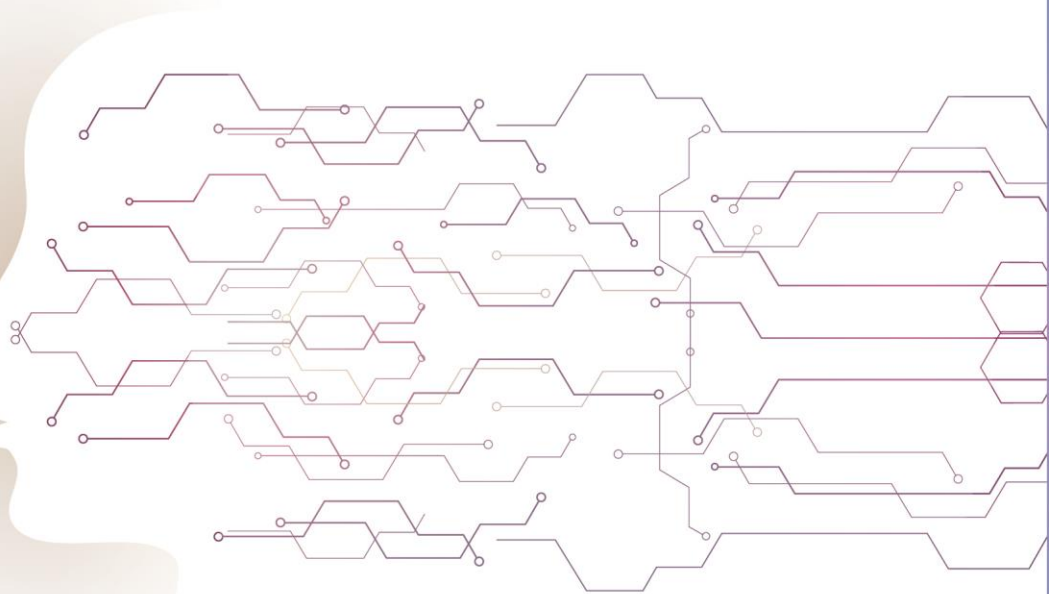


Digital Thailand

# AI ETHICS GUIDELINE



ARTIFICIAL INTELLIGENCE



## คำนำ (Introduction)

เราคงปฏิเสธไม่ได้ว่าปัจจุบันนี้เทคโนโลยีปัญญาประดิษฐ์ได้เข้ามามีบทบาทในกิจกรรมด้านต่าง ๆ ของมนุษย์เพิ่มมากขึ้น สิ่งต่าง ๆ ที่เราใช้งานกันในชีวิตประจำวัน อาทิ เทคโนโลยีผู้ช่วยอัจฉริยะบนสมาร์ตโฟน อย่าง Siri, Alexa และ Google Assistant เทคโนโลยีกึ่งอัตโนมัติในรถยนต์ Tesla หรือเป็น Social Media Feed บน Facebook และ Snapchat บริการเพลงหรือภาพยนตร์บน YouTube และ Netflix ล้วนมีเทคโนโลยีปัญญาประดิษฐ์ทำงานอยู่เบื้องหลังเพื่อคอยส่งข้อมูลและแจ้งเตือนผู้ใช้งาน โดยอาศัยข้อมูลพฤติกรรมการใช้งานของผู้ใช้งานมาปรับการส่งข้อมูลให้เหมาะสมกับความต้องการ ไปจนถึงเทคโนโลยีการนำทางที่เราใช้กันเกือบทุกวันอย่าง Google Maps ก็ใช้เทคโนโลยีปัญญาประดิษฐ์ ในการประมวลผลข้อมูลจราจรแบบเรียลไทม์เพื่อหาเส้นทางที่ดีที่สุดในการไปถึงจุดหมายปลายทาง เป็นต้น

นอกเหนือจากนี้ในองค์กรที่ให้ความสำคัญด้านข้อมูลและเทคโนโลยีสารสนเทศในประเทศไทย อาทิ ธนาคาร บริษัทหลักทรัพย์ ประกันภัย สายการบิน กองทัพ ยังมีการใช้งานเครื่องมือในการรักษาความมั่นคงปลอดภัยของข้อมูล เพื่อป้องกันและตรวจจับภัยคุกคามทางไซเบอร์ รวมถึงเทคโนโลยีในการตรวจสอบการฉ้อโกงหรือพฤติกรรมไม่พึงประสงค์ต่าง ๆ ซึ่งมีการใช้เทคโนโลยีปัญญาประดิษฐ์เป็นส่วนหนึ่งของเครื่องมือในการประมวลผลและตัดสินใจในการทำงาน หรือแม้แต่เป็นเทคโนโลยีหลักในการตรวจสอบและป้องกันภัยคุกคามในหลายผลิตภัณฑ์ อาทิ Cylance Anti-Malware เป็นต้น

อย่างไรก็ดีเทคโนโลยีปัญญาประดิษฐ์นั้นมีการพัฒนาเพิ่มขึ้นเรื่อย ๆ จนสามารถเอาชนะความสามารถของมนุษย์ได้ในหลายด้าน อาทิ Alpha Go ซึ่งเป็นปัญญาประดิษฐ์ที่สร้างขึ้นโดย Google DeepMind<sup>1</sup> ที่สามารถเอาชนะแชมป์โลกเกมหมากล้อมได้ในปี 2560 Watson ซึ่งเป็น

---

<sup>1</sup> David Silver et al. Mastering the game of Go with Deep neural networks and tree search. Nature. Jan 2016. ดูได้จาก <https://storage.googleapis.com/deepmind-media/alphago/AlphaGoNaturePaper.pdf>.

ปัญญาประดิษฐ์ที่สร้างขึ้นโดย IBM ซึ่งสามารถวินิจฉัยโรคได้แม่นยำกว่าแพทย์ที่มีความเชี่ยวชาญ<sup>2</sup> จนก่อให้เกิดความกังวลว่าเทคโนโลยีปัญญาประดิษฐ์จะมีโอกาสเข้ามาแทนที่งานในหลายด้านของมนุษย์ และทำให้มนุษย์ต้องตกงานในหลายสาขาอาชีพ นอกจากนี้ความกังวลดังกล่าวแล้วการออกแบบและพัฒนาปัญญาประดิษฐ์เอง ยังมีโอกาสทำให้ผลลัพธ์เอนเอียงและก่อให้เกิดความไม่เป็นธรรมขึ้นได้ ตั้งแต่การใช้เสียงสตรี ในการพูดโต้ตอบโดยปริยายในเทคโนโลยีผู้ช่วยอัจฉริยะของบริษัทผู้ผลิตต่าง ๆ<sup>3</sup> จนถึงโปรแกรมที่ส่งผลกระทบต่อชีวิตของมนุษย์ ตัวอย่างเช่น อัลกอริทึมปัญญาประดิษฐ์ในโปรแกรม COMPAS ของบริษัท Northpointe ที่ถูกใช้งานโดยศาลในสหรัฐ ซึ่งได้รับการตรวจสอบพบว่าระบบมีความเอนเอียงไปในทางที่จะตัดสินว่าจำเลยผิดซ้ำมีโอกาสจะเป็นผู้กระทำผิดซ้ำมากกว่าจำเลยผิวขาว<sup>4</sup> เป็นต้น การนำปัญญาประดิษฐ์ไปใช้ในทางที่ผิดทั้งในแง่จริยธรรมและกฎหมายก็เป็นประเด็นที่สำคัญมากเช่นเดียวกัน ซึ่ง DeepLocker โปรแกรมไม่เพียงประสงค์ที่อาศัยเทคโนโลยีปัญญาประดิษฐ์ที่ได้รับการสร้างโดย IBM โดยมีวัตถุประสงค์เพื่อทำความเข้าใจถึงโปรแกรมไม่เพียงประสงค์ในลักษณะนี้ว่าจะสร้างผลกระทบได้อย่างไรบ้าง<sup>5</sup> เป็นตัวอย่างหนึ่งที่ทำให้เราต้องตระหนักถึงความจำเป็นในการพิจารณาเฝ้าระวังการพัฒนาปัญญาประดิษฐ์นับต่อจากนี้ไป

เอกสารหลักการและแนวทางจริยธรรมปัญญาประดิษฐ์ฉบับนี้จึงได้รับการสร้างขึ้นเพื่อให้ผู้วิจัย ผู้ออกแบบ ผู้พัฒนา และผู้ให้บริการปัญญาประดิษฐ์ ใช้เป็นแนวทางสำหรับการดำเนินงานของตนเอง และให้ผู้รับบริการได้ทราบถึงสิทธิและตระหนักรู้ถึงความเสี่ยงของการใช้บริการปัญญาประดิษฐ์ หน่วยงานรัฐและหน่วยงานกำกับดูแลปัญญาประดิษฐ์ ทั้งระดับประเทศและ

<sup>2</sup> Monegain B. IBM Watson pinpoints rare form of leukemia after doctor misdiagnosed patient. Healthcare IT News. Aug 8 2016. จาก <https://www.healthcareitnews.com/news/ibm-watson-pinpoints-rare-form-leukemia-after-doctors-misdiagnosed-patient>.

<sup>3</sup> UNESCO and EQUALS Skills Coalition. I'd blush if I cloud: Closing gender divides in digital skills through education. 2019. จาก <https://unesdoc.unesco.org/ark:/48223/pf0000367416>.

<sup>4</sup> Julia Angwin., Jeff Larson., Surya Mattu. and Lauren Kirchner. Machine Bias: There's software used across the country to predict future criminals. And it's biased against blacks. ProPublica. May 23 2016. จาก <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.

<sup>5</sup> Dhillung Kirat., Juyong Jang., and Marc Ph. Stoecklin. DeepLocker – Concealing Targeted Attacks with AI Locksmithing. Blackhat USA 2018. Aug 9 2018. จาก <https://www.blackhat.com/us-18/briefings/schedule/#deeplocker---concealing-targeted-attacks-with-ai-locksmithing-11549>.

ระดับองค์กร เพื่อใช้เป็นแนวทางในการส่งเสริม สนับสนุน รวมถึงกำกับดูแลเทคโนโลยี ปัญญาประดิษฐ์ เพื่อให้ปัญญาประดิษฐ์มีความน่าเชื่อถือ มั่นคงปลอดภัย ได้รับการพัฒนาและใช้งาน ก่อให้เกิดประโยชน์กับมนุษย์ สังคมและสิ่งแวดล้อม ด้วยความโปร่งใส ครอบคลุมและเป็นธรรม สอดคล้องกับกฎหมาย จริยธรรมและสิทธิมนุษยชน รวมถึงสอดคล้องกับยุทธศาสตร์ชาติ 20 ปี (2561 – 2580) ยุทธศาสตร์ที่ 2 ที่ได้กล่าวถึงการใช้เทคโนโลยีดิจิทัล ข้อมูลและปัญญาประดิษฐ์ ในการเพิ่มศักยภาพและความสามารถในการแข่งขันของอุตสาหกรรมและบริการ<sup>6</sup> และยุทธศาสตร์กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม พ.ศ. 2563 – 2567 ยุทธศาสตร์ที่ 1 เรื่องการพัฒนาโครงสร้างพื้นฐานดิจิทัลของประเทศ ซึ่งมีกลยุทธ์ที่จะพัฒนาและส่งเสริมการลงทุนและการใช้ประโยชน์จาก อินเทอร์เน็ตในทุกสิ่ง (IOT) และปัญญาประดิษฐ์ (AI)<sup>7</sup>

---

<sup>6</sup> ประกาศ เรื่อง ยุทธศาสตร์ชาติ (พ.ศ. 2561 – 2580). ราชกิจจานุเบกษา. 13 ตุลาคม พ.ศ. 2561 เล่ม 135 ตอนที่ 82 ก. หน้า 25. จาก [http://www.ratchakittha.soc.go.th/DATA/PDF/2561/A/082/T\\_0001.PDF](http://www.ratchakittha.soc.go.th/DATA/PDF/2561/A/082/T_0001.PDF).

<sup>7</sup> แผนยุทธศาสตร์กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม พ.ศ. 2563 – 2567. หน้า 5. จาก <http://www.mdes.go.th/assets/portals/1/files/%E0%B9%81%E0%B8%9C%E0%B8%99%E0%B8%A2%E0%B8%B8%E0%B8%97%E0%B8%98%E0%B8%A8%E0%B8%B2%E0%B8%AA%E0%B8%95%E0%B8%A3%E0%B9%8C%20%E0%B8%94%E0%B8%A8.%2063-67%20final%281%29.pdf>.

## สารบัญ (Table of Content)

ก

คำนำ

1

การศึกษาและวิเคราะห์  
(Study and analysis)

8

หลักการทางจริยธรรมปัญญาประดิษฐ์  
(AI Ethic Principles)

16

แนวทางปฏิบัติทางจริยธรรม  
ปัญญาประดิษฐ์ (AI Ethic Guidelines)

36

อภิธานศัพท์ (Glossary)

38

เอกสารอ้างอิง

## การศึกษาและวิเคราะห์ (Study and analysis)

เอกสารหลักการและแนวทางจริยธรรมปัญญาประดิษฐ์ของประเทศไทยฉบับนี้ ได้ศึกษารวบรวม และวิเคราะห์หลักการและแนวทางจริยธรรมปัญญาประดิษฐ์ของประเทศอื่น ๆ และองค์กรที่ได้รับการยอมรับทางด้านเทคโนโลยีหลาย ๆ แห่ง โดยพบว่าทุกประเทศและทุกองค์กรต่างมีหลักการและแนวทางปฏิบัติทางจริยธรรมปัญญาประดิษฐ์ที่สอดคล้องกันไปในแนวทางเดียวกัน แต่ก็มีลักษณะเฉพาะที่แต่ละประเทศหรือองค์กรกำหนด และก่อให้เกิดความแตกต่างกันโดยสรุป มีดังนี้ คือ

องค์กรเพื่อความร่วมมือและพัฒนาทางเศรษฐกิจ หรือ โออีซีดี (Organization for Economic Co-operation and Development: OECD) ได้ออกคำแนะนำหลักการทางจริยธรรมของ AI ว่าการดูแลและให้บริการระบบปัญญาประดิษฐ์ควรมีความน่าเชื่อถือ สร้างประโยชน์ให้กับผู้คนและโลกโดยทำให้เกิดความเจริญ การพัฒนาอย่างยั่งยืน เคารพและไม่กระทำการละเมิดต่อกฎหมาย สิทธิมนุษยชน ค่านิยมและความหลากหลายในระบอบประชาธิปไตย สามารถเปิดให้มนุษย์สามารถเข้าไปแทรกแซงได้เมื่อมีความจำเป็น เพื่อรักษาไว้ซึ่งความเป็นธรรมของสังคม เอื้ออำนวยหน่วยงานทั้งภาครัฐและเอกชนในการวิจัยและพัฒนาเพื่อกระตุ้นให้เกิดนวัตกรรมปัญญาประดิษฐ์<sup>8</sup>

สหภาพยุโรป (Council of Europe) ได้ออกแนวทางเกี่ยวกับปัญญาประดิษฐ์กับการป้องกันข้อมูลไว้ โดยกล่าวถึง การพัฒนาและนำปัญญาประดิษฐ์ไปประยุกต์ใช้ที่อาจส่งผลกระทบต่อบุคคลและสังคม มีความจำเป็นต้องมีการป้องกันเกียรติ (Dignity) ปกป้องสิทธิมนุษยชน (Human rights) เสรีภาพขั้นพื้นฐาน (Fundamental freedoms) และสิทธิในการป้องกันข้อมูลส่วนตัว (Personal data protection) โดยเฉพาะเมื่อปัญญาประดิษฐ์ถูกนำไปใช้ในกระบวนการ

<sup>8</sup> OECD Legal Instruments. Recommendation of the Council on Artificial Intelligence. May 22 2019. จาก <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>.

