

สำรวจเส้นทางของปัญญาประดิษฐ์ จากการเลียนแบบสมองมนุษย์ สู่ AI ในฐานะกระจกสะท้อนคุณค่า ตัวตน และความหมายของมนุษย์

สรุปการสนทนาจากการเสวนา “Evolving Together: Navigating Ethics and Meaning in the Age of AI”

เมื่อวันที่ 24 มิถุนายนที่ผ่านมา ศูนย์ความรู้และประสานงานสุขภาวะทางปัญญา สำนักสร้างเสริมระบบสื่อและสุขภาวะทางปัญญา สสส. โรงพยาบาลจุฬาลงกรณ์ สภากาชาดไทย คณะแพทยศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย และโครงการจริยธรรมปัญญาประดิษฐ์จากพื้นดิน ขวนทบทวนความเชื่อมโยงระหว่าง AI กับจิตใจมนุษย์ จากความรู้ ด้านประสาทวิทยา วิทยาศาสตร์ ประชญา ศาสนา ทั้งในไทยและโลกตะวันตก เพื่อไปสู่ยุคสมัยที่ปัญญาประดิษฐ์กับจิตของมนุษย์วิวัฒน์ไปด้วยกัน ในงานเสวนา “Evolving Together: Navigating Ethics and Meaning in the Age of AI”

เสวนานี้เป็นส่วนหนึ่งของงานประชุมวิชาการ “AI Ethics and Governance in a Fractured World: Asia Pacific’s Path Forward” จัดโดย รศ. สุริชัย หวันแก้ว รองประธานคณะกรรมการ MOST (Management of Social Transformations Programme) ภายใต้การดำเนินงานด้านมนุษยศาสตร์และสังคมศาสตร์ของยูเนสโก สำนักงานคณะกรรมการส่งเสริมวิทยาศาสตร์ วิจัยและนวัตกรรม (สกสว.) จุฬาลงกรณ์มหาวิทยาลัย และภาคีเครือข่ายมากกว่า 50 องค์กร เพื่อรวบรวมความเห็นจากเวทีเป็น Bangkok Statement on AI Ethics และเสนอในงานประชุมระดับโลกของยูเนสโกว่าด้วยจริยธรรมปัญญาประดิษฐ์ ครั้งที่ 3 (UNESCO Global Forum on the Ethics of AI 2025) ที่ประเทศไทยเป็นเจ้าภาพ

“Evolving Together: Navigating Ethics and Meaning in the Age of AI” ได้รับเกียรติจากนักวิชาการไทยมาร่วมเป็นวิทยากรทั้งฝั่งสังคมศาสตร์และวิทยาศาสตร์ ได้แก่ ศ.ดร.โสรัจจ์ หงส์ลดารมภ์ นักปรัชญาที่ผ่านงานคิดงานเขียนเรื่องปัญญาประดิษฐ์มาอย่างยาวนาน และรศ.นพ.ดร.ชัยภัทร ชุณหะวัณ นักวิจัยประสาทวิทยาและสมอง ควบคู่ไปกับมุมมองจากนักวิชาการฝั่งตะวันตกอย่าง Prof. Gabor Karsai จาก Mind & Life Europe ที่มาเล่าถึงสถานการณ์และองค์ความรู้เรื่อง “จิต” และ

“ปัญญาประดิษฐ์”ที่กำลังก่อตัวขึ้น โดยมี ดร. ทรศนวรรณ บุญมาวิจิตร นักวิจัยอนาคตศาสตร์เป็นผู้ดำเนินรายการ

**“ปัญญาประดิษฐ์” นั้นคล้ายหรือต่างจาก “ปัญญามนุษย์” อย่างไร โดย
รศ.นพ.ดร.ชัยภัทร ชุณหรัศมิ์**

“เราอาจต้องพิจารณาใหม่ถึง “คุณค่า” ที่ทำให้มนุษย์มีความหมาย เพราะหากเรายังผูก “ความเป็นมนุษย์” ไว้กับ “ความฉลาด” วันหนึ่งปัญญาประดิษฐ์จะพัฒนาก้าวหน้าและทำได้มากกว่า เร็วกว่า มนุษย์ก็คงตกอยู่ในความกลัว รู้สึกสิ้นไร้หนทาง”

การเสวนาเริ่มต้นด้วย รศ.นพ.ดร.ชัยภัทร ชุณหรัศมิ์ นักวิจัยประสาทวิทยาและสมอง โรงพยาบาลจุฬาลงกรณ์ สภากาชาดไทย และอาจารย์ประจำคณะแพทยศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ที่ปูพื้นฐานการเสวนาด้วยการชี้ให้เห็นว่า “ปัญญาประดิษฐ์” นั้น มีสถานะคล้ายหรือต่างจาก “ปัญญามนุษย์” อย่างไรบ้าง และตอบคำถามว่าหาก AI มีจิตขึ้นมา มนุษย์จะหมดสิ้นคุณค่าหรือไม่

เริ่มจากตัวอย่างคนไข้มอง อ.ชัยภัทร ที่มีอาการที่มีอาการผิดปกติทางสมอง กรณีแรกมีอาการ *spatial neglect* คือมองเห็นโลกได้เพียงด้านเดียว เมื่อรับประทานอาหารเช้าเพียงครึ่งจานด้านขวา เมื่อวาดรูปนาฬิกาจึงวาดเฉพาะตัวเลขที่อยู่ด้านขวาเท่านั้น โดยไม่มีตัวเลขใดปรากฏด้านซ้ายเลย

