

จริยธรรมปัญญาประดิษฐ์ในการบริหารทรัพยากรมนุษย์กับการกำกับดูแลการตัดสินใจโดยมนุษย์ในบริบทประเทศไทย

Ethical Artificial Intelligence in Human Resource Management and Human Oversight of Decision-Making in the Thai Context

จิราพัชร มันทาพันธ์,¹ จักรพันธ์ กิตตินรรัตน์,^{1*} อรรถนัฐ อชีรวาวัฒน์,² อาภากรณ์ หาโสีะ³ และ นิพนธ์ สุวรรณกุล³
Jirapat Mantapan,^{1} Jakkrapan Kittinorarat,^{1*} Orranarth Acherayawathana,² Apaporn Hasoh³ and Nipphon Suwannakut³*

Received 9 February 2026, Revised 20 February 2026, Accepted 23 February 2026

ABSTRACT

Artificial intelligence (AI) innovations have transformed human resource management, from recruitment, performance evaluation and succession planning to employee capability development, enhancing the efficiency and accuracy of decision-making in organizations. However, the use of AI in human resource management also generates important ethical questions, such as bias, discrimination, privacy violations and a lack of transparency in decision-making processes. This article presents an overview of AI use in human resource management in the Thai context, together with an analysis of relevant ethical frameworks, with the aim of explaining the potential and limitations of these technologies, identifying structural ethical issues across the employee lifecycle and proposing responsible AI practices based on human oversight of decision-making. The main proposition of the article underlines the importance of maintaining the human role as reviewer and final decision-maker, alongside the design of governance frameworks and the development of HR professionals' competencies, in order to balance technological efficiency with the protection of employees' dignity and labour rights.

Keywords: Artificial intelligence, Human resource management, AI ethics, Human oversight of decision making

บทคัดย่อ

นวัตกรรมปัญญาประดิษฐ์ได้เข้ามาเปลี่ยนแปลงการบริหารทรัพยากรมนุษย์ ตั้งแต่การสรรหา การประเมินผลงาน การวางแผนสืบทอดตำแหน่ง ไปจนถึงการพัฒนาศักยภาพพนักงาน ช่วยเพิ่มประสิทธิภาพและความแม่นยำของการตัดสินใจในองค์กร แต่ก็สร้างคำถามด้านจริยธรรมสำคัญ เช่น อคติ การเลือกปฏิบัติ การละเมิดความเป็นส่วนตัว และความไม่โปร่งใส

^{1*} หลักสูตรบริหารธุรกิจบัณฑิต มหาวิทยาลัยราชภัฏพระนคร แขวงอนุสาวรีย์ เขตบางเขน จ.กรุงเทพฯ 10220

Business Administration, Phranakhon Rajabhat University Anusawari, Bang Khen, Bangkok 10220, Thailand.

² หลักสูตรเศรษฐศาสตร์ธุรกิจดิจิทัล มหาวิทยาลัยราชภัฏพระนคร แขวงอนุสาวรีย์ เขตบางเขน จ.กรุงเทพฯ 10220

Digital Business Economics Program, Phranakhon Rajabhat University, Anusawari, Bang Khen, Bangkok 10220, Thailand.

³ นักวิชาการอิสระ

Independent Academic

* Corresponding author: Tel. 06-1636-5416 E-mail address: jakkrapan@pnru.ac.th

ของกระบวนการตัดสินใจ บทความนี้นำเสนอภาพรวมของการใช้ปัญญาประดิษฐ์ในการบริหารทรัพยากรมนุษย์ในบริบทประเทศไทย ร่วมกับการวิเคราะห์กรอบจริยธรรมที่เกี่ยวข้อง โดยมุ่งอธิบายศักยภาพและข้อจำกัดของเทคโนโลยีระบุประเด็นจริยธรรมเชิงโครงสร้างในวงจรชีวิตการบริหารบุคคล และเสนอแนวทางใช้ปัญญาประดิษฐ์อย่างมีความรับผิดชอบบนฐานของการกำกับดูแลการตัดสินใจโดยมนุษย์

ผลการวิเคราะห์ พบว่า ปัญญาประดิษฐ์มีศักยภาพสูงในการลดภาระงานธุรการและเพิ่มการตัดสินใจเชิงข้อมูล แต่มีข้อจำกัดสำคัญด้านอคติเชิงโครงสร้าง ความทับซ้อนของโมเดล และความเสี่ยงต่อการละเมิดความเป็นส่วนตัว ประเด็นจริยธรรมหลักครอบคลุมทั้งอคติในข้อมูลฝึกสอน สิทธิในการขอคำอธิบาย การเฝ้าระวังในที่ทำงาน และช่องว่างด้านความรับผิดชอบ บทความเสนอโมเดล "HERO: Human-Ethical Review & Oversight" เป็นกรอบแนวคิดที่คำนึงถึงวัฒนธรรมองค์กรแบบลำดับขั้น และปรากฏการณ์อคติจากระบบอัตโนมัติ ในสังคมไทย เพื่อสร้างสมดุลระหว่างประสิทธิภาพทางเทคโนโลยีกับการคุ้มครองศักดิ์ศรีและสิทธิแรงงานของพนักงาน

คำสำคัญ: ปัญญาประดิษฐ์ การบริหารทรัพยากรมนุษย์ จริยธรรมปัญญาประดิษฐ์ วงจรชีวิตพนักงาน บริบทไทย

บทนำ

นวัตกรรมปัญญาประดิษฐ์ (Artificial Intelligence: AI) กำลังปฏิรูประบบการทำงานขององค์กรอย่างลึกซึ้งในศตวรรษที่ 21 โดยเฉพาะในด้านการบริหารทรัพยากรมนุษย์ (Human Resource Management: HRM) ซึ่งเป็นกลไกหลักในการขับเคลื่อนองค์กรสู่เป้าหมายเชิงกลยุทธ์ (พัฒนาศิณานู วรณรัตน์, 2567) ด้วยขีดความสามารถในการประมวลผลข้อมูลขนาดใหญ่ วิเคราะห์รูปแบบซ่อนเร้น และสร้างคำแนะนำเชิงกลยุทธ์ได้อย่างรวดเร็วและแม่นยำ ปัญญาประดิษฐ์จึงกลายเป็นเครื่องมือสำคัญที่ช่วยยกระดับประสิทธิภาพงานการบริหารทรัพยากรมนุษย์ตั้งแต่การสรรหาบุคลากร (Recruitment) การประเมินผลงาน (Performance Evaluation) การวางแผนสืบทอดตำแหน่ง (Succession Planning) ไปจนถึงการพัฒนาศักยภาพพนักงาน (Employee Development) (ฉัตรธีรพล ปรีชญาณภัสศิริ, 2562; Kayusi, Chavula, Omwenga, Juma, Kayus, González Vallejo, & Mishra, 2025)

สถานการณ์การนำปัญญาประดิษฐ์มาใช้ในภาคธุรกิจไทยแสดงให้เห็นแนวโน้มการขยายตัวอย่างต่อเนื่อง ผลการสำรวจ Thailand Digital Outlook จัดทำโดยสำนักงานคณะกรรมการดิจิทัลเพื่อเศรษฐกิจและสังคมแห่งชาติ พบว่าอัตราการใช้งานปัญญาประดิษฐ์ในภาคธุรกิจเพิ่มขึ้นจากร้อยละ 40.5 ในปี พ.ศ. 2565 เป็นร้อยละ 66.1 ในปี พ.ศ. 2568 โดยธุรกิจขนาดกลางและขนาดใหญ่มีส่วนการใช้งานสูงถึงร้อยละ 73.8 และร้อยละ 72.9 ตามลำดับ ขณะที่ผลสำรวจ AI Readiness Measurement ปี 2567 ของศูนย์ธรรมาภิบาลปัญญาประดิษฐ์ (AI Governance Center: AIGC) ภายใต้อำนาจสำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์ ระบุว่าองค์กรไทยที่ประยุกต์ใช้ปัญญาประดิษฐ์แล้วมีสัดส่วนร้อยละ 17.8 และอีกร้อยละ 73.3 มีแผนจะนำมาใช้ในอนาคตอันใกล้ การขยายตัวอย่างรวดเร็วเช่นนี้ทำให้ประเด็นจริยธรรมและการกำกับดูแลการใช้ปัญญาประดิษฐ์ในงานบริหารทรัพยากรมนุษย์มีความเร่งด่วนยิ่งขึ้น

