นหรือเกี่ยวข้อง 4. รับและพิจารณาเรื่องร้องเรียน ข้อขัดแย้ง หรือประเด็นต่าง ๆ ที่เกี่ยวข้องกับจริยธรรมด้านปัญญาประดิษฐ์ของ สวทช. 5. ทบทวนนโยบาย และแนวปฏิบัติฯ ตลอดจนให้ข้อเสนอแนะต่อผู้บริหารอย่างน้อยปีละ 1 ครั้ง 6. เสนอแต่งตั้งคณะทำงานหรือผู้ทรงคุณวุฒิเพิ่มเติม เพื่อช่วยปฏิบัติงานได้ตามความเหมาะสม
2	คณะทำงานที่เกี่ยวข้อง หรือผู้ทรงคุณวุฒิ (ถ้ามี)	ปฏิบัติงานตามที่คณะกรรมการจริยธรรมปัญญาประดิษฐ์ สวทช. มอบหมาย
3	ผู้อำนวยการศูนย์/ผู้อำนวยการหน่วยเฉพาะทาง (ของนักวิจัย)	กำกับดูแลการปฏิบัติงานของผู้ใต้บังคับบัญชาให้เป็นไปตามนโยบายและแนวปฏิบัติจริยธรรมด้านปัญญาประดิษฐ์ ของ สวทช.
4	ผู้บังคับบัญชา (ของนักวิจัย) หรือหัวหน้ากลุ่มโครงการวิจัย	1. ให้คำปรึกษา คำแนะนำ ด้านจริยธรรมปัญญาประดิษฐ์แก่นักวิจัย ผู้ช่วยวิจัย หรือผู้ปฏิบัติงานที่อยู่ในการกำกับดูแล

ลำดับ	ผู้รับผิดชอบ	บทบาท หน้าที่ และความรับผิดชอบ
		2. ติดตาม ตรวจสอบ และพิจารณากิจกรรมด้านจริยธรรมปัญญาประดิษฐ์ของโครงการที่อยู่ในการกำกับดูแล เพื่อความเหมาะสมและสอดคล้องกับแนวปฏิบัติฯ ที่กำหนด 3. รายงานผลการประเมินกิจกรรมด้านจริยธรรมปัญญาประดิษฐ์ของโครงการที่อยู่ในการกำกับดูแล ที่ตรวจสอบพบความเสี่ยงที่อาจก่อให้เกิดการละเมิดด้านจริยธรรม ให้คณะกรรมการจริยธรรมปัญญาประดิษฐ์ของ สวทช. ทราบ 4. แสดงความคิดเห็น ข้อเสนอแนะ และรายงานประเด็นต่าง ๆ ที่เกี่ยวข้อง ต่อคณะกรรมการจริยธรรมปัญญาประดิษฐ์ของ สวทช. เพื่อนำไปสู่การพัฒนาการดำเนินงานด้านปัญญาประดิษฐ์
5	หัวหน้าโครงการ (ผู้วิจัย ผู้ออกแบบ และผู้พัฒนาปัญญาประดิษฐ์)	1. ขอรับคำปรึกษาด้านจริยธรรมปัญญาประดิษฐ์ในกรณีต่าง ๆ จากผู้บังคับบัญชา หรือคณะกรรมการประจำองค์กรที่เกี่ยวข้อง ตัวอย่างเช่น กรณีที่แนวปฏิบัติฯ มีขั้นตอนที่ต้องอาศัยการพิจารณาเลือกแนวทางดำเนินงานที่เหมาะสม 2. ประเมินกิจกรรมด้านจริยธรรมปัญญาประดิษฐ์ ของโครงการที่ตนเองรับผิดชอบ ตามหัวข้อที่กำหนดในแนวปฏิบัติสำหรับผู้มีส่วนเกี่ยวของในกลุ่มวิชาการ (บทที่ 4) 3. รายงานผลการประเมินกิจกรรมให้หัวหน้ากลุ่มโครงการวิจัย ทราบ ทั้งนี้ ในกรณีที่ตรวจสอบพบความเสี่ยงที่อาจก่อให้เกิดการละเมิดด้านจริยธรรม จะต้องรายงานให้คณะกรรมการจริยธรรมปัญญาประดิษฐ์ของ สวทช. ทราบด้วย เพื่อคณะกรรมการฯ จะพิจารณาการบริหารความเสี่ยง และให้คำแนะนำแนวทางการดำเนินงานต่อไป 4. รับและพิจารณานำผลสะท้อนกลับ (feedback) หรือความคิดเห็นจากผู้ที่มีส่วนเกี่ยวข้องต่อปัญญาประดิษฐ์ มาพัฒนาปรับปรุงระบบ 5. แสดงความคิดเห็น ข้อเสนอแนะ และรายงานประเด็นต่าง ๆ ที่เกี่ยวข้อง ต่อคณะกรรมการจริยธรรมปัญญาประดิษฐ์ของ สวทช. เพื่อนำไปสู่การพัฒนาการดำเนินงานด้านปัญญาประดิษฐ์

บทที่ 5 กรณีศึกษา ที่เกี่ยวข้องกับจริยธรรมด้านปัญญาประดิษฐ์

สืบเนื่องจากในปัจจุบัน เทคโนโลยีปัญญาประดิษฐ์เข้ามามีบทบาทในชีวิตประจำวันของมนุษย์มากขึ้น ทั้งในด้านการแพทย์ อุตสาหกรรม เศรษฐกิจ หรือการใช้เป็นเครื่องมือในการรักษาความมั่นคงปลอดภัยของข้อมูลขององค์กรต่าง ๆ เช่น ธนาคาร สายการบิน กองทัพ เป็นต้น อย่างไรก็ตาม การพัฒนาของเทคโนโลยีปัญญาประดิษฐ์ที่รวดเร็วก็อาจนำมาซึ่งความกังวลว่า ความก้าวหน้าของเทคโนโลยีปัญญาประดิษฐ์นี้อาจส่งผลกระทบต่อมนุษย์โดยเฉพาะในด้านจริยธรรม หากหน่วยงานที่กำกับดูแลไม่มีการกำหนดแนวปฏิบัติด้านจริยธรรม หรือวิธีการควบคุมปัญญาประดิษฐ์ที่เหมาะสม ผู้พัฒนาปัญญาประดิษฐ์ไม่คำนึงถึงมาตรฐานด้านจริยธรรม หรือผู้ใช้งานปัญญาประดิษฐ์ไม่ตระหนักถึงความเสี่ยงจากการใช้งาน รวมถึงการนำปัญญาประดิษฐ์ไปใช้ในทางที่ผิด ดังกรณีศึกษาที่เกี่ยวข้องกับจริยธรรมปัญญาประดิษฐ์ เรื่องการรู้จำใบหน้า (face recognition) และดีปเฟก (deepfake) ที่นำไปสู่ประเด็นถกเถียงโต้แย้งด้านจริยธรรมและด้านกฎหมายที่เกี่ยวข้อง

ตัวอย่างที่ 1 การรู้จำใบหน้า (face recognition)

ระบบรู้จำใบหน้า คือ ระบบตรวจหาใบหน้าของมนุษย์และจัดจำแนกใบหน้าดังกล่าวเพื่อประโยชน์ตามที่ได้ออกแบบไว้ โดยทั่วไประบบรู้จำใบหน้าจะประกอบด้วย 2 ขั้นตอน คือ การตรวจจับใบหน้า (face detection) ซึ่งเป็นการระบุตำแหน่งของใบหน้าในภาพหรือในภาพเคลื่อนไหว และการรู้จำใบหน้า (face recognition) ซึ่งเป็นการระบุว่าใบหน้าดังกล่าวตรงกับบุคคลใดหรือบุคคลประเภทใด เทคโนโลยีรู้จำใบหน้านั้นถูกใช้งานแพร่หลายสำหรับงานหลายแบบ เช่น การยืนยันอัตลักษณ์ของบุคคล ซึ่งอาจเป็นส่วนหนึ่งของการพิสูจน์สิทธิการเข้าถึงพื้นที่ควบคุมในอาคารหรือสิทธิการทำธุรกรรมทางการเงิน มีความพยายามจะใช้เทคโนโลยีดังกล่าวในงานการยุติธรรมโดยหน่วยงานบังคับใช้กฎหมาย แต่ก็ยังมีข้อถกเถียงถึงความเหมาะสมอยู่

ในรายงานวิชาการ Gender shades: Intersectional accuracy disparities in commercial gender classification ของ Joy Buolamwini และ Timnit Gebru [15] แสดงให้เห็นถึงการทำงานของปัญญาประดิษฐ์ที่มีความไม่เท่าเทียมกันต่อเชื้อชาติและเพศ ทำให้ผลการวิเคราะห์การจดจำใบหน้าผิดพลาด โดยนักวิจัยได้นำภาพใบหน้าคนมาวิเคราะห์ด้วยโปรแกรมของ IBM, Microsoft และ Face++ ผลการทดสอบพบว่าความผิดพลาดเกิดขึ้นกับใบหน้าของหญิงผิวสีเข้มมากที่สุด โดยมีอัตราความผิดพลาดคิดเป็น 1 ใน 3 ขณะที่การวิเคราะห์ใบหน้าของชายผิวสีอ่อนพบความผิดพลาดเป็นสัดส่วนน้อยมาก คือเพียงร้อยละ 0.8

[15] Buolamwini, J., & Gebru, T. (2018). *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification* Proceedings of the 1st Conference on Fairness, Accountability and Transparency, Proceedings of Machine Learning Research. http://proceedings.mlr.press

Classifier	Metric	All	F	M	Darker	Lighter	DF	DM	LF	LM
MSFT	PPV(%)	93.7	89.3	97.4	87.1	99.3	79.2	94.0	98.3	**100**
	Error Rate(%)	6.3	10.7	2.6	12.9	0.7	**20.8**	6.0	1.7	0.0
	TPR (%)	93.7	96.5	91.7	87.1	99.3	92.1	83.7	**100**	98.7
	FPR (%)	6.3	8.3	3.5	12.9	0.7	**16.3**	7.9	1.3	0.0
Face++	PPV(%)	90.0	78.7	99.3	83.5	95.3	65.5	**99.3**	94.0	99.2
	Error Rate(%)	10.0	21.3	0.7	16.5	4.7	**34.5**	0.7	6.0	0.8
	TPR (%)	90.0	98.9	85.1	83.5	95.3	98.8	76.6	**98.9**	92.9
	FPR (%)	10.0	14.9	1.1	16.5	4.7	**23.4**	1.2	7.1	1.1
IBM	PPV(%)	87.9	79.7	94.4	77.6	96.8	65.3	88.0	92.9	**99.7**
	Error Rate(%)	12.1	20.3	5.6	22.4	3.2	**34.7**	12.0	7.1	0.3
	TPR (%)	87.9	92.1	85.2	77.6	96.8	82.3	74.8	**99.6**	94.8
	FPR (%)	12.1	14.8	7.9	22.4	3.2	**25.2**	17.7	5.20	0.4

Table 4: Gender classification performance as measured by the positive predictive value (PPV), error rate (1-PPV), true positive rate (TPR), and false positive rate (FPR) of the 3 evaluated commercial classifiers on the PPB dataset. All classifiers have the highest error rates for darker-skinned females (ranging from 20.8% for Microsoft to 34.7% for IBM).