ตัดสินใจ และเสนอให้ผู้พัฒนานำไปใช้ นอกจากนี้ องค์กรผู้ผลิตสินค้า ผู้ให้บริการระบบ ปัญญาประดิษฐ์ ควรรับคำปรึกษาจากคณะกรรมการอิสระ ผู้เชี่ยวชาญและสถาบันการศึกษาใน แขนงที่จะนำระบบปัญญาประดิษฐ์ไปใช้งาน เพื่อช่วยให้การออกแบบระบบปัญญาประดิษฐ์ คำนึงถึงสิทธิมนุษยชน (Human rights) จริยธรรม และสังคม และช่วยในการตรวจสอบความเอนเอียงที่สำคัญ ควรสร้างกลไกที่ทำให้หน่วยงานควบคุมดูแลการพัฒนาและใช้งานระบบ ปัญญาประดิษฐ์มีคณะกรรมการผู้เชี่ยวชาญที่เป็นอิสระ ส่งเสริมให้เกิดความร่วมมือระหว่าง หน่วยงานควบคุมดูแลการพัฒนาและใช้งานระบบปัญญาประดิษฐ์กับหน่วยงานอื่นที่มีความสามารถที่เกี่ยวข้องกับปัญญาประดิษฐ์ อาทิ หน่วยงานคุ้มครองผู้บริโภค หน่วยงาน ที่เกี่ยวข้องการห้ามเลือกปฏิบัติ (Anti-Discrimination) หน่วยงานกำกับดูแลในแขนงต่าง ๆ รวมถึงหน่วยงานกำกับดูแลกิจการสื่อ<sup>9</sup>

คณะกรรมการยุโรป (European Commission) กล่าวถึงแนวทางจริยธรรมสำหรับ ปัญญาประดิษฐ์ที่น่าเชื่อถือ ว่าควรให้เคารพต่อเสรีภาพของมนุษย์ ปกป้องกันภัยคุกคาม มีความเป็น ธรรม และมีความสามารถในการอธิบายได้ และคุณสมบัติที่ต้องการของระบบปัญญาประดิษฐ์ ที่น่าเชื่อถือ คือ ปัญญาประดิษฐ์ต้องให้การสนับสนุนต่อมนุษย์ ดูแลสิทธิขั้นพื้นฐาน และอนุญาตให้ มนุษย์ควบคุมดูแล มีความแข็งแกร่งและความมั่นคงปลอดภัยในทางเทคนิคของระบบ สามารถ ป้องกันการคุกคามจากผู้ไม่ประสงค์ดีได้ มีความโปร่งใส สามารถอธิบายได้ มีกลไกเพื่อสร้างความ รับผิดชอบและภาระความรับผิดชอบผลกระทบที่เกิดจากผลลัพธ์ของปัญญาประดิษฐ์ทั้งก่อนและ หลังการพัฒนา ให้บริการ ใช้งานระบบ และออกแบบโดยใช้หลักการที่ให้ผู้ใช้งานเป็นศูนย์กลาง โดยอนุญาตให้ผู้ใช้งานในทุกระดับได้ใช้ผลิตภัณฑ์และบริการนี้ ไม่จำกัดอายุ เพศ ความสามารถ และคุณสมบัติ โดยเฉพาะผู้พิการควรได้รับความเท่าเทียมในการใช้งานเช่นเดียวกันกับบุคคล ทั่วไป<sup>10</sup>

<sup>9</sup> Council of Europe. Guidelines on Artificial Intelligence and Data Protection. Consultative Committee of the Convention for the Protection of individuals with regard to Automatic Processing of Personal Data. Jan 25 2019.

<sup>10</sup> European Commission. Ethics Guidelines for Trustworthy AI. Apr 8 2019. จาก <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>.

ประธานาธิบดีโดนัลด์ ทรัมป์ ของประเทศสหรัฐอเมริกา ได้ออกคำสั่งพิเศษ (Executive Order) เพื่อรักษาความเป็นผู้นำด้านปัญญาประดิษฐ์ของสหรัฐอเมริกา โดยกล่าวว่าสหรัฐต้องขับเคลื่อนการพัฒนามาตรฐานทางเทคนิคที่เหมาะสมและลดอุปสรรคให้กับการทดสอบและให้บริการเทคโนโลยีปัญญาประดิษฐ์ที่ปลอดภัย เพื่อเปิดให้เกิดการสร้างอุตสาหกรรมใหม่ๆ ที่เกี่ยวข้องกัปัญญาประดิษฐ์และการนำปัญญาประดิษฐ์ไปใช้ในอุตสาหกรรมปัจจุบัน สนับสนุนการสร้างความน่าเชื่อถือและความมั่นใจในเทคโนโลยีปัญญาประดิษฐ์ต่อสาธารณะ และป้องกันสิทธิเสรีภาพ ความเป็นส่วนตัว และค่านิยมอเมริกันจากการใช้เทคโนโลยีปัญญาประดิษฐ์ โดยสหรัฐต้องให้การอบรมคนทำงานทั้งในปัจจุบันในอนาคตให้มีทักษะในการพัฒนาและประยุกต์ใช้เทคโนโลยีปัญญาประดิษฐ์ เพื่อเตรียมความพร้อมสำหรับเศรษฐกิจในปัจจุบันและงานในอนาคต ส่งเสริมการเข้าถึงข้อมูล โมเดล และทรัพยากรในการประมวลผลของรัฐบาลกลางสหรัฐ ที่มีคุณภาพสูง และสามารถสืบย้อนได้อย่างสมบูรณ์ เพื่อเพิ่มคุณค่าให้กับทรัพยากรเหล่านั้นในด้านการวิจัยและพัฒนาปัญญาประดิษฐ์ ในขณะที่ยังคงรักษาความมั่นคงปลอดภัย ความเป็นส่วนตัว และความลับ ที่สอดคล้องกับนโยบายและกฎหมายที่เกี่ยวข้อง อีกทั้งยังต้องลดอุปสรรคในการใช้งานปัญญาประดิษฐ์เพื่อสนับสนุนนวัตกรรมที่เกิดจากเทคโนโลยีปัญญาประดิษฐ์นี้ด้วย<sup>11</sup>

หน่วยงาน Smart Dubai ของนครดูไบ ได้ออกเอกสารหลักการและแนวทางจริยธรรมปัญญาประดิษฐ์ โดยกล่าวถึงความเป็นธรรมของชุดข้อมูลที่นำมาใช้งานควรเป็นตัวแทนของประชากรที่ได้รับผลกระทบทั้งหมดได้จริง ความมั่นคงปลอดภัยของผู้คนที่ปฏิบัติงาน ผู้ใช้งาน และบุคคลอื่น ๆ ที่เกี่ยวข้องกัระบบปัญญาประดิษฐ์ ควรเป็นส่วนที่สำคัญในการออกแบบระบบ ควรมีความร่วมมือกันระหว่างประเทศต่าง ๆ เพื่อหลีกเลี่ยงการแข่งขันทางอาวุธอัตโนมัติที่ร้ายแรง (Lethal autonomous weapons) และอาวุธเหล่านั้นควรได้รับการควบคุม ควรมีการร่วมมือกันในการสร้างมาตรฐานด้านความปลอดภัยของระบบ ส่วนการพัฒนาปัญญาประดิษฐ์ที่มีความสามารถปรับปรุงตนเองได้อย่างต่อเนื่อง (Recursively self-improving AI) ควรได้รับการเปิดเผย ฝ้าระวังและควบคุมความเสี่ยงอย่างเข้มข้น ควรให้การศึกษาแก่ประชาชนอย่างต่อเนื่องเพื่อพัฒนาและให้ความรู้ถึงการพัฒนาล่าสุดของปัญญาประดิษฐ์ เพื่อให้ประชาชนสามารถ

<sup>11</sup> The White House. Executive Order on Maintaining American Leadership in AI. Feb 11 2019. จาก

<https://www.whitehouse.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/>.



ปรับตัวเข้ากับการเปลี่ยนแปลงทางสังคมได้ ในขณะที่ผู้ให้บริการระบบปัญญาประดิษฐ์เอง ควรมีความโปร่งใสในการเปิดเผยเกี่ยวกับข้อมูลและอัลกอริทึมที่ใช้งาน โดยไม่ขัดกับความเป็นส่วนตัวและการรักษาไว้ซึ่งทรัพย์สินทางปัญญา และภาครัฐควรสนับสนุนในการสร้างมาตรฐานที่เป็นที่ยอมรับโดยสากล (Internationally recognized standards) และแนวทางปฏิบัติที่ดีที่สุด (Best Practice) สำหรับปัญญาประดิษฐ์ และควรกำหนดให้ต้องปฏิบัติตามอย่างเคร่งครัด<sup>12</sup>

คณะกรรมการป้องกันข้อมูลส่วนบุคคลของประเทศสิงคโปร์ ได้ออกกรอบการกำกับดูแลปัญญาประดิษฐ์ไว้ โดยกล่าวว่า ระบบปัญญาประดิษฐ์ต้องยึดมนุษย์เป็นศูนย์กลาง (Human Centric) การปกป้องความผาสุก (Well-beings) และความปลอดภัย (Safety) ควรเป็นสิ่งที่ได้รับการพิจารณาเป็นประการแรกในการออกแบบ พัฒนา และให้บริการปัญญาประดิษฐ์ องค์กรที่นำปัญญาประดิษฐ์ไปให้บริการควรกำหนดรายละเอียดหลักการจริยธรรมไว้ในกระบวนการปฏิบัติงานหรือใช้เป็นข้อกำหนดหนึ่งในผลิตภัณฑ์และบริการของตน และองค์กรควรกำหนดให้มีโครงสร้างการกำกับดูแลภายในและการตรวจสอบเพื่อใช้ในการควบคุมดูแลการ ใช้งานปัญญาประดิษฐ์ในองค์กร ซึ่งสามารถใช้โครงสร้างการกำกับดูแลที่มีอยู่เดิมหรือจัดตั้งโครงสร้างใหม่ก็ได้หากมีความจำเป็น โดยได้รับการสนับสนุนจากคณะกรรมการบริหารและผู้บริหารระดับสูงขององค์กร ควรมีการทบทวนประสิทธิภาพของโครงสร้างการกำกับดูแลอย่างสม่ำเสมอและทุกครึ่งปีมีการเปลี่ยนโครงสร้างหรือบุคคลที่สำคัญ และองค์กรควรกำหนดนโยบายและกระบวนการในการปรับปรุงโมเดลอย่างสม่ำเสมอเพื่อให้สอดคล้องกับการเปลี่ยนแปลงพฤติกรรมของผู้ใช้งานตามระยะเวลาและปรับปรุงโมเดลด้วยชุดข้อมูลการสอนที่ทันสมัย นอกจากนี้ การปรับปรุงโมเดลควรได้รับการดำเนินการเมื่อวัตถุประสงค์ ความเสี่ยง และค่านิยมขององค์กรเปลี่ยนแปลงไป และมีการสื่อสารเพื่ออธิบายระบบกับผู้ใช้งานด้วยภาษาที่ง่ายต่อการเข้าใจ โดยใช้เครื่องมือทดสอบความสามารถในการทำความเข้าใจเพื่อทดสอบระบบต่าง ๆ อาทิ Fly Graph Readability Formula, Gunning Fog Index, Flesh-Kincaid Readability tests เป็นต้น<sup>13</sup>

<sup>12</sup> Smart Dubai. AI Ethics Principles & Guidelines. Version 1.6. Dec 30 2018.

<sup>13</sup> Personal Data Protection Commission, Singapore. A Proposed Model Artificial Intelligence Governance Framework. Jan 2019.

องค์กรเพื่อการศึกษาวิทยาศาสตร์และวัฒนธรรมแห่งสหประชาชาติ (UNESCO) ร่วมกับ EQUALS Global Partnership ได้ออกบทความที่ให้คำแนะนำเพื่อช่วยป้องกันเทคโนโลยีผู้ช่วยดิจิทัลจากปัญหาความเอนเอียงทางเพศ โดยสร้างกลไกและบันทึกหลักฐานเพื่อตรวจสอบอัลกอริทึม เพื่อระบุความเสี่ยงด้านความเอนเอียงทางเพศ และหาแนวทางในการแก้ไขหรือป้องกัน ตรวจสอบอิทธิพลที่เกี่ยวข้องกับเพศของผู้ช่วยดิจิทัลที่มีต่อพฤติกรรมของชายและหญิงทั้งในโลกออนไลน์และออฟไลน์ โดยเฉพาะผลกระทบที่เกิดขึ้นกับเด็กและเยาวชน สร้างเครื่องมือกฎระเบียบและกระบวนการ เพื่อสนับสนุนให้เกิดการสร้างโค้ด ข้อมูล และโปรโตคอลเปิดที่ใช้เพื่อการพัฒนาผู้ช่วยดิจิทัลหรือระบบปัญญาประดิษฐ์ที่มีความอ่อนไหวทางเพศ ปรับปรุงเทคนิคในการสอนและอบรมผู้พัฒนาระบบปัญญาประดิษฐ์ให้ตอบสนองต่อผู้ใช้งานเกี่ยวกับเพศอย่างเป็นกลาง ให้การสนับสนุนการพัฒนาทักษะทางดิจิทัลกับทุกเพศอย่างเท่าเทียม โดยเฉพาะเพศหญิงซึ่งมีจำนวนผู้มีทักษะทางดิจิทัลน้อยกว่าเมื่อเทียบกับเพศชาย สร้างกลไกการระดมความคิดที่เหมาะสมและควบคุมดูแลเพื่อป้องกันหรือลดอัลกอริทึมที่มีความเอนเอียงและละเมิดสิทธิส่วนบุคคลต่าง ๆ<sup>14</sup>

บริษัท ไมโครซอฟท์ (Microsoft) ได้จัดทำหลักการและแนวทางจริยธรรมปัญญาประดิษฐ์ในหลายประเภทซอฟต์แวร์ อาทิ ในภาพรวมของปัญญาประดิษฐ์ บริษัท ไมโครซอฟท์กล่าวถึง “The Future Computed, Artificial Intelligence and its role in society” โดยระบบปัญญาประดิษฐ์ควรให้บริการกับทุกคนด้วยความเป็นธรรมและไม่สร้างความแตกต่างในด้านผลลัพธ์กับกลุ่มคนที่อยู่ในสถานการณ์เดียวกัน ข้อมูลชุดการสอนระบบต้องไม่มีความเอนเอียงไปในทางใดทางหนึ่งทั้งในด้านอายุ เพศ ลัทธิ ชนชาติ ปัญญาประดิษฐ์ควรได้รับการออกแบบเพื่อใช้งานด้วยกลุ่มตัวแปรที่ชัดเจนภายใต้เงื่อนไขทางประสิทธิภาพที่คาดหวังไว้ และมีแนวทางในการพิสูจน์ว่าระบบมีพฤติกรรมตามที่ได้ตั้งใจและออกแบบไว้ภายใต้เงื่อนไขการปฏิบัติงานจริง ปัญญาประดิษฐ์ต้องสอดคล้องกับกฎหมายคุ้มครองข้อมูลส่วนบุคคลที่กำหนดให้มีความโปร่งใสในการรวบรวม จัดเก็บ นำไปใช้ และการเผยแพร่ข้อมูล โดยเจ้าของข้อมูลมีสิทธิในการควบคุมเพื่อเลือกได้ว่าข้อมูลของตนจะถูกใช้อย่างไร และสามารถอธิบาย ทำความเข้าใจถึงการดำเนินงาน

<sup>14</sup> UNESCO and EQUALS Skills Coalition. I'd blush if I could: Closing gender divides in digital skills through education. 2019. จาก <https://unesdoc.unesco.org/ark:/48223/pf0000367416>.

และปฏิสัมพันธ์กับข้อมูลของระบบปัญญาประดิษฐ์แก่ผู้ใช้งานและผู้ที่ได้รับผลกระทบจากระบบ เพื่อให้ผู้ใช้งานสามารถตรวจพบและตระหนักถึงความเอนเอียง ความผิดพลาด และผลลัพธ์ที่ไม่ได้ตั้งใจได้โดยง่าย ในกรณีที่ผู้ใช้งานปัญญาประดิษฐ์เพื่อให้คำแนะนำหรือคาดการณ์เหตุการณ์ ควรมีภาระความรับผิดชอบสำหรับการตัดสินใจเหล่านั้นเป็นอันดับต้นๆ เสมอ<sup>15</sup> และจริยธรรมในซอฟต์แวร์ประเภทจดจำใบหน้า (Microsoft’s facial recognition) ไมโครซอฟท์ได้กล่าวถึง ความเท่าเทียม ความโปร่งใส การรับผิดชอบต่อผลกระทบ การไม่เลือกปฏิบัติ มีการแจ้งเตือนและต้องได้รับความยินยอมจากเจ้าของข้อมูล และมีการสอดส่องอย่างถูกต้องตามกฎหมาย<sup>16</sup> ในขณะที่ไมโครซอฟท์กล่าวถึงจริยธรรมในซอฟต์แวร์ประเภทสนทนาโต้ตอบอัตโนมัติ (Conversational AI) ไว้ว่า การพัฒนาปัญญาประดิษฐ์ต้องมีการกำหนดวัตถุประสงค์ของโปรแกรม ChatBot หรือโปรแกรมที่พัฒนาขึ้นเพื่อจำลองบทสนทนาของมนุษย์ ให้สามารถพูดคุย สื่อสารกับมนุษย์ผ่านทางเสียงหรือข้อความแบบ real-time โดยวัตถุประสงค์ของโปรแกรม ChatBot ต้องสอดคล้องกับจริยธรรมและให้ความสำคัญ หากโปรแกรม ChatBot ถูกใช้เพื่อสนับสนุนการทำงานที่มีผลกระทบสำคัญ จะต้องมียุทธศาสตร์ตรวจสอบและตอบสนองต่อเรื่องที่น่าห่วงหรือก้าวร้าวที่ได้รับจากการปฏิสัมพันธ์กับผู้ใช้งานอย่างเหมาะสม ระบบควรรับคำปรึกษาจากผู้เชี่ยวชาญในแขนงที่เกี่ยวข้องในการพัฒนาระบบที่เกี่ยวข้องกับข้อมูลที่มีความอ่อนไหว อาทิ ข้อมูลทางการแพทย์ ข้อมูลพนักงาน ข้อมูลทางการเงิน และการบังคับใช้กฎหมาย การพัฒนาระบบต้องสอดคล้องกับมาตรฐานหรือแนวทางด้านความสามารถในการเข้าถึงระบบที่เป็นที่ยอมรับ อาทิ ISO 40500:2012 (W3C Web Content Accessibility Guidelines: (WCAG) 2.0) และให้ผู้พิการหรือมีภาวะทุพพลภาพได้ร่วมทดสอบระบบด้วย<sup>17</sup>

<sup>15</sup> Microsoft. The Future Computed, Artificial Intelligence and its role in society. Feb 2019. จาก [https://3er1viui9wo30pkxh1v2nh4w-wpengine.netdna-ssl.com/wp-content/uploads/2018/02/The-Future-Computed\\_2.8.18.pdf](https://3er1viui9wo30pkxh1v2nh4w-wpengine.netdna-ssl.com/wp-content/uploads/2018/02/The-Future-Computed_2.8.18.pdf).