อย่างไรก็ตาม ศักยภาพดังกล่าวมิได้มีเพียงด้านบวก การนำปัญญาประดิษฐ์มาใช้ในการบริหารทรัพยากรมนุษย์ทำให้เกิดความเสี่ยงและข้อจำกัดด้านจริยธรรมหลายประการ ได้แก่ อคติที่ซ่อนอยู่ในข้อมูล (Hidden Bias) ความไม่โปร่งใสของกระบวนการตัดสินใจ (Lack of Transparency) การละเมิดความเป็นส่วนตัวของข้อมูลส่วนบุคคล (Personal Data Privacy Violation) และความไม่ชัดเจนว่าผู้ใดต้องรับผิดชอบเมื่อเกิดความผิดพลาด (Accountability Gap) ประเด็นเหล่านี้มีความสำคัญอย่างยิ่งในงานบริหารทรัพยากรมนุษย์ เพราะการตัดสินใจแต่ละครั้งส่งผลกระทบต่อชีวิต การทำงาน และศักดิ์ศรีของพนักงาน

สำหรับประเทศไทย ได้เริ่มมีการขับเคลื่อนกรอบกำกับดูแลปัญญาประดิษฐ์อย่างมีจริยธรรมอย่างเป็นระบบมากขึ้น โดยมีการจัดทำทั้งแนวปฏิบัติภายในหน่วยงานและกรอบนโยบายระดับชาติที่สอดคล้องกัน สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติได้จัดทำแนวปฏิบัติจริยธรรมด้านปัญญาประดิษฐ์ (AI Ethics Guidelines) บนหลักการสำคัญ 7 ประการ ได้แก่ ความเป็นส่วนตัว (Privacy) การกำกับดูแลการตัดสินใจโดยมนุษย์ (Human Oversight) ความโปร่งใสและ

การอธิบายได้ (Transparency and Explainability) ความเป็นธรรม (Fairness) ความปลอดภัย (Safety) ความรับผิดชอบ (Accountability) และการยึดมนุษย์เป็นศูนย์กลาง (Human-centricity) (สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ, 2566) ขณะที่สำนักงานคณะกรรมการการดิจิทัลเพื่อเศรษฐกิจและสังคมแห่งชาติ (Office of the National Digital Economy and Society Commission) ได้ออกแนวทางจริยธรรมปัญญาประดิษฐ์แห่งชาติ (Thailand Artificial Intelligence Ethics Guideline) เพื่อเป็นกรอบจริยธรรมปัญญาประดิษฐ์ในระดับชาติ (สำนักงานคณะกรรมการการดิจิทัลเพื่อเศรษฐกิจและสังคมแห่งชาติ, 2564) ซึ่งมีความสอดคล้องกับแนวนโยบายและแนวปฏิบัติด้านธรรมาภิบาลปัญญาประดิษฐ์ (AI Governance) ของสำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์ (Electronic Transactions Development Agency: ETDA) ที่มุ่งยกระดับมาตรฐานจริยธรรมปัญญาประดิษฐ์ของไทยให้เทียบเคียงแนวนิยมสากลในระดับนานาชาติ (สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์, 2567)

สหภาพยุโรปได้จัดทำกฎหมายปัญญาประดิษฐ์ (EU Artificial Intelligence Act) ซึ่งจำแนกระบบปัญญาประดิษฐ์ตามระดับความเสี่ยง (Risk-based Classification) โดยกำหนดให้ระบบปัญญาประดิษฐ์ที่ใช้ในกระบวนการจ้างงานและบริหารบุคลากรเป็นระบบปัญญาประดิษฐ์ความเสี่ยงสูง (High-risk AI System) ที่ต้องปฏิบัติตามข้อกำหนดเข้มงวดด้านการบริหารความเสี่ยง ความโปร่งใส และการกำกับดูแลโดยมนุษย์ (European Union, 2024) ขณะเดียวกัน กฎหมายคุ้มครองข้อมูลส่วนบุคคลของสหภาพยุโรป (General Data Protection Regulation: GDPR) ได้รับรองสิทธิของบุคคลในการได้รับคำอธิบายเกี่ยวกับการตัดสินใจอัตโนมัติ (Right to Explanation) และสิทธิในการคัดค้านการประมวลผลดังกล่าว (Goodman & Flaxman, 2017; Edwards & Veale, 2017) แนวโน้มเหล่านี้ก่อให้เกิดแรงกดดันเชิงโครงสร้างให้ประเทศต่าง ๆ รวมถึงประเทศไทย พัฒนากฎหมาย AI ที่สามารถสร้างสมดุลระหว่างการส่งเสริมนวัตกรรมและการคุ้มครองสิทธิขั้นพื้นฐานของประชาชน (เชมภัทร ทฤษฎิกุล และวิษญาดา อามานิกวิวัฒน์, 2568)

บทความนี้ใช้กรอบแนวคิดจริยธรรมปัญญาประดิษฐ์ที่มีความรับผิดชอบ (Responsible AI Ethics Framework) เป็นฐานในการวิเคราะห์ โดยบูรณาการหลักการจากแนวปฏิบัติจริยธรรมด้านปัญญาประดิษฐ์ของสำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ กฎหมายปัญญาประดิษฐ์ของสหภาพยุโรป (EU AI Act) และแนวคิดการกำกับดูแลการตัดสินใจโดยมนุษย์ (Human Oversight) จากผู้ตรวจการคุ้มครองข้อมูลยุโรป (European Data Protection Supervisor: EDPS) ในด้านวิธีวิทยา บทความนี้เป็นบทความวิชาการเชิงวิเคราะห์ (Analytical Article) ที่ใช้วิธีการสังเคราะห์เอกสาร (Documentary Synthesis) โดยรวบรวมและวิเคราะห์วรรณกรรมวิชาการ เอกสารนโยบาย กรอบกฎหมาย และแนวปฏิบัติด้านจริยธรรมปัญญาประดิษฐ์ทั้งในระดับประเทศและระดับสากล จากนั้นสังเคราะห์องค์ความรู้เพื่อนำเสนอข้อเสนอเชิงแนวคิดและเชิงปฏิบัติสำหรับการใช้ปัญญาประดิษฐ์ในการบริหารทรัพยากรมนุษย์อย่างมีจริยธรรมในบริบทไทย

1. นวัตกรรมปัญญาประดิษฐ์กับการประยุกต์ใช้ในการบริหารทรัพยากรมนุษย์

การพูดถึงปัญญาประดิษฐ์ในงานทรัพยากรมนุษย์มักเริ่มต้นด้วยน้ำเสียงเชิงบวก เน้นความสามารถของระบบในการจัดการข้อมูลจำนวนมากได้รวดเร็วและ ไม่มีอคติ ประหนึ่งเป็นผู้ตัดสินที่ปราศจากอคติโดยธรรมชาติแต่เมื่อพิจารณาอย่างใกล้ชิดจะพบว่าปัญญาประดิษฐ์ไม่ได้ลอยตัวอยู่เหนือโครงสร้างสังคม หากกลับสะท้อนและขยายความลำเอียงในข้อมูลในอดีตอย่างแนบเนียน (Barocas & Selbst, 2016; Raghavan *et al.*, 2020) นั่นทำให้การวิเคราะห์บทบาทของ AI ใน HR จำเป็นต้องมองออกจากรอบ เทคโนโลยีช่วยเพิ่มประสิทธิภาพ ไปสู่คำถามว่าเทคโนโลยีดังกล่าวกำลังจัดระเบียบความสัมพันธ์เชิงอำนาจ (Power Relations) ในองค์กรใหม่อย่างไร

ในระดับแนวคิด ปัญญาประดิษฐ์ใน HR ปัจจุบันส่วนใหญ่ยังเป็นปัญญาประดิษฐ์แบบจำกัด (Narrow AI) คือระบบที่ทำงานได้เฉพาะด้าน เช่น การคัดกรองเรซูเม่ ใช้แชทบอตสัมภาษณ์เบื้องต้น การวิเคราะห์ประสิทธิภาพงาน การทำนายความเสี่ยงลาออก หรือการแนะนำหลักสูตรพัฒนาทักษะ (ฉัตรธีรพล ปรัชญาลภัสศิริ, 2562; พัฒน์ดิณีนากู วรณรัตน์, 2567) การเดินหน้าไปในทิศทางนี้ทำให้ฝ่าย HR ลดภาระงานธุรการ หันมาเน้นบทบาทเชิงกลยุทธ์มากขึ้น และเปิดพื้นที่ใหม่ให้กับการใช้ข้อมูลเชิงลึกเพื่อออกแบบประสบการณ์พนักงานอย่างละเอียด (Kayusi *et al.*, 2025) หากมองจากแง่นี้