(Source: Buolamwini, J., & Gebru, T. (2018). *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification* Proceedings of the 1st Conference on Fairness, Accountability and Transparency, Proceedings of Machine Learning Research)

ดังนั้น ในฐานะผู้ออกแบบและผู้พัฒนาระบบการจดจำใบหน้าจึงควรคำนึงมาตรฐานด้านจริยธรรมเรื่องความเท่าเทียม ให้มีชุดข้อมูลที่หลากหลายและมากเพียงพอที่ให้ระบบเรียนรู้ เพื่อพัฒนาให้ปัญญาประดิษฐ์มีประสิทธิภาพในการวิเคราะห์ได้อย่างถูกต้องและเท่าเทียมกันมากยิ่งขึ้น อีกทั้งเป็นการป้องกันผลกระทบที่อาจตามมาจากการวิเคราะห์ที่ผิดพลาด เช่น การนำไปใช้งานที่เกี่ยวข้องกับการระบุผู้ต้องหาอาชญากรรม หรือการระบุตัวตนเพื่อรับสิทธิประโยชน์ต่าง ๆ ในสังคม

นอกจากนี้ สำนักพิมพ์ Nature ได้สำรวจข้อมูลมุมมองของนักวิจัยต่อประเด็นจริยธรรมที่เกี่ยวข้องกับเทคโนโลยีรู้จำใบหน้า [16] ซึ่งจากผลการสำรวจพบว่า นักวิจัยมีความกังวลเกี่ยวกับประเด็นดังกล่าว โดยปัญหาด้านจริยธรรมที่พบส่วนใหญ่ คือ การนำรูป หรือชุดข้อมูลไปใช้โดยไม่มีการขออนุญาตบุคคลในรูป ซึ่งมากกว่าครึ่งหนึ่งของผู้ตอบแบบสำรวจให้ความคิดเห็นว่านักวิจัยสามารถนำรูปที่อยู่ในอินเทอร์เน็ตหรือออนไลน์มาใช้สอนและทดสอบระบบรู้จำใบหน้าได้อย่างอิสระ นักวิจัยบางท่านให้เหตุผลว่า จะเป็นความยากลำบากต่อการสอนและทดสอบระบบรู้จำใบหน้า ให้สามารถวิเคราะห์ผลได้อย่างถูกต้อง หากไม่มีชุดข้อมูลที่มีขนาดใหญ่เพียงพอ นอกจากนี้ ยังมีประเด็นเกี่ยวกับการใช้เทคโนโลยีรู้จำใบหน้ากับกลุ่มเปราะบาง โดยนักวิจัยส่วนใหญ่ให้ความคิดเห็นว่า ถึงแม้กลุ่มคนเปราะบางเหล่านี้จะได้ให้ความยินยอม (consent) ในการใช้ข้อมูลแล้ว แต่ก็ยังคงเป็นคำถามทางด้านจริยธรรมอยู่เสมอว่า นักวิจัยได้รับความยินยอมและสามารถนำข้อมูลของกลุ่มเปราะบางมาใช้ได้อย่างแท้จริงหรือไม่ ทั้งนี้ ยังได้สำรวจมุมมองของนักวิจัยต่อการนำการรู้จำใบหน้าไปใช้ในรูปแบบต่าง ๆ ซึ่งพบว่า นักวิจัยส่วนใหญ่ไม่สามารถยอมรับได้ หรือรู้สึก

[16] Facial-recognition research needs an ethical reckoning. (2020). *Nature*, *587*(7834), 330-330. https://doi.org/10.1038/d41586-020-03256-7

ไม่สบายใจ เมื่อการรู้จำใบหน้าถูกนำไปใช้เพื่อการระบุตัวตนของบุคคล การใช้ประเมินบุคลิกภาพและอารมณ์ของผู้สมัครงาน และการติดตามผู้คนในที่สาธารณะ แต่ค่อนข้างยอมรับได้ หากเป็นการนำไปใช้เพื่อระบุหาผู้ต้องสงสัยในทางอาชญากรรม การใช้ตรวจสอบตัวตนของนักท่องเที่ยวในสนามบิน หรือการใช้การรู้จำใบหน้าเพื่อปลดล็อกหน้าจอโทรศัพท์

ดังนั้น นักวิจัยจึงควรคำนึงถึงขอบเขตของข้อมูลสาธารณะว่า ข้อมูลใดเป็นข้อมูลสาธารณะ หรือข้อมูลส่วนบุคคล รวมถึงคำนึงถึงผลกระทบจากการวิจัยที่อาจเกิดขึ้น และสื่อสารให้กลุ่มคนผู้เป็นเจ้าของข้อมูลเข้าใจ อย่างไรก็ตาม ถึงแม้ในปัจจุบันมีการจัดทำแนวปฏิบัติด้านจริยธรรมปัญญาประดิษฐ์แล้ว แต่แนวปฏิบัติดังกล่าวอาจประกอบด้วยหลักการทางจริยธรรมปัญญาประดิษฐ์ที่ไม่สามารถนำมาใช้ในทางปฏิบัติได้ง่ายนัก

ประเด็นทางจริยธรรมด้านปัญญาประดิษฐ์สำหรับ กรณีศึกษาการรู้จำใบหน้าที่ควรคำนึงถึง

- ในการวิจัยและพัฒนามักใช้ภาพซึ่งเจ้าของภาพตั้งใจอัปโหลดสู่พื้นที่สาธารณะด้วยตนเอง ซึ่งหากเป็นการใช้ข้อมูลที่เป็นสาธารณะอย่างแท้จริง และไม่มีการทำให้สามารถระบุตัวตนของเจ้าของภาพได้นั้น ก็น่าจะสามารถทำได้ แต่ในกรณีที่มีการเก็บข้อมูลและจงใจเก็บชื่อมาด้วยเพื่อให้สามารถระบุตัวบุคคลได้จากใบหน้านั้น อาจต้องตั้งข้อสงสัยว่าเป็นการละเมิดข้อมูลส่วนบุคคลหรือไม่ นอกจากนี้ ควรคำนึงถึงคุณภาพของข้อมูลดิบที่นำมาใช้สอนเครื่องว่าสามารถเชื่อถือได้มากน้อยแค่ไหน
- ในบริบทประเทศไทยนั้น หากมีโปรแกรมที่ใช้ประโยชน์จากการรู้จำใบหน้า ก็ควรอยู่บนพื้นฐานของข้อมูลที่หลากหลาย คำนึงถึงการเก็บข้อมูลอย่างครอบคลุม (inclusiveness) เพื่อให้สามารถเป็นตัวแทนของประชากรทั้งหมดได้ ต้องมีความพยายามลดความเหลื่อมล้ำของข้อมูล โดยรวบรวมใบหน้าที่มีรูปร่างลักษณะต่าง ๆ จากกลุ่มตัวอย่างที่มีความหลากหลายทางเศรษฐกิจและสังคม เพื่อให้ทุกคนมีโอกาสได้ประโยชน์จากโปรแกรมดังกล่าวเท่าเทียมกัน อีกทั้ง ควรแจ้งให้ผู้ใช้งานทราบถึงข้อจำกัดของโปรแกรมหรือปัญญาประดิษฐ์นั้นและให้ผู้ใช้งานตระหนักอยู่เสมอว่าผลจากการทำงานของปัญญาประดิษฐ์นั้นอาจจะไม่ถูกต้องหรือเท่าเทียมกันเสมอไป

ตัวอย่างที่ 2 ดีปเฟก (deepfake)

ดีปเฟก (deepfake) เป็นการใช้ปัญญาประดิษฐ์เพื่อสังเคราะห์ภาพ เสียง และภาพเคลื่อนไหวของบุคคล ให้เคลื่อนไหวและพูดในสิ่งที่ผู้สร้างกำหนด โดยที่บุคคลเจ้าของหน้าหรือเสียงอาจไม่เคยทำหรือพูดสิ่งดังกล่าวจริง ได้อย่างแนบเนียน เทคโนโลยีนี้ได้รับความสนใจอย่างมากในการผลิตวิดีโอสร้างความบันเทิง อย่างไรก็ตามดีปเฟก สามารถเป็นดาบสองคมได้ หากถูกนำไปใช้ในทางที่ผิด โดยมีความกังวลเกี่ยวกับการปลอมแปลงวิดีโอ การปลอมแปลงเสียง การสร้างสื่อที่ไม่เหมาะสมและละเมิดสิทธิส่วนบุคคล รวมถึงการถูกนำไปใช้ในการบิดเบือนข่าวสาร (disinformation) หรือการสร้างข่าวเท็จ (misinformation) เพื่อวัตถุประสงค์ต่าง ๆ ส่งผลให้เกิดความเสียหายตามมาได้ เช่น การเสื่อมเสียชื่อเสียงและความน่าเชื่อถือของตัวบุคคล หรือความเสียหายทางธุรกิจ เป็นต้น