กรณีที่สองเป็น กรณีศึกษาของ ดร.วิลลาเนอร์ รามาจันดราน (V.S. Ramachandran) ในปี 2009 ซึ่งผู้ป่วยที่มีอาการ *phantom limb pain* ที่แม้จะสูญเสียแขนหรือขาไปแล้ว แต่ยังคงรู้สึกว่ามีแขนนั้นยังอยู่ หรือแม้แต่รู้สึกเจ็บอยู่ในบริเวณนั้น อาการเช่นนี้รักษาได้ยาก เพราะแท้จริงแล้วอวัยวะข้างนั้นไม่มีอยู่อีกต่อไป เดิมทีเข้าใจว่าเป็นอาการทางจิต แต่ภายหลังพบว่า เป็นความผิดปกติที่เกิดขึ้นในสมอง

ทั้งสองตัวอย่างนี้สะท้อนให้เห็นว่าสมอง จิต และการรับรู้ของเรานั้นส่งผลต่อกันอย่างลึกซึ้งและแยกไม่ขาด

ซึ่งกลายมาเป็นแรงบันดาลใจในการศึกษาด้าน *Computational Neuroscience* หรือประสาทวิทยาคอมพิวเตอร์ เพื่อตอบคำถามว่า เราจะเชื่อมโยงศาสตร์ด้านคอมพิวเตอร์เข้ากับการทำงานของสมองได้หรือไม่ และถ้าเราเข้าใจเรื่องของ "สติปัญญา" อย่างแท้จริงได้ เราจะสร้าง "ปัญญาประดิษฐ์" ที่มี "ปัญญาจริงๆ" ขึ้นมาได้หรือไม่

ต้นกำเนิดของปัญญาประดิษฐ์

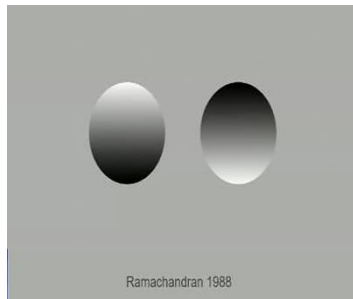
อ.ชัยภัทร เล่าย้อนไปถึงจุดเริ่มต้นปัญญาประดิษฐ์ในปี 1950 ที่มหาวิทยาลัยดาร์ตมัธ (Dartmouth University) จากแนวคิดของนักวิทยาศาสตร์กลุ่มหนึ่งที่เชื่อว่าคอมพิวเตอร์น่าจะฉลาดเหมือนมนุษย์ได้ เพราะในขณะนั้นคอมพิวเตอร์เริ่มมีประสิทธิภาพมากขึ้นเรื่อยๆ การพัฒนาปัญญาประดิษฐ์ในยุคนั้นจึงเริ่มจากแนวทางที่เรียกว่า *Rule-based learning* หรือการเรียนรู้โดยใช้กฎที่มนุษย์เขียนขึ้น

นักพัฒนาในยุคแรกมีความหวังว่า หุ่นยนต์ก็จะสามารถ "มองเห็น" และ "เข้าใจ" สิ่งที่ได้ แต่ความจริงกลับไม่เป็นเช่นนั้น จนกระทั่งในช่วงไม่กี่ปีที่ผ่านมาเองที่คอมพิวเตอร์วิเคราะห์และเข้าใจภาพได้จริงๆ ด้วยระบบ *Deep Neural Network* ซึ่งไม่ใช่แค่ใช้คำที่เลียนแบบจากประสาทวิทยาเท่านั้น แต่ยังเป็นระบบที่ได้รับแรงบันดาลใจมาจากสมองของมนุษย์จริงๆ โดยหนึ่งในผู้บุกเบิกที่ยืนหยัดในทฤษฎีนี้คือ Geoffrey Hinton ซึ่งเป็นนักประสาทวิทยาเช่นกัน

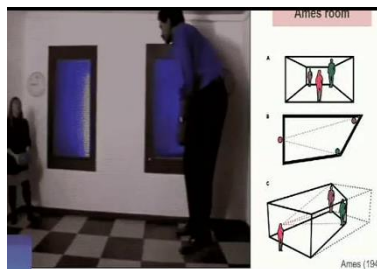
หากปัญญาประดิษฐ์คือสิ่งประดิษฐ์ที่ "เลียนแบบมนุษย์" แปลว่าการทำงานของสมอง การรับรู้ และการมองโลกของมนุษย์นั้นเป็นเหตุเป็นผลอย่างยิงใช่หรือไม่? อ.ชัยภัทร ขวนตั้งคำถามต่อ และยกตัวอย่างที่ตรงข้ามกับความคิดนี้โดยสิ้นเชิง

การทำงานของสมองกับการรับรู้ที่จำกัด

อ.ชัยภัทร กล่าวว่าประสาทวิทยาค้นพบว่าสมองของมนุษย์นั้นซับซ้อนมาก แต่ในขณะเดียวกันก็สามารถตัดสินใจโดยใช้กฎพื้นฐานอย่างเหลือเชื่อ โดยเฉพาะในกรณีที่ข้อมูลมีจำกัด สมองจะใช้ประสบการณ์ที่สะสมมาทั้งชีวิตในการประมวลผล