ปัญญาประดิษฐ์จึงมีลักษณะเป็นเครื่องมือที่เอื้อให้ฝ่ายทรัพยากรมนุษย์สามารถมุ่งเน้นบทบาทเชิงกลยุทธ์ได้มากยิ่งขึ้น มากกว่าจะเป็นภัยคุกคามโดยตรง

อย่างไรก็ตาม การเปลี่ยนผ่านสู่ HR เชิงข้อมูล (data-driven HR) ซึ่งหมายถึงการบริหารบุคลากรที่ใช้ข้อมูลและสถิติเป็นฐานหลักในการตัดสินใจ แทนที่จะพึ่งพาดุลยพินิจส่วนบุคคลเพียงอย่างเดียว ก็แฝงความเสี่ยงที่จะลดทอนมนุษย์ให้เหลือเพียง วัตถุดิบข้อมูล ที่ต้องถูกจับตามอง วัดผล และจัดอันดับอยู่ตลอดเวลา การใช้ AI คัดกรองผู้สมัครบนพื้นฐานของ คำสำคัญและแบบแผนในเรซูเมอาจช่วยลดเวลา แต่ก็เสี่ยงทำให้ผู้สมัครที่มีเส้นทางชีวิตไม่ตรงตาม มาตรฐาน เช่น คนเปลี่ยน อาชีพ หรือผู้ที่มีประสบการณ์ไม่เป็นทางการ ถูกกันออกไปก่อนจะได้แสดงศักยภาพจริง ในแง่นี้ AI ไม่ได้แค่ช่วย HR เห็นได้ ไกลขึ้น หากยังมีอำนาจกำหนดว่าใคร ควรค่าแก่การถูกเห็น ตั้งแต่ต้นทาง

ในขั้นตอนการประเมินผลงานและการพัฒนา ระเบียบวิเคราะห์ข้อมูลจากแพลตฟอร์มการทำงาน การบันทึกเวลา การมีส่วนร่วมในกิจกรรม และแม้แต่ปฏิสัมพันธ์บนระบบสื่อสารภายใน เพื่อวัด ความมีส่วนร่วม และ ประสิทธิภาพ (ฉัตรธีรพล ปรชญาณภัสศิริ, 2562) เครื่องมือเหล่านี้เปิดโอกาสให้เห็นพฤติกรรมที่มนุษย์อาจมองไม่เห็น เช่น รูปแบบการสื่อสารที่ บ่งชี้ความสอดคล้องของทีม หรือสัญญาณเริ่มต้นของการหมดไฟ แต่ขณะเดียวกันก็เปลี่ยนพื้นที่การทำงานให้กลายเป็นพื้นที่ การเฝ้าระวังอย่างถาวร ถ้าองค์กรไม่มีกติกาที่ชัดเจนและโปร่งใส พนักงานอาจรู้สึกถูกจับผิดมากกว่าถูกดูแล

AI ยังเข้ามามีบทบาทในการวิเคราะห์ข้อมูลทรัพยากรมนุษย์เพื่อทำนายแนวโน้มลาออก ระบุ พนักงานดาวเด่น และวางแผนสืบทอดตำแหน่ง โดยใช้ตัวชี้วัดการทำงาน การฝึกอบรม และเส้นทางอาชีพในอดีตเป็นฐานข้อมูล (พัฒนศิณีนากู วรณรัตน์, 2567) ในเชิงบวก สิ่งนี้ช่วยให้การจัดการบุคลากรเชิงกลยุทธ์ ทันเวลา มากขึ้น แต่ในอีกด้านหนึ่ง การทำนายล่วงหน้าว่าใคร เสี่ยงลาออก หรือ มีศักยภาพต่ำ หากไร้การกำกับ อาจกลายเป็นคำทำนายที่ตอกย้ำตัวเอง (self-fulfilling prophecy) เมื่อ HR และผู้จัดการปฏิบัติต่อบุคคลนั้นตามภาพที่ระบบวาดให้ จนทำให้เขาหรือเธอไม่มีโอกาส พิสูจน์ตนเอง

ในบริบทไทย การรับ AI เข้าไปใน HR ยังเชื่อมโยงกับโครงการระดับชาติเช่น Thailand 4.0 ยุทธศาสตร์ AI แห่งชาติ และความพยายามผลักดันประเทศเป็น ฮับ AI ของภูมิภาค สิ่งนี้ก่อให้เกิดแรงกดดันทางนโยบายให้ภาคธุรกิจเร่ง นำเทคโนโลยีมาใช้ โดยอาจจะละเลยการตั้งคำถามเชิงจริยธรรมและผลกระทบต่อแรงงานระดับล่างและแรงงานสูงอายุที่มี ทักษะดิจิทัลจำกัด (สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์, 2567; World Bank, 2024) จึงอาจกล่าวได้ว่า AI ใน HR ไม่ได้เป็นเพียงเทคโนโลยีเสริมประสิทธิภาพ แต่เป็นส่วนหนึ่งของโครงการสังคมที่กำลังคัดกรองว่าใครคือ แรงงานแห่ง อนาคต และใครเสี่ยงถูกทิ้งไว้ข้างหลัง

2. จริยธรรมปัญญาประดิษฐ์ในการบริหารทรัพยากรมนุษย์

เมื่อมองผ่านกรอบจริยธรรม ปัญญาประดิษฐ์ใน HR ทำให้คำถามดั้งเดิมว่า อะไรคือการปฏิบัติที่เป็นธรรม ซับซ้อน ขึ้นอย่างมาก เพราะการตัดสินใจไม่ได้มาจากผู้บังคับบัญชาที่มองเห็นได้ และรับผิดชอบได้ชัดเจนเท่านั้น หากยังมาจาก อัลกอริทึมที่ซ่อนตัวอยู่ในโค้ด แพลตฟอร์มของบริษัทเทคโนโลยี และชุดข้อมูลในอดีตที่พนักงานไม่มีโอกาสตรวจสอบ (Hongladarom, 2021) ภายใต้สถานการณ์นี้ การอ้างว่าเป็นการตัดสินใจของระบบอาจกลายเป็นวาทกรรมที่เสี่ยงทำให้ ความรับผิดชอบเลื่อนไปจากตัวบุคคลและองค์กร

ประเด็นแรกที่น่าสนใจ คือ อคติเชิงโครงสร้าง (structural bias) หมายถึงอคติที่ฝังอยู่ในข้อมูลและระบบตั้งแต่ต้น โดยที่ผู้ใช้อาจไม่รู้ตัว ระบบ AI ที่เรียนรู้จากประวัติการจ้างงานในอดีตมักทำให้รูปแบบการเลือกปฏิบัติที่มีอยู่แล้วฝังแน่น ยิ่งขึ้น ไม่ว่าจะเป็นมิติเพศ อายุ สถาบันการศึกษา หรือภูมิภาค (Barocas & Selbst, 2016; Raghavan *et al.*, 2020) ใน บริบทไทย ความเหลื่อมล้ำด้านโอกาสทางการศึกษาระหว่างกรุงเทพฯ กับต่างจังหวัด รวมถึงช่องว่างทักษะดิจิทัลของ ผู้สูงอายุและแรงงานกลุ่มเปราะบาง ทำให้กลุ่มเหล่านี้เสี่ยงถูกกันออกจากตลาดแรงงานที่พึ่งพาประสิทธิภาพและ AI เป็น กลไกหลัก (Ayele *et al.*, 2025) หากไม่มีมาตรการถ่วงดุลอย่างจริงจัง ระบบสรรหาที่อ้างว่า วัดจากความสามารถล้วน ๆ อาจกลายเป็นเครื่องผลิตซ้ำความเหลื่อมล้ำโดยไม่รู้ตัว