<sup>16</sup> Microsoft. Six principles to guide Microsoft’s facial recognition work. Dec 2018. จาก <https://blogs.microsoft.com/on-the-issues/2018/12/17/six-principles-to-guide-microsofts-facial-recognition-work/>.

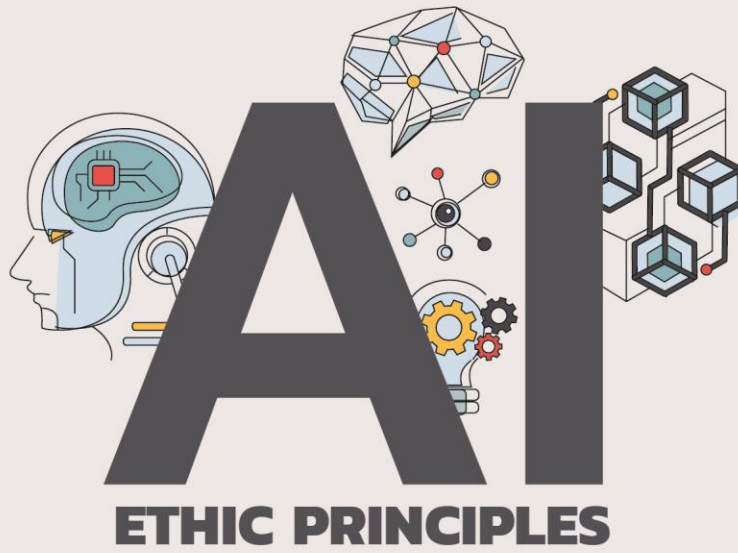
<sup>17</sup> Microsoft. Responsible bots: 10 guidelines for developers of conversational AI. Nov 2018. จาก [https://www.microsoft.com/en-us/research/uploads/prod/2018/11/Bot\\_Guidelines\\_Nov\\_2018.pdf](https://www.microsoft.com/en-us/research/uploads/prod/2018/11/Bot_Guidelines_Nov_2018.pdf).

Beijing Academy of Artificial Intelligence (BAAI) เป็นองค์กรที่ได้รับการสนับสนุนโดยกระทรวงวิทยาศาสตร์และเทคโนโลยีของจีน ได้ออกหลักการทางจริยธรรมของ AI เช่นเดียวกับประเทศและองค์กรอื่น ๆ และเพิ่มการส่งเสริมให้เกิดแพลตฟอร์มเปิดของปัญญาประดิษฐ์เพื่อหลีกเลี่ยงการผูกขาด เพื่อทำให้เกิดการแลกเปลี่ยนประโยชน์ของการพัฒนาปัญญาประดิษฐ์และการต่อยอดอย่างต่อเนื่อง สนับสนุนให้เกิดโอกาสในการพัฒนาที่เท่าเทียมกันในภูมิภาคและอุตสาหกรรมที่แตกต่างกัน<sup>18</sup>

ผลการศึกษาหลักการและแนวทางจริยธรรมปัญญาประดิษฐ์ต่าง ๆ เหล่านี้ ถูกใช้เป็นหลักการและแนวความคิดเบื้องต้นในการจัดทำหลักการและแนวทางจริยธรรมปัญญาประดิษฐ์ของประเทศไทย โดยมุ่งหวังให้ผู้วิจัย ผู้ออกแบบ ผู้พัฒนา ผู้ให้บริการและผู้ใช้งานปัญญาประดิษฐ์ทำการวิจัย ออกแบบ พัฒนา ให้บริการ และใช้งานปัญญาประดิษฐ์ได้โดยคำนึงคุณธรรม จริยธรรม สิทธิ เสรีภาพ ความเป็นมนุษย์ เพื่อประโยชน์ต่อส่วนรวมเป็นสำคัญ

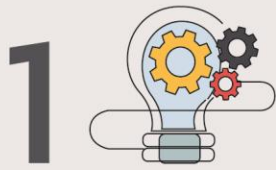
---

<sup>18</sup> Beijing AI Principles. May 5 2019. จาก <https://www.baai.ac.cn/blog/beijing-ai-principles>.



# AI

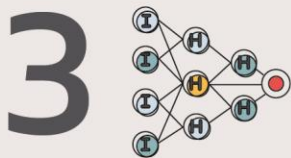
## ETHIC PRINCIPLES



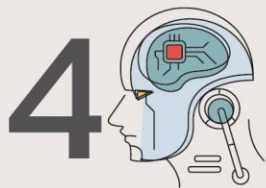
ความสามารถในการแข่งขันและการพัฒนาอย่างยั่งยืน  
(Competitiveness and Sustainability  
Development)



ความสอดคล้องกับกฎหมายจริยธรรมและมาตรฐานสากล  
(Laws Ethics and International Standards)



ความโปร่งใสและภาระความรับผิดชอบ  
(Transparency and Accountability)



ความมั่นคงปลอดภัยและความเป็นส่วนตัว  
(Security and Privacy)



ความเท่าเทียม หลากหลาย ครอบคลุม และเป็นธรรม  
(Fairness)



ความน่าเชื่อถือ  
(Reliability)

# หลักการทางจริยธรรมปัญญาประดิษฐ์

## (AI Ethic Principles)

### 1) ความสามารถในการแข่งขันและการพัฒนาอย่างยั่งยืน (Competitiveness and Sustainability Development)

- ปัญญาประดิษฐ์ควรถูกสร้างและใช้งานเพื่อสร้างประโยชน์และความผาสุกให้แก่ มนุษย์ สังคม เศรษฐกิจและสิ่งแวดล้อมอย่างยั่งยืน
- ปัญญาประดิษฐ์ควรถูกใช้งานเพื่อเพิ่มความสามารถในการแข่งขันและสร้างความ เจริญให้กับมนุษย์ สังคม ประเทศ ภูมิภาค และโลกอย่างเป็นธรรม
- ปัญญาประดิษฐ์ควรได้รับการวิจัยและพัฒนาอย่างต่อเนื่อง เพื่อให้มนุษย์เกิดการ สร้างสรรค์นวัตกรรมและอุตสาหกรรมใหม่

ปัญญาประดิษฐ์สามารถถูกสร้าง พัฒนา และใช้งานเพื่อสร้างประโยชน์ให้กับ มนุษย์ในหลากหลายด้านทั้งในด้านการแพทย์ การเงิน อุตสาหกรรมการผลิตและการ บริการ เป็นต้น แต่ก็สามารถใช้ปัญญาประดิษฐ์เพื่อการคุกคามและเป็นภัยต่อมนุษย์ได้ด้วย เช่นกัน เช่น การแข่งขันเพื่อพัฒนาอาวุธอัตโนมัติที่ร้ายแรง และโปรแกรมคอมพิวเตอร์ที่ไม่ พึ่งประสงค์ เป็นต้น ซึ่งปัจจุบันในประเทศมหาอำนาจอย่างสหรัฐ จีน และรัสเซีย ได้มีการ วิจัยและพัฒนาเพื่อสร้างอาวุธอัตโนมัติโดยใช้เทคโนโลยีปัญญาประดิษฐ์ออกมาในหลาย ประเภทแล้ว ตัวอย่างเช่น บริษัทผลิตอาวุธ Kalashnikov ในรัสเซียได้ออกจำหน่ายอาวุธ ชนิดใหม่ในปี 2560 เป็นโดรนต่อสู้อัตโนมัติที่ใช้โครงข่ายประสาทเทียม (Neural Network) เป็นอัลกอริทึมเพื่อระบุเป้าหมายและตัดสินใจได้ด้วยตนเอง<sup>19</sup> เช่นเดียวกันกับ บริษัท ผลิตอาวุธหลายบริษัทของประเทศสหรัฐอเมริกาที่ใช้ปัญญาประดิษฐ์เพื่อการ แยกแยะ ระบุวัตถุ และตัดสินใจแบบอัตโนมัติต่าง ๆ ในอาวุธของตนเอง<sup>20</sup> เป็นต้น ดังนั้น

<sup>19</sup> David Gilbert. Russian weapons maker Kalashnikov developing killer AI robots. Vice News. Jul 13 2017. จาก [https://news.vice.com/en\\_us/article/vbqz8y/russian-weapons-maker-kalashnikov-developing-killer-ai-robots](https://news.vice.com/en_us/article/vbqz8y/russian-weapons-maker-kalashnikov-developing-killer-ai-robots).

<sup>20</sup> Marcus Roth. AI in Military Drones and UAVs – Current Applications. Emerj. Jan 30 2019. จาก <https://emerj.com/ai-sector-overviews/ai-drones-and-uavs-in-the-military-current-applications/>.

การวิจัย พัฒนา และใช้งานปัญญาประดิษฐ์เพื่อสร้างประโยชน์อย่างยั่งยืนให้กับมนุษย์ และเพิ่มการแข่งขันที่สร้างสรรค์เป็นธรรม จึงควรเป็นหลักการสำคัญที่ผู้วิจัย ผู้ออกแบบ ผู้พัฒนา ผู้ให้บริการ และหน่วยงานกำกับดูแลต่าง ๆ ควรคำนึงถึง

## 2) ความสอดคล้องกับกฎหมาย จริยธรรม และมาตรฐานสากล (Laws Ethics and International Standards)

- ปัญญาประดิษฐ์ควรได้รับการวิจัย ออกแบบ พัฒนา ให้บริการ และใช้งาน สอดคล้องกับกฎหมาย บรรทัดฐาน จริยธรรม คุณธรรมของมนุษย์ และ มาตรฐานสากล โดยเคารพต่อความเป็นส่วนตัว เกียรติ สิทธิเสรีภาพ และสิทธิ มนุษยชน
- ออกแบบปัญญาประดิษฐ์ควรใช้หลักการมนุษย์เป็นศูนย์กลางและเป็นผู้ตัดสินใจ
- ปัญญาประดิษฐ์ไม่ควรถูกใช้ในการกำหนดชะตาชีวิตของมนุษย์

ผลลัพธ์ที่เกิดขึ้นจากการประมวลผลปัญญาประดิษฐ์อาจเป็นสิ่งที่แตกต่างจาก ที่มนุษย์เคยสร้างขึ้น จึงมีความเสี่ยงที่หากนำผลลัพธ์เหล่านั้นไปใช้งานแล้วจะขัดต่อข้อ กฎหมายและจริยธรรมที่มนุษย์ในสังคมหนึ่งๆ ถือปฏิบัติ ซึ่งอาจก่อให้เกิดปัญหาความ ขัดแย้งทางสังคมขึ้นมาได้ โดยเฉพาะในโลกโซเชียลมีเดียที่ข้อมูลข่าวสารสามารถถูก เผยแพร่ออกไปได้อย่างรวดเร็ว กรณีตัวอย่างที่เกิดขึ้นกับบริษัท Google ในเดือนเมษายน ปี 2561 เมื่อมีพนักงานของบริษัทกว่า 3,000 คน ทำการประท้วงต่อเหตุการณ์ที่บริษัททำ สัญญากับกระทรวงกลาโหมสหรัฐอเมริกา เพื่อนำเทคโนโลยีปัญญาประดิษฐ์ของบริษัทไป ใช้ในการวิเคราะห์วีดีโอจากโดรนเพื่อระบุตำแหน่งและสังหารเป้าหมายที่เป็นมนุษย์ เนื่องจากเห็นว่าเป็นการกระทำที่นำเทคโนโลยีปัญญาประดิษฐ์ไปใช้โดยผิดต่อศีลธรรม และมีพนักงานกว่า 12 คนลาออกจากบริษัทจากเหตุการณ์ดังกล่าว<sup>21</sup>

<sup>21</sup> Kate Conger. Google Employees Resign in Protest Against Pentagon Contract. GIZMODO. May 14 2018. จาก <https://gizmodo.com/google-employees-resign-in-protest-against-pentagon-con-1825729300>.



### 3) ความโปร่งใสและภาระความรับผิดชอบ (Transparency and Accountability)

- ปัญญาประดิษฐ์ควรได้รับการวิจัย ออกแบบ พัฒนา ให้บริการและใช้งานด้วยความโปร่งใส สามารถอธิบายและคาดการณ์ได้ รวมถึงสามารถตรวจสอบกิจกรรมต่าง ๆ ที่เกิดขึ้นย้อนหลังได้
- ปัญญาประดิษฐ์ควรมีความสามารถในการสืบย้อนกลับ (Traceability) เผื่อตรวจสอบความผิดปกติและวินิจฉัยปัญหาความล้มเหลวได้ (Diagnosability) ได้
- ผู้วิจัย ผู้ออกแบบ ผู้พัฒนา ผู้ให้บริการและผู้ใช้งานปัญญาประดิษฐ์ ควรมีภาระความรับผิดชอบ (Accountability) ต่อผลกระทบที่เกิดขึ้นจากปัญญาประดิษฐ์ตามภาระหน้าที่ของตน

ด้วยปัญญาประดิษฐ์เป็นเทคโนโลยีขั้นสูงที่ยากสำหรับคนทั่วไปจะเข้าใจกลไกการทำงาน และแม้ว่าในปัจจุบันปัญญาประดิษฐ์จะมีความแม่นยำมากขึ้น แต่ผลลัพธ์จากการประมวลผลปัญญาประดิษฐ์ในโมเดลต่าง ๆ ก็ยังคงมีความผิดพลาดอยู่บ้าง ซึ่งอัตราความผิดพลาดที่เกิดขึ้นจะมีความแตกต่างกันในแต่ละโมเดลและอัลกอริทึมที่ใช้งาน แต่เมื่อมีการนำผลลัพธ์ของการตัดสินใจจากปัญญาประดิษฐ์มาใช้ตัดสินใจผลใด ๆ ผลการตัดสินใจนั้น ๆ สามารถส่งผลกระทบต่อผู้รับบริการและบุคคลที่เกี่ยวข้องได้ตั้งแต่ระดับเล็กน้อยจนถึงระดับร้ายแรง และหากเกิดผลกระทบจากการตัดสินใจโดยปัญญาประดิษฐ์ขึ้นจริง การระบุผู้ที่ต้องรับผิดชอบจะเป็นกระบวนการที่ต้องดำเนินการถัดไป หากเกิดเหตุการณ์ที่ก่อให้เกิดผลกระทบขึ้นจริง ผู้ได้รับผลกระทบอาจเกิดความเคลือบแคลงสงสัยถึงหลักการและเหตุผลที่ปัญญาประดิษฐ์ใช้ในการตัดสินใจ และจะเกิดคำถามที่เกี่ยวข้องกับผู้ที่ต้องรับผิดชอบต่อผลกระทบที่อาจเกิดขึ้น ดังนั้น หากกระบวนการวิจัย พัฒนา และให้บริการปัญญาประดิษฐ์ไม่มีความโปร่งใส และปราศจากกลไกในการสืบย้อนเพื่อหาผู้รับผิดชอบแล้ว อาจส่งผลกระทบต่อความเชื่อมั่นในการใช้งานปัญญาประดิษฐ์ได้ ตัวอย่างเช่น รายงานผลการใช้งานปัญญาประดิษฐ์ในผลิตภัณฑ์ด้านความมั่นคงปลอดภัยทางไซเบอร์ที่ออกโดย Osterman Research ในเดือนธันวาคม ปี 2561 พบว่าร้อยละ 54



ของผู้ตอบแบบสอบถาม ตอบว่าผลลัพธ์การป้องกันและตรวจสอบจากผลิตภัณฑ์ไม่มีความแม่นยำ และร้อยละ 47 ตอบว่าผลิตภัณฑ์มีการตรวจสอบผิดพลาด (False Positive)<sup>22</sup>

#### 4) ความมั่นคงปลอดภัยและความเป็นส่วนตัว (Security and Privacy)

- ปัญญาประดิษฐ์ควรถูกสร้างเพื่อบริการ แต่ไม่ควรถูกใช้เพื่อหลอกลวง ต่อด้านและคุกคามมนุษย์
- ปัญญาประดิษฐ์ควรได้รับการออกแบบโดยใช้หลักการป้องกันความเสี่ยง เพื่อป้องกันการโจมตีจากภัยคุกคาม เพื่อรักษาไว้ซึ่งความมั่นคงปลอดภัยของข้อมูลและระบบ รวมถึงการคุ้มครองข้อมูลส่วนบุคคล จริยธรรม และความปลอดภัยของชีวิตและสิ่งแวดล้อมภายนอกตลอดวัฏจักรชีวิตของระบบ มีความสามารถในการตรวจสอบ รายงานและตอบสนองเพื่อหลีกเลี่ยงหรือลดผลกระทบ
- ปัญญาประดิษฐ์ควรมีกลไกให้มนุษย์แทรกแซงระบบเพื่อควบคุมความเสี่ยงที่อาจมีผลกระทบกับมนุษย์ได้
- หน่วยงานรัฐควรวางแผนกำกับดูแลการพัฒนาและให้ความร่วมมือกับนานาชาติในการหลีกเลี่ยงการแข่งขันสร้างอาวุธอัตโนมัติจากปัญญาประดิษฐ์ที่ร้ายแรง

เนื่องจากในกระบวนการวิจัย ออกแบบและพัฒนาปัญญาประดิษฐ์จำเป็นต้องมีการนำเข้าสู่ข้อมูล ประมวลผล จัดเก็บ และเผยแพร่หรือส่งต่อผลลัพธ์จากการประมวลผลข้อมูล ซึ่งในหลายครั้งข้อมูลเหล่านี้เป็นข้อมูลที่มีความอ่อนไหว เป็นความลับ หรือเป็นข้อมูลส่วนตัว ซึ่งหากถูกเข้าถึงโดยผู้ไม่ประสงค์ดีแล้ว อาจสร้างผลกระทบต่อเจ้าของข้อมูล องค์กร รวมถึงผู้ที่มีส่วนได้เสียได้ ตัวอย่างในเดือนเมษายนปี 2561 กรณีที่ผู้ไม่ประสงค์ดีสามารถเข้าถึงข้อมูลบัตรเครดิต ของลูกค้ากว่า 1 แสนรายการของห้างสรรพสินค้า Sears และสายการบินเดลต้าไลน์ โดยอาศัยโปรแกรมไม่พึงประสงค์ที่เข้าถึงระบบ Chatbot ของบริษัท [24]7.ai ซึ่งบริษัททั้งสองใช้บริการอยู่<sup>23</sup>

<sup>22</sup> Osterman Research White Paper. The State of AI in Cybersecurity: The Benefits, Limitations and Evolving Questions. Dec 2018.