ในรายงานวิชาการ Regulating deep-fakes: Legal and ethical considerations ได้นำเสนอกรณีศึกษาการใช้ดีปเฟกในทางที่ผิดและผลกระทบที่เกิดขึ้น เช่น การนำไปสร้างสื่อลามกอนาจาร เป็นต้น โดยผู้วิจัยได้วิเคราะห์ลักษณะด้านต่าง ๆ ของดีปเฟก ได้แก่ ประโยชน์ ความกังวลหลัก ผลที่ตามมาโดยไม่คาดคิด รวมถึงกฎหมายที่เกี่ยวข้อง ดังแสดงในตารางด้านล่าง [17]

	Examples	Advantages	Major Concerns	Unintended Consequences	Legal Responses
Deep Fake Pornography (e.g., revenge porn)	One's face transferred onto porn actor's naked body	Could mean opportunities to create more videos, if created with consent	Invasion into autonomy and sexual privacy; humiliation and abuse	Humiliation; exploitation; physical, mental or financial abuse of individuals or corporations	*Public law* (criminal law, administrative action) *Private law* (torts)
Political Campaigns	Speeches of politicians, news reports, information about socially significant events	Could promote freedom of speech	Damage to reputation, distortion of democratic discourse, hostile governments, impact on election results	Eroding of trust in institutions; deepening social divisions and polarisation; damage to national security and international relations	*Public law* (constitutional law, administrative law, criminal law) *Private law* (defamation, libel, slander, torts; copyright law)
Reduction of Transaction Costs	Translating video records into multiple languages	Facilitation of social interactions, creation of new business models	Ownership of IPRs to the content; privacy	Emergence of new data silos	*Private law*: contract and tort law
Creative and Original Deep Fakes	Nicolas Cage scenes, parody memes	Promotion of creativity and science, free speech	Ownership of IPRs, privacy	Bullying among children	*Private law* (fair use and copyright law, contract law, tort law) *Public law*: constitutional law

Table 1. Taxonomy of deep fakes and their characteristics

(Source: Meskys, E., Liaudanskas, A., Kalpokiene, J., & Jurcys, P. (2020). Regulating deep fakes: Legal and ethical considerations. *Journal of Intellectual Property Law & Practice*, 15(1), 24–31.)

ในรายงานวิชาการ Anticipating and addressing the ethical implications of deepfakes in the context of elections แสดงให้เห็นว่าดีปเฟกที่ถูกนำมาใช้ในการเลือกตั้ง สามารถก่อให้เกิดผลทางด้านลบต่อ ผู้ชม ผู้ลงสมัคร และความโปร่งใสของการเลือกตั้งได้ เช่น การสร้างข่าวหลอกลวง การทำลายชื่อเสียง

[17] Meskys, E., Liaudanskas, A., Kalpokiene, J., & Jurcys, P. (2020). Regulating deep fakes: Legal and ethical considerations. Journal of Intellectual Property Law & Practice, 15(1), 24–31. https://doi.org/10.1093/jiplp/jpz167

การทำให้ประชาชนเกิดความไม่เชื่อมั่นในกระบวนการเลือกตั้ง เป็นต้น ทั้งนี้ ผู้วิจัยได้เสนอแนวทางการลดผลกระทบจากดีปเฟกในมุมมองทั่วไป เช่น การสร้างความรู้เท่าทันสื่อ การพิสูจน์ความจริง เป็นต้น [18]

ในทางตรงกันข้าม ดีปเฟกก็สามารถใช้เป็นวิธีการป้องกันข้อมูลส่วนบุคคลได้เช่นกัน ดังตัวอย่างในรายงานวิชาการ Deepfakes for Medical Video De-Identification: Privacy Protection and Diagnostic Information Preservation [19] ซึ่งผู้วิจัยได้เปรียบเทียบวิธีการลบหรืออำพรางข้อมูลระบุตัวตน (de-identification) แบบดั้งเดิมที่ใช้ในทางการแพทย์ เช่น การเบลอหน้าของผู้ป่วย กับการใช้ดีปเฟก ในวิดีโอของผู้ป่วยโรคพาร์กินสัน เพื่อไม่ให้สามารถระบุตัวตนของผู้ป่วยได้ ซึ่งผลการศึกษาพบว่า ข้อมูลจากการใช้ดีปเฟกนั้นมีความน่าเชื่อถือและสามารถตรวจจับจุดสำคัญได้ไม่แตกต่างกับข้อมูลก่อนการอำพราง และที่สำคัญสามารถช่วยแก้ปัญหาข้อจำกัดด้านจริยธรรมเกี่ยวกับการใช้ข้อมูลร่วมกัน (data sharing) ทำให้สามารถสร้างชุดข้อมูลของวิดีโอทางการแพทย์ที่เป็นโอเพนซอร์ส (open source) ที่มีคุณภาพสูง สอดคล้องกับจริยธรรมในการปกป้องข้อมูลส่วนบุคคล ซึ่งชุดข้อมูลดังกล่าวจะช่วยสนับสนุนและเป็นประโยชน์ต่องานวิจัยทางการแพทย์ในอนาคตได้

ประเด็นทางจริยธรรมด้านปัญญาประดิษฐ์สำหรับ กรณีศึกษาดีปเฟกที่ควรคำนึงถึง

- วิวัฒนาการของ generative adversarial network (GAN) ที่แพร่หลาย ทำให้ในอนาคตทุกคนจะสามารถผลิตดีปเฟกได้ และอาจมีการนำไปใช้ในทางที่ผิด เช่น การนำไปใช้กลั่นแกล้งรังแก หรือ ส่งผลให้เกิดการฆ่าตัวตายได้
- ดีปเฟกเป็นประเด็นที่มีความสำคัญมากในทางกฎหมาย โดยมีกฎหมายอาญาที่เกี่ยวข้องค่อนข้างมาก เนื่องจากดีปเฟกอาจถูกนำมาใช้สร้างข้อมูลหรือพยานหลักฐานเท็จที่ส่งผลในทางการเมืองและความมั่นคงได้ รวมถึงยังมีประเด็นเรื่องการพิสูจน์สิทธิในอัตลักษณ์ในทางแพ่งอีกด้วย เนื่องจากความยากในการพิสูจน์หรือตรวจสอบ แม้ในปัจจุบันจะมีการใช้เทคโนโลยีตรวจสอบดีปเฟกและมีการควบคุมตนเองของผู้พัฒนา (self-regulation) แล้วก็ตาม ดังนั้น ในเบื้องต้นอาจต้องกำหนดกรอบในการนำเทคโนโลยีดีปเฟกมาใช้ และควรพิจารณาวัตถุประสงค์ในการใช้ที่ชัดเจน
- ควรมีการให้ความรู้และสร้างความตระหนักเรื่องดีปเฟก แก่ประชาชนทั่วไปให้ทราบว่า การปลอมแปลงโดยใช้ดีปเฟกนั้นสามารถทำได้โดยง่าย เพื่อเป็นการสร้างภูมิคุ้มกันให้คนในสังคม ทำให้นักวิจัยรวมถึงประชาชนทั่วไปทราบถึงขอบเขตและข้อจำกัดของเทคโนโลยีปัญญาประดิษฐ์

---

[18] Diakopoulos, N., & Johnson, D. Anticipating and addressing the ethical implications of deepfakes in the context of elections. *New Media & Society*, *0*(0), 1461444820925811. https://doi.org/10.1177/1461444820925811

[19] Zhu, B., Fang, H., Sui, Y., & Li, L. (2020). Deepfakes for Medical Video De-Identification: Privacy protection and diagnostic information preservation.

เอกสารอ้างอิง

Ad Hoc Expert Group (AHEG). (2020). *First Version of a Draft Text of a Recommendation on the Ethics of Artificial Intelligence*. Retrieved 21 March 2021 from https://ircai.org/project/pdf-attachment-recommendation-eng/

Bostrom, N. (1998). How long before superintelligence? *International Journal of Futures Studies*, *2*. https://www.nickbostrom.com/superintelligence.html

Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., Dafoe, A., Scharre, P., Zeitzoff, T., Filar, B., Anderson, H., Roff, H., Allen, G., Steinhardt, J., Flynn, C., hÉigeartaigh, S., Beard, S., Belfield, H., Farquhar, S., & Amodei, D. (2018). The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation.

Buolamwini, J., & Gebru, T. (2018). *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification* Proceedings of the 1st Conference on Fairness, Accountability and Transparency, Proceedings of Machine Learning Research. http://proceedings.mlr.press

Commission, E. (2019). *Ethics guidelines for trustworthy AI*. Retrieved 22 March from https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

Diakopoulos, N., & Johnson, D. Anticipating and addressing the ethical implications of deepfakes in the context of elections. *New Media & Society*, *0*(0), 1461444820925811. https://doi.org/10.1177/1461444820925811

DLA Piper. (2021). *The Future Regulation of Technology: EU AI Regulation Handbook*. https://www.dlapiper.com/~/media/files/insights/publications/2021/05/ai-regs-handbook.pdf

European Commission. (2021). *Europe fit for the Digital Age: Commission proposes new rules and actions for excellence and trust in Artificial Intelligence*. https://ec.europa.eu/commission/presscorner/detail/en/IP_21_1682

Facial-recognition research needs an ethical reckoning. (2020). *Nature*, *587*(7834), 330-330. https://doi.org/10.1038/d41586-020-03256-7

High-Level Expert Group on Artificial Intelligence. (2019). *A definition of Artificial Intelligence: main capabilities and scientific disciplines*.