ประเด็นที่สอง คือ ความทึบแสงและสิทธิในการขอคำอธิบาย คำว่า ความทึบแสงของโมเดล (model opacity) หรือที่มักเรียกว่า กล่องดำ (black box) หมายถึงสภาพที่คนภายนอกมองไม่เห็นวาระบบ AI ใช้ข้อมูลปัจจัยใด ให้น้ำหนักอย่างไร และคิดอย่างไรก่อนจะตัดสินใจ ส่วน สิทธิในการขอคำอธิบาย (Right to Explanation) หมายถึงสิทธิของบุคคลที่จะได้รับคำชี้แจงที่เข้าใจได้ว่าระบบตัดสินใจเกี่ยวกับตัวเขาอย่างไรและเพราะอะไร ลักษณะกล่องดำนี้ทำให้ทั้ง HR และพนักงานไม่อาจเข้าใจเหตุผลแท้จริงเบื้องหลังการตัดสินใจ เช่น ใครถูกคัดออกจากการสมัครงาน หรือเหตุใดคะแนนประเมินผลงานจึงต่ำกว่าที่คาด (Hongladarom, 2021) การถกเถียงในยุโรปรอบ Right to Explanation ภายใต้ GDPR สะท้อนความพยายามทางกฎหมายในการ เปิดกล่องดำ เพื่อยืนยันว่าบุคคลมีสิทธิรับรู้เหตุผลหลักที่ใช้ตัดสินใจของตนเอง (Goodman & Flaxman, 2017; Edwards & Veale, 2017) แนวคิดนี้เชื่อมโยงกับสิ่งที่เรียกว่า ปัญญาประดิษฐ์ที่อธิบายได้ (Explainable AI: XAI) ซึ่งคือการออกแบบระบบ AI ให้สามารถบอกได้ว่าปัจจัยสำคัญอะไรบางที่นำไปสู่ผลลัพธ์ โดยอธิบายในภาษาที่คนทั่วไปเข้าใจได้ ไม่จำเป็นต้องเปิดเผยรายละเอียดเชิงเทคนิคทั้งหมด แต่ต้องเพียงพอให้ผู้ได้รับผลกระทบสามารถตั้งคำถามและทักท้วงได้ ในประเทศไทย พระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล (Personal Data Protection Act: PDPA) และแนวทางจริยธรรม AI ของหน่วยงานรัฐเริ่มหยิบประเด็นนี้ขึ้นมาพูดถึง โดยร่างหลักการกฎหมาย AI เสนอให้ผู้ได้รับผลกระทบจากระบบ AI มีสิทธิได้รับแจ้งผล การอธิบาย และการอุทธรณ์ หากแต่ในทางปฏิบัติ การเปลี่ยนหลักการเหล่านี้ให้เป็นสิทธิเชิงรูปธรรมของลูกจ้างยังเป็นหนทางยาวไกล (Kaminski & Malgieri, 2021)

ประเด็นที่สาม คือ ความเป็นส่วนตัวและการเฝ้าระวัง กล่าวคือ การที่ AI ต้องอาศัยข้อมูลส่วนบุคคลเชิงลึกตั้งแต่ระดับประสิทธิภาพการทำงานไปจนถึงพฤติกรรมการใช้อินเทอร์เน็ต ทำให้เส้นแบ่งระหว่างการดูแลกับการละเมิดกลายเป็นพื้นที่พร่าเลือน (สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ, 2566) การใช้ข้อมูลเพื่อตรวจจับการหมดไฟหรือสัญญาณความเครียดอาจช่วยป้องกันปัญหาสุขภาพจิตได้ ทว่าหากไม่มีการสื่อสารที่โปร่งใสและขอความยินยอมอย่างแท้จริง พนักงานอาจรู้สึกถูกรังแกติดตามทุกฝีก้าว มากกว่าจะได้รับการ ใส่ใจ จากองค์กร

สุดท้าย ประเด็นความรับผิดชอบและความเป็นมนุษย์ ทำให้เห็นการปะทะกันระหว่างตรรกะประสิทธิภาพของเทคโนโลยีกับตรรกะศักดิ์ศรีของมนุษย์อย่างชัดเจน แม้ว่าร่างกฎหมาย AI ของไทยและกรอบสากลต่างย้ำว่า AI เป็นเพียงเครื่องมือ มนุษย์ต้องรับผิดชอบผลลัพธ์เสมอ แต่ในความเป็นจริง การกระจายความรับผิดชอบระหว่างผู้พัฒนา ผู้จัดหา และผู้ใช้กลับซับซ้อน โดยเฉพาะเมื่อองค์กรมักพึ่งพาซอฟต์แวร์สำเร็จรูปจากต่างประเทศ หากไม่ระบุความรับผิดชอบเชิงกฎหมายและเชิงจริยธรรมให้ชัดเจน ความเสียหายที่เกิดกับพนักงานจากการตัดสินใจของ AI อาจกลายเป็นอุบัติเหตุทางเทคนิคที่ไม่มีใครยอมรับผิดจริงจัง.

กล่าวโดยสรุป จริยธรรม AI ใน HR จึงไม่ได้เป็นเพียงการปิดช่องว่างทางเทคนิค เช่น ปรับสูตรโมเดลให้เป็นธรรมขึ้น แต่ต้องเป็นการตั้งคำถามต่อโครงสร้างอำนาจ วิธีคิด และวัฒนธรรมองค์กรที่ทำให้การตัดสินใจของ AI ได้รับสถานะความจริงเหนือประสบการณ์ของมนุษย์ที่ได้รับผลกระทบโดยตรง

ตาราง 1 เปรียบเทียบกรอบกฎหมายและแนวปฏิบัติด้านปัญญาประดิษฐ์ ระหว่างสหภาพยุโรปกับประเทศไทย

ประเด็น	EU AI Act / GDPR	PDPA / ร่างกฎหมาย AI ของไทย
การจำแนกระดับความเสี่ยง	จำแนก 4 ระดับ — ระบบ AI ในงาน HR จัดเป็นความเสี่ยงสูง (High-risk) มีข้อบังคับเข้มงวด	ร่างหลักการกฎหมาย AI ปี 2568 เสนอแนวทางจำแนกตามความเสี่ยง คาดว่าจะมีผลบังคับใช้ในปี 2569
สิทธิในการขอคำอธิบาย (Right to Explanation)	GDPR มาตรา 22 รับรองสิทธิไม่ยอมรับการตัดสินใจอัตโนมัติ + สิทธิขอคำอธิบาย (Goodman & Flaxman, 2017)	PDPA กำหนดสิทธิเจ้าของข้อมูลในการเข้าถึงและคัดค้านการประมวลผล แต่ยังไม่มียกข้อยกเว้นเฉพาะเรื่องการตัดสินใจอัตโนมัติ
การกำกับดูแลโดยมนุษย์ (Human Oversight)	กำหนดให้ระบบ High-risk ต้องมีมนุษย์กำกับดูแลอย่างมีประสิทธิภาพ (European Union, 2024)	แนวปฏิบัติของ สวทช. ระบุหลักการกำกับดูแลโดยมนุษย์เป็น 1 ใน 7 หลักการ แต่ยังไม่มียกข้อยกเว้นทางกฎหมาย

ตาราง 1 เปรียบเทียบกรอบกฎหมายและแนวปฏิบัติด้านปัญญาประดิษฐ์ ระหว่างสหภาพยุโรปกับประเทศไทย (ต่อ)

ประเด็น	EU AI Act / GDPR	PDPA / ร่างกฎหมาย AI ของไทย
บทลงโทษ	ปรับสูงสุด 35 ล้านยูโร หรือร้อยละ 7 ของรายได้ทั่วโลก	PDPA ปรับสูงสุด 5 ล้านบาท + โทษจำคุกสูงสุด 1 ปี; ร่างกฎหมาย AI ยังไม่กำหนดบทลงโทษชัดเจน
หน่วยงานกำกับดูแล	European AI Office + หน่วยงานกำกับดูแลระดับประเทศสมาชิก	AIGC ภายใต้ ETDA + สำนักงานคณะกรรมการดิจิทัลฯ (ยังอยู่ระหว่างพัฒนาหลัก)

ตารางที่ 1 เปรียบเทียบกรอบกฎหมายและแนวปฏิบัติด้านจริยธรรมปัญญาประดิษฐ์ระหว่างสหภาพยุโรปกับประเทศไทย โดยแสดงให้เห็นถึงช่องว่างสำคัญของบริบทไทยที่ยังขาด (1) ผลบังคับใช้ทางกฎหมาย (2) บทลงโทษที่ชัดเจน และรุนแรง และ (3) กลไกหน่วยงานกำกับดูแลที่เป็นรูปธรรม ซึ่งเป็นประเด็นที่จำเป็นสำหรับการคุ้มครองสิทธิแรงงานในกระบวนการตัดสินใจโดย AI ในงานบริหารทรัพยากรมนุษย์