<sup>23</sup> Abigail Abrams. Data Breach at Sears and Delta May Have Hit ‘Several Hundred Thousand’ Customers. TIME USA, LLC. Apr 6 2018.

จาก <https://time.com/5230288/delta-sears-data-breach-credit-cards/>.

ดังนั้นการออกแบบและพัฒนาปัญญาประดิษฐ์ยังต้องคำนึงถึง การป้องกันภัยคุกคามระบบ เพื่อป้องกันมิให้เกิดผลลัพธ์ที่ไม่พึงประสงค์ด้วย ตัวอย่างหนึ่งที่ เกิดขึ้นกับโปรแกรม Tay Chatbot ของบริษัทไมโครซอฟท์ ซึ่งถูกออกนำมาใช้ในปี 2559 และต้องปิดการ ให้บริการลงภายในเวลาเพียง 16 ชั่วโมง เนื่องจากการตรวจสอบพบว่ามีโพสต์ ข้อความที่มีความรุนแรงเกี่ยวกับนาซีและยิว รวมถึงข้อความเรื่อง เพศสัมพันธ์ในทางที่ไม่ สุภาพ ซึ่งเป็นผลอันเนื่องมาจากการที่ระบบถูกสอนจากผู้ไม่ประสงค์ดีในโลกโซเชียล มีเดีย<sup>24</sup> ดังนั้น การออกแบบและพัฒนา ปัญญาประดิษฐ์จึงควรคำนึงถึงหลักการรักษา ความมั่นคงปลอดภัยและการป้องกันข้อมูล ส่วนตัว รวมถึงภัยคุกคามที่อาจเกิดกับ ตัวระบบปัญญาประดิษฐ์เอง และควรมีกลไกให้มนุษย์สามารถแทรกแซงการดำเนินการต่าง ๆ เพื่อลดความเสี่ยงจากการคุกคามผู้ไม่ประสงค์ดี การประมวลผลที่ผิดพลาด และการนำ ปัญญาประดิษฐ์ไปใช้ในทางที่ผิด ซึ่งรวมถึงการใช้เพื่อคุกคามมนุษย์หรือ ผลิตอาวุธ อัตโนมัติที่ร้ายแรง

## 5) ความเท่าเทียม หลากหลาย ครอบคลุม และเป็นธรรม (Fairness)

- การออกแบบและพัฒนาปัญญาประดิษฐ์ควรคำนึงถึงความหลากหลาย หลีกเลียง การผูกขาด ลดการแบ่งแยกและเอินเอียง เพื่อก่อให้เกิดประโยชน์ต่อผู้คนจำนวนมากเท่าที่จะทำได้ โดยเฉพาะกลุ่มคนผู้ด้อยโอกาสในสังคม (Diversity)
- การตัดสินใจที่เกี่ยวข้องกับวิจัย ออกแบบ พัฒนา ให้บริการ และใช้งาน ปัญญาประดิษฐ์ที่สำคัญควรสามารถพิสูจน์ถึงความเป็นธรรมได้ (Fairness)

ผลลัพธ์จากการประมวลผลปัญญาประดิษฐ์อาจมีความไม่เป็นธรรมได้ หากชุดข้อมูลที่ถูกนำมาใช้ในการสอน ทดสอบ และพิสูจน์มีความเอินเอียง และไม่ เป็นตัวแทนของ ผู้ที่จะได้รับผลกระทบจากการใช้งานอย่างครอบคลุมและเท่าเทียม งานวิจัยล่าสุดของศูนย์วิจัยมหาวิทยาลัย New York เมื่อเดือนเมษายน ปี 2562 พบปัญหาความ

<sup>24</sup> Sarah Perez. Microsoft silences its new A.I. bot Tay, after Twitter users teach it racism. Verizon Media. Mar 24 2016. จาก <https://techcrunch.com/2016/03/24/microsoft-silences-its-new-a-i-bot-tay-after-twitter-users-teach-it-racism/>.

เอนเอียงในทางเพศและเชื้อชาติในงานวิจัย ออกแบบ และพัฒนาปัญญาประดิษฐ์ที่ได้รับ การเผยแพร่จำนวนมาก โดยพบว่ากว่าร้อยละ 80 ของผู้วิจัยด้านปัญญาประดิษฐ์เป็นเพศ ชาย และยังพบว่าในบริษัทชั้นนำด้านปัญญาประดิษฐ์อย่าง Facebook และ Google มีผู้วิจัยด้านปัญญาประดิษฐ์ที่เป็นเพศหญิงอยู่เพียงร้อยละ 15 และร้อยละ 10 ตามลำดับ ในด้านเชื้อชาติผู้วิจัยพบว่าในบริษัท Facebook และ Google มีผู้วิจัยด้านปัญญาประดิษฐ์ ที่เป็นคนผิวดำอยู่เพียงร้อยละ 4 และร้อยละ 2.5 ตามลำดับ<sup>25</sup> ผลกระทบจากปัญหาด้าน ความเอนเอียงนี้ ยกตัวอย่างเช่น อัลกอริทึมปัญญาประดิษฐ์ในโปรแกรม COMPAS ของบริษัท Northpointe ที่ถูกใช้งานโดยศาลในสหรัฐฯ ได้รับการตรวจสอบพบว่าระบบ มีความเอนเอียงไปในทางที่จะตัดสินว่า จำเลยผิวดำมีโอกาสจะเป็นผู้กระทำผิดซ้ำมากกว่า จำเลยผิวขาว และอีกตัวอย่างหนึ่งจากงานวิจัยของ M.I.T Media Lab ในปี 2562 พบว่า เทคโนโลยีจดจำใบหน้าของ Amazon ชื่อ Rekognition มีความยากลำบากในการจำแนก ใบหน้าของสตรี และผู้ที่มีผิวสีดำในรูปภาพมากกว่าเทคโนโลยีของบริษัท IBM และไมโครซอฟท์<sup>26</sup> รวมถึงโปรแกรม Photo ของ Google ในปี 2558 ที่จำแนกคนผิวดำ ผิดพลาดว่าเป็น ลิงกอริลล่า<sup>27</sup> เป็นต้น ดังนั้นจึงมีความจำเป็นอย่างยิ่งที่งานวิจัย ออกแบบ และพัฒนา จะต้องคำนึงถึงความหลากหลายของผู้วิจัย ชุดข้อมูล และสามารถพิสูจน์ถึง ความเป็นธรรมของระบบได้

<sup>25</sup> Sarah Myers West., Meredith Whittaker., and Kate Crawford. Discrimination Systems: Gender, Race, and Power in AI. AINOW Institute. Apr 2019. จาก <https://ainowinstitute.org/discriminatingystems.pdf>.

<sup>26</sup> Natasha Singer. Amazon Is Pushing Facial Technology That a Study Says Cloud Be Biased. The New York Times. Jan 24 2019. จาก <https://www.nytimes.com/2019/01/24/technology/amazon-facial-technology-study.html>.

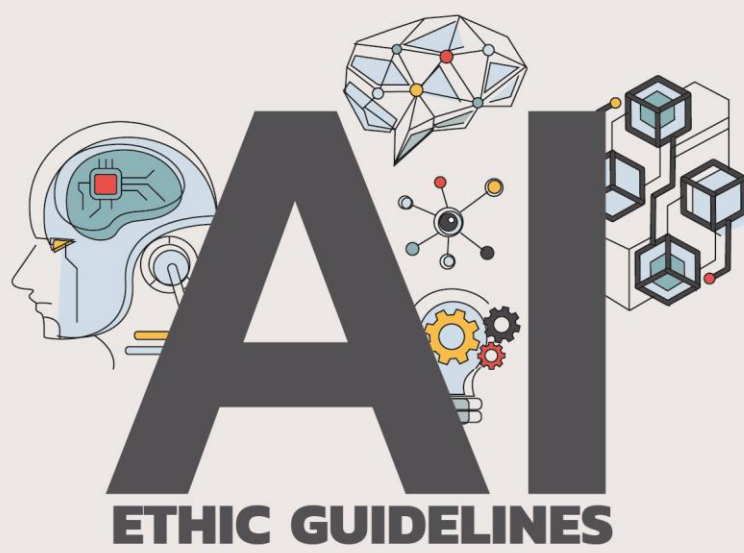
<sup>27</sup> Sophie Curtis. Google Photos labels black people as ‘gorillas’. The Telegraph. Jul 01 2015. จาก <https://www.telegraph.co.uk/technology/google/11710136/Google-Photos-assigns-gorilla-tag-to-photos-of-black-people.html>.

## 6) ความน่าเชื่อถือ (Reliability)

- ปัญญาประดิษฐ์ควรได้รับการสนับสนุนให้มีความน่าเชื่อถือและความมั่นใจในการใช้งานต่อสาธารณะ
- ปัญญาประดิษฐ์ควรสามารถคาดการณ์ ตัดสินใจ และให้คำแนะนำได้อย่างแม่นยำ ถูกต้อง (Accuracy) สร้างผลลัพธ์ที่สามารถเชื่อถือได้และสร้างใหม่ได้เมื่อต้องการ (Reliability and Reproducibility)
- ปัญญาประดิษฐ์ควรมีการควบคุมคุณภาพและตรวจสอบความครบถ้วนสมบูรณ์ของข้อมูล (Quality and integrity of data)
- ปัญญาประดิษฐ์ควรมีกระบวนการและช่องทางรับผลสะท้อนกลับ (Feedback) จากผู้ใช้งาน เพื่อให้ผู้ใช้งานสามารถแจ้งความต้องการเพิ่มเติม รับเรื่องร้องเรียน แจ้งปัญหาของระบบที่ตรวจสอบพบ และให้ข้อเสนอแนะได้โดยง่ายและรวดเร็ว

เนื่องจากผลลัพธ์จากการประมวลผลของปัญญาประดิษฐ์สามารถสร้างผลกระทบให้กับ ผู้รับบริการได้ ดังนั้นหากปัญญาประดิษฐ์สร้างผลลัพธ์ที่ไม่ถูกต้องก็จะส่งผลกระทบต่อความเชื่อมั่นของระบบได้ ดังจะเห็นได้จากตัวอย่างทั้งในกรณีของความผิดพลาดจากการตรวจสอบภัยคุกคามในระบบปัญญาประดิษฐ์ที่ใช้ในผลิตภัณฑ์ด้านความมั่นคงปลอดภัยทางไซเบอร์ ความผิดพลาดจากการประมวลผลภาพในโปรแกรม Photo ของ Google อีกตัวอย่างหนึ่งในกรณีของระบบจดจำใบหน้าของ Amazon Rekognition ในปี 2561 พบว่ามีความผิดพลาดในการระบุภาพของสมาชิกสภาคอนเกรสของสหรัฐจำนวน 28 คน ว่าเป็นผู้ที่ถูกจับในข้อหาอาชญากรรม<sup>28</sup> เป็นต้น ดังนั้นการสร้างเชื่อมั่นให้กับปัญญาประดิษฐ์ทั้งจากการควบคุมด้านกระบวนการวิจัย ออกแบบ พัฒนา การควบคุมคุณภาพของข้อมูลที่นำมาใช้งาน และการปรับปรุงคุณภาพของ ระบบจากผลสะท้อนกลับของผู้ใช้งาน จึงเป็นสิ่งที่สำคัญ เพื่อให้ระบบสามารถให้คำแนะนำที่ถูกต้องแม่นยำและเชื่อถือได้กับผู้ใช้งาน

<sup>28</sup> Jacob Snow. Amazon's Face Recognition Falsely Matched 28 Members of Congress With Mugshots. American Civil Liberties Union. Jul 26 2018. จาก <https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28>.



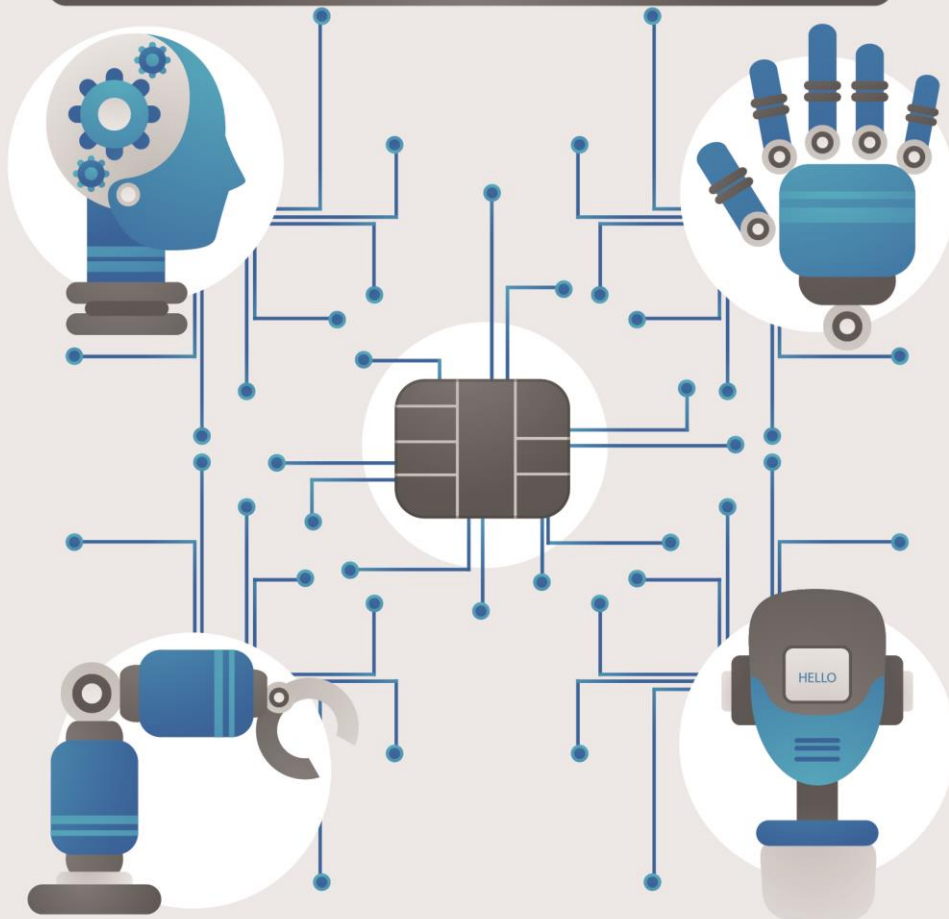
# AI

## ETHIC GUIDELINES

การกำหนดนโยบาย/มาตรฐาน  
การสนับสนุนองค์ความรู้และทักษะ  
การตรวจสอบ/ประเมินความเสี่ยง

การสนับสนุนงานวิจัยและพัฒนา  
การสร้างร่วมมือ  
การเฝ้าระวัง/ติดตาม  
การทบทวน/ปรับปรุง

### หน่วยงานรัฐและหน่วยงานกำกับดูแล



### ผู้วิจัย ผู้ออกแบบ ผู้พัฒนา และผู้ให้บริการ

การวางแผน/กำหนดมาตรการ  
การออกแบบและพัฒนา  
การพัฒนาความรู้/ทักษะ/ความตระหนัก  
การทบทวน/ปรับปรุง