IBM Cloud Education. (2020). *Data Science*. https://www.ibm.com/cloud/learn/data-science-introduction

Martinez-Plumed, F., Gutierrez, E. G., & Hernandez-Orallo, J. (2020). *AI Watch Assessing Technology Readiness Levels for Artificial Intelligence*.

O'Carroll, B. (2017). *What are the 3 types of AI? A guide to narrow, general, and super artificial intelligence*. Retrieved 22 March from https://codebots.com/artificial-intelligence/the-3-types-of-ai-is-the-third-even-possible

Society of Automotive Engineers International. (2021). *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*. https://www.sae.org/standards/content/j3016_202104

U.S. Department of Transportation. (2018). *A Framework for Automated Driving System Testable Cases and Scenarios*. National Highway Traffic Safety Administration (NHTSA). https://www.nhtsa.gov/sites/nhtsa.gov/files/documents/13882-automateddrivingsystems_092618_v1a_tag.pdf

Yang, G.-Z., Cambias, J., Cleary, K., Daimler, E., Drake, J., Dupont, P. E., Hata, N., Kazanzides, P., Martel, S., & Patel, R. V. (2017). Medical robotics—Regulatory, ethical, and legal considerations for increasing levels of autonomy. *Science Robotics*, *2*(4), 8638.

Zhu, B., Fang, H., Sui, Y., & Li, L. (2020). Deepfakes for Medical Video De-Identification: Privacy protection and diagnostic information preservation.

สำนักงานแถลงขาวสหประชาชาติ กรุงเทพฯ. กฎบัตรสหประชาชาติ. https://treaties.mfa.go.th/pdf/สนธิสัญญาที่น่าสนใจ/UN_Charter_Thai.pdf

สำนักงานเจรจาการค้าบริการและการลงทุน กรมเจรจาการค้าระหว่างประเทศ. (2018). กฎหมายคุ้มครองข้อมูลส่วนบุคคลของสหภาพยุโรป. http://www.moc.go.th/images/633/GDPR-3-5.pdf

สำนักงานคณะกรรมการสิทธิมนุษยชนแห่งชาติ. กติการะหว่างประเทศว่าด้วยสิทธิพลเมืองและสิทธิทางการเมือง. http://www.nhrc.or.th/Human-Rights-Knowledge/International-Human-Rights-Affairs/International-Law-of-human-rights/ICCPR_th.aspx

สุภาภรณ์ เกียรติสิน, มนัสศิริ จันสุทธิรางกูร, ปรัชญ์ สง่างาม, & ยุทธพงศ์ อุณหทวีทรัพย์. (2021). เอกสารแนวปฏิบัติจริยธรรมปัญญาประดิษฐ์ (*Thailand AI Ethics Guideline)*. สำนักงานคณะกรรมการดิจิทัลเพื่อเศรษฐกิจและสังคมแห่งชาติ กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม.

ภาคผนวก

ที่มาและความสำคัญ

ปัญญาประดิษฐ์ ได้ถูกพัฒนาขึ้นและมีความก้าวหน้าอย่างก้าวกระโดด โดยมีอัตราการเพิ่มขึ้นของเทคโนโลยีปัญญาประดิษฐ์อย่างรวดเร็วในระยะหลายปีที่ผ่านมา ซึ่งเป็นผลเนื่องมาจากคุณสมบัติของปัญญาประดิษฐ์ที่เอื้อประโยชน์ให้กับมนุษย์ในแทบทุกด้าน โดยเฉพาะด้านการเงิน ธุรกิจ การแพทย์และสาธารณสุข และการดำรงชีวิตประจำวัน เนื่องด้วยความชาญฉลาดและความสามารถของปัญญาประดิษฐ์ที่กำลังจะเป็นตัวขับเคลื่อนเศรษฐกิจโลก จึงทำให้นานาประเทศต่างให้ความสนใจและแข่งขันกันเพื่อเป็นผู้นำด้านปัญญาประดิษฐ์ โดยการประกาศนโยบายสนับสนุนการวิจัยและพัฒนาปัญญาประดิษฐ์ เพื่อเพิ่มขีดความสามารถในการแข่งขันกับต่างประเทศ ร่วมถึงส่งเสริมให้มีการพัฒนาที่ก้าวทันกับเทคโนโลยีโลก ทั้งนี้ เมื่อพิจารณาถึงการมีอำนาจทางด้านปัญญาประดิษฐ์ของกลุ่มผู้มีส่วนเกี่ยวข้อง เช่น นักวิจัย ผู้ออกแบบและพัฒนาปัญญาประดิษฐ์ ตลอดจนภาครัฐที่กำกับดูแลการวิจัยและพัฒนา จะสามารถแบ่งได้เป็น 3 กลุ่ม ดังนี้

1. ผู้มีอำนาจทางความรู้ ได้แก่ นักวิทยาศาสตร์ นักวิจัย
2. ผู้มีอำนาจทางธุรกิจและสังคม ได้แก่ ผู้ประกอบการรายใหญ่ กลุ่มการค้า และกลุ่มไม่แสวงหากำไร แต่มีอิทธิพลทางสังคม เช่น Partnership on AI และ Global Partnership on Artificial Intelligence
3. ผู้มีอำนาจทางการปกครอง ได้แก่ ผู้มีอำนาจรัฐ ระดับประเทศ ระดับการเมือง และการปกครอง

ในทางกลับกัน ปัญญาประดิษฐ์ก็เป็นเสมือนดาบสองคม เนื่องด้วยคุณสมบัติเด่นของปัญญาประดิษฐ์ที่ชาญฉลาดจนสามารถเอาชนะมนุษย์ ก็อาจนำภัยคุกคามมาสู่มนุษย์ได้ หากไม่มีการควบคุมการวิจัย การออกแบบ หรือการพัฒนา ให้เป็นไปตามกรอบมาตรฐานและหลักการจริยธรรม โดยผลกระทบอาจเกิดจากการประมวลผล การตัดสินใจที่ผิดพลาด การให้ผลลัพธ์ที่มีอคติ หรือการทำให้เกิดความไม่เท่าเทียม นอกจากนี้ยังมีความกังวลเกี่ยวกับประเด็นเรื่องความปลอดภัยทางไซเบอร์ และธรรมาภิบาลข้อมูล เนื่องจากการพัฒนาปัญญาประดิษฐ์อาจต้องใช้ข้อมูลส่วนบุคคลและฐานข้อมูลสำคัญอื่น ๆ การจัดเก็บและการใช้ข้อมูลจึงต้องมีความมั่นคงปลอดภัย สามารถสร้างความเชื่อมั่นให้กับผู้ใช้งานปัญญาประดิษฐ์ได้ ตัวอย่างปัญญาประดิษฐ์ที่ทำให้หลายประเทศเริ่มกังวลและตระหนักถึงผลกระทบที่อาจเกิดขึ้นกับมนุษย์และสังคมในอนาคตได้ เช่น หุ่นยนต์ Sophia ที่มีใบหน้าคล้ายมนุษย์และได้รับสถานะพลเมืองประเทศซาอุดีอาระเบีย, ผลงานเพลง Hello World ที่แต่งโดยอาศัยปัญญาประดิษฐ์, โครงการ The Next Remdbrandt ที่ให้ปัญญาประดิษฐ์วาดภาพ, ซูเปอร์คอมพิวเตอร์เทียนเหอ-2 ของจีนซึ่งมีพลังประมวลผลมากกว่าสมองมนุษย์, AlphaGo ที่เล่นหมากกระดานชนะมนุษย์, ระบบ Watson ที่ให้คำวินิจฉัยทางการแพทย์, และโปรแกรม COMPAS ที่ใช้คำนวณโอกาสการกระทำผิดกฎหมายซ้ำ เป็นต้น

อย่างไรก็ตาม การนำปัญญาประดิษฐ์ไปใช้ในทางที่ไม่พึงประสงค์นั้น ไม่จำเป็นต้องเกิดจากกลุ่มผู้มีอำนาจทางความรู้ หรือนักวิทยาศาสตร์เท่านั้น แต่อาจเกิดจากกลุ่มนักธุรกิจหรือกลุ่มการเมืองได้เช่นกัน ดังนั้น เราจึงควรรู้เท่าทัน ไม่นำปัญญาประดิษฐ์ไปใช้งานในทางผิด และตระหนักถึงความเสี่ยงที่อาจเกิดขึ้นได้

จากการใช้งาน ซึ่งในปัจจุบันมีหลายองค์กรในหลายประเทศทั่วโลกที่ศึกษาผลกระทบที่อาจเกิดขึ้นจากเทคโนโลยีปัญญาประดิษฐ์ และจัดทำแนวปฏิบัติในการนำเอาหลักการทางจริยธรรมเข้าไปร่วมอยู่ในกระบวนการการออกแบบ พัฒนา และนำปัญญาประดิษฐ์ไปใช้ประโยชน์แล้ว ด้วยเหตุนี้ จึงเป็นที่มาของการจัดทำแนวปฏิบัติจริยธรรมด้านปัญญาประดิษฐ์ของ สวทช. โดยมีหลักการสำคัญ คือ เป็นแนวปฏิบัติเพื่อให้เกิดความสมดุลระหว่างการพัฒนาที่รวดเร็ว กับการรู้เท่าทันการนำไปใช้ในทางที่ผิดและเลวร้ายต่อมนุษย์ โดยคาดหวังว่า ผู้วิจัย ออกแบบ และพัฒนาปัญญาประดิษฐ์ ตลอดจนผู้ที่สนใจที่ศึกษาแนวปฏิบัติฉบับนี้จะได้รับความเข้าใจเกี่ยวกับจริยธรรมปัญญาประดิษฐ์ และสามารถนำไปปฏิบัติให้เกิดประโยชน์ต่อการดำเนินงานด้านปัญญาประดิษฐ์อย่างมีจริยธรรมได้อย่างแท้จริง