3. การกำกับดูแลการตัดสินใจโดยมนุษย์ (Human Oversight)

การกำกับดูแลการตัดสินใจโดยมนุษย์ (Human Oversight) ในบทความนี้ หมายถึงกระบวนการที่มนุษย์ยังคงมีอำนาจตระหนักและแทรกแซงได้อย่างแท้จริงในการตรวจสอบ ตีความ และปรับเปลี่ยนหรือยับยั้งผลลัพธ์การตัดสินใจของระบบปัญญาประดิษฐ์ (AI) โดยไม่จำกัดอยู่เพียงการรับรองผลลัพธ์ของระบบแบบไร้การพิจารณาเชิงวิเคราะห์ ซึ่งแนวคิดดังกล่าวมักถูกนำเสนอในฐานะกลไกรับประกันความปลอดภัย โดยชี้ให้เห็นว่าการมีมนุษย์เข้ามาเกี่ยวข้องในกระบวนการบริหารทรัพยากรมนุษย์ (HR) จะป้องกันความเสี่ยงจากการตัดสินใจอัตโนมัติโดยสมบูรณ์ อย่างไรก็ตาม เอกสารของผู้ตรวจการคุ้มครองข้อมูลยุโรป (European Data Protection Supervisor: EDPS) ได้วิเคราะห์อย่างรอบคอบว่าการปรากฏตัวของมนุษย์ในกระบวนการดังกล่าวมิได้รับประกันการหลีกเลี่ยงการตัดสินใจอัตโนมัติในทางปฏิบัติ หากมนุษย์มีส่วนร่วมเพียงการอนุมัติผลลัพธ์ของระบบโดยปราศจากการใช้วิจารณญาณวิเคราะห์ที่เป็นสาระสำคัญ ดังนั้น แนวคิด Human Oversight จึงอาจกลายเป็นกลไกเชิงภาพลักษณ์ที่ลดทอนความกังวลของสาธารณะ ในขณะที่อำนาจการตัดสินใจที่แท้จริงยังคงถูกขับเคลื่อนโดยระบบอัลกอริทึม (European Data Protection Supervisor, 2025)

EDPS ชี้ให้เห็นปรากฏการณ์ที่เรียกว่าอคติจากระบบอัตโนมัติ (Automation Bias) ซึ่งหมายถึงแนวโน้มที่มนุษย์จะเชื่อคำแนะนำของระบบมากเกินไป แม้ในกรณีที่ตนเองรู้สึกว่าคุณสมบัติอาจไม่ถูกต้อง โดยเฉพาะเมื่อระบบถูกวางกรอบว่าแม่นยำกว่ามนุษย์หรือผ่านการรับรองแล้ว ในบริบทการสรรหา การประเมิน หรือการเลิกจ้าง HR อาจมองว่าการทำตามคำแนะนำหรือคำแนะนำของระบบเป็นทางเลือกที่ปลอดภัยที่สุดในเชิงความรับผิดชอบ แม้จะรู้สึกไม่สบายใจกับผลลัพธ์ก็ตาม เมื่อรวมกับวัฒนธรรมลำดับชั้นและการหลีกเลี่ยงความขัดแย้งในองค์กรไทย การหักทวงผลของระบบ โดยเฉพาะระบบที่ผู้บริหารระดับสูงผลักดัน อาจถูกรับรู้ว่าเป็นพฤติกรรมที่ไม่เหมาะสมในบริบทองค์กร (Hongladarom, 2021).

การพูดถึงมนุษย์ในวงจร (Human in the Loop) คือรูปแบบที่กำหนดให้มนุษย์ต้องอนุมัติทุกครั้งก่อนที่ AI จะมีผลต่อคน เช่น ต้องมี HR อนุมัติก่อนรับหรือปฏิเสธผู้สมัคร ขณะที่มนุษย์คอยกำกับวงจร (Human on the Loop) คือรูปแบบที่ปล่อยให้ระบบทำงานอัตโนมัติในระดับหนึ่ง แต่มีมนุษย์คอยเฝ้าดูและพร้อมเข้าหยุดหรือแก้ไขเมื่อพบปัญหา ทั้งสองแนวทางนี้จะมีความหมายจริงก็ต่อเมื่อก้าวพ้นการออกแบบขั้นตอนเชิงเทคนิคไปสู่การสร้างบรรยากาศในองค์กรที่เปิดให้คนกล้าตั้งคำถามและกล้าไม่ทำตามระบบเมื่อจำเป็น เช่น การกำหนดให้ HR ต้องบันทึกเหตุผลทั้งกรณีเห็นด้วยและไม่เห็นด้วยกับคำแนะนำของระบบ การเปิดพื้นที่ให้มีการอภิปรายข้ามสายงานเกี่ยวกับผลกระทบของ AI ต่อแรงงาน และการสร้างภูมิคุ้มกันด้านจริยธรรมให้กับ HR ผ่านการฝึกอบรมเรื่อง Automation Bias ล้วนเป็นตัวอย่างของมาตรการที่ทำให้ Human Oversight มีความหมายจริงมากกว่าจะเป็นเพียงคำประกอบในเอกสารนโยบาย

อีกมิติหนึ่งที่ถูกกล่าวยกย่องคือ ภาระที่โยนให้มนุษย์ เอกสารของ EDPS และ Global Privacy Assembly (2024) เตือนว่าการฝากความหวังเรื่องความปลอดภัยและความเป็นธรรมไว้ที่ Human Oversight เพียงอย่างเดียว เป็นการมองข้ามความจำเป็นในการออกแบบระบบให้ปลอดภัยและเป็นธรรมตั้งแต่ต้น หากระบบยังเต็มไปด้วยอคติและขาดการ

ทดสอบอย่างจริงจัง มนุษย์ที่อยู่ปลายทางจะถูกบังคับให้ตัดสินใจบนข้อมูลที่บิดเบี้ยว การบอกให้ มนุษย์รับผิดชอบ จึงอาจเป็นการผลักภาระจากผู้ออกแบบและผู้กำกับดูแลระบบไปสู่ผู้ใช้งานด้านหน้าโดยไม่เป็นธรรม (European Data Protection Supervisor, 2025)

อย่างไรก็ดี แนวคิด Human Oversight ไม่ได้ควรถูกปฏิเสธ หากแต่ต้องถูก แปลความใหม่ ให้หมายถึงการสร้างกลไกที่ให้นมนุษย์มีอำนาจจริงในการขัดขวาง แก้ไข และวิพากษ์การตัดสินใจของ AI โดยตั้งอยู่บนข้อมูลที่ครบถ้วนและเข้าใจได้ การนำเทคนิค Explainable AI มาช่วยอธิบายปัจจัยที่ส่งผลต่อการตัดสินใจ การกำหนดให้พนักงานมีสิทธิขอคำชี้แจงและอุทธรณ์ การมีกระบวนการพิจารณาซ้ำโดยคณะกรรมการที่หลากหลาย และการเปิดโอกาสให้สหภาพแรงงานหรือผู้แทนพนักงานเข้ามามีส่วนร่วมในการกำกับดูแลระบบ AI ล้วนเป็นแนวทางที่เปลี่ยน Human Oversight จาก ตรายาง ความเป็นอิสระ ไปสู่พื้นที่ของการถกเถียงและต่อรองเชิงอำนาจอย่างแท้จริง (European Union, 2024)

ในบริบทไทยที่กำลังร่างกฎหมาย AI และผลักดันแนวทางธรรมาภิบาล AI ระดับชาติ การนิยาม Human Oversight ให้มีมิติสิทธิแรงงานและสิทธิมนุษยชนอย่างชัดเจนจึงเป็นเดิมพันสำคัญ หาก Human Oversight ถูกตีความเพียงว่า มีคนเซ็นอนุมัติอยู่ กฎหมาย AI อาจกลายเป็นเพียงเครื่องมือเพิ่มความชอบธรรมให้การตัดสินใจอัตโนมัติในตลาดแรงงาน มากกว่าจะเป็นเกราะคุ้มครองคนทำงานจากอำนาจใหม่ของอัลกอริทึม