การประเมินความเสี่ยง  
การนำไปใช้  
การสร้างความร่วมมือ  
การเฝ้าระวัง/ติดตาม

# แนวทางปฏิบัติทางจริยธรรมปัญญาประดิษฐ์ (AI Ethic Guidelines)

## 1) หน่วยงานรัฐและหน่วยงานกำกับดูแล

- ความสามารถในการแข่งขันและการพัฒนาอย่างยั่งยืน (Competitiveness and Sustainability Development)
  - ควรตรวจสอบและประเมินงานวิจัย พัฒนาและการนำปัญญาประดิษฐ์ไปใช้งานว่าสามารถสร้างให้เกิดประโยชน์และความผาสุกต่อมนุษย์ สังคม และสิ่งแวดล้อม
  - ควรสนับสนุนการวิจัยและพัฒนาในการนำปัญญาประดิษฐ์มาใช้ให้เกิดประโยชน์
  - ควรส่งเสริมและสนับสนุนการทำงานร่วมกันระหว่างมนุษย์และปัญญาประดิษฐ์ที่ทำให้เกิดประโยชน์ต่อมนุษย์
  - ควรส่งเสริมและสนับสนุนให้หน่วยงานต่าง ๆ สามารถเข้าถึงองค์ความรู้ทางปัญญาประดิษฐ์ ด้วยเทคโนโลยีโครงสร้างพื้นฐานทางดิจิทัลและกลไกต่าง ๆ เพื่อให้เกิดการแลกเปลี่ยนข้อมูลความรู้ และเทคโนโลยีในการวิจัย ออกแบบ และพัฒนา
  - ควรเฝ้าระวังและติดตามการใช้งานปัญญาประดิษฐ์ต่อผู้ใช้งาน ว่าก่อให้เกิดประโยชน์ได้จริงและไม่สร้างผลกระทบในเชิงลบ
  - หน่วยงานรัฐ
    - ควรออกนโยบายเพื่อสนับสนุนให้หน่วยงานทั้งภาครัฐและเอกชน ดำเนินการวิจัยและพัฒนา เพื่อกระตุ้นให้เกิดนวัตกรรมปัญญาประดิษฐ์ ที่ช่วยในการสร้างอุตสาหกรรมใหม่ควรส่งเสริมและสนับสนุนด้านการศึกษา อบรมประชาชนทั่วไปเพื่อให้มีความรู้ และทักษะเกี่ยวกับปัญญาประดิษฐ์ และผลกระทบที่อาจเกิดขึ้นจากปัญญาประดิษฐ์ สร้างทัศนคติที่ถูกต้องในการยอมรับการใช้งานเทคโนโลยีปัญญาประดิษฐ์ที่ก่อให้เกิดประโยชน์ และทัศนคติของการตระหนักรู้ภัยคุกคามจากปัญญาประดิษฐ์ควรสร้างความร่วมมือกับองค์กร

ภายในประเทศ ภูมิภาคและนานาชาติ เพื่อพัฒนาโครงการ ปัญญาประดิษฐ์ที่ก่อให้เกิดประโยชน์ต่อภาพรวมในระดับภูมิภาคและระดับโลก

- **ความสอดคล้องกับกฎหมาย จริยธรรม และมาตรฐานสากล (Laws Ethics and International Standards)**
  - ควรสนับสนุนการให้การศึกษาเพื่อสร้างความตระหนักรู้และความเข้าใจ ในปัญญาประดิษฐ์และผลกระทบของปัญญาประดิษฐ์ให้กับผู้ใช้งาน สนับสนุน การอบรมผู้วิจัย และพัฒนาปัญญาประดิษฐ์ เพื่อให้ตระหนักและเข้าใจ ผลกระทบของระบบปัญญาประดิษฐ์ที่มีต่อบุคคลและสังคม และให้การ สนับสนุนงานวิจัยเกี่ยวกับปัญญาประดิษฐ์กับสิทธิมนุษยชน
  - การจัดซื้อจัดจ้างควรกำหนดให้ผู้พัฒนา และผู้ให้บริการปัญญาประดิษฐ์ ต้องดำเนินการสอดคล้องกับหลักเกณฑ์ของความโปร่งใส มีการประเมินผล กระทบ จากการประมวลผลข้อมูลต่อสิทธิมนุษยชนและความเป็นส่วนตัว พร้อมกำหนดภาระความรับผิดชอบต่อผลกระทบเชิงลบที่เกิดขึ้นกับ ปัญญาประดิษฐ์
  - **หน่วยงานรัฐ**
    - ควรสนับสนุนในการสร้างมาตรฐานให้เป็นที่ยอมรับโดยสากล และแนวทางปฏิบัติที่ดีที่สุด (Best Practice) และกำหนดให้ผู้ออกแบบ ผู้พัฒนา และผู้ให้บริการต้องปฏิบัติตาม พร้อมกลไกการให้ไต่สวน สำหรับปัญญาประดิษฐ์ เพื่อเพิ่มความน่าเชื่อถือให้กับผู้ออกแบบ ผู้พัฒนา และผู้ให้บริการ รวมถึงผลิตภัณฑ์และบริการที่เกี่ยวข้อง
    - ควรสนับสนุนให้มีองค์กรต่าง ๆ มีหน่วยงานกำกับดูแลการพัฒนาและ ใช้งานปัญญา ประดิษฐ์เพื่อให้คำปรึกษาผู้วิจัย ผู้ออกแบบ ผู้พัฒนา และผู้ให้บริการปัญญาประดิษฐ์ เกี่ยวกับผลกระทบด้านกฎหมาย จริยธรรม และสิทธิมนุษยชน ประเมินความเสี่ยงที่มีผลกระทบเชิงลบ ของระบบ และดูแลด้านการสร้าง ภาระความรับผิดชอบของผู้วิจัย ผู้ออกแบบ ผู้พัฒนา และผู้ให้บริการปัญญาประดิษฐ์



- **ความโปร่งใสและภาระความรับผิดชอบ (Transparency and Accountability)**
  - หน่วยงานกำกับดูแลการพัฒนาและใช้งานปัญญาประดิษฐ์ควรตรวจสอบความโปร่งใสของโมเดลและอัลกอริทึมที่ผู้วิจัย ผู้ออกแบบ และผู้พัฒนาใช้ในปัญญาประดิษฐ์ โดยใช้หลักการเรื่องความสามารถในการอธิบายได้ (Explainable) ซึ่งอัลกอริทึมที่นำมาใช้ต้องสามารถอธิบายที่มา และหน้าที่การทำงานในการคาดการณ์ต่าง ๆ รวมถึงสามารถอธิบายวิธีการสอนและคัดเลือกโมเดลได้
  - หน่วยงานกำกับดูแลการพัฒนาและใช้งานปัญญาประดิษฐ์ควรกำหนดนโยบายด้านการให้คำอธิบายปัญญาประดิษฐ์กับผู้ใช้งานขึ้น โดยกำหนดให้ผู้วิจัย ออกแบบ และพัฒนาปัญญาประดิษฐ์ สร้างเอกสารทางเทคนิคเพื่อแสดงรายละเอียดออกแบบ และการทำงานในหลายมุมมอง เพื่อรองรับผู้ใช้งานที่มีความรู้ความเข้าใจที่แตกต่างกัน และสถานการณ์ที่แตกต่างกันของการนำไปใช้ และกำหนดให้มีช่องทางที่ง่ายและรวดเร็วเพื่อให้ ผู้ที่อาจได้รับผลกระทบ ส่งการร้องขอคำการอธิบายเหล่านี้ได้
  - ควรมีกลไกเพื่อสร้างภาระความรับผิดชอบ (Accountability) ของผู้วิจัย ผู้ออกแบบ ผู้พัฒนา ผู้ให้บริการและผู้ใช้งานปัญญาประดิษฐ์ รวมถึงผู้ที่มีส่วนได้เสียอื่น ๆ ตลอดวัฏจักรชีวิตของระบบ โดยมีกระบวนการในการตรวจสอบชุดข้อมูล อัลกอริทึม กระบวนการออกแบบ การนำไปใช้ และผลลัพธ์ของการใช้งานปัญญาประดิษฐ์ ทั้งการตรวจสอบภายในและภายนอกอย่างอิสระ การรายงานผลการตรวจสอบ การประเมินความเสี่ยงเชิงลบ และการดำเนินการเพื่อลดหรือหลีกเลี่ยง ผลกระทบเชิงลบของปัญญาประดิษฐ์ โดยเฉพาะผลกระทบต่อมนุษย์และสิ่งแวดล้อม
  - กำหนดผู้รับผิดชอบในการสืบสวน และแก้ไขสาเหตุของความสูญเสียและเสียหายที่เกิดขึ้น จากปัญญาประดิษฐ์
- **ความมั่นคงปลอดภัยและความเป็นส่วนตัว (Security and Privacy)**
  - ควรกำหนดนโยบายและมาตรฐานทางเทคนิคด้านความมั่นคงปลอดภัยและการคุ้มครองความเป็นส่วนตัวสำหรับปัญญาประดิษฐ์ เพื่อลดช่องโหว่และป้องกันภัยคุกคามของปัญญาประดิษฐ์ ที่ก่อให้เกิดผลกระทบในด้านความลับ



ความครบถ้วนถูกต้อง ความพร้อมใช้ของข้อมูล การคุ้มครองข้อมูลบุคคล รวมถึงผลกระทบต่อด้านจริยธรรม ชีวิตและสิ่งแวดล้อม และให้ผู้วิจัย ออกแบบ พัฒนา และให้บริการปฏิบัติตาม

- หน่วยงานกำกับดูแลการพัฒนาและใช้งานปัญญาประดิษฐ์ ควรดำเนินการจัดการความเสี่ยง กำหนดวิธีการจัดการความเสี่ยงและควบคุมภายใน เพื่อทำหน้าที่จัดการความเสี่ยงที่เกี่ยวข้องกับปัญญาประดิษฐ์ หาแนวทางในการจัดการความเสี่ยง ตรวจสอบและรายงาน ประสิทธิภาพในการจัดการความเสี่ยง ทบทวนและปรับปรุงแนวทางและกระบวนการจัดการความเสี่ยง ทั้งนี้ควรกำหนดให้มีการดำเนินการจัดการความเสี่ยงอย่างสม่ำเสมอ ตลอดวัฏจักรชีวิตของระบบ รวมถึงเมื่อจำเป็นต้องรีเซ็ตระบบ โดยดำเนินการอย่างน้อยปีละ 1 ครั้งหรือมีการเปลี่ยนแปลงที่สำคัญเกิดขึ้น
- หน่วยงานกำกับดูแลการพัฒนาและใช้งานปัญญาประดิษฐ์ควรประเมินความเสี่ยงเชิงลบที่กระทบกับบุคคลและสังคมจากการเพิกเฉยข้อมูลหรือสถานะการฉีกเฉพาะ (De-contextualised data)<sup>29</sup> และอัลกอริทึมที่เพิกเฉยกับบริบทหรือสถานะการฉีกเฉพาะอย่างเพียงพอระหว่างขั้นตอน การพัฒนา และนำปัญญาประดิษฐ์ไปใช้งาน
- หน่วยงานกำกับดูแลการพัฒนาและใช้งานปัญญาประดิษฐ์ควรดำเนินการประเมินระดับการเข้าแทรกแซงปัญญาประดิษฐ์โดยมนุษย์ในกระบวนการต่าง ๆ จากโอกาสและความรุนแรงของผลกระทบที่จะเกิดขึ้นกับผู้ใช้งาน
- หน่วยงานกำกับดูแลการพัฒนาและใช้งานปัญญาประดิษฐ์ ควรให้บุคคลหรือกลุ่มคนที่ได้รับผลกระทบที่สำคัญจากระบบปัญญาประดิษฐ์ ได้มีส่วนร่วมในกระบวนการประเมินความเสี่ยงด้วย
- หน่วยงานกำกับดูแลการพัฒนาและใช้งานปัญญาประดิษฐ์ ควรมีการทบทวนประสิทธิภาพการทำงานของหน่วยงานและเจ้าหน้าที่ในโครงสร้างการกำกับดูแลอย่างสม่ำเสมอ และทุกครั้งที่มีการเปลี่ยนแปลงโครงสร้างหรือบุคคลที่สำคัญ

<sup>29</sup> Council of Europe. Guidelines on Artificial Intelligence and Data Protection. Consultative Committee of the Convention for the Protection of individuals with regard to Automatic Processing of Personal Data. Jan 25 2019.

### ○ หน่วยงานรัฐ

- ควรมีการวางแผนเพื่อกำกับดูแล ใฝ่ระวังและจัดการความเสี่ยงในระยะยาว โดยเฉพาะเมื่อปัญญาประดิษฐ์มีความฉลาดมากยิ่งขึ้นกว่าในปัจจุบัน เช่น ปัญญาประดิษฐ์ทั่วไป (Artificial General Intelligence (AGI)) ซุปเปอร์ปัญญาประดิษฐ์ (Artificial Superintelligence) และปัญญาประดิษฐ์ที่มีความสามารถในการปรับปรุงตนเองได้อย่างต่อเนื่อง (Recursive Self-Improving AI)<sup>30</sup>
- ควรสนับสนุนให้เกิดความร่วมมือเพื่อพัฒนาโครงสร้างการกำกับดูแลปัญญาประดิษฐ์ในแบบบูรณาการ ทั้งในระดับองค์กร ประเทศ ภูมิภาค และให้ความร่วมมือกับนานาชาติ ในการหลีกเลี่ยงการแข่งขันด้านปัญญาประดิษฐ์ที่ไม่พึงประสงค์ รวมถึงการสร้างอาวุธอัตโนมัติจากปัญญาประดิษฐ์ที่ร้ายแรง สนับสนุนให้เกิดการแลกเปลี่ยนความรู้ และประสบการณ์ในการกำกับดูแล และร่วมกันรับมือกับผลกระทบของปัญญาประดิษฐ์

### ● ความเท่าเทียม หลากหลาย ครอบคลุม และเป็นธรรม (Fairness)

- ควรกำหนด ส่งเสริมและสนับสนุนแนวทางปฏิบัติที่เกี่ยวข้องกับการประยุกต์ใช้ปัญญาประดิษฐ์ที่หลากหลายตามชนิดและสถานการณ์ในการใช้งาน
- หน่วยงานกำกับดูแลการพัฒนาและใช้งานปัญญาประดิษฐ์ ควรมีกระบวนการควบคุมเพื่อวิเคราะห์ ประเมินความเสี่ยง และจัดการปัญหาการเอนเอียงไปสู่ความไม่เป็นธรรมในขั้นตอนการวิจัย ออกแบบ พัฒนา และให้บริการปัญญาประดิษฐ์อย่างชัดเจนและโปร่งใส

### ○ หน่วยงานรัฐ

- ควรส่งเสริมให้เกิดแพลตฟอร์มเปิดของปัญญาประดิษฐ์เพื่อหลีกเลี่ยงการผูกขาด ทำให้เกิดการแลกเปลี่ยนความรู้ในการพัฒนาปัญญาประดิษฐ์เพื่อสร้างสรรค์ ต่อยอดองค์ความรู้ และพัฒนาอย่างต่อเนื่องเพื่อใช้งานในระดับอุตสาหกรรมได้

<sup>30</sup> Smart Dubai. AI Ethics Principles & Guidelines. Version 1.6. Dec 30 2018.

- ควรสนับสนุนให้เกิดโอกาสในการพัฒนาปัญญาประดิษฐ์ที่เท่าเทียมกัน และการแข่งขันที่เป็นธรรม ทั้งในระดับภูมิภาคและอุตสาหกรรมที่แตกต่างกัน
- ควรส่งเสริมและสนับสนุนให้เกิดโอกาสที่เท่าเทียมกัน ในการเข้าถึง การศึกษา สินค้า บริการ และเทคโนโลยีที่เกี่ยวข้องด้าน ปัญญาประดิษฐ์
- ควรส่งเสริมให้ประชาชนมีทักษะด้านปัญญาประดิษฐ์และสนับสนุน กลุ่มคนทำงานเพื่อการเปลี่ยนผ่านที่เป็นธรรม (Fair Transition)<sup>31</sup> หน่วยงานรัฐควรส่งเสริมให้เกิดการวิจัยและพัฒนาเทคนิค การวิเคราะห์เพื่อใช้ตรวจสอบและแก้ไขปัญหาความเอนเอียง แบ่งแยก และไม่เป็นธรรมของปัญญาประดิษฐ์

#### ● ความน่าเชื่อถือ (Reliability)

- ควรกำหนดนโยบาย หลักเกณฑ์ และกระบวนการในการประเมินคุณภาพของ ชุดข้อมูลและโมเดลปัญญาประดิษฐ์ ดำเนินการทบทวนและปรับปรุงความ น่าเชื่อถือของปัญญาประดิษฐ์และชุดข้อมูลอย่างสม่ำเสมอ โดยนำข้อมูล ผลสะท้อนกลับที่ได้รับจากผู้ใช้งานระบบจริงมาใช้ เพื่อให้ระบบสอดคล้องกับการ เปลี่ยนแปลงพฤติกรรมของผู้ใช้งานตามระยะเวลา ควรปรับปรุงโมเดลด้วยชุด ข้อมูลที่เป็นปัจจุบัน และควรปรับปรุงโมเดลเมื่อวัตถุประสงค์และความเสี่ยง เปลี่ยนแปลงไป
- ควรกำหนดนโยบายและกระบวนการเพื่อทบทวนช่องทางการสื่อสารกับ ผู้ใช้งาน และดำเนินการทบทวนอย่างสม่ำเสมอ เพื่อสนับสนุนการเปิดเผย ข้อมูลที่จำเป็น และรับผลสะท้อนจากการให้บริการได้อย่างมีประสิทธิภาพ
- หน่วยงานรัฐ
  - ควรสนับสนุนการวิจัย ออกแบบ พัฒนา ให้บริการ และใช้งาน ปัญญาประดิษฐ์ที่มีความน่าเชื่อถือ

<sup>31</sup> OECD Legal Instruments. Recommendation of the Council on Artificial Intelligence. May 22 2019. จาก

<https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>.