กฎหมาย นโยบาย และแนวปฏิบัติต่าง ๆ ที่เกี่ยวข้อง

ความก้าวหน้าของเทคโนโลยีปัญญาประดิษฐ์นำมาซึ่งประโยชน์ต่าง ๆ มากมาย แต่หากผู้วิจัย ออกแบบ พัฒนา หรือผู้ใช้งานปัญญาประดิษฐ์ ไม่รู้เท่าทันหรือไม่ตระหนักถึงความเสี่ยงจากการใช้งาน ก็อาจทำให้เกิดผลกระทบและความเสียหาย จากการประมวลผลและการตัดสินใจที่ผิดพลาดของปัญญาประดิษฐ์ หรือการนำไปใช้ในทางไม่พึงประสงค์ได้ กฎหมาย นโยบาย และแนวปฏิบัติต่าง ๆ เป็นกลไกหนึ่งที่ทำให้เกิดภาระความรับผิดชอบต่อปัญญาประดิษฐ์ และสามารถใช้กำกับดูแลปัญญาประดิษฐ์ได้ โดยการดำเนินงานวิจัย ออกแบบ พัฒนา ประยุกต์ใช้ และถ่ายทอดเทคโนโลยี จะต้องอยู่ภายใต้กรอบกฎหมายภายในประเทศ และกฎหมายระหว่างประเทศที่เกี่ยวข้อง ซึ่งภายในประเทศไทยมีกฎหมายที่ถือเป็นแนวทางที่ผู้เกี่ยวข้องจะต้องปฏิบัติตาม เช่น

1. รัฐธรรมนูญแห่งราชอาณาจักรไทย พ.ศ. 2560

ตัวอย่างเช่น มาตรา 27 (วรรค 3) ว่าด้วยการไม่เลือกปฏิบัติโดยไม่เป็นธรรมต่อบุคคล

<u>ตัวอย่างแนวทางการดำเนินงาน:</u> ปัญญาประดิษฐ์จะต้องถูกออกแบบมาโดยคำนึงถึงความเสมอภาคเท่าเทียมกัน ไม่เลือกปฏิบัติ ทั้งในขั้นตอนการนำปัญญาประดิษฐ์ไปใช้งาน เพื่อให้สามารถรองรับความต้องการและการเข้าถึงของประชาชนทุกคนในสังคม เช่น กลุ่มคนด้อยโอกาส ผู้พิการและผู้ทุพพลภาพ ตลอดจนขั้นตอนการพัฒนาปัญญาประดิษฐ์โดยใช้ข้อมูลที่หลากหลาย เพื่อให้ได้ผลลัพธ์ที่ถูกต้อง ไม่เอนเอียง หรือเกิดอคติ

2. พระราชบัญญัติคุ้มครองผู้บริโภค พ.ศ. 2522

ตัวอย่างเช่น มาตรา 4 ว่าด้วยสิทธิที่ผู้บริโภคได้รับความคุ้มครอง

<u>ตัวอย่างแนวทางการดำเนินงาน:</u> ผู้วิจัยหรือพัฒนาปัญญาประดิษฐ์จะต้องให้ข้อมูลเกี่ยวกับคุณสมบัติและการทำงานของระบบปัญญาประดิษฐ์ที่เพียงพอแก่ผู้ใช้งาน ตลอดจนคำนึงถึงความปลอดภัยจากการใช้งาน แนวทางการชดเชยความเสียหาย ในกรณีที่เกิดอันตรายหรือผลกระทบ เพื่อให้สอดคล้องและเป็นไปตามสิทธิที่ผู้ใช้งานควรได้รับความคุ้มครองในฐานะผู้บริโภค

3. พระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562

ตัวอย่างเช่น หมวด 2 การคุ้มครองข้อมูลส่วนบุคคล ส่วนที่ 2 การเก็บรวบรวมข้อมูลส่วนบุคคล (มาตรา 22 – 26) และส่วนที่ 3 การใช้หรือเปิดเผยข้อมูลส่วนบุคคล (มาตรา 27 – 29)

ตัวอย่างแนวทางการดำเนินงาน: หากมีการนำข้อมูลส่วนบุคคลมาใช้ในระบบปัญญาประดิษฐ์ ผู้วิจัยและพัฒนาจะต้องดำเนินงานให้สอดคล้องกับ พ.ร.บ. คุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562 เช่น การเก็บรวบรวมข้อมูลต้องทำเท่าที่จำเป็นและเหมาะสม มีการขอความยินยอมจากเจ้าของข้อมูลที่ชัดเจน เข้าใจง่าย และการใช้และการเปิดเผยข้อมูลต้องเป็นไปตามวัตถุประสงค์ที่ได้แจ้งไว้

4. พระราชบัญญัติการรักษาความมั่นคงปลอดภัยทางไซเบอร์ พ.ศ. 2562

ตัวอย่างเช่น มาตรา 58 ว่าด้วยการรับมือกับภัยคุกคามทางไซเบอร์

ตัวอย่างแนวทางการดำเนินงาน: ผู้วิจัยและพัฒนาระบบปัญญาประดิษฐ์จะต้องมีการวางแผนการป้องกัน การรับมือ และการลดความเสี่ยงทางไซเบอร์ มีการพัฒนาความรู้ความสามารถของบุคคลากรและให้ความรู้ ความตระหนักถึงภัยไซเบอร์อยู่เสมอ

นอกจากกฎหมายภายในประเทศข้างต้นแล้ว ยังมีกฎหมายภายในประเทศและกฎหมายระหว่างประเทศอื่นที่เกี่ยวข้องกับงานด้านปัญญาประดิษฐ์ ที่ผู้วิจัย ผู้ออกแบบ และผู้พัฒนา จะต้องคำนึงถึง ในกรณีที่มีการนำปัญญาประดิษฐ์ไปใช้ในทางการค้า หรือมีแผนเปิดให้บริการด้านปัญญาประดิษฐ์ในต่างประเทศ ตัวอย่างเช่น

1. **พระราชบัญญัติความรับผิดต่อความเสียหายที่เกิดขึ้นจากสินค้าที่ไม่ปลอดภัย พ.ศ. 2551** โดยผู้วิจัยจะต้องเข้าใจถึงขอบเขตความรับผิดชอบในสินค้าที่มีปัญญาประดิษฐ์เป็นองค์ประกอบ เช่น ในการทำสัญญาการอนุญาตให้ใช้สิทธิ (licensing) ควรเขียนขอบเขตความรับผิดชอบระหว่างผู้วิจัยกับบริษัทที่ผลิตสินค้าให้ชัดเจน เพื่อให้มีภาระรับผิดชอบอยู่ในขอบเขตที่เหมาะสม ตัวอย่างเช่น การกำหนดให้ผู้วิจัยไม่มีภาระความรับผิดชอบในการแจ้งให้ผู้ใช้งานทราบถึงรายละเอียดการนำสินค้าไปใช้งาน และบริษัทผู้ผลิตสินค้าจะเป็นผู้รับผิดชอบชดใช้ค่าเสียหาย หากมีความเสียหายจากการใช้งานสินค้าเกิดขึ้น เป็นต้น

2. **พระราชบัญญัติการบริหารงานและการให้บริการภาครัฐผ่านระบบดิจิทัล พ.ศ. 2562** เช่น มีการจัดทำธรรมาภิบาลข้อมูล (data governance)

3. **กฎหมายกำกับเฉพาะกิจการหรือกลุ่มอุตสาหกรรม** ซึ่งอาจกำหนดเรื่องการใช้ข้อมูลส่วนบุคคล ความปลอดภัย หรือหลักการไม่เลือกปฏิบัติที่เจาะจงกับลักษณะกิจการ เช่น ในกิจการสุขภาพ ประเทศไทยมี พ.ร.บ. สุขภาพแห่งชาติ พ.ศ. 2550 ซึ่งมาตรา 7 กำหนดหลักไว้ว่าข้อมูลด้านสุขภาพของบุคคลเป็นความลับส่วนบุคคล ผู้ใดจะนำไปเปิดเผยในประการที่น่าจะทำให้บุคคลนั้นเสียหายไม่ได้ หรือในกิจการโทรคมนาคม ประเทศไทยมีประกาศคณะกรรมการกิจการโทรคมนาคมแห่งชาติ เรื่อง มาตรการคุ้มครองสิทธิของผู้ใช้บริการโทรคมนาคมเกี่ยวกับข้อมูลส่วนบุคคล สิทธิในความเป็นส่วนตัว และเสรีภาพในการสื่อสารถึงกันโดยทางโทรคมนาคม

4. **หลักกฎหมายระหว่างประเทศว่าด้วยการไม่ใช้กำลัง ของกฎบัตรสหประชาชาติ** (ข้อ 2 (4)) [28]

5. **หลักกฎหมายระหว่างประเทศด้านสิทธิมนุษยชน: กติการะหว่างประเทศว่าด้วยสิทธิพลเมืองและสิทธิทางการเมือง (International Covenant on Civil and Political Rights - ICCPR)** เช่น หลักการไม่เลือกปฏิบัติ (ข้อ 2 (1)) และ สิทธิในความเป็นส่วนตัว (ข้อ 17 (1)) เป็นต้น [29]

6. **ระเบียบการคุ้มครองข้อมูลทั่วไป หรือ General Data Protection Regulation (GDPR)** ซึ่งเป็นกฎหมายที่ให้ความสำคัญกับการคุ้มครองข้อมูลส่วนบุคคลของพลเมืองสหภาพยุโรปและภายในสหภาพยุโรป [30]

7. **ข้อแนะนำ (recommendation) หรือแนวปฏิบัติ (guideline) ซึ่งเป็นมาตรฐานหรือบรรทัดฐานระหว่างประเทศที่ไม่มีผลผูกพันทางกฎหมาย** เช่น Recommendation on the Ethics of Artificial Intelligence จากองค์การเพื่อการศึกษา วิทยาศาสตร์ และวัฒนธรรมแห่งสหประชาชาติ (UNESCO) [31] และร่าง Data Privacy Guidelines for the Development and Operation of Artificial Intelligence Solutions จากคณะมนตรีสิทธิมนุษยชนแห่งสหประชาชาติ (UNHRC) [32]