4. โมเดลการทำงานร่วมกันระหว่างปัญญาประดิษฐ์กับมนุษย์ในวงจรชีวิตพนักงาน

โมเดลการทำงานร่วมกันระหว่าง AI กับมนุษย์ (human-AI collaboration model) ในบทความนี้ หมายถึงการออกแบบบทบาทและขอบเขตหน้าที่ของ AI และมนุษย์ในแต่ละช่วงของเส้นทางการทำงานของพนักงาน โดยยึดหลักว่า AI ทำหน้าที่ ช่วยขยายมุมมอง ไม่ใช่ ทำแทนทั้งหมด และการตัดสินใจที่กระทบชีวิตคนต้องมีมนุษย์รับผิดชอบขั้นสุดท้ายเสมอ

เมื่อหันมามองในระดับเชิงปฏิบัติ คำถามสำคัญคือ จะออกแบบการทำงานร่วมกันระหว่าง AI กับมนุษย์ใน HR อย่างไรให้ไม่ตกหลุมทั้ง เทคโนโลยีฟิเวอร์ ที่เชื้อระบบมากเกินไป และ เทคโนโลยีโฟเบีย ที่ปฏิเสธระบบอย่างสิ้นเชิง การเสนอโมเดลการทำงานร่วมกันจึงไม่ควรเป็นเพียงการแจกแจงขั้นตอน แต่ต้องสะท้อนทัศนะทางจริยธรรมต่อความเป็นมนุษย์ในที่ทำงานด้วย

ในขั้นตอนการสรรหา โมเดลที่สมเหตุสมผลคือให้ AI ทำหน้าที่ ตะแกรงหยาบ ในการคัดกรองเรซูเม่จำนวนมาก และจัดกลุ่มผู้สมัครตามเกณฑ์พื้นฐาน เช่น ประสบการณ์ ทักษะ และข้อจำกัดด้านเวลา จากนั้น HR และผู้จัดการสายงานเป็นผู้ดำเนินการสัมภาษณ์เชิงลึก ประเมินแรงจูงใจ ชีวิตการทำงานที่ผ่านมา และความสอดคล้องกับวัฒนธรรมองค์กร พร้อมทั้งมีสิทธิ เรียกกลับ เรซูเม่ที่ระบบคัดออกหากเห็นว่ามีศักยภาพที่ระบบมองไม่เห็น แนวทางนี้ตั้งอยู่บนสมมติฐานว่า คุณค่าของคนทำงานไม่ได้จำกัดอยู่ที่สิ่งที่ระบุในแบบฟอร์ม และ HR มีหน้าที่ปกป้องพื้นที่สำหรับ ผู้สมัครที่ไม่เข้าแพตเทิร์นให้ยังมีโอกาสเข้าสู่กระบวนการ

ในขั้นตอนการประเมินผลงานและพัฒนา AI สามารถเป็น เลนส์ขยาย ให้เห็นรูปแบบการทำงานที่ต่อเนื่องและยาวนาน ซึ่งมนุษย์มักมองข้าม เช่น การลดลงของการมีส่วนร่วมในโปรเจกต์ หรือความไม่สมดุลระหว่างปริมาณงานกับผลลัพธ์ แต่การตัดสินใจปรับเป้าหมาย ย้ายตำแหน่ง หรือจัดแผนพัฒนารายบุคคล ควรตั้งอยู่บนการสนทนาระหว่างหัวหน้าและพนักงาน โดยใช้ข้อมูลจาก AI เป็น จุดเริ่มบทสนทนา ไม่ใช่ คำพิพากษา ที่ได้แย้งไม่ได้ การออกแบบเช่นนี้เปลี่ยน AI จากผู้ชี้ชะตาเป็นผู้แจ้งเตือน ที่ต้องผ่านกระบวนการตีความร่วมกันก่อนนำไปสู่การตัดสินใจที่มีผลกระทบต่องานคน

สำหรับการบริหารความเสี่ยงลาออกและการรักษาพนักงาน การใช้ AI ทำนายความเสี่ยงอาจมีประโยชน์อย่างมากในเชิงการป้องกัน แต่ก็เสี่ยงสร้าง トラ ให้กับพนักงานบางคนว่าเป็น คนเสี่ยงลาออก จนถูกกันออกจากโอกาสที่สำคัญ โมเดลที่รับผิดชอบคือการใช้สัญญาณจาก AI เป็นเหตุให้หัวหน้าหรือ HR เปิดบทสนทนาเพื่อทำความเข้าใจบริบท เช่น ภาระครอบครัว ปัญหาสุขภาพ หรือความอึดอัดกับบทบาทงาน มากกว่าที่จะใช้เป็นเหตุผลหลักในการลดการลงทุนในตัวบุคคล การหันกลับมาใช้ AI เป็น เครื่องมือรับฟัง ไม่ใช่ เครื่องมือคัดทิ้ง เป็นจุดชี้วัดสำคัญว่าระบบนี้ถูกใช้เพื่อผลประโยชน์ของใคร

ส่วนในขั้นตอนที่มีความเสี่ยงสูงสุดอย่างการลงโทษทางวินัยและการเลิกจ้าง โมเดลการทำงานร่วมกันที่มีจริยธรรมต้องถือหลักว่า AI ไม่ควรมีบทบาทเกินกว่าการรวบรวมข้อเท็จจริงที่ตรวจสอบได้ เช่น ประวัติการมาสาย การละเมิดนโยบาย

ที่ถูกบันทึกไว้อย่างเป็นทางการ หรือสถิติผลงานในช่วงเวลาหนึ่ง แต่การตีความข้อเท็จจริงเหล่านั้นในบริบทของชีวิตมนุษย์แต่ละคน รวมถึงการชั่งน้ำหนักระหว่างผลประโยชน์ขององค์กรกับผลกระทบต่อพนักงาน ต้องเป็นงานของมนุษย์ภายใต้กรอบกฎหมายแรงงานและหลักมนุษยธรรม การยอมให้ AI มีบทบาทตัดสินใจขั้นสุดท้ายในเรื่องเหล่านี้เท่ากับยอมให้อัลกอริทึมเข้าไปแตะเส้นแดงของศักดิ์ศรีความเป็นมนุษย์โดยตรง

ตารางที่ 2 บทบาทของปัญญาประดิษฐ์และมนุษย์ในวงจรชีวิตพนักงาน

ขั้นตอน	บทบาทปัญญาประดิษฐ์	บทบาทมนุษย์	ระดับการกำกับ
การสรรหา (Recruitment)	คัดกรองเรซูเม่เบื้องต้น จัดกลุ่มผู้สมัครตามเกณฑ์พื้นฐาน	สัมภาษณ์เชิงลึก ประเมินแรงจูงใจ มีสิทธิเรียกคืนเรซูเม่ที่ถูกคัดออก	Human in the Loop — ต้องอนุมัติทุกขั้นตอน
การประเมินผลงาน (Performance Evaluation)	วิเคราะห์รูปแบบการทำงาน ตรวจจับสัญญาณความไม่สมดุล	ตีความข้อมูลร่วมกับพนักงาน ใช้ข้อมูลเป็นจุดเริ่มบทสนทนา	Human in the Loop — การตัดสินใจปรับเป้าหมาย/ย้ายตำแหน่งต้องผ่านมนุษย์
การพัฒนาและฝึกอบรม (Development)	แนะนำหลักสูตรฝึกอบรม ประเมินช่องว่างทักษะ	ออกแบบแผนพัฒนารายบุคคล พิจารณาบริบทชีวิต	Human on the Loop — มนุษย์กำกับและปรับแก้เมื่อจำเป็น
การบริหารความเสี่ยงลาออก (Retention Risk)	ทำนายความเสี่ยงลาออก ตรวจจับสัญญาณเตือนล่วงหน้า	เปิดบทสนทนาเพื่อทำความเข้าใจบริบท ห้ามใช้เป็นเหตุผลโอกาส	Human in the Loop — ห้ามดำเนินการโดยอัตโนมัติ
การลงโทษทางวินัยและเลิกจ้าง (Discipline & Termination)	รวบรวมข้อเท็จจริงที่ตรวจสอบได้เท่านั้น	ตีความข้อเท็จจริงในบริบทชีวิตมนุษย์ ชั่งน้ำหนักผลกระทบ ตัดสินใจภายใต้กฎหมายแรงงาน	Human in the Loop เท่านั้น — AI ห้ามมีบทบาทตัดสินใจขั้นสุดท้าย