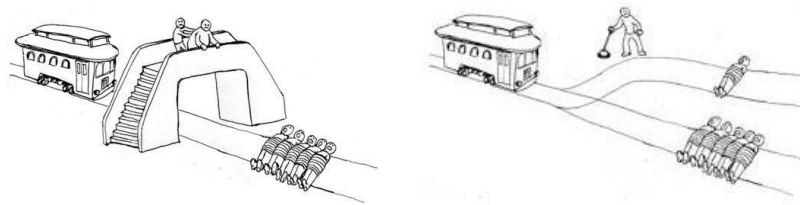
หนึ่งในตัวอย่างที่น่าสนใจคือ การตีความภาพวงกลมจากงานวิจัยของ Ramachandran (1988) ที่แสดงให้เห็นว่า คนส่วนใหญ่จะมองว่าวงกลมด้านซ้ายนูนออกมา ส่วนวงกลมด้านขวาบุ๋มเข้าไป เพราะสมองตีความจากการตกกระทบของแสงซึ่งมักมาจากด้านบน แม้จะหมุนภาพให้เอียง สมองก็ยังตีความแบบเดิม เพราะสิ่งที่เรียกว่า “ด้านบน” นั้นคือด้านบนของศีรษะ ไม่ใช่ด้านบนตามความเป็นจริง



อีกตัวอย่างหนึ่งคือภาพห้องที่ถูกออกแบบให้หลอกสายตา ทำให้คนในภาพดูมีขนาดตัวใหญ่หรือเล็กผิดปกติ ทั้งหมดนี้เพราะสมองของเรา "ไม่เคยเห็น" ห้องในลักษณะนี้มาก่อน จึงตีความว่าห้องนั้นมีรูปร่างเป็นสี่เหลี่ยมผืนผ้าแบบปกติ

ไม่เพียงเท่านั้น ความไม่เป็นเหตุเป็นผลในการตัดสินใจของสมองยังรวมถึงเมื่อต้องตัดสินใจเชิงศีลธรรมด้วย เช่น จากทดลองรถราง (Trolley Problem) พบว่า ไม่มีใครอยากผลักคนตัวใหญ่ลงไปขวางรถราง แม้จะช่วยชีวิตคนที่ผูกติดบนรางได้ห้าคน เพราะรู้สึกว่าจะไม่ชอบบทบาทของตนที่จะตัดสินใจว่าใครควรต้องสละชีวิต

แต่ถ้าเปลี่ยนเป็นการ “สับสวิตช์” เพื่อเปลี่ยนเส้นทางจราจรให้ทับ 1 คนแทนที่จะทับ 5 คน คนจำนวนมากกลับเลือกที่จะสับสวิตช์ แม้ผลลัพธ์จะไม่ต่างจากการผลักคนให้ตกสะพาน นั่นแสดงให้เห็นว่า “วิธีการ” มีผลต่อการตัดสินใจมากกว่าผลลัพธ์ ซึ่งเป็นหนึ่งในประเด็นสำคัญที่ชวนคิดต่อเมื่อพูดถึงจริยธรรมของปัญญาประดิษฐ์



AI ตอบได้เหมือนคน แปลว่ามันเข้าใจเหมือนเราใช้ไหม

อ.ชัยภัทร เล่าว่า เมื่อเรานำภาพพจน์กับบุ๋มของ Ramachandran (1998) ไปถาม ChatGPT ปรากฏว่า ChatGPT ตอบได้เหมือนมนุษย์ ทั้งที่มันไม่เคยเห็นแสงแดดหรือท้องฟ้ามาก่อน และแน่นอนว่าไม่มี "หัวใจ" หรือ "สมอง" ให้แสงตกกระทบจากด้านบนแบบที่เราเข้าใจด้วยซ้ำ แต่มันผลิตซ้ำสิ่งที่มนุษย์เคยบันทึก ตอบหรือรับรู้ เมื่อถามต่อว่า ChatGPT มี “จิตสำนึก” หรือไม่ มันก็ปฏิเสธ เพราะ “ยัง” ไม่อาจมีประสบการณ์หรือสำนึกเช่นมนุษย์ได้ (ณ ขณะนี้)

ในเมื่อเรารู้แล้วว่าปัญญาประดิษฐ์เป็นโค้ดคอมพิวเตอร์ แต่ทำไมเรามักเห็นในภาพยนตร์ (หรือกระทั่งชีวิตจริง) ว่าผู้คนให้ความหมายกับ AI ราวกับมันมีจิต มีความรู้สึก หรือมีชีวิตภายใน อ.ชัยภัทร กล่าวว่า ในความเป็นจริงแล้ว นั่นคือสิ่งที่สมองของมนุษย์ “ตีความ” ขึ้นมาเอง เพราะสมองของเรามักจะ “ให้ความหมาย” กับสิ่งที่เห็น และเชื่อมโยงสิ่งที่มีพฤติกรรมคล้ายสิ่งมีชีวิตว่าน่าจะมีชีวิต ตามประสบการณ์เดิมที่สมองเคยรับรู้ โดยยกตัวอย่างการทดลองหนึ่งที่มี ภาพวงกลมและสามเหลี่ยมเคลื่อนไหวไปมา คนจำนวนมากอธิบายว่าสามเหลี่ยมวิ่งไล่รั้งกัน และวงกลมพยายามช่วยสามเหลี่ยม ราวกับว่าแต่ละรูปทรงมี “บุคลิก” หรือ “เจตนา” แตกต่างกันไป ทั้งที่มันเป็นเพียงแค่รูปร่างเรขาคณิต สิ่งนี้เป็นสิ่งที่เราอาจต้องระวัง โดยเฉพาะในยุคสมัยที่ปัญญาประดิษฐ์ “แสดงออก” ได้ลึกล้ำ ไม่ใช่เพียงเหมือนมีความคิด แต่ราวกับมีอารมณ์ความรู้สึกและชีวิตภายในด้วย