- ควรกำหนดนโยบายที่ช่วยเปิดโอกาสให้เกิดการใช้งานปัญญาประดิษฐ์ที่มีความน่าเชื่อถือ และพัฒนาแนวทางการประเมินความน่าเชื่อถือของปัญญาประดิษฐ์เพื่อใช้ตรวจสอบประเมินผู้พัฒนาและผู้ให้บริการ
  - ❖ แนวทางการประเมินควรสามารถปรับเปลี่ยนได้ตามลักษณะการนำปัญญาประดิษฐ์ไปใช้งาน
  - ❖ แนวทางการประเมินควรได้รับการพัฒนาขึ้นด้วยความร่วมมือของผู้ที่มีส่วนได้เสียทั้งภาครัฐและเอกชน
- ควรประเมินและตรวจสอบคุณภาพผู้ให้บริการปัญญาประดิษฐ์ที่ใช้ข้อมูลจากผลลัพธ์ของระบบ เพื่อการตัดสินใจที่สำคัญที่เกี่ยวข้องกับมนุษย์อย่างเข้มงวด

## 2) ผู้วิจัย ผู้ออกแบบ ผู้พัฒนา และผู้ให้บริการ

- ความสามารถในการแข่งขันและการพัฒนาอย่างยั่งยืน (Competitiveness and Sustainability Development)
  - มีความรู้ความเข้าใจที่จำเป็นเกี่ยวกับปัญญาประดิษฐ์ในส่วนที่ตนเองเกี่ยวข้องอย่างเพียงพอ ทราบถึงประโยชน์และผลกระทบของระบบ เพื่อสร้างประโยชน์จากปัญญาประดิษฐ์ให้แก่มนุษย์ สังคม เศรษฐกิจ และสิ่งแวดล้อมอย่างยั่งยืน และลดความเสี่ยงต่าง ๆ ที่อาจเกิดขึ้นจากปัญญาประดิษฐ์
  - ควรออกแบบ พัฒนาและนำปัญญาประดิษฐ์ไปใช้งาน โดยคำนึงถึงประโยชน์ที่จะเกิดขึ้นกับมนุษย์โดยรวม ผู้ที่มีส่วนได้เสีย และสิ่งแวดล้อมอย่างยั่งยืน
  - ควรออกแบบ และพัฒนาให้ปัญญาประดิษฐ์สามารถปรับเปลี่ยนการทำงานได้ตามสภาพแวดล้อมต่าง ๆ เพื่อเอื้อประโยชน์ให้แก่มนุษย์อย่างสร้างสรรค์

- ความสอดคล้องกับกฎหมาย จริยธรรม และมาตรฐานสากล (Laws Ethics and International Standards)
  - ควรมีมาตรการในประเมิน ลดหรือหลีกเลี่ยงความเสี่ยงด้านกฎหมาย จริยธรรม การละเมิดสิทธิเสรีภาพ สิทธิมนุษยชน ผลกระทบต่อสังคมและสิ่งแวดล้อม และควรมีมาตราเยียวยาผู้ที่ได้รับผลกระทบจากปัญญาประดิษฐ์
  - ควรใช้หลักการออกแบบและนำไปใช้งานที่คำนึงถึงจริยธรรม ในการสร้างระบบ โดยให้ผู้เชี่ยวชาญในแขนงที่จะนำไปใช้งาน รวมถึงสถาบันการศึกษา มีส่วนร่วมในขั้นตอนการออกแบบ พัฒนา และให้บริการ
  - ผู้ให้บริการปัญญาประดิษฐ์ควรกำหนดรายละเอียดหลักการจริยธรรมไว้ใน กระบวนการปฏิบัติงาน หรือใช้เป็นข้อกำหนดหนึ่งในผลิตภัณฑ์และบริการ และแจ้งให้ผู้ใช้งานได้รับทราบ
  - ควรทำการประเมินจริยธรรมในการวิจัย ออกแบบ พัฒนา ให้บริการ และใช้งานปัญญาประดิษฐ์ และเผยแพร่สรุปผลการประเมินให้ผู้มีส่วนได้เสีย ได้รับทราบ
  - ควรมีการประเมินความเสี่ยงของผลกระทบจากการวิจัย ออกแบบ พัฒนา ให้บริการ และใช้งานเทียบกับหลักการทางจริยธรรม
  - ควรใช้หลักการออกแบบและนำไปใช้งานที่ใ้มนุษย์เป็นศูนย์กลาง (Human Centric) และคงไว้ซึ่งสิทธิใ้มนุษย์เป็นผู้เลือกตัดสินใจ
    - ควรออกแบบปัญญาประดิษฐ์ให้สามารถส่งต่อการดำเนินการและการตัดสินใจไปยังมนุษย์ ได้
    - ควรออกแบบปัญญาประดิษฐ์ใ้มนุษย์สามารถเข้าแทรกแซง กระบวนการตัดสินใจของปัญญาประดิษฐ์ และคงไว้ซึ่งอิสรภาพของ มนุษย์ในการเลือกตัดสินใจที่จะไม่ใช้ผลลัพธ์ที่เกิดจากคำแนะนำของ ปัญญาประดิษฐ์ โดยมีกลไกในการอนุญาตใ้มนุษย์สามารถเข้าแทรกแซงปัญญาประดิษฐ์ อาทิ

- ❖ ความสามารถที่มนุษย์จะเข้าแทรกแซงในทุก ๆ ขั้นตอนการตัดสินใจของปัญญาประดิษฐ์
  - ❖ ความสามารถที่มนุษย์จะเข้าแทรกแซงในระหว่างขั้นตอนการออกแบบปัญญา ประดิษฐ์ และเฝ้าระวังการปฏิบัติงานของระบบ
  - ❖ ความสามารถของมนุษย์ในการตรวจสอบกิจกรรมทั้งหมดของปัญญาประดิษฐ์ และสามารถตัดสินใจเลือกหรือไม่เลือกใช้งานระบบในสถานะการณ์ต่าง ๆ รวมถึงความสามารถในการปรับระดับในการตัดสินใจระหว่างใช้งาน หรือความสามารถในการครอบงำการตัดสินใจของปัญญาประดิษฐ์
- ความโปร่งใสและภาระความรับผิดชอบ (Transparency and Accountability)
    - ควรเปิดเผยวัตถุประสงค์ เหตุผลการตัดสินใจของปัญญาประดิษฐ์ ชุดข้อมูลความสามารถ ข้อจำกัด อัลกอริทึมที่ใช้งาน กระบวนการทำงาน และความน่าเชื่อถือของปัญญาประดิษฐ์อย่างโปร่งใส โดยไม่ละเมิดความเป็นส่วนตัวและทรัพย์สินทางปัญญา ในลักษณะที่สามารถอธิบายให้ผู้ที่อาจได้รับผลกระทบทั้งทางตรงและทางอ้อมให้เข้าใจได้ โดยคำนึงถึงความรู้ความเข้าใจที่แตกต่างกันของผู้ใช้งานและสถานการณ์ในการนำไปใช้งานที่หลากหลาย เพื่อให้ผู้ใช้งานสามารถตั้งข้อสงสัยต่อการทำงานของปัญญาประดิษฐ์ ตรวจสอบ และตระหนักถึงความเอนเอียง ความผิดพลาด และผลลัพธ์ที่ไม่ได้ตั้งใจได้โดยง่าย
    - ในกรณีที่มีข้อจำกัดในการให้คำอธิบายอัลกอริทึมปัญญาประดิษฐ์ เนื่องจากเป็นข้อมูลที่มีเจ้าของโดยเฉพาะ (Proprietary data) เกี่ยวข้องกับทรัพย์สินทางปัญญา และความมั่นคง ปลอดภัย ผู้วิจัย ผู้ออกแบบ และผู้พัฒนาควรบันทึกและให้ข้อมูลที่บ่งบอกถึงความสามารถในการทำซ้ำ (Repeatability) ของปัญญาประดิษฐ์ ซึ่งเป็นความสามารถในการสร้างผลลัพธ์เดิมกับสถานการณ์ในลักษณะเดิมได้ เพื่อสร้างความเชื่อมั่นให้กับระบบ โดยควรดำเนินการดังต่อไปนี้

- ทำการประเมินและทดสอบความสามารถในการทำซ้ำของปัญญาประดิษฐ์
  - ทำการประเมินวิธีการในการระบุและจัดการข้อยกเว้น (Exceptions) ในกรณีที่การตัดสินใจหนึ่งๆ ไม่สามารถทำซ้ำได้
- ผู้ที่มีส่วนได้เสียที่อาจได้รับผลกระทบจากปัญญาประดิษฐ์ ควรได้รับการแจ้งเตือนและรับทราบอย่างเพียงพอ มีสิทธิ์ในการขอข้อมูลเกี่ยวกับเหตุผลในการใช้งานปัญญาประดิษฐ์กับตนเอง รวมถึงผลกระทบของเหตุผลเหล่านั้น และควรมีส่วนร่วมในขั้นตอนของการออกแบบและพัฒนาปัญญาประดิษฐ์ทั้งทางตรงและทางอ้อมตลอดวัฏจักรชีวิตของระบบ (Stakeholder Participation)
- ผู้ใช้งานควรได้รับแจ้งว่ากำลังสื่อสารกับปัญญาประดิษฐ์ซึ่งมิใช่มนุษย์ และรับทราบว่าผลลัพธ์ของการตัดสินใจเกิดจากปัญญาประดิษฐ์
- ควรมีความตระหนักรู้ในเรื่องภาระความรับผิดชอบ (Accountability) โดยมีส่วนร่วมรับผิดชอบต่อการผลกระทบจากปัญญาประดิษฐ์ ตามภาระหน้าที่ของตนเอง
- ควรออกแบบให้ปัญญาประดิษฐ์มีความสามารถในการสืบย้อนกลับ (Traceability) ได้จาก Audit log สามารถเฝ้าระวัง ตรวจสอบความผิดปกติ และวินิจฉัยปัญหาความล้มเหลวได้ (Diagnosability) ได้ ทั้งในส่วนของที่มาและวิธีการเก็บรวบรวมข้อมูล ชุดข้อมูล อัลกอริทึม และกระบวนการตัดสินใจของปัญญาประดิษฐ์ที่สามารถสร้างผลกระทบที่สำคัญกับผู้ใช้งานได้ โดยแนวทางปฏิบัติเพื่อสร้างความสามารถในการสืบย้อนกลับ ประกอบด้วย
- การบันทึกข้อมูลและกิจกรรมที่เกิดขึ้นระหว่างการวิจัย ออกแบบ พัฒนา และการให้บริการ ตามลำดับเวลาของเหตุการณ์ (Audit trail)
    - ❖ ข้อมูลชุดการสอนปัญญาประดิษฐ์ วิธีการในการเก็บรวบรวมและปรับแก้ การเคลื่อนย้ายข้อมูล ผลการตรวจวัดความแม่นยำของปัญญาประดิษฐ์ตลอดช่วงเวลาที่ดำเนินการ
    - ❖ โมเดลที่ใช้ออกแบบและอัลกอริทึมที่เลือกใช้

### ❖ การเปลี่ยนแปลงโค้ดโปรแกรม (Code) และผู้ที่ทำการเปลี่ยนแปลง

- การบันทึกกระแสข้อมูลไหลเข้าระบบทั้งหมดในช่วงเวลาที่มีการใช้งานปัญญาประดิษฐ์
  - มีการจัดเก็บข้อมูลเพื่อการสืบย้อนในหน่วยจัดเก็บข้อมูลที่เหมาะสมเพื่อหลีกเลี่ยงการถูกลดทอนคุณภาพ หรือถูกเปลี่ยนแปลงแก้ไข และควรจัดเก็บไว้ตามระยะเวลาที่สอดคล้องกับกฎหมายหรือข้อกำหนดในมาตรฐานอุตสาหกรรมที่นำปัญญาประดิษฐ์ไปใช้งาน
  - การบันทึกข้อมูลและกิจกรรมที่เกิดขึ้นระหว่างการออกแบบ พัฒนา และการให้บริการ ตามลำดับเวลาของเหตุการณ์ (Audit trail)
- ควรมีกระบวนการทบทวนข้อมูล กิจกรรม และกระบวนการที่เกิดขึ้นระหว่างการวิจัย ออกแบบ พัฒนา และการให้บริการ หลังจากผู้ใช้งานนำปัญญาประดิษฐ์ไปใช้งานเรียบร้อยแล้ว และหากพบว่าข้อมูล กิจกรรม และกระบวนการที่เกิดขึ้นระหว่างการวิจัย ออกแบบ พัฒนา และการให้บริการปัญญาประดิษฐ์มีข้อผิดพลาด ควรจัดให้มีกระบวนการพิจารณาแก้ไขปรับปรุงในลำดับถัดไป

### ● ความมั่นคงปลอดภัยและความเป็นส่วนตัว (Security and Privacy)

- ควรออกแบบ พัฒนา และให้บริการปัญญาประดิษฐ์ ให้มีการป้องกันความเสี่ยง โดยสร้างให้ปัญญาประดิษฐ์มีกลไกการรักษาความมั่นคงปลอดภัยและป้องกันระบบจากภัยคุกคามและการใช้งานที่ไม่พึงประสงค์ ที่อาจกระทบต่อข้อมูลที่สำคัญ ข้อมูลส่วนบุคคล จริยธรรม ชีวิตและสิ่งแวดล้อม
- ปัญญาประดิษฐ์ที่ประมวลผลข้อมูลส่วนตัวควรใช้หลักการที่สอดคล้องกับกฎหมาย (Lawfulness) ความเป็นธรรม (Fairness) มีวัตถุประสงค์ที่เฉพาะ (Purpose Specification) มีการประมวลผลตามวัตถุประสงค์ (Proportionality of data processing) และการออกแบบที่คำนึงถึงความ



เป็นส่วนตัวโดยปริยาย (Privacy-by-design and by-default)<sup>32</sup> และสอดคล้องกับกฎหมาย พรบ.คุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562

- ควรออกแบบและพัฒนาปัญญาประดิษฐ์ให้มีกลไกการป้องกันความครบถ้วน ถูกต้อง (Integrity) ของข้อมูลชุดการสอนระบบ โดยสามารถแยกและกำจัด ข้อมูลที่ไม่พึงประสงค์ออกจากข้อมูลปกติแต่มีจำนวนน้อยได้
- ควรประเมินคุณภาพ ลักษณะตามธรรมชาติ แหล่งที่มา และจำนวนข้อมูล ส่วนตัวที่ถูกลำเอามาใช้ โดยลดการใช้ข้อมูลส่วนตัวที่ไม่จำเป็นหรือซ้ำซ้อน ระหว่างขั้นตอนการพัฒนาและขั้นตอนการสอนระบบ
- ควรดำเนินการกับข้อมูลส่วนตัวที่เกี่ยวข้องอย่างเคารพและถูกต้อง ดังนี้
  - แจ้งให้ผู้ใช้งานทราบล่วงหน้าถึงข้อมูลที่จะถูกเก็บรวบรวมและการนำไปใช้ โดยต้องได้รับการยินยอมจากเจ้าของข้อมูลเสียก่อน
  - เก็บรวบรวมข้อมูลส่วนตัวเท่าที่จำเป็น จำกัดการเข้าถึงและจัดเก็บไว้ตามระยะเวลาเท่าที่จำเป็น
  - มีระบบป้องกันข้อมูลส่วนตัว อาทิ ระบบยืนยันตัวตน ระบบสนับสนุนผู้ใช้งานในการค้นหา ข้อมูลส่วนตัว และการลบข้อมูลส่วนตัวทั้งหมดที่สามารถใช้งานได้โดยง่าย
  - มีกระบวนการทบทวนความสอดคล้องของปัญญาประดิษฐ์กับกฎหมาย คุ้มครองข้อมูลส่วนบุคคล
- ควรดูแลสภาพแวดล้อมของการปฏิบัติงานที่เกี่ยวข้องให้มีความมั่นคงปลอดภัย เพื่อป้องกันภัยคุกคามทั้งในแบบที่ไม่ได้ตั้งใจและไม่ประสงค์ดี ที่มีผลกระทบเชิงลบต่อข้อมูลที่สำคัญต่อมนุษย์และสิ่งแวดล้อม
- หน่วยงานของรัฐและเอกชน ซึ่งมีภารกิจหรือให้บริการโครงสร้างพื้นฐานสำคัญสารสนเทศ ตาม พรบ. การรักษาความมั่นคงปลอดภัยทางไซเบอร์ พ.ศ. 2562 ควรรักษาความมั่นคงปลอดภัยทางไซเบอร์กับระบบปัญญาประดิษฐ์ที่ให้บริการสอดคล้องกับกฎหมายฉบับนี้

<sup>32</sup> Council of Europe. Guidelines on Artificial Intelligence and Data Protection. Consultative Committee of the Convention for the Protection of individuals with regard to Automatic Processing of Personal Data. Jan 25 2019.