8. **ร่างกฎหมายปัญญาประดิษฐ์ หรือ Artificial Intelligence Act (AIA) ของสหภาพยุโรป** ซึ่งใช้ควบคุมปัญญาประดิษฐ์ที่มีการจำหน่าย/มีผู้ใช้งานอยู่ในสหภาพยุโรป หรืออาจมีผลกระทบจากการใช้งานเกิดขึ้นในสหภาพยุโรป [33]

9. **มาตรฐานอุตสาหกรรมและมาตรฐานองค์กรวิชาชีพ** เช่น ISO และ IEEE ที่เกี่ยวข้องกับปัญญาประดิษฐ์ ตัวอย่างเช่น

9.1. ISO/IEC 27701:2019 Security techniques – Extension to ISO/IEC 27001 and ISO/IEC 27002 for privacy information management – Requirements and guidelines เป็นมาตรฐานสำหรับการบริหารจัดการข้อมูลส่วนบุคคลอย่างมั่นคงปลอดภัย มีประสิทธิภาพ และลดความเสี่ยงอันเนื่องมาจากการละเมิดความเป็นส่วนบุคคลของทั้งพนักงานและลูกค้าขององค์กร [34]

---

[28] สำนักงานแถลงข่าวสหประชาชาติ กรุงเทพฯ. กฎบัตรสหประชาชาติ. https://treaties.mfa.go.th/pdf/สนธิสัญญาที่น่าสนใจ/UN_Charter_Thai.pdf

[29] สำนักงานคณะกรรมการสิทธิมนุษยชนแห่งชาติ. กติการะหว่างประเทศว่าด้วยสิทธิพลเมืองและสิทธิทางการเมือง. http://www.nhrc.or.th/Human-Rights-Knowledge/International-Human-Rights-Affairs/International-Law-of-human-rights/ICCPR_th.aspx

[30] สำนักงานเจรจาการค้าบริการและการลงทุน กรมเจรจาการค้าระหว่างประเทศ. (2018). กฎหมายคุ้มครองข้อมูลส่วนบุคคลของสหภาพยุโรป. http://www.moc.go.th/images/633/GDPR-3-5.pdf

[31] https://unesdoc.unesco.org/ark:/48223/pf0000373434

[32] https://www.ohchr.org/EN/Issues/Privacy/SR/Pages/CFI_data_privacy_guidelines.aspx

[33] https://dpoblog.eu/the-artificial-intelligence-act-aia-a-brief-overview

[34] https://www.iso.org/standard/71670.html

9.2. ISO/IEC 20547-3:2020 Information technology – Big data reference architecture – Part 3: Reference architecture เป็นมาตรฐานที่เกี่ยวข้องกับสถาปัตยกรรมข้อมูลขนาดใหญ่ (Big Data) [35]

9.3. ISO/IEC TR 24028:2020 Information technology – Artificial intelligence – Overview of trustworthiness in artificial intelligence เป็นมาตรฐานที่อธิบายภาพกว้างและสำรวจประเด็นที่เกี่ยวข้องกับความน่าเชื่อถือของระบบปัญญาประดิษฐ์ [36]

9.4. ISO/IEC TR 24027:2021 Information technology – Artificial intelligence – Bias in AI systems and AI aided decision making เป็นมาตรฐานเกี่ยวกับอคติในระบบปัญญาประดิษฐ์และระบบการตัดสินใจที่มีปัญญาประดิษฐ์แนะนำ [37]

9.5. ISO 40500:2012 Information technology – W3C Web Content Accessibility Guidelines (WCAG) 2.0 เป็นแนวทางการออกแบบเนื้อหาเว็บที่คำนึงถึงการเข้าถึงโดยกลุ่มผู้ใช้ทุกกลุ่ม [38]

9.6. IEEE 7010-2020 - IEEE Recommended Practice for Assessing the Impact of Autonomous and Intelligent Systems on Human Well-Being แนวปฏิบัติสำหรับการประเมินผลกระทบของระบบอัตโนมัติและระบบอัจฉริยะต่อสวัสดิภาพของมนุษย์ [39]

สวทช. ได้เล็งเห็นถึงความสำคัญของการพัฒนางานวิจัยด้านปัญญาประดิษฐ์ให้มีความก้าวหน้า เพิ่มขีดความสามารถของเทคโนโลยีปัญญาประดิษฐ์และการถ่ายทอดไปสู่การใช้ประโยชน์เพื่อการพัฒนาและสร้างความสามารถในการแข่งขันของประเทศ ภายใต้การดำเนินงานที่สอดคล้องกับจริยธรรมที่เป็นสากล กรอบมาตรฐานของสังคม และกฎหมายที่เกี่ยวข้อง ด้วยเหตุนี้ สวทช. จึงได้จัดทำนโยบายจริยธรรมด้านปัญญาประดิษฐ์ พ.ศ. 2564 เพื่อใช้ดูแลการดำเนินงานด้านปัญญาประดิษฐ์ ภายใน สวทช. ให้มีการดำเนินงานอย่างเหมาะสม ก่อให้เกิดความสมดุลระหว่างความก้าวหน้าของเทคโนโลยีปัญญาประดิษฐ์ที่รวดเร็วกับการรู้เท่าทันการนำปัญญาประดิษฐ์ไปใช้งาน และลดผลกระทบความเสียหายที่อาจเกิดขึ้นจากปัญญาประดิษฐ์ นอกจากนี้ สวทช. ยังได้จัดทำนโยบายและแนวปฏิบัติต่าง ๆ ที่เกี่ยวข้องกับการบริหารจัดการข้อมูลให้มีประสิทธิภาพ เหมาะสม และสอดคล้องกับกฎหมายที่เกี่ยวข้อง เช่น นโยบายการคุ้มครองข้อมูลส่วนบุคคล นโยบายการรักษาความลับ แนวปฏิบัติในการรักษาความลับ และนโยบายธรรมาภิบาลข้อมูล เป็นต้น ดังนั้น นอกจากจะต้องปฏิบัติตามกฎหมายที่เกี่ยวข้องแล้ว นักวิจัย ผู้ช่วยวิจัย ของ สวทช. หรือผู้มีส่วน

---

[35] https://www.iso.org/standard/71277.html

[36] https://www.iso.org/standard/77608.html

[37] https://www.iso.org/standard/77607.html

[38] https://www.iso.org/standard/58625.html

[39] https://standards.ieee.org/content/ieee-standards/en/standard/7010-2020.html

เกี่ยวข้องกับงานวิจัย และพัฒนาด้านปัญญาประดิษฐ์ที่ดำเนินงานภายใน สวทช. ยังจำเป็นต้องอยู่ภายใต้นโยบายและแนวปฏิบัติต่าง ๆ ของ สวทช. ที่เกี่ยวข้องดังกล่าวข้างต้นนี้ด้วยเช่นกัน

ตารางที่ 5 ตัวอย่าง เกณฑ์มาตรฐานสากลที่ใช้ในการจำแนกระดับความสามารถและความแม่นยำของปัญญาประดิษฐ์ประเภทต่าง ๆ

เทคโนโลยี	จำนวนระดับ	รายละเอียด
ระบบผู้เชี่ยวชาญ (expert system)[1]	3	Level 1: Narrow static and certified knowledge การจัดระบบองค์ความรู้เชิงสถิต สามารถอธิบาย ให้เหตุผล และข้อสรุปที่มีความซับซ้อนได้ Level 2: Narrow dynamic and uncertain knowledge การกลั่นกรององค์ความรู้เชิงกล สามารถให้เหตุผลภายใต้ความไม่แน่นอน และให้ความเข้าใจที่สามารถนำไปปฏิบัติได้จริง (actionable insight) Level 3: Broad knowledge, common sense and meta-cognition มีองค์ความรู้ที่กว้าง มีการรู้คิด และมีปฏิกิริยาตอบสนอง
เครื่องแปลภาษา (machine translation) [1]	5	Level 1: Machine-assisted human translation (MAHT) การแปลภาษาที่มีมนุษย์เป็นผู้แปล โดยใช้คอมพิวเตอร์เป็นเครื่องมือช่วยเพิ่มความรวดเร็วในกระบวนการแปลภาษา Level 2: Human-assisted machine translation (HAMT) ข้อความถูกดัดแปลงโดยมนุษย์ก่อน ระหว่าง หรือหลังจากที่ถูกแปลโดยคอมพิวเตอร์ Level 3: Fully automatic (automated) machine translation (FAMT) การแปลภาษาอัตโนมัติ ซึ่งมีคุณภาพการแปลไม่สูง และไม่มีมนุษย์มาเกี่ยวข้อง Level 4: Fully automatic high-quality machine translation in restricted and controlled domains (FAHQMTr) การแปลภาษาอัตโนมัติ ซึ่งมีคุณภาพการแปลสูง และไม่มีมนุษย์มาเกี่ยวข้อง ภายใต้โดเมนที่ถูกจำกัดและถูกควบคุม