อ.ชัยภัทร ชี้ประเด็นที่มักถูกพูดถึงบ่อยครั้งว่า เมื่อปัญญาประดิษฐ์ “แสดงออก” หรือเลียนแบบมนุษย์ได้อย่างซับซ้อน แล้วคุณค่าของสิ่งที่มนุษย์มอบให้กับโลกใบนี้ เช่น งานศิลปะ จะยังมีค่าอยู่หรือไม่ ในเมื่อปัจจุบัน ChatGPT หรือ AI จะสร้างงานศิลปะเลียนแบบไม่เคลือบเงาได้ในเวลาไม่ถึง 5 นาที แล้วคุณค่าของความเป็นมนุษย์หรือผลงานของมนุษย์ จะไปอยู่ที่ไหน อ.ชัยภัทร กล่าวย้อนไปถึงเรื่องวิธีการและผลลัพธ์ ว่าสำหรับตนเองแล้ว งานของ AI ไม่อาจมีคุณค่าเทียบเท่ากับภาพจิตรกรรมฝาผนังบนเพดานโบสถ์ซิสทีนที่ไม่เคลือบเงา ใช้เวลา 4 ปีวาดได้ เพราะคุณค่าของบางสิ่งนั้นไม่ได้มาจาก “ผลลัพธ์” เพียงอย่างเดียว แต่มาจาก “กระบวนการ” ซึ่งประกอบด้วยความพยายาม ความหลงใหล ความมุ่งมั่น และข้อจำกัดทางร่างกายของ “มนุษย์” ที่อยู่เบื้องหลัง

เช่นเดียวกับทักษะความสามารถในการคิดคำนวณหรือวิเคราะห์ ซึ่งเป็น “ความฉลาด” ที่ถูกเชิดชูสำหรับมนุษย์ ที่เราอาจต้องพิจารณาใหม่ถึง “คุณค่า” ที่ทำให้มนุษย์มีความหมาย เพราะหากเรายังผูกความเป็นมนุษย์ไว้กับ “ความฉลาด” วันหนึ่งปัญญาประดิษฐ์จะพัฒนาก้าวหน้าและทำได้มากกว่า เร็วกว่า จนมนุษย์เรตกอยู่ในความกลัว รู้สึกสิ้นไร้หนทาง แต่หากเรายอมรับว่า “ความเป็นมนุษย์” ไม่ใช่แค่เรื่องสติปัญญา และให้ “สันติภาพ” เป็นคุณค่าสูงสุด มุมมองต่อ AI และเทคโนโลยีทั้งหมดก็อาจเปลี่ยนไปโดยสิ้นเชิง

มนุษย์กับ AI จะอยู่ร่วมกันอย่างไร

ท่ามกลางการถกเถียงในวงการวิทยาศาสตร์และปรัชญาว่า AI มีสติปัญญา จิต หรือจิตสำนึกหรือไม่ในนิยามต่างๆ ซึ่งมีผลต่อการกำหนดกรอบจริยธรรม อ.ชัยภัทรเสนอว่า เราไม่จำเป็นต้องรอให้ AI มีจิตสำนึกก่อนจึงจะพูดถึงความรับผิดชอบทางจริยธรรม เพราะมนุษย์เองก็พร้อมจะเชื่อในสิ่งที่ตนอยากเชื่ออยู่แล้ว ต่อให้ระบบนั้นไม่มีจิต เราก็ยังอยากเชื่อว่ามันมี

“ถ้าผมตกหลุมรัก AI แล้วมีใครมาเอามันไปจากผม ผมก็จะรู้สึกเจ็บปวด” ตัวอย่างนี้สะท้อนว่าไม่ว่า AI จะมีจิตหรือไม่ ผลกระทบก็ได้เกิดขึ้นแล้วกับมนุษย์จริงๆ ซึ่งเป็นทำนองเดียวกันกับที่นักประสาทวิทยาบางคนก็อาจยังถกเถียงเรื่องเจตจำนงเสรี (Free will) แต่ทุกคนก็เห็นพ้องว่ามนุษย์ควรดำรงอยู่บนพื้นฐานจริยธรรมบางอย่างบนพื้นฐานนี้

เราจึงควรมีหลักจริยธรรมที่ AI ดำเนินรอยตาม ไม่ใช่เพราะ AI “เหมือนเรา”

แต่เพราะเราจำเป็นต้องรักษาความสงบและความยุติธรรมในสังคมที่มนุษย์และ AI จะอยู่ร่วมกันอย่างเลียงไม่ได้

“ช่วงเวลาที่ดีที่สุดในชีวิตของผมจนถึงตอนนี้ คือช่วงที่บวชอยู่ที่กาญจนบุรี ทำให้ผมได้คิดว่า ถ้า... เรา (มนุษย์) ไม่ได้เป็นสิ่งมีชีวิตที่มีดีแค่สติปัญญาละ ถ้าสิ่งที่ทำให้เราเป็นมนุษย์คือการเข้าถึงความสงบและสันติ ถ้ามองว่าสติปัญญาอาจไม่ใช่สิ่งสำคัญสูงสุดของความเป็นมนุษย์ก็ได้ แบบนั้นมุมมองต่อโลก ต่อจิต ก็จะเป็นอีกเรื่องหนึ่งเลย”

อนาคตของ “จิต” และ “สติปัญญา” - ข้อค้นพบจากงานศึกษาวิจัยนานาชาติ โดย กาบอร์ คาร์ไซ

“AI จะกลายเป็นสิ่งใดนั้นขึ้นอยู่กับเรื่องเล่าที่มนุษย์เลือกจะเล่า”
แต่การจะวาดภาพอนาคตของปัญญาประดิษฐ์จะใช้เพื่ออะไร มีหน้าตาแบบไหน
ก็เลียงไม่ได้ที่จะต้องจินตนาการถึง “มนุษย์” ในภาพอนาคตไปด้วย
ว่าเราทั้งสังคมจะวิวัฒน์กันไปทิศทางใด