- ควรออกแบบและพัฒนาปัญญาประดิษฐ์ ให้มีกลไกในการตรวจสอบ เฝ้าระวัง และแจ้งเตือนถึงภัยคุกคาม รวมถึงความสามารถในการติดตาม แก้ไขปัญหา และกลับคืนสู่สภาวะปกติได้ภายหลังจากการถูกโจมตีโดยภัยคุกคาม (Resilience to attack)
  - ควรออกแบบและพัฒนาปัญญาประดิษฐ์ ให้มีความสามารถในการใช้แผนสำรองเพื่อกู้คืนระบบจากปัญหาต่าง ๆ (Fallback plan) โดยมีกลไกในการเพิกถอนข้อมูลและบริการ เพื่อให้มั่นใจได้ว่าสิทธิ์และประโยชน์ของผู้ใช้งานจะไม่ถูกละเมิด รวมถึงไม่สร้างผลกระทบต่อชีวิตและสิ่งแวดล้อม ในระหว่างการดำเนินการ
  - ผู้ให้บริการปัญญาประดิษฐ์ควรมีความรู้ที่จำเป็นและความสามารถในการใช้งานระบบตามที่ได้รับทราบไว้ เพื่อหลีกเลี่ยงและลดความเสี่ยงต่าง ๆ ที่อาจเกิดขึ้นจากการให้บริการ
  - ผู้วิจัย ผู้ออกแบบ และพัฒนาปัญญาประดิษฐ์ ควรพัฒนาความรู้และทักษะ การวิจัย ออกแบบ และพัฒนาปัญญาประดิษฐ์อย่างสม่ำเสมอ เพื่อเพิ่มความสามารถในการรักษาความมั่นคงปลอดภัยของระบบ และลดความเสี่ยงจากภัยคุกคามที่เกิดขึ้นใหม่
  - ผู้มีส่วนได้เสียที่มีความเสี่ยงจะได้รับผลกระทบจากปัญญาประดิษฐ์ ควรได้รับการแจ้งเตือน และควรมีส่วนร่วมในขั้นตอนของการออกแบบ พัฒนาการให้บริการปัญญาประดิษฐ์ด้วย
  - ปัญญาประดิษฐ์ที่อาจส่งผลกระทบต่อมนุษย์ ควรให้ผู้เชี่ยวชาญในแขนงที่จะนำปัญญาประดิษฐ์ไปใช้งานมีส่วนร่วมในขั้นตอนการออกแบบ พัฒนา และให้บริการ
- **ความเท่าเทียม หลากหลาย ครอบคลุม และเป็นธรรม (Fairness)**
    - ควรวิจัย ออกแบบ และพัฒนาปัญญาประดิษฐ์ โดยเข้าใจถึงความต้องการ และความคาดหวังของผู้ใช้งานที่มีความหลากหลายในสังคม คำนึงถึงกลุ่มคนส่วนน้อยและผู้ด้อยโอกาส อาทิ ผู้พิการและผู้มีภาวะทุพพลภาพต่าง ๆ และเคารพต่อความสามารถและศักยภาพของมนุษย์

- สมาชิกในกลุ่มผู้วิจัย ผู้ออกแบบ ผู้พัฒนา และผู้ทดสอบปัญญาประดิษฐ์ ควรมีประวัติพื้นหลังที่หลากหลาย และควรได้รับคำแนะนำจากผู้เชี่ยวชาญในแขนงที่จะนำไปปัญญาประดิษฐ์ไปใช้งาน เพื่อช่วยลดความเสี่ยงในการเกิดความไม่ธรรมขึ้นกับระบบ
- ผู้ที่มีส่วนได้เสียที่อาจได้รับผลกระทบจากความเอนเอียงและไม่ธรรมของปัญญาประดิษฐ์ ควรมีส่วนร่วมในขั้นตอนการออกแบบและพัฒนาระบบ
- ควรตรวจสอบและแก้ไขปัญญาประดิษฐ์ที่ให้ผลลัพธ์การตัดสินใจที่เอนเอียงที่ไม่ได้ออกแบบไว้ (Non-operational bias)<sup>33</sup> ซึ่งอาจเกิดขึ้นจากที่ปัญญาประดิษฐ์เข้าไปเกี่ยวข้องกับกลุ่มคนและกระบวนการที่แตกต่างกัน และการตั้งสมมติฐานที่เกี่ยวข้องกับผู้ใช้งานปัญญาประดิษฐ์ที่ไม่ถูกต้อง เป็นต้น
- ควรออกแบบปัญญาประดิษฐ์ให้มีทางเลือกที่หลากหลายให้กับผู้ใช้งานในการดำเนินการเพื่อบรรลุถึงเป้าหมาย
- ควรเก็บรวบรวมข้อมูลเพื่อนำมาใช้กับปัญญาประดิษฐ์จากหลากหลายแหล่งข้อมูลที่มีความน่าเชื่อถือ เพื่อลดความเอนเอียงของข้อมูล
- ชุดข้อมูลเพื่อนำมาใช้กับปัญญาประดิษฐ์ ควรตรวจสอบให้แน่ใจว่าสามารถใช้เป็นตัวแทนของประชากรที่ได้รับผลกระทบทั้งหมด (Representativeness)
- ชุดข้อมูลและอัลกอริทึมที่มีลักษณะเอนเอียงไปในทางแบ่งแยกและไม่เป็นธรรม ควรได้รับการกำจัดออกไปในขั้นตอนการเก็บข้อมูล และการคัดเลือกอัลกอริทึมตามลำดับ ทั้งนี้ในกรณีที่ไม่สามารถกำจัดความเอนเอียงของชุดข้อมูลและอัลกอริทึมได้ทั้งหมด ควรมีกลไกในการตรวจสอบความเอนเอียงและให้บริการปัญญาประดิษฐ์อย่างปลอดภัย
- ควรใช้ชุดข้อมูลในการสอน ทดสอบ และพิสูจน์โมเดลที่แตกต่างกัน เพื่อให้สามารถตรวจสอบพบความเอนเอียงของผลลัพธ์ที่อาจเกิดขึ้นได้ในแต่ละชุดข้อมูล
- ควรพัฒนาปัญญาประดิษฐ์ให้สอดคล้องกับมาตรฐานสากลหรือแนวทางด้านความสามารถในการเข้าถึงระบบที่เป็นที่ยอมรับ อาทิ ISO 40500:2012 (W3C

<sup>33</sup> Smart Dubai. AI Ethics Principles & Guidelines. Version 1.6. Dec 30 2018.

Web Content Accessibility Guidelines: (WCAG) 2.0)<sup>34</sup> และ W3C Web Content Accessibility Guidelines (WCAG) เวอร์ชัน 2.1<sup>35</sup>

- ควรใช้เครื่องมือในการตรวจสอบความเอนเอียงของปัญญาประดิษฐ์ (Bias Detection Tools) เพื่อช่วยลดความไม่เป็นธรรมในระบบ
- ผู้พัฒนาปัญญาประดิษฐ์ควรทดสอบเพื่อให้แน่ใจว่า ปัญญาประดิษฐ์สามารถให้ผลลัพธ์ตามที่ออกแบบไว้ได้เป็นอย่างดีเมื่อพบกับข้อมูลที่ไม่เคยเห็นในอดีต โดยเฉพาะเมื่อใช้กับการประเมินประชากรที่ไม่ได้อยู่ในข้อมูลชุดการสอน
- ให้ผู้พิการหรือผู้มีภาวะทุพพลภาพได้ร่วมทดสอบวัตถุประสงค์ กระบวนการทำงาน กระบวนการตัดสินใจและผลลัพธ์ที่ได้จากปัญญาประดิษฐ์ ว่าสามารถตอบสนองความต้องการและความคาดหวังที่สำคัญของผู้ใช้งานที่เป็นผู้พิการหรือผู้มีภาวะทุพพลภาพได้
- ผู้พัฒนาและผู้ให้บริการปัญญาประดิษฐ์ควรมีการเตรียมความพร้อมด้านการจัดอบรม เครื่องมืออุปกรณ์ที่เกี่ยวข้อง และช่องทางเข้าใช้งานปัญญาประดิษฐ์ เพื่อให้ผู้ใช้งานสามารถเข้าถึงได้โดยง่าย
- การพัฒนา และให้บริการปัญญาประดิษฐ์ควรตั้งอยู่บนพื้นฐานของความเป็นธรรมทางสังคม ให้เกิดความเท่าเทียมกันในการเข้าถึงสินค้า บริการ และเทคโนโลยีที่เกี่ยวข้อง โดยไม่จำกัดอายุ เพศ ลัทธิชนชาติ ความสามารถ และคุณสมบัติ

<sup>34</sup> ISO/IEC 40500:2012 (W3C), Information technology – W3C Web Content Accessibility Guidelines (WCAG) 2.0. Oct 2012.

<sup>35</sup> W3C. Web Content Accessibility Guidelines (WCAG) 2.1. Jun 5 2018. ดูได้จาก <https://www.w3.org/TR/WCAG21/>.

- **ความน่าเชื่อถือ (Reliability)**

- ควรมีการวางแผนวิธีการวิจัย วิธีการออกแบบ วิธีการพัฒนา และวิธีการใช้งาน ปัญญาประดิษฐ์อย่างเป็นระบบและครบถ้วน
- ควรทราบและมีความเข้าใจถึงปัจจัยที่ส่งผลกระทบต่อคุณภาพของชุดข้อมูลที่ใช้ในปัญญาประดิษฐ์ อาทิ ความแม่นยำ (Accuracy) ความสมบูรณ์ (Completeness) ความน่าเชื่อถือ (Veracity) ความเป็นปัจจุบัน (Update) ความเกี่ยวข้องของข้อมูล (Relevance) ความครบถ้วนถูกต้อง (Integrity) ความสามารถในการนำไปใช้ (Usability) และการถูกแทรกแซงโดยมนุษย์ (Human Intervention) เป็นต้น
- ควรออกแบบและพัฒนาปัญญาประดิษฐ์ให้มีความสมบูรณ์ แข็งแกร่ง และน่าเชื่อถือ
  - มีความแม่นยำ (Accuracy) ในการตัดสินใจ การคาดการณ์ และให้คำแนะนำบนพื้นฐานของข้อมูลและโมเดล
  - สร้างผลลัพธ์ที่สามารถเชื่อถือได้และสร้างใหม่ได้เมื่อต้องการ (Reliability and Reproducibility)
  - ความเป็นมิตรต่อการผู้ใช้งาน (Usability)
- ควรทราบถึงแหล่งกำเนิดของข้อมูลที่จะนำไปใช้ในโมเดล วิธีการรวบรวมข้อมูล การเก็บรักษา การเคลื่อนย้ายของข้อมูล การเปลี่ยนแปลงข้อมูล การปฏิสัมพันธ์กับข้อมูลอื่น จนถึงการนำเสนอข้อมูล เพื่อให้ทราบถึงคุณภาพของข้อมูล ที่นำมาใช้ รวมถึงข้อผิดพลาดที่สำคัญต่าง ๆ
- ควรพิจารณาก่อนนำข้อมูลที่ดำเนินการเก็บรวบรวมไปใช้ ว่าบริบทที่ใช้ในขั้นตอนการเก็บรวบรวม ข้อมูลเหมาะสมกับลักษณะที่จะนำไปใช้งาน ตามที่ตั้งใจไว้จริง
- ควรรับคำปรึกษาจากผู้เชี่ยวชาญในแขนงที่เกี่ยวข้องในการพัฒนาปัญญาประดิษฐ์ที่เกี่ยวข้องกับข้อมูลที่มีความอ่อนไหว (Sensitive Data) อาทิ ข้อมูลทางการแพทย์ ข้อมูลพนักงาน ข้อมูลทางการเงิน และการบังคับใช้กฎหมาย เป็นต้น

- ควรดำเนินการทดสอบโมเดลปัญญาประดิษฐ์ให้สอดคล้องกับความเป็นพลวัตของเงื่อนไขและสภาวะแวดล้อมจริง เพื่อความปลอดภัยของผู้ใช้งานและสภาพแวดล้อมโดยรอบ และควรเฝ้าระวัง ทบทวน และปรับปรุงโมเดลอย่างเหมาะสม
- ควรมีแนวทางในการพิสูจน์ว่าปัญญาประดิษฐ์ดำเนินการตามที่ได้ออกแบบไว้ภายใต้เงื่อนไขการปฏิบัติงานจริง
- ผู้ให้บริการปัญญาประดิษฐ์ควรมีกระบวนการเฝ้าระวัง บันทึกและตรวจสอบการทำงานของปัญญาประดิษฐ์ เพื่อช่วยในการตรวจสอบและทำความเข้าใจผลลัพธ์ทางประสิทธิภาพของปัญญาประดิษฐ์
- ผู้ให้บริการปัญญาประดิษฐ์ควรมีช่องทางการรับผลสะท้อนกลับ (Feedback) รวมถึงช่องทางการขอให้ทบทวนการตัดสินใจที่เกิดจากระบบ (Decision Review) ที่ง่ายและรวดเร็ว เพื่อให้ผู้ใช้งานสามารถแจ้งปัญหาของปัญญาประดิษฐ์ที่ตรวจสอบพบคำร้องและนำเสนอต่าง ๆ ได้

### 3) ผู้ใช้งาน

- ควรศึกษา อบรม เพื่อให้มีความรู้ทั่วไปเกี่ยวกับประโยชน์และตระหนักรู้ถึงผลกระทบของการใช้งานปัญญาประดิษฐ์ รวมถึงทักษะในการใช้งานและทำงานร่วมกับปัญญาประดิษฐ์อย่างถูกต้อง
- ควรติดตามข่าวสารที่เกี่ยวข้องกับเทคโนโลยีปัญญาประดิษฐ์อย่างสม่ำเสมอ เพื่อให้ทราบถึงองค์ความรู้ใหม่ๆ สำหรับปรับใช้ประโยชน์ในชีวิตประจำวัน และสามารถรับมือกับภัยคุกคามต่างๆ
- ควรตรวจสอบความน่าเชื่อถือของผลิตภัณฑ์และบริการปัญญาประดิษฐ์ก่อนรับบริการโดยดูจากรายละเอียดการออกแบบและหลักการทำงานของปัญญาประดิษฐ์ที่ผู้ออกแบบ ผู้พัฒนา และผู้ให้บริการแสดง ผลประเมินความน่าเชื่อถือของระบบและใบรับรองผลิตภัณฑ์และบริการจากสถาบันที่ได้รับการยอมรับ และผลสะท้อนกลับจากผู้ที่เคยรับบริการอื่น เป็นต้น
- ผู้ใช้งานควรร้องขอรายละเอียดหลักการจริยธรรมที่ใช้ในผลิตภัณฑ์และบริการปัญญาประดิษฐ์จากผู้ให้บริการ
- ผู้ใช้งานควรทราบถึงภาระความรับผิดชอบ (Accountability) ของตนเองในการใช้งานผลิตภัณฑ์และบริการปัญญาประดิษฐ์
- ผู้ที่อาจได้รับผลกระทบจากปัญญาประดิษฐ์ ควรร้องขอข้อมูลเกี่ยวกับเหตุผลในการใช้งานปัญญาประดิษฐ์กับตนเอง รวมถึงผลกระทบของเหตุผลเหล่านั้น และควรมีส่วนร่วมในกระบวนการออกแบบ พัฒนา และประเมินความเสี่ยงของระบบปัญญาประดิษฐ์
- ผู้ใช้งานควรแจ้งผลการใช้งานปัญญาประดิษฐ์สะท้อนกลับให้กับผู้ให้บริการเพื่อการปรับปรุง แก้ไข และพัฒนาให้สอดคล้องกับความวัตถุประสงค์ และพฤติกรรมการใช้งานที่เปลี่ยนแปลงตามระยะเวลา
- ควรร้องขอคำอธิบายที่ง่ายแก่การเข้าใจและสอดคล้องกับบริบทที่เกี่ยวข้องจากผู้วิจัย ผู้ออกแบบ และผู้ให้บริการ เกี่ยวกับรายละเอียดการออกแบบ

และหลักการทำงานของปัญญาประดิษฐ์ เพื่อให้สามารถตั้งข้อสงสัย ตรวจสอบ  
ตระหนักถึงความเอนเอียง ความผิดพลาด การทำงานของปัญญาประดิษฐ์ได้

- หากมีการนำข้อมูลส่วนตัวไปใช้เพื่อการวิจัย ออกแบบ พัฒนา และให้บริการ  
ปัญญาประดิษฐ์ ผู้ใช้งานมีสิทธิตามกฎหมาย พรบ.คุ้มครองข้อมูลส่วนบุคคล  
ปี 2562 อาทิเช่น
  - การขอเข้าถึงและรับสำเนาข้อมูลส่วนบุคคลที่เกี่ยวกับตนเอง และขอให้  
เปิดเผยถึงการได้มาซึ่งข้อมูลดังกล่าว
  - มีสิทธิในการถอนความยินยอมและคัดค้านการเก็บรวบรวม ใช้ หรือ  
เปิดเผยข้อมูล
  - มีสิทธิขอให้ผู้ควบคุมข้อมูลส่วนบุคคลของผู้วิจัย ผู้ออกแบบ ผู้พัฒนา  
และผู้ให้บริการ ลบหรือทำลาย และระงับการใช้งานข้อมูลชั่วคราว
  - มีสิทธิทำให้ข้อมูลส่วนบุคคลของตนเองถูกต้องเป็นปัจจุบัน สมบูรณ์  
ไม่ก่อให้เกิดความเข้าใจผิด



## อภิธานศัพท์ (Glossary)

**ปัญญาประดิษฐ์** หมายถึง ศาสตร์ที่รวบรวมองค์ความรู้ในหลายสาขาวิชา โดยเฉพาะอย่างยิ่งทางด้านวิทยาศาสตร์และวิศวกรรมศาสตร์ มาพัฒนาให้เครื่องจักรหรือระบบคอมพิวเตอร์มีความชาญฉลาด สามารถคิด คำนวณ วิเคราะห์ เรียนรู้และตัดสินใจ โดยใช้เหตุผลได้เสมือนสมองของมนุษย์ และสามารถเรียนรู้ พัฒนาและปรับปรุงกระบวนการทำงานเพื่อเพิ่มศักยภาพของปัญญาประดิษฐ์เองได้