ตารางที่ 2 กำหนดบทบาทที่ชัดเจนของปัญญาประดิษฐ์และมนุษย์ในแต่ละขั้นตอนของวงจรชีวิตพนักงาน โดยยึดหลักการแบ่งระดับการกำกับดูแล 3 ระดับ ได้แก่ Human in the Loop (มนุษย์ต้องอนุมัติทุกขั้นตอน) Human on the Loop (มนุษย์กำกับและแทรกแซงได้เมื่อจำเป็น) และการห้าม AI ตัดสินใจขั้นสุดท้ายในประเด็นที่มีผลกระทบรุนแรงต่อชีวิตการทำงาน เพื่อสร้างสมดุลระหว่างประสิทธิภาพทางเทคโนโลยีและการคุ้มครองสิทธิแรงงาน

โมเดล HERO: Human-Ethical Review & Oversight

โมเดล HERO เป็นกรอบแนวคิดที่พัฒนาขึ้นจากการสังเคราะห์หลักจริยธรรมปัญญาประดิษฐ์สากลและหลักการจากแนวปฏิบัติของไทย โดยคำนึงถึงลักษณะเฉพาะของวัฒนธรรมองค์กรแบบลำดับชั้น (Hierarchical Culture) และปรากฏการณ์อคติจากระบบอัตโนมัติ (Automation Bias) ในสังคมไทย ประกอบด้วยองค์ประกอบหลัก 4 ประการ

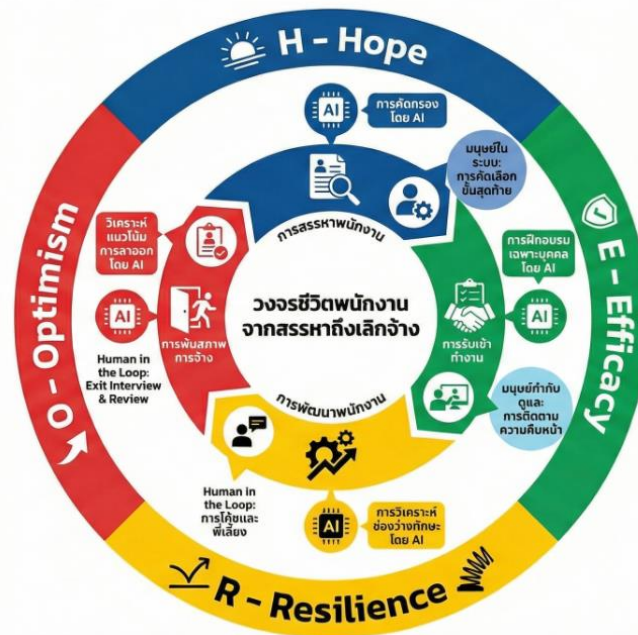
H — Human Authority (อำนาจของมนุษย์) กำหนดให้มนุษย์มีอำนาจจริงในการทักท้วง หยุด และเปลี่ยนแปลงการตัดสินใจของระบบปัญญาประดิษฐ์ ไม่ใช่เพียงกดยืนยัน โดยกำหนดให้บุคลากรทรัพยากรมนุษย์ต้องบันทึกเหตุผลทั้งกรณีที่เห็นด้วยและไม่เห็นด้วยกับคำแนะนำของระบบ

E — Explainability (การอธิบายได้) ระบบปัญญาประดิษฐ์ต้องอธิบายปัจจัยสำคัญที่ส่งผลต่อการตัดสินใจในภาษาที่เข้าใจได้ พนักงานมีสิทธิขอคำชี้แจงและอุทธรณ์ผลการตัดสินใจทุกกรณี โดยนำเทคนิคปัญญาประดิษฐ์ที่อธิบายได้ (Explainable AI: XAI) มาประยุกต์ใช้

R — Review Mechanism (กลไกการทบทวน) จัดให้มีกระบวนการพิจารณาซ้ำโดยคณะกรรมการที่หลากหลาย เปิดพื้นที่ให้สหภาพแรงงานหรือผู้แทนพนักงานเข้ามามีส่วนร่วมในการกำกับดูแลระบบ และจัดตั้งคณะกรรมการจริยธรรม ปัญญาประดิษฐ์ที่มีอำนาจหยุดการใช้ระบบที่มีปัญหา

O — Organizational Culture Safeguards (กลไกคุ้มครองเชิงวัฒนธรรมองค์กร) สร้างบรรยากาศที่เปิดให้บุคลากรกล้าตั้งคำถามต่อผลลัพธ์ของระบบ จัดฝึกอบรมเรื่องอคติจากระบบอัตโนมัติ (Automation Bias Awareness Training) และออกแบบช่องทางร้องเรียนที่ปลอดภัย เพื่อลดอิทธิพลของวัฒนธรรมลำดับชั้นที่อาจขัดขวางการทักท้วง

โมเดล HERO และ AI ในวงจรชีวิตพนักงาน



รูปภาพที่ 1 โมเดลการทำงานร่วมกันระหว่าง AI กับมนุษย์ตลอดวงจรชีวิตพนักงาน
ที่มา: จักรพันธ์ กิตตินรรัตน์ (2569)

โมเดลการทำงานร่วมกันระหว่าง AI กับมนุษย์ตลอดวงจรชีวิตพนักงานที่น่าเชื่อถือได้ ต้องสะท้อนหลักการอย่างน้อยสามประการ คือ (1) การยอมรับข้อจำกัดของ AI และใช้มันเป็นเครื่องมือขยายสายตา ไม่ใช่แทนสายตา (2) การรักษาพื้นที่ของการสนทนาและการอธิบายเหตุผลร่วมกันระหว่างคนทำงาน และ (3) การให้มนุษย์เป็นผู้แบกรับความรับผิดชอบสุดท้ายบนฐานของข้อมูลที่โปร่งใสและตรวจสอบได้

สรุป

บทความนี้มุ่งทำความเข้าใจจริยธรรมของการใช้ปัญญาประดิษฐ์ในการบริหารทรัพยากรมนุษย์ควบคู่กับบทบาทของมนุษย์ในการกำกับดูแลการตัดสินใจในบริบทประเทศไทย ผลการวิเคราะห์สรุปได้ ดังนี้

1. ศักยภาพและข้อจำกัดของปัญญาประดิษฐ์ในการบริหารทรัพยากรมนุษย์ ปัญญาประดิษฐ์แบบจำกัด (Narrow AI) ที่ใช้ในงานบริหารทรัพยากรมนุษย์ในปัจจุบัน มีศักยภาพในการลดภาระงานธุรการ เพิ่มความรวดเร็วในการคัดกรองผู้สมัคร วิเคราะห์ประสิทธิภาพการทำงาน ทำนายความเสี่ยงลาออก และแนะนำหลักสูตรพัฒนาทักษะ ช่วยให้ฝ่ายทรัพยากรมนุษย์สามารถมุ่งเน้นบทบาทเชิงกลยุทธ์ได้มากขึ้น อย่างไรก็ตาม ปัญญาประดิษฐ์มีข้อจำกัดสำคัญ ได้แก่ (1) ไม่ได้ปราศจากอคติ เพราะสะท้อนและขยายความลำเอียงในข้อมูลฝึกสอน (2) เสี่ยงลดทอนมนุษย์ให้เหลือเพียงข้อมูลเชิงปริมาณ และ (3) อาจกลายเป็นกลไกคัดกรองที่กั้นผู้ที่มีเส้นทางชีวิตไม่ตรงตามมาตรฐานออกจากโอกาส ในบริบทไทยซึ่งมีความเหลื่อมล้ำด้านทักษะดิจิทัลและโอกาสทางการศึกษาระหว่างพื้นที่ ข้อจำกัดเหล่านี้มีความรุนแรงเป็นพิเศษ

2. ประเด็นจริยธรรมที่เป็นหัวใจของการนำปัญญาประดิษฐ์มาใช้ในการบริหารทรัพยากรมนุษย์ ประเด็นจริยธรรมหลัก 4 ประการที่พบจากการวิเคราะห์ ได้แก่ (1) อคติเชิงโครงสร้าง (Structural Bias) ที่ฝังอยู่ในข้อมูลและระบบ ทำให้เกิดการผลิตซ้ำความเหลื่อมล้ำในมิติเพศ อายุ สถาบันการศึกษา และภูมิภาค (2) ความทึบแสงของโมเดล (Model Opacity) ที่ทำให้ทั้งบุคลากรทรัพยากรมนุษย์และพนักงานไม่สามารถเข้าใจเหตุผลเบื้องหลังการตัดสินใจ ซึ่งนำไปสู่ความจำเป็นในการรับรองสิทธิในการขอคำอธิบาย (3) ความเป็นส่วนตัวและการเฝ้าระวัง (Privacy and Surveillance) ที่เส้นแบ่งระหว่างการดูแลกับการละเมิดมีความพร่าเลือนมากขึ้น และ (4) ช่องว่างด้านความรับผิดชอบ (Accountability Gap) ที่การกระจายความรับผิดชอบระหว่างผู้พัฒนา ผู้จัดหา และผู้ที่มีความซับซ้อนจนเสี่ยงทำให้ไม่มีผู้รับผิดชอบที่แท้จริง