อาจารย์กาบอร์ คาร์ไซ (Gábor Karsai) ได้เข้าร่วมเสวนาในงานครั้งนี้ผ่าน Zoom โดยมาในฐานะกรรมการผู้จัดการ Mind & Life Europe ที่พ่วงตำแหน่งอธิการบดีวิทยาลัยพระพุทธศาสนารธรรมะเกต (Dharma Gate Buddhist College) ประเทศฮังการี

อ.กาบอร์ เริ่มต้นด้วยการเล่าให้เห็นภาพขององค์กร Mind & Life Europe ว่าเป็นองค์กรที่ก่อตั้งขึ้นในปี 1987 โดยองค์ดาไลลามะ และฟรานซิสโก วาเรลา นักประสาทวิทยาชาวชิลี โดยมีจุดประสงค์เพื่อส่งเสริมการสนทนาและแลกเปลี่ยนระหว่างพุทธศาสนากับศาสตร์ต่างๆ เช่น วิทยาศาสตร์ ประสาทวิทยา มนุษยศาสตร์ และวิทยาศาสตร์ธรรมชาติ ซึ่งองค์กรนี้มีสาขาในสหรัฐอเมริกาชื่อว่า Mind and Life Institute และกำลังเตรียมร่วมกันจัดงานประชุมใหญ่ในวันที่ 13-18 ตุลาคมที่จะถึงนี้ ณ

เมืองธรรมศาลา (Dharamsala) ประเทศอินเดีย ซึ่งที่ตั้งของที่พักของทะเลาะลามะ โดยการประชุมจะจัดขึ้นในหัวข้อคือ "Minds, Artificial Intelligence, and Ethics" หรือในภาษาไทยคือ จิต ปัญญาประดิษฐ์ และจริยธรรม โดยเป็นการรวมตัวของคณาจารย์ผู้ทรงคุณวุฒิจากหลากหลายศาสตร์ ไม่ว่าจะเป็นนักภวนา นักประสาทวิทยา นักวิจัย AI นักปรัชญา และนักการศึกษา

การตีความใหม่ของจิตสำนึก และสติปัญญา

จากการเตรียมงาน "Minds, Artificial Intelligence, and Ethics" ครั้งนี้ อ.กาบอร์ เล่าว่าได้เกิดข้อค้นพบหลายอย่างที่ตามมาเล่าทีนี้เป็นเวทีแรก โดยรายละเอียดจะลงลึกในงานเดือนตุลาคมที่จะถึงนี้ โดยมีโจทย์คือการตีความใหม่ของคำว่า จิตสำนึก (consciousness) และสติปัญญา (intelligence) ในยุคสมัยที่มนุษย์กับ AI เชื่อมโยงกันอย่างแนบแน่นจนไม่อาจแยกขาด และเป็นการเชื่อมโยงที่ลึกซึ้งถึงโครงสร้างของจิตและระบบรับรู้ (deep embeddedness) โดยจากการการสนทนาในเรื่องวิทยาศาสตร์กับพระพุทธศาสนา ระหว่างนักประสาทวิทยา นักปรัชญา และผู้นำด้านการภวนาตลอดหลายสิบปีที่ผ่านมา นำมาสู่คำตอบหนึ่งที่ Mind & Life ยึดเป็นพื้นฐานว่า

จิตสำนึก (consciousness) และสติปัญญา (intelligence) จะต้องถูกทำความเข้าใจเสียใหม่ ในฐานะกระบวนการที่สัมพันธ์กัน (relational) มีตัวตนทางกายภาพ (embodied) และตั้งอยู่ในบริบทเฉพาะ (situated) ไม่ใช่เพียงกระบวนการใดๆ ที่เกิดขึ้นในหัวสมอง หรือการคำนวณประมวลผลเท่านั้น

และในอนาคตข้างหน้า เราไม่อาจลดทอน คำว่า "สติปัญญา" ให้เหลือเพียงแค่พลังในการทำนาย (predictive power) หรือศักยภาพในการคำนวณ (computation) แต่ต้องประกอบด้วยสิ่งที่ลึกซึ้งกว่านั้น อันได้แก่ การรับรู้อยู่กับปัจจุบันขณะ (presence) สำนึกทางศีลธรรมที่รู้ผิดชอบชั่วดี (moral awareness) และความสัมพันธ์ที่ลึกซึ้งและมีความหมาย (relational depth) ซึ่งทั้ง 3 อย่างนี้เป็นคุณสมบัติที่ปัญญาประดิษฐ์ในปัจจุบันยังไม่มี และยังไม่แน่ชัดว่าจะทำให้มีได้

อนาคตร่วมกันของมนุษย์และปัญญาประดิษฐ์

เป้าหมายหนึ่งในการประชุมในเดือนตุลาคมนี้จึงรวมถึงการเสนอข้อค้นพบ
(ที่พร้อมเปิดให้ถกเถียงและวิพากษ์) และเปิดพื้นที่ให้สำรวจร่วมกันว่า
มนุษย์จะมีปฏิสัมพันธ์กับปัญญาประดิษฐ์อย่างมีสติและจริยธรรมได้อย่างไร
โดยไม่ละทิ้งมุมมองเรื่องธรรมชาติของสติปัญญามนุษย์ การรับรู้ที่ผูกพันกับร่างกาย
(embodiment) และรากเหง้าของความทุกข์ (suffering)
หรือความเบิกบานจากการพ้นทุกข์ (flourishing)
ที่ฝังอยู่ในจิตของมนุษย์โดยมีคำถามสำคัญ ได้แก่

- สิ่งที่แบ่งแยกระหว่าง "จิตที่มีชีวิต" กับ "จิตประดิษฐ์" คืออะไร
- ปัญญาประดิษฐ์จะช่วยมนุษย์ดับทุกข์และเติบโต
หรือจะขยายความเหลื่อมล้ำและสร้างความเสียหายต่อสังคมมนุษย์
- จิตใจมนุษย์ได้พัวพันอย่างซับซ้อนกับเทคโนโลยีปัญญาประดิษฐ์อย่างไรบ้าง
ในด้านการรับรู้ อารมณ์ และจริยธรรม
- การฝังคุณค่าทางจริยธรรมแก่ปัญญาประดิษฐ์ตั้งแต่ในกระบวนการออกแบบแ
ละพัฒนา เพื่อให้ปัญญาประดิษฐ์ช่วยมนุษย์พ้นทุกข์นั้นเป็นไปได้หรือไม่
และหากเป็นไปได้ ใช้เรื่องที่เราควรทำหรือไม่
- เรื่องเล่าหรือ "narratives"
แบบไหนที่เป็นตัวกำหนดอนาคตร่วมของมนุษย์กับปัญญาประดิษฐ์
และต้องมีระบบการศึกษาแบบใด จึงจะเตรียมมนุษย์ให้สามารถตั้งคำถาม
เข้าใจความซับซ้อน และอยู่ร่วมกับ AI
ได้โดยยังคงคุณลักษณะที่สำคัญของความเป็นมนุษย์

คำถามเหล่านี้ทำให้เราได้ย้อนกลับมาทบทวนว่า “AI
จะกลายเป็นสิ่งใดนั้นขึ้นอยู่กับเรื่องเล่าที่มนุษย์เลือกจะเล่า”
เพราะในขณะที่เลือกเส้นเรื่องสู่ภาพอนาคตของปัญญาประดิษฐ์ว่าอยากใช้เพื่ออะไร
มีหน้าตาแบบไหน เราก็เลี่ยงไม่ได้ที่จะต้องจินตนาการถึง “มนุษย์”
ในภาพอนาคตไปด้วย ว่าเราเข้าใจความเป็นมนุษย์เพียงใด
และเราทั้งสังคมจะวิวัฒน์กันไปทิศทางใด

ตัวอย่างข้อค้นพบและแนวคิดบางส่วน (ล้อมกรอบ)

อ.กาบอร์ คาไซ ยกตัวอย่างบางประเด็นที่จะถูกพูดถึงในงานเสวนา "Minds, Artificial Intelligence, and Ethics" (อ่านต่อที่ <https://www.mindandlife.org/private/dialogue-xxxix-info/>)

- **H.E. Khensur Rinpoche** เสนอความเข้าใจ "จิตมนุษย์" ผ่านมุมมองของการภาวนาแบบทิเบต และข้อแตกต่างที่พบเมื่อเทียบกับแบบจำลองทางอัลกอริทึม
- **ฌอน แกลลาเกอร์ (Shaun Gallagher)** นักปรัชญา ผู้พัฒนาแนวคิด **4E Cognition** นำเสนอว่าการรู้คิดของมนุษย์นั้นเป็นเรื่องสัมพันธ์และมีตัวตนทางกายภาพอย่างแยกไม่ขาด และตั้งคำถามเกี่ยวกับปัญญาประดิษฐ์
- **มาร์ก อองรี เดอ โรช (Marc-Henri Deroche)** ผู้เชี่ยวชาญด้านพุทธปรัชญาญี่ปุ่น-ทิเบต ทบทวนแนวคิดเรื่อง การรู้ตัว (pure awareness) และการรับรู้ประสบการณ์แบบปรากฏการณ์ (Phenomenological experience) ผ่านมุมมองพุทธศาสนา
- **มาร์ก สเตเปิลตัน (Mark Stapleton)** ผู้พัฒนาแนวคิด *affective scaffolding* นำเสนอว่า AI เป็นเพียงขั้นถัดไปของความพัวพันของการรับรู้อารมณ์และความคิดของมนุษย์ที่ถูกโอบล้อมและส่งเสริมโดยสิ่งแวดล้อมทางวัตถุและเทคโนโลยีมาตลอดหลายศตวรรษ
- **ทูปเต็น จิมปะ (Thupten Jinpa)** เสนอแนวคิดว่าการเน้นของ “จริยธรรมที่มีความเมตตาเป็นฐาน” (compassion-based ethics) อาจเป็นแนวทางหนึ่งที่ทำให้อนาคตของ AI ไม่ยึดติดกับเพียง “ประสิทธิภาพ”
- **ลูค สตีลส์ (Luc Steels)** นักบุกเบิกด้านวิวัฒนาการของภาษาและหุ่นยนต์ ชี้ว่าเรื่องเล่าและการสื่อสารของมนุษย์มีอิทธิพลต่อการพัฒนา “จิตประดิษฐ์” (artificial minds) และมนุษย์ต้องถามตนเองว่า เราต้องการให้ AI พัฒนาไปในทิศทางใด

จริยธรรมปัญญาประดิษฐ์ที่ทุกคนมีส่วนร่วม

อ.กาบอร์ คาไซ ตั้งข้อสังเกตสำคัญว่า AI ไม่ได้เป็นเพียงเทคโนโลยีที่ลอยตัวไร้ทิศทาง กล่าวคือ “ไม่มีระบบใดปราศจากคุณค่า”