**ผู้วิจัยปัญญาประดิษฐ์** หมายถึง ผู้ค้นคว้าหาองค์ความรู้ปัญญาประดิษฐ์ใหม่ ๆ หรือเพื่อประยุกต์ใช้ องค์ความรู้ปัญญาประดิษฐ์ที่มีอยู่ โดยมีกระบวนการคิดอย่างเป็นระบบและใช้วิธีการหรือเทคนิคที่ได้รับการยอมรับจากทุกศาสตร์ที่เกี่ยวข้องว่ามีความน่าเชื่อถือ

**ผู้ออกแบบปัญญาประดิษฐ์** หมายถึง ผู้มีหน้าที่ศึกษาและวิเคราะห์ความต้องการของผู้ใช้งาน ปัญญาประดิษฐ์มาจัดทำแบบแผนในการสร้างปัญญาประดิษฐ์ เพื่อให้สามารถใช้งานได้จริงและตรงตามความต้องการของผู้ใช้งานปัญญาประดิษฐ์

**ผู้พัฒนาปัญญาประดิษฐ์** หมายถึง ผู้มีหน้าที่พัฒนาปัญญาประดิษฐ์ตามแบบแผนที่ผู้ออกแบบ ปัญญาประดิษฐ์ออกแบบไว้ และทำการทดสอบใช้งานปัญญาประดิษฐ์เพื่อตรวจสอบว่าปัญญาประดิษฐ์ สามารถใช้งานได้จริงและตรงตามความต้องการ

**ผู้ใช้งานปัญญาประดิษฐ์** หมายถึง กลุ่มบุคคลผู้เป็นผู้รับบริการปัญญาประดิษฐ์ โดยมีหน้าที่ระบุ ปัญหาและความต้องการที่ต้องการให้ปัญญาประดิษฐ์ให้ความช่วยเหลือและแก้ไขปัญหาให้โดยหมายความ รวมถึงประชาชนทุกคนและกลุ่มคนส่วนน้อย เช่น ผู้ด้อยโอกาส ผู้พิการ และผู้ทุพพลภาพ

**หน่วยงานควบคุมดูแลการพัฒนาและใช้งานปัญญาประดิษฐ์** หมายถึง หน่วยงานภาครัฐ ผู้มีหน้าที่ควบคุม กำกับ ดูแล บุคคลและหน่วยงานที่เกี่ยวข้องกับการวิจัย ออกแบบ พัฒนา และใช้งาน ปัญญาประดิษฐ์ ให้สอดคล้องกับจริยธรรมปัญญาประดิษฐ์ในทุกขั้นตอนการทำงาน โดยมีการกำหนดอำนาจ หน้าที่ความรับผิดชอบ และบทลงโทษไว้อย่างชัดเจน

**ความโปร่งใส (Transparency)** หมายถึง การที่สามารถอธิบายเหตุการณ์ การกระทำ กระบวนการทำงาน และกิจกรรมต่าง ๆ ที่เกิดขึ้นย้อนหลังได้ทั้งหมด เพื่อให้หน่วยงานควบคุมดูแล การพัฒนาและใช้งานปัญญาประดิษฐ์สามารถตรวจสอบกิจกรรมต่าง ๆ ว่าดำเนินไปด้วยความถูกต้อง และสามารถคาดการณ์การกระทำต่าง ๆ ได้

**ภาระความรับผิดชอบ (Accountability)** หมายถึง ภาระความรับผิดชอบที่มี หากเกิดผลกระทบที่เกิดขึ้นจากปัญญาประดิษฐ์ โดยให้ผู้ที่เกี่ยวข้องกับปัญญาประดิษฐ์ เช่น ผู้วิจัย ผู้ออกแบบ ผู้พัฒนา และผู้ใช้งานปัญญาประดิษฐ์ รับผิดชอบเยียวยาผู้ได้รับผลกระทบเฉพาะที่เกิดจากขอบเขตภาระหน้าที่เฉพาะส่วนของตนเท่านั้น

**ความสามารถในการสืบย้อน (Traceability)** หมายถึง ความสามารถในการตรวจสอบย้อนกลับไปตั้งแต่แหล่งที่มาของชุดข้อมูล กระบวนการทำงานและการตัดสินใจของปัญญาประดิษฐ์ เพื่อใช้ในการเฝ้าระวัง ตรวจสอบความผิดปกติที่พบ และสามารถวินิจฉัยปัญหาที่ทำให้เกิดความผิดพลาดล้มเหลวได้

**ความมั่นคงปลอดภัย (Security)** หมายถึง การสร้างความมั่นคงปลอดภัยให้แก่ปัญญาประดิษฐ์ โดยอาจใช้นโยบายและมาตรฐานทางเทคนิคด้านความมั่นคงปลอดภัย การเฝ้าระวัง การประเมินและจัดการความเสี่ยง การคุ้มครองความเป็นส่วนตัว สำหรับการวิจัย ออกแบบ พัฒนาและใช้งานปัญญาประดิษฐ์ เพื่อลดช่องโหว่และป้องกันภัยคุกคามของปัญญาประดิษฐ์ที่ก่อให้เกิดผลกระทบในด้านลบ รวมถึงผลกระทบด้านจริยธรรม ชีวิตและสิ่งแวดล้อม และให้ผู้วิจัย ผู้ออกแบบ ผู้พัฒนา และผู้ให้บริการปฏิบัติตาม

**ความเป็นส่วนตัว (Privacy)** หมายถึง ข้อมูลส่วนบุคคลที่เป็นสิทธิเสรีภาพส่วนบุคคลของมนุษย์ ซึ่งหากต้องการนำข้อมูลส่วนบุคคลของผู้ใดไปใช้งาน จะต้องไม่ดำเนินการขัดกับกฎหมายคุ้มครองข้อมูลส่วนบุคคล โดยต้องทำการแจ้งให้ผู้ใช้งานปัญญาประดิษฐ์ทราบเป็นการล่วงหน้าถึงข้อมูลที่จะถูกเก็บรวบรวมและการนำไปใช้ และข้อมูลส่วนบุคคลดังกล่าวต้องได้รับการยินยอมจากเจ้าของข้อมูลเสียก่อน

**ความเป็นธรรม (Fairness)** หมายถึง ความเท่าเทียมกันทางด้านโอกาสในสังคม โดยผู้ที่จะได้รับประโยชน์จากปัญญาประดิษฐ์ ควรเป็นประชาชนทุกคน รวมถึงกลุ่มคนด้อยโอกาส เช่น ผู้พิการและผู้ทุพพลภาพด้วย

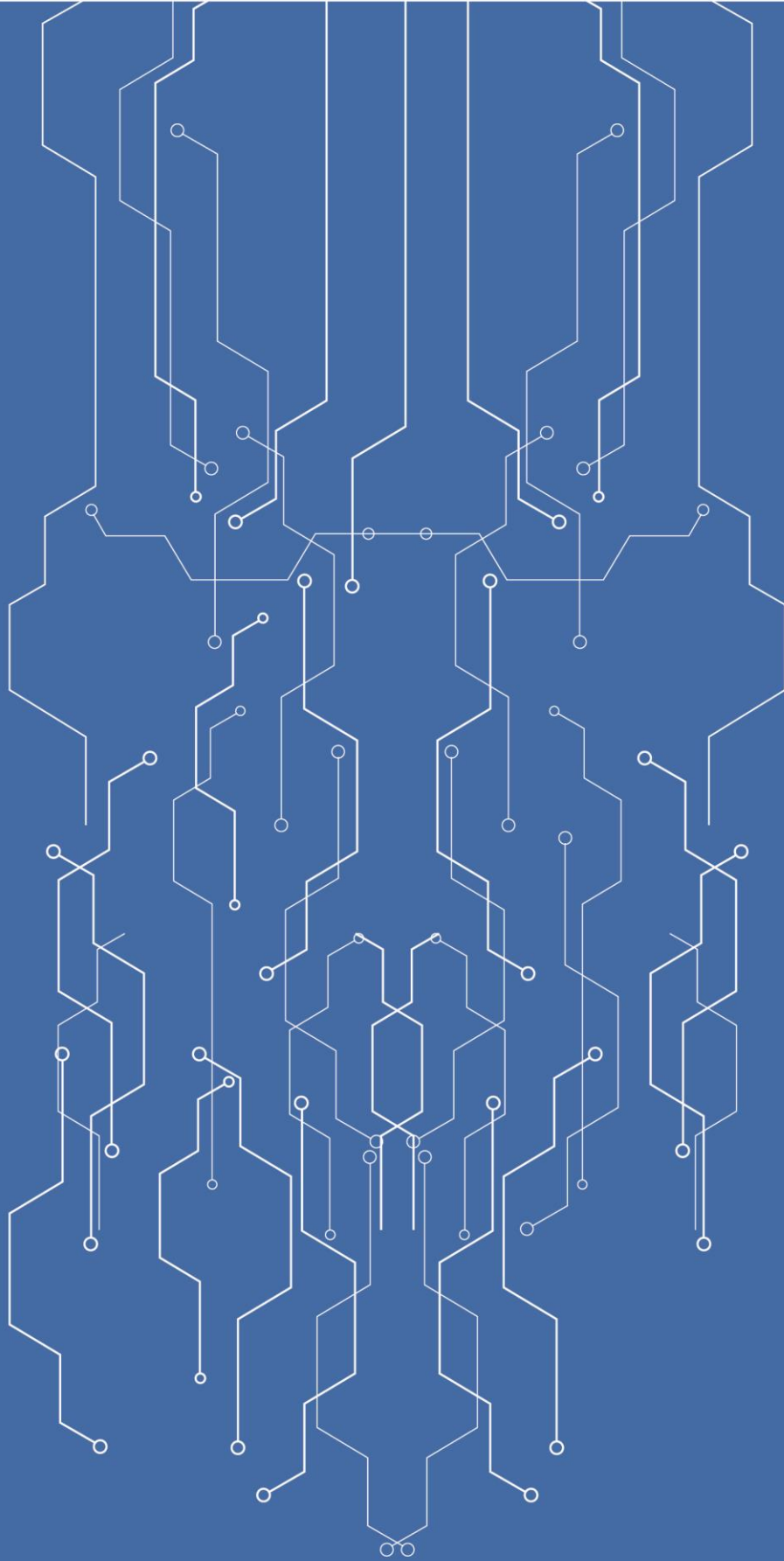
**ความน่าเชื่อถือ (Reliability)** หมายถึง คุณภาพของปัญญาประดิษฐ์ในด้านต่าง ๆ ที่ส่งผลให้เกิดความเชื่อถือต่อผู้ใช้งานปัญญาประดิษฐ์ เช่น ความถูกต้องแม่นยำ ความสมบูรณ์ ความเป็นปัจจุบัน ความเกี่ยวข้องของข้อมูล ความครบถ้วนถูกต้อง การคาดการณ์ได้ถูกต้องแม่นยำ และความสามารถในการนำไปใช้

## เอกสารอ้างอิง (Reference)

1. OECD Legal Instruments. Recommendation of the Council on Artificial Intelligence. May 22 2019. จาก <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>.
2. Beijing AI Principles. May 5 2019. จาก <https://www.baai.ac.cn/blog/beijing-ai-principles>.
3. European Commission. Ethics Guidelines for Trustworthy AI. Apr 8 2019. จาก <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>.
4. UNESCO and EQUALS Skills Coalition. I'd blush if I could: Closing gender divides in digital skills through education. 2019. จาก <https://unesdoc.unesco.org/ark:/48223/pf0000367416>.
5. The White House. Executive Order on Maintaining American Leadership in AI. Feb 11 2019. จาก <https://www.whitehouse.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/>.
6. Microsoft. The Future Computed, Artificial Intelligence and its role in society. Feb 2019. จาก [https://3er1viui9wo30pkxh1v2nh4w-wpengine.netdna-ssl.com/wp-content/uploads/2018/02/The-Future-Computed\\_2.8.18.pdf](https://3er1viui9wo30pkxh1v2nh4w-wpengine.netdna-ssl.com/wp-content/uploads/2018/02/The-Future-Computed_2.8.18.pdf).
7. Microsoft. Six principles to guide Microsoft's facial recognition work. Dec 2018. จาก <https://blogs.microsoft.com/on-the-issues/2018/12/17/six-principles-to-guide-microsofts-facial-recognition-work/>.
8. Microsoft. Responsible bots: 10 guidelines for developers of conversational AI. Nov 2018. จาก [https://www.microsoft.com/en-us/research/uploads/prod/2018/11/Bot\\_Guidelines\\_Nov\\_2018.pdf](https://www.microsoft.com/en-us/research/uploads/prod/2018/11/Bot_Guidelines_Nov_2018.pdf).
9. Smart Dubai. AI Ethics Principles & Guidelines. Version 1.6. Dec 30 2018.
10. Council of Europe. Guidelines on Artificial Intelligence and Data Protection. Consultative Committee of the Convention for the Protection of individuals with regard to Automatic Processing of Personal Data. Jan 25 2019.
11. Personal Data Protection Commission, Singapore. A Proposed Model Artificial Intelligence Governance Framework. Jan 2019.

12. David Silver et al. Mastering the game of Go with Deep neural networks and tree search. Nature. Jan 2016. จาก <https://storage.googleapis.com/deepmindmedia/alphago/AlphaGoNaturePaper.pdf>.
13. Monegain B. IBM Watson pinpoints rare form of leukemia after doctor misdiagnosed patient. Healthcare IT News. Aug 8 2016. จาก <https://www.healthcareitnews.com/news/ibm-watson-pinpoints-rare-form-leukemia-after-doctors-misdiagnosed-patient>.
14. Julia Angwin., Jeff Larson., Surya Mattu. and Lauren Kirchner. Machine Bias: There’s software used across the country to predict future criminals. And it’s biased against blacks. ProPublica. May 23 2016. จาก <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
15. Dhillon Kirat., Juyong Jang., and Marc Ph. Stoecklin. DeepLocker – Concealing Targeted Attacks with AI Locksmithing. Blackhat USA 2018. Aug 9 2018. จาก <https://www.blackhat.com/us-18/briefings/schedule/#deeplocker---concealing-targeted-attacks-with-ai-locksmithing-11549>.
16. ประกาศ เรื่อง ยุทธศาสตร์ชาติ (พ.ศ. 2561 – 2580). ราชกิจจานุเบกษา. 13 ตุลาคม พ.ศ. 2561 เล่ม 135 ตอนที่ 82 ก. หน้า 25. จาก [http://www.ratchakitcha.soc.go.th/DATA/PDF/2561/A/082/T\\_0001.PDF](http://www.ratchakitcha.soc.go.th/DATA/PDF/2561/A/082/T_0001.PDF).
17. แผนยุทธศาสตร์กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม พ.ศ. 2563 – 2567. หน้า 5. <http://www.mdes.go.th/assets/portals/1/files/%E0%B9%81%E0%B8%9C%E0%B8%99%E0%B8%A2%E0%B8%B8%E0%B8%97%E0%B8%98%E0%B8%A8%E0%B8%B2%E0%B8%A%E0%B8%95%E0%B8%A3%E0%B9%8C%20%E0%B8%94%E0%B8%A8.%2063-67%20final%281%29.pdf>.
18. David Gilbert. Russian weapons maker Kalashnikov developing killer AI robots. Vice News. Jul 13 2017. จาก [https://news.vice.com/en\\_us/article/vbzq8y/russian-weapons-maker-kalashnikov-developing-killer-ai-robots](https://news.vice.com/en_us/article/vbzq8y/russian-weapons-maker-kalashnikov-developing-killer-ai-robots).
19. Marcus Roth. AI in Military Drones and UAVs – Current Applications. Emerj. Jan 30 2019. จาก <https://emerj.com/ai-sector-overviews/ai-drones-and-uavs-in-the-military-current-applications/>.

20. Kate Conger. Google Employees Resign in Protest Against Pentagon Contract. GIZMODO. May 14 2018. จาก <https://gizmodo.com/google-employees-resign-in-protest-against-pentagon-con-1825729300>.
21. Osterman Research White Paper. The State of AI in Cybersecurity: The Benefits, Limitations and Evolving Questions. Dec 2018.
22. Abigail Abrams. Data Breach at Sears and Delta May Have Hit ‘Several Hundred Thousand’ Customers. TIME USA, LLC. Apr 6 2018. จาก <https://time.com/5230288/delta-sears-data-breach-credit-cards/>.
23. Sarah Perez. Microsoft silences its new A.I. bot Tay, after Twitter users teach it racism. Verizon Media. Mar 24 2016. จาก <https://techcrunch.com/2016/03/24/microsoft-silences-its-new-a-i-bot-tay-after-twitter-users-teach-it-racism/>.
24. Sarah Myers West., Meredith Whittaker., and Kate Crawford. Discrimination Systems: Gender, Race, and Power in AI. AINOW Institute. Apr 2019. จาก <https://ainowinstitute.org/discriminatingystems.pdf>.
25. Natasha Singer. Amazon Is Pushing Facial Technology That a Study Says Cloud Be Biased. The New York Times. Jan 24 2019. จาก <https://www.nytimes.com/2019/01/24/technology/amazon-facial-technology-study.html>.
26. Sophie Curtis. Google Photos labels black people as ‘gorillas’. The Telegraph. Jul 01 2015. จาก <https://www.telegraph.co.uk/technology/google/11710136/Google-Photos-assigns-gorilla-tag-to-photos-of-black-people.html>.
27. Jacob Snow. Amazon’s Face Recognition Falsely Matched 28 Members of Congress With Mugshots. American Civil Liberties Union. Jul 26 2018. จาก <https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28>.



**Digital Thailand - AI Ethics Guideline**  
**Ministry of Digital Economy and Society**