3. แนวทางการใช้ปัญญาประดิษฐ์อย่างมีความรับผิดชอบบนฐานการกำกับดูแลโดยมนุษย์ในบริบทไทย โดยบทความเสนอโมเดล HERO (Human-Ethical Review & Oversight) เป็นกรอบแนวคิดสำหรับบริบทไทย ประกอบด้วย 4 องค์ประกอบ คือ อำนาจของมนุษย์ (Human Authority) การอธิบายได้ (Explainability) กลไกการทบทวน (Review Mechanism) และกลไกคุ้มครองเชิงวัฒนธรรมองค์กร (Organizational Culture Safeguards) โมเดลนี้เน้นย้ำว่าการกำกับดูแลโดยมนุษย์ต้องมีความหมายอย่างแท้จริง กล่าวคือ มนุษย์ต้องมีอำนาจจริงในการทักท้วงและเปลี่ยนแปลงการตัดสินใจของระบบ ไม่ใช่เพียงกดยืนยัน โดยคำนึงถึงวัฒนธรรมลำดับชั้นและปรากฏการณ์อคติจากระบบอัตโนมัติในสังคมไทยที่อาจบั่นทอนประสิทธิภาพของการกำกับดูแลในทางปฏิบัติ

ข้อเสนอแนะเชิงนโยบาย

1. หน่วยงานกำกับดูแลของไทยควรจัดประเภทระบบปัญญาประดิษฐ์ที่ใช้ในกระบวนการจ้างงาน ประเมินผลงาน และเลิกจ้างให้เป็นระบบความเสี่ยงสูง (High-risk AI System) ในร่างกฎหมายปัญญาประดิษฐ์ โดยกำหนดข้อบังคับด้านความปลอดภัย การประเมินอคติ และการกำกับดูแลโดยมนุษย์ที่มีผลบังคับทางกฎหมาย
2. ควรบัญญัติสิทธิของลูกจ้างในการได้รับแจ้ง ขอคำอธิบาย และอุทธรณ์ผลการตัดสินใจที่เกิดจากระบบปัญญาประดิษฐ์ไว้อย่างชัดเจนในกฎหมาย โดยเชื่อมต่อกับสิทธิตามพระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล
3. ควรพัฒนากลไกประเมินผลกระทบด้านจริยธรรมปัญญาประดิษฐ์ (AI Ethics Impact Assessment) ที่เหมาะสมกับบริบทไทย ก่อนนำระบบปัญญาประดิษฐ์มาใช้ในการบริหารทรัพยากรมนุษย์

ข้อแนะนำสำหรับฝ่ายทรัพยากรมนุษย์

1. องค์กรควรจัดตั้งคณะกรรมการจริยธรรมปัญญาประดิษฐ์ที่มีตัวแทนจากหลายฝ่าย มีอำนาจหยุดการใช้ระบบที่มีปัญหา และกำหนดให้การตัดสินใจที่กระทบชีวิตการทำงานอย่างรุนแรงต้องผ่านกระบวนการกำกับดูแลโดยมนุษย์อย่างแท้จริง (Human in the Loop)
2. บุคลากรทรัพยากรมนุษย์ต้องได้รับการพัฒนาสมรรถนะใหม่ ครอบคลุมความเข้าใจเทคโนโลยี การอ่านและตีความข้อมูล การตระหนักถึงอคติจากระบบอัตโนมัติ (Automation Bias) และความไวต่อมิติทางจริยธรรม
3. ควรจัดทำนโยบายการใช้ปัญญาประดิษฐ์ในงานบริหารทรัพยากรมนุษย์ที่ระบุขอบเขตการใช้งาน บทบาทของมนุษย์ ช่องทางร้องเรียนที่ปลอดภัย และการประกันสิทธิของพนักงานในการรับคำอธิบายและอุทธรณ์ผลการตัดสินใจอย่างชัดเจน

เอกสารอ้างอิง

- เขมภักทร ทฤษฎีคุณ และ วิษญาดา อามณกิจวิวัฒน์. (2568, 27 มีนาคม). การเดินทางกฎหมายปัญญาประดิษฐ์ของไทย: การพัฒนาที่อยู่บนทางสองแพร่ง. <https://tdri.or.th/2025/03/navigating-thailand-ai-law-article/>
- ฉัตรธีรพล ปรัชญาลภัสศิริ. (2562). การศึกษานวัตกรรมปัญญาประดิษฐ์เพื่อการบริหารทุนมนุษย์ในประเทศไทย: กรณีศึกษาองค์กรกลุ่มธุรกิจยานยนต์ [สารนิพนธ์ปริญญาโทบริหารธุรกิจ, มหาวิทยาลัยมหิดล]. หอสมุดมหาวิทยาลัยมหิดล.

- พัฒนาศินีนาถ วรรณรัตน์. (2567). การศึกษานวัตกรรมปัญญาประดิษฐ์เพื่อการบริหารทรัพยากรมนุษย์: กรณีศึกษาของบริษัทในกลุ่มธุรกิจอุตสาหกรรมไมซ์ (MICE) [สารนิพนธ์ปริญญามหาบัณฑิต, มหาวิทยาลัยมหิดล]. หอสมุดมหาวิทยาลัยมหิดล.
- สำนักงานคณะกรรมการดิจิทัลเพื่อเศรษฐกิจและสังคมแห่งชาติ. (2564). *Thailand artificial intelligence ethics guideline*. กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม.
- สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์. (2567). *แนวทางและธรรมาภิบาลการใช้ปัญญาประดิษฐ์* [รายงานนโยบาย]. <https://www.etda.or.th>
- สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ. (2565). *แนวปฏิบัติจริยธรรมด้านปัญญาประดิษฐ์*. <https://www.nstda.or.th/ori/wp-content/uploads/2022/08/20220831-aw-book-ai-ethics-guideline.pdf>
- Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *California Law Review*, 104(3), 671–732.
- Kayusi, F., Chavula, P., Keri Omwenga, M., Juma, L., Agura Kayus, B., Gonzalez Vallejo, R., & Mishra, R. (2025). AI-Driven HR Analytics: Transforming Talent Management and Employee Engagement. *Revista Multidisciplinaria Voces De América Y El Caribe*, 2(1), 558-582. <https://doi.org/10.69821/REMUUVAC.v2i1.214>
- Edwards, L., & Veale, M. (2017). Slave to the algorithm? Why a "right to an explanation" is probably not the remedy you are looking for. *Duke Law & Technology Review*, 16(1), 18–84.
- European Data Protection Supervisor. (2025, September 23). *TechDispatch #2/2025 — Human oversight of automated decision-making*. https://www.edps.europa.eu/data-protection/our-work/publications/techdispatch/2025-09-23-techdispatch-22025-human-oversight-automated-making_en
- Global Privacy Assembly. (2024). *Resolution on human oversight of automated decisions*. <https://globalprivacyassembly.org>
- Goodman, B., & Flaxman, S. (2017). European Union regulations on algorithmic decision-making and a "right to explanation". *AI Magazine*, 38(3), 50–57.
- Hongladarom, S. (2021). The Thailand national AI ethics guideline: An analysis. *Journal of Information, Communication and Ethics in Society*, 19(4), 480–493.
- Kaminski, M. E., & Malgieri, G. (2021). Algorithmic impact assessments under the GDPR: Producing multi-layered explanations. *International Data Privacy Law*, 11(2), 125–145.
- Raghavan, M., Barocas, S., Kleinberg, J., & Levy, K. (2020). Mitigating bias in algorithmic hiring: Evaluating claims and practices. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 469–481. <https://doi.org/10.1145/3351095.3372828>
- World Bank. (2024). *South Asia development update, April 2024: Jobs for resilience* (Chapter 2). <https://openknowledge.worldbank.org/handle/10986/41330>