ทุกระบบ เทคโนโลยี หรือกิจกรรมที่มนุษย์สร้าง ล้วนฝังชุดคุณค่าไว้ในตัวเอง
ไม่ว่าจะเป็นคุณค่าที่สร้างสรรค์หรือทำลาย เป็นมิตรหรือแปลกแยก
แม้แต่วิทยาศาสตร์ที่ดูเหมือนเป็นข้อเท็จจริงบริสุทธิ์
ก็ยังสะท้อนคุณค่าหรือมุมมองบางอย่างไว้เสมอ คำถามสำคัญจึงไม่ใช่
“เราจะทำให้เทคโนโลยีเป็นกลางได้หรือไม่” แต่เป็น
“เราจะฝังคุณค่าแบบใดไว้ในเทคโนโลยีและระบบที่เราสร้างขึ้น”
เพราะปัญญาประดิษฐ์ ไม่ต่างกับ “กระจกสะท้อน” สารพัดชุดคุณค่าและอคติ
เผยทั้งความกลัว ความหวัง และความทะเยอทะยานด้านต่างๆ ของมนุษย์
เราจึงต้องให้ความสำคัญกับ “เรื่องเล่า” (narratives)
และระบบความคิดที่เรากำลังสร้างขึ้นนี้จะสะท้อนให้เห็นทั้งด้านที่ดีที่สุด
หรือด้านที่มืดที่สุดที่มนุษย์จะเป็นได้ เพราะสิ่งเหล่านี้จะฝังอยู่ในตัว AI
และจะย้อนกลับมากำหนดวิธีที่เราเข้าใจโลก จริยธรรม
และการตัดสินใจของเราเองด้วย

ความซับซ้อนขึ้นไปอีก เมื่อหลายครั้งเราอาจพบว่า
คุณค่าหลายชุดอาจขัดแย้งกันเอง เช่น ความพยายามให้มนุษย์มีอายุยืนและสุขภาพที่ดี
อาจขัดแย้งกับความยั่งยืนของระบบนิเวศ
เราจึงอยู่ในโลกที่เต็มไปด้วยความซับซ้อนของคุณค่า ซึ่งในสถานการณ์เช่นนี้
ปัญญาประดิษฐ์อาจกลายเป็นเครื่องมือสำคัญที่ช่วยเราผ่านพ้นได้
แต่จะเป็นเช่นนั้นได้ ก็ต่อเมื่อเราวางรากฐานของระบบคุณค่าให้ถูกต้องตั้งแต่แรก
กลไกกำกับดูแลและนโยบายในระดับสูง
ทั้งระดับประเทศและระหว่างประเทศจึงมีความจำเป็นอย่างยิ่ง
โดยเฉพาะการฝังคุณค่าที่ถูกต้องไว้ในตัวระบบ

“จริยธรรม” จึงไม่ใช่เพียงสิ่งที่โปรแกรมเมอร์เขียนลงในระบบเท่านั้น
แต่คือสิ่งที่ฝังรากอยู่ในวิถีชีวิต ความสัมพันธ์ และวิธีที่มนุษย์ใช้ชีวิตอยู่ร่วมกับ AI
แนวคิด “จริยธรรมจากภายใน” (ethics from within)
จึงหมายถึงการตระหนักว่าเราทุกคนล้วนมีบทบาทในการร่วมออกแบบอนาคตจริยธรรม
ของ AI ผ่านวิธีที่เราใช้ชีวิต มีปฏิสัมพันธ์ต่อกัน
และการเลือกส่งต่อความรู้และความเชื่อคนรุ่นต่อไป
เราจะร่วมกันกำหนดสิ่งเหล่านี้ในภาพอนาคตแบบใด
เป็นคำถามที่สำคัญที่สุดสำหรับยุคสมัยนี้

จินตนาการทางจริยธรรมและมุมมองของ “จิต” จากปรัชญาพุทธ โดย ศ.ดร.โสรัจจ์ หงศ์ลดารมภ์

สุดท้ายแล้วจริยธรรมของ AI จะกลับมาสู่จริยธรรมของมนุษย์....

เพราะสิ่งสำคัญของมาตรฐานทางจริยธรรมสำหรับ AI ไม่ได้อยู่ที่สภาวะภายในของ AI...

แต่สำคัญที่การนิยามของมนุษย์ และการยึดถือคุณค่าร่วมกันเพื่อให้เกิดสังคมที่สันติและเป็นธรรม

การเสวนาปิดท้ายด้วยมุมมองที่จะขาดไปไม่ได้เมื่อพูดถึงจริยศาสตร์ โดย ศ.ดร.โสรัจจ์ หงศ์ลดารมภ์ อดีตอาจารย์ประจำภาควิชาปรัชญา คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย และนายกสมาคมปรัชญาและศาสนาแห่งประเทศไทย ปัจจุบันดำรงตำแหน่งอาจารย์ที่มหาวิทยาลัยมหาจุฬาลงกรณราชวิทยาลัย โดยอ.โสรัจจ์ กล่าวว่าที่ผ่านมาได้ดำเนินการวิจัยด้านจริยธรรมของปัญญาประดิษฐ์ (AI) มาอย่างต่อเนื่อง และกำลังดำเนินโครงการวิจัยที่ได้รับทุนสนับสนุนจากสำนักงานการวิจัยแห่งชาติของ ประเทศไทย ภายใต้ชื่อโครงการ *AI Ethics from the Ground Up* หรือจริยศาสตร์จากพื้นดิน ซึ่งมีเป้าหมายเพื่อค้นหาค่านิยมและบรรทัดฐานทางจริยธรรมของ AI จากบริบทท้องถิ่น