เทคโนโลยี	จำนวนระดับ	รายละเอียด
		Level 5: Fully automatic high-quality machine translation in unrestricted domains (FAHQMTu) การแปลภาษาอัตโนมัติ ซึ่งมีคุณภาพการแปลสูง และไม่มีมนุษย์มาเกี่ยวข้อง ภายใต้โดเมนที่ไม่ถูกจำกัด
การรู้จำเสียง (speech recognition) [1]	4	Level 1: Limited voice commands คำสั่ง หรือการสั่งงานด้วยเสียงที่กำหนดไว้ล่วงหน้า ในระบบการรู้จำ ซึ่งใช้ไวยากรณ์ที่เฉพาะเจาะจง Level 2: Large-vocabulary continuous speech recognition systems ระบบการรู้จำเสียงภายใต้โดเมนที่จำกัด ซึ่งมีคลังคำศัพท์ขนาดใหญ่สำหรับคำและวลีในภาษาพูด (ทั้งแบบเป็นทางการและไม่เป็นทางการ) มีปฏิสัมพันธ์กับผู้ใช้งาน มีความแม่นยำและถูกต้องของข้อมูลอยู่ในระดับสูง สามารถตรวจจับจุดเริ่มต้นและสิ้นสุดของคำพูดได้เอง (endpoint detection) และไม่จำกัดเวลารอรับคำพูด (speech timeout) เป็นต้น Level 3: Free speech recognition in restricted contexts คลังคำศัพท์ที่ไม่จำกัด (ทั้งแบบเป็นทางการและไม่เป็นทางการ) รับเสียงจากแหล่งที่อยู่ไกลได้ ระบุผู้พูดได้ ถอดความจากเสียงและวิดีโอได้ และสามารถจัดการกับปัญหาเสียงรบกวน เสียงสะท้อน สำเนียง และคำพูดที่ไม่มีระเบียบแบบแผน (disorganised speech) ได้ Level 4: Native-level free speech recognition in unrestricted contexts รู้จำการออกเสียงหลากหลายภาษาได้เหมือนเจ้าของภาษา ในสภาวะแวดล้อมที่รบกวนการทำงานอย่างหนัก (adversarial environment) มีการประมวลผลการพูด/การเปล่งเสียงที่ซับซ้อน มีเทคโนโลยีรู้จำจากคำพูดตามธรรมชาติ ซึ่งอาจพูดเพียงบางส่วนจากประโยคเต็ม พูดวนซ้ำ พูดผิดตอนต้นแล้วพูดใหม่ (spontaneous speech) เป็นต้น

เทคโนโลยี	จำนวนระดับ	รายละเอียด
การจดจำใบหน้า (facial recognition) [1]	3	Level 1 - Recognition under ideal situations การรู้จำอัตลักษณ์ เพศ และอายุ จากใบหน้าตรงเต็มหน้าที่อยู่นิ่ง และมีคุณภาพสูง ภายใต้สถานการณ์ที่ถูกควบคุมความเข้มของแสง กล้อง และบุคคล Level 2 - Recognition under partially controlled situations การรู้จำอัตลักษณ์ เพศ และอายุ จากใบหน้าตรง ซึ่งมีคุณภาพต่ำ ภายใต้สถานการณ์ที่อาจถูกควบคุมน้อยกว่า กล่าวคือ มีการควบคุมความเข้มของแสงและกล้อง แต่ไม่ควบคุมบุคคล เช่น การลงทะเบียนที่สนามบินหรือสถานีรถไฟ Level 3 - Recognition under uncontrolled situations การรู้จำอัตลักษณ์ เพศ และอายุ จากภาพ/วิดีโอที่แสดงบางส่วนของใบหน้า ซึ่งมีความละเอียดต่ำ และมีท่าทางที่หลากหลาย ในสถานการณ์ที่ไม่มีการควบคุมกล้อง ความเข้มของแสง หรือบุคคล ทั้งนี้ ระบบการจดจำต้องมีความแม่นยำต่อลักษณะของบุคคล เช่น เชื้อชาติ เพศ รวมถึง การเปลี่ยนแปลงของทรงผม หนวดเครา น้ำหนัก และความชรา
ยานยนต์ไร้คนขับ (autonomous vehicle) [2]	6	Level 0: No driving automation รถยนต์โดยทั่วไปในปัจจุบัน ซึ่งไม่มีระบบที่สามารถควบคุมตัวรถได้โดยอัตโนมัติ มนุษย์ต้องเป็นผู้ขับและควบคุมรถด้วยตัวเอง Level 1: Driver assistance รถยนต์ที่มีระบบช่วยเหลือผู้ขับรถ โดยที่มนุษย์ยังเป็นผู้ดูแลและควบคุมรถอยู่ เช่น การถือพวงมาลัย และการเหยียบเบรก เป็นต้น ตัวอย่างในกลุ่มนี้ เช่น ระบบควบคุมความเร็วอัตโนมัติแบบแปรผัน (adaptive cruise control) ให้เหมาะสมต่อสภาพการจราจรและมีความปลอดภัยมากขึ้น และระบบ Lane Keeping Assistance ช่วยประคองพวงมาลัยเมื่อรถออกนอกเลนโดยไม่ตั้งใจ เป็นต้น

เทคโนโลยี	จำนวนระดับ	รายละเอียด
		Level 2: Partial driving automation รถยนต์ที่มีระบบที่สามารถเร่งความเร็ว เหยียบเบรก หรือควบคุมพวงมาลัยได้ด้วยตัวเอง โดยที่มนุษย์ยังสามารถเข้าควบคุมรถเมื่อใดก็ได้ที่ต้องการ ตัวอย่างเช่น Tesla Autopilot และ Super Cruise ของบริษัท Cadillac ก็จัดอยู่ในกลุ่มนี้ Level 3: Conditional driving automation รถยนต์ที่มีความสามารถตรวจจับสภาวะแวดล้อมต่าง ๆ และสามารถตัดสินใจได้ แต่มนุษย์ยังจำเป็นต้องเตรียมตัวและมีความพร้อมในการเข้าควบคุมรถด้วยตัวเองตลอดเวลา หากระบบไม่สามารถดำเนินการได้ หรืออยู่ในสถานการณ์ฉุกเฉิน Level 4: High driving automation รถยนต์ที่มีระบบที่สามารถควบคุมรถได้โดยอัตโนมัติทั้งหมด แต่มนุษย์ยังคงมีทางเลือกในการเข้าควบคุมรถด้วยตัวเองได้ ทั้งนี้ มีข้อกำหนดทางกฎหมาย ซึ่งจำกัดพื้นที่ในการใช้ระบบขับเคลื่อนอัตโนมัติของรถยนต์ในกลุ่มนี้ Level 5: Full driving automation รถยนต์ที่ไม่ต้องการการควบคุมใด ๆ โดยมนุษย์ มนุษย์ทำหน้าที่เพียงสตาร์ทเครื่องและระบุจุดหมายปลายทางเท่านั้น ดังนั้น รถยนต์ในกลุ่มนี้จึงอาจไม่มีแม้แต่พวงมาลัย แป้นคันเร่ง หรือเบรก อีกทั้ง สามารถใช้ระบบขับเคลื่อนอัตโนมัติได้ในทุกสถานการณ์โดยไม่ละเมิดกฎหมายใด ๆ
หุ่นยนต์ทางการแพทย์ (medical robotics) [3]	6	Level 0: No autonomy ไม่มีระบบอัตโนมัติ ซึ่งรวมถึงหุ่นยนต์ควบคุมระยะไกล (tele-operated robot) หรือ กายอุปกรณ์เทียม (prosthetic device) ซึ่งตอบสนองและทำงานตามที่ผู้ใช้งานกำหนด หุ่นยนต์ผ่าตัดที่มี motion scaling ก็จัดอยู่ในกลุ่มนี้ เนื่องจากเป็นการทำงานตามการเคลื่อนไหวของศัลยแพทย์

เทคโนโลยี	จำนวนระดับ	รายละเอียด
		Level 1: Robot assistance มีการให้คำแนะนำเชิงกล หรือการช่วยเหลือโดยหุ่นยนต์ระหว่างการทำงาน ในขณะที่มนุษย์เป็นผู้ควบคุมระบบอย่างต่อเนื่อง ตัวอย่างเช่น หุ่นยนต์ผ่าตัดที่มี virtual fixture หรือกายอุปกรณ์เสริมสำหรับขา (lower-limb device) ที่มีการควบคุมความสมดุล Level 2: Task autonomy หุ่นยนต์สามารถทำงานได้โดยอัตโนมัติในงานที่มีความจำเพาะ โดยมีการเริ่มการทำงานโดยมนุษย์ แต่มนุษย์ไม่จำเป็นต้องควบคุมการทำงานอย่างต่อเนื่อง ตัวอย่างเช่น การเย็บแผล ศัลยแพทย์จะกำหนดบริเวณที่เย็บแผล จากนั้น หุ่นยนต์จะทำหน้าที่เย็บแผลโดยอัตโนมัติ ในขณะที่ศัลยแพทย์จะดูแลและเข้าแทรกแซงการทำงานเมื่อจำเป็น Level 3: Conditional autonomy ระบบจะสร้างการทำงานรูปแบบต่าง ๆ ซึ่งมนุษย์สามารถเลือกมาใช้งานได้ หรือมนุษย์อาจยอมให้ระบบเลือกรูปแบบการทำงานให้อัตโนมัติ หุ่นยนต์ผ่าตัดในกลุ่มนี้ สามารถทำหน้าที่ได้โดยปราศจากการดูแลอย่างใกล้ชิด หรือกายอุปกรณ์เสริมสำหรับขา ซึ่งสามารถรับรู้ถึงความต้องการในการเคลื่อนที่ของผู้สวมใส่ และสามารถปรับเปลี่ยนได้โดยอัตโนมัติ โดยปราศจากการสั่งงานโดยตรงจากผู้สวมใส่ Level 4: High autonomy หุ่นยนต์ที่สามารถตัดสินใจทางการแพทย์ได้ แต่ยังอยู่ภายใต้การควบคุมโดยแพทย์ ตัวอย่างเช่น robotic resident ซึ่งสามารถทำหน้าที่ผ่าตัดได้ โดยอยู่ภายใต้การควบคุมของศัลยแพทย์