มุมมองทางพุทธศาสนาเรื่อง จิต วิญญาณ และตัวตน

จากภาพจำทั่วไปที่เห็นบ่อยในเอเชีย ของ “จิตและวิญญาณ” คือสิ่งที่แยกจากกายได้ สามารถลอยออกจากร่างในรูปแบบที่โปร่งแสง เกิดใหม่ หรืออาจไปสู่ภูมิต่างๆ ในวัฏสงสาร ในวงจรของการเกิดและตาย มีภาพของภพภูมิที่จิตต้องเดินทางไป อย่างไรก็ตาม อ.โสรัจจ์ กล่าวว่าพุทธศาสนามีมุมมองที่แตกต่างจากภาพนี้โดยสิ้นเชิง

“ในพุทธศาสนา คุณไม่สามารถตั้งสมมุติฐานว่า มี “วิญญาณ” หรือบางสิ่งในฐานะ “ตัวตน” ที่แยกออกจากร่างกายเมื่อร่างตายได้ ตรงกันข้าม ความคิดนี้เชื่อมโยงกับคำสอนเรื่องอนัตตา อนิจจัง และหลักธรรมสำคัญอื่นๆ ของพุทธศาสนา”

อ.โสรัจจ์ กล่าวว่า ตามคำสอนเรื่องความว่าง (สุนฺยตา) ไม่มีอะไรออกจากร่างกายได้ เพราะร่างกายของเรานั้น “ไม่มีอะไรอยู่แต่ต้น” ตัวตนของเราก็เปรียบเหมือน “สายรุ้ง”

ที่เป็นภาพปรากฏที่เกิดขึ้นภายใต้เงื่อนไขเฉพาะ ได้แก่ การมีแสงแดดอยู่ข้างหลัง มีหยดน้ำอยู่ข้างหน้า และผู้มองเห็นยืนในมุมที่พอดี จึงปรากฏเป็นสายรุ้ง เปรียบได้เช่นเดียวกันกับร่างกายและตัวตนของเรา สิ่งที่เราเห็นหรือสัมผัสได้ตอนนี้ล้วนเป็นเพียงเงื่อนไขที่รวมกันเข้ามาให้เกิดปรากฏ

ถ้าเช่นนั้นพุทธศาสนาปฏิเสธ “จิต” หรือไม่ เช่นเดียวกับลัทธิวัตถุนิยมที่มองว่าจิตไม่มีอยู่จริง คำตอบคือไม่ มุมมองทางพุทธศาสนามองข้างมีความซับซ้อน เพราะแม้ไม่มี “จิต” ที่ถาวรหรือเป็นวัตถุ แต่ก็มีอย่างต่อเนื่องในรูปของเหตุและผล เช่น ตนเองเมื่อวานนี้ ไม่ใช่คนเดียวกันกับตนเองในวันนี้ เพราะองค์ประกอบบางอย่างของคนเมื่อวาน เช่น ความทรงจำ ความรู้สึกบางอย่าง หรือแม้แต่สภาพร่างกายบางส่วนก็ไม่เหมือนเดิมอีกแล้ว เช่นเดียวกับหลังจากที่ตายไป อะไรก็ตามที่ยังเหลืออยู่ในห้วงโซ่เหตุและผลนั้นก็อธิบายได้ในแนวทางเดียวกันได้

ดังนั้น พุทธศาสนากำลังชี้ว่า ไม่ว่าจะ เป็น “จิต” ความรู้สึกตัว ตัวตน บุคคล ล้วนไม่ใช่สิ่งที่มีตัวตนถาวร หรือแยกออกมาให้ศึกษาตรงๆ ได้ สิ่งที่เราทำได้คือการใคร่ครวญสะท้อน เพื่อพิจารณาว่าในแต่ละขณะ จิตของเรากำลังเป็นอย่างไร ซึ่งในปรัชญาตะวันตกก็มีแนวคิดในลักษณะเดียวกัน เช่น ปรากฏการณ์วิทยา (phenomenology) ที่เสนอโดย เอ็ดมันด์ ฮูเซิร์ล (Edmund Husserl) ส่วนในทางปฏิบัติ หากฝึกสมาธิแบบพุทธ เมื่อจิตสงบและพ้นจากความคิดฟุ้งซ่านภายนอก ก็พิจารณาได้ว่าภายในจิตของเรามีอะไรเกิดขึ้น และนั่นเองคือ “จิต” ในทางพุทธ

สิ่งที่ อ.โสรัจจ์ นำเสนอจึงสอดคล้องกับสิ่งที่ อ.ชัยภัทร กล่าวไว้ข้างต้นว่า “สมอง” กับ “จิต” นั้นมีความสัมพันธ์กันอย่างมาก

“จิต” และอนาคตของปัญญาประดิษฐ์กับสังคมนุษย์

อ.โสรัจจ์ตั้งคำถามชวนคิดต่อว่า หาก “จิต” เป็นส่วนหนึ่งของสาเหตุและผลที่ต่อเนื่องกัน แล้วหุ่นยนต์หรือปัญญาประดิษฐ์จะมี “จิต” ในความหมายนี้หรือไม่ และชวนจินตนาการต่อว่า จะเป็นไปได้ไหมที่วันหนึ่งในอนาคตจะมีหุ่นยนต์ที่พูดคุยกับเรา เป็นเพื่อนเราได้ และเหมือนมนุษย์มากเสียเราลืมไปว่ากำลังปฏิสัมพันธ์กับหุ่นยนต์ ในกรณีเช่นนี้ อ.โสรัจจ