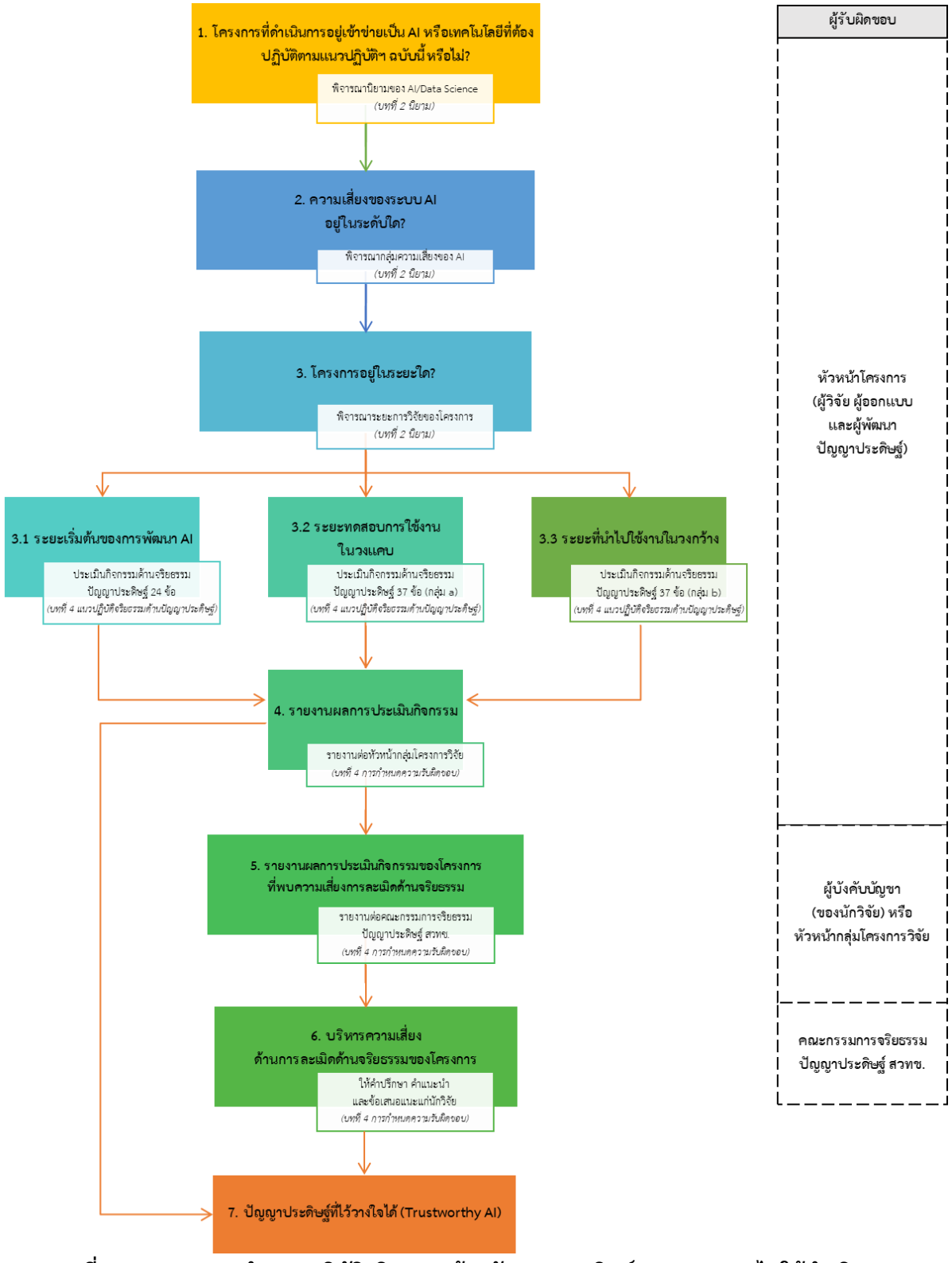
เทคโนโลยี	จำนวนระดับ	รายละเอียด
		Level 5: Full autonomy ระบบอัตโนมัติโดยสมบูรณ์ ไม่ต้องมีการควบคุมใด ๆ โดยมนุษย์ หุ่นยนต์สามารถดำเนินการผ่าตัดได้ตลอดทั้งกระบวนการ
ผู้ช่วยเสมือน (virtual assistant) [1]	3	Level 1 - Predetermined written queries in one domain คำถามถูกจำกัด หรือจะต้องมีคำสำคัญ (keyword) ที่ระบบจดจำได้ เพื่อค้นหาหัวข้อและข้อมูลที่เกี่ยวข้อง คำตอบมีลักษณะเป็นข้อความตามเทมเพลตที่กำหนด หรือมีการบันทึกเอาไว้ล่วงหน้า Level 2 - Multi-domain spoken queries สามารถใช้ภาษาต่าง ๆ ในการส่งข้อความ และการสั่งงานด้วยเสียงได้ คำถามอาจมีความหลากหลาย และคำตอบจะถูกสร้างขึ้น และไม่ถูกเก็บไว้ Level 3 - Fully open-ended, with user modelling, routine learning and anticipation เป็นระดับที่มีความก้าวหน้ามากที่สุด ไม่จำเพาะต่อประเภทของการมีปฏิสัมพันธ์ และคำถามจากผู้ใช้งาน ระบบอาจมีลักษณะการทำงานในเชิงรุกมากกว่าเป็นเพียงการโต้ตอบ

แหล่งอ้างอิง:

[1] JRC Technical report: AI Watch Assessing Technology Readiness Levels for Artificial Intelligence, EU Commission (Martinez-Plumed et al., 2020)

[2] Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (Society of Automotive Engineers International, 2021)

[3] Medical robotics - Regulatory, ethical, and legal considerations for increasing levels of autonomy, American Association for the Advancement of Science (Yang et al., 2017)



ภาพที่ 2 แนวทางการนำแนวปฏิบัติจริยธรรมด้านปัญญาประดิษฐ์ ของ สวทช. ไปใช้ดำเนินการ

ตารางที่ 6 ตัวอย่างรายชื่อกรอบแนวทางการดำเนินงานด้านจริยธรรมปัญญาประดิษฐ์ของหน่วยงานต่าง ๆ

กรอบแนวทางการดำเนินงาน	หน่วยงานที่จัดทำ	ปีที่จัดทำ	แหล่งที่มา
Tenets	Partnership on AI	2016	https://partnershiponai.org/
Top 10 Principles for Ethical Artificial Intelligence	UNI Global Union	2017	http://www.thefutureworldofwork.org/media/35420/uni_ethical_ai.pdf
AI Policy Principles	Information Technology Industry Council	2017	https://www.itic.org/resources/AI-Policy-Principles-FullReport2.pdf
Asilomar AI Principles	Future of Life Institute	2017	https://futureoflife.org/ai-principles/?cn-reloaded=1
AI Principles	Microsoft	2018	https://www.microsoft.com/en-us/ai/responsible-ai
AI at Google: Our Principles	Google	2018	https://ai.google/responsibilities/
AI in the UK: Ready, Willing and Able?	UK House of Lords, Select Committee on Artificial Intelligence	2018	https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf
Toronto Declaration: Protecting the Right to Equality and Non-Discrimination in Machine Learning Systems	Amnesty International, Access Now	2018	https://www.accessnow.org/cms/assets/uploads/2018/08/The-Toronto-Declaration_ENG_08-2018.pdf
Algorithmic Impact Assessments:	AI Now Institute	2018	https://ainowinstitute.org/aiareport2018.pdf

กรอบแนวทางการดำเนินงาน	หน่วยงานที่จัดทำ	ปีที่จัดทำ	แหล่งที่มา
A Practical Framework for Public Agency Accountability			
IBM Everyday Ethics for AI	IBM	2019	https://www.ibm.com/design/ai/ethics/everyday-ethics/
Artificial Intelligence Principles and Ethics	Smart Dubai	2019	https://www.digitaldubai.ae/initiatives/ai-principles-ethics
Principles to Promote Fairness, Ethics, Accountability and Transparency (FEAT) in the Use of Artificial Intelligence and Data Analytics in Singapore's Financial Sector	Monetary Authority of Singapore	2019	https://www.mas.gov.sg/publications/monographs-or-information-paper/2018/FEAT
Social Principles of Human-Centric Artificial Intelligence	Japanese Cabinet Office, Council for Science, Technology and Innovation	2019	https://www8.cao.go.jp/cstp/english/humancentricai.pdf
EU's Ethics Guidelines for Trustworthy AI	European Commission	2019	https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai
OECD Principles on Artificial Intelligence	OECD	2019	https://www.oecd.org/digital/artificial-intelligence/ai-principles/
Ethically Aligned Design: A Vision for Prioritizing Human Well-Being with Autonomous and Intelligent Systems	IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems	2019	https://ethicsinaction.ieee.org/

กรอบแนวทางการดำเนินงาน	หน่วยงานที่จัดทำ	ปีที่จัดทำ	แหล่งที่มา
Beijing AI Principles	Beijing Academy of Artificial Intelligence	2019	https://www.baai.ac.cn/news/beijing-ai-principles-en.html
A guide to using artificial intelligence in the public sector	UK Office for AI, Government Digital Service and the Alan Turing Institute	2019	https://www.gov.uk/government/collections/a-guide-to-using-artificial-intelligence-in-the-public-sector
Guidelines for AI procurement	UK Office for AI and World Economic Forum	2020	https://www.gov.uk/government/publications/guidelines-for-ai-procurement
The Aletheia Framework ™	University of California Presidential Working Group on AI	2020	https://www.rolls-royce.com/sustainability/ethics-and-compliance/the-aletheia-framework.aspx
The future regulation of technology: EU AI Regulation Handbook	DLA Piper	2021	https://www.dlapiper.com/~/media/files/insights/publications/2021/05/ai-regs-handbook.pdf
Responsible AI #AIForAll Approach Document for India	National Institution for Transforming India	2021	https://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf https://www.niti.gov.in/sites/default/files/2021-08/Part2-Responsible-AI-12082021.pdf
Recommendations to Guide the University of	University of California Presidential	2021	https://www.ucop.edu/ethics-compliance-audit-

กรอบแนวทางการดำเนินงาน	หน่วยงานที่จัดทำ	ปีที่จัดทำ	แหล่งที่มา
California's Artificial Intelligence Strategy	Working Group on AI		services/compliance/uc-ai-working-group-final-report.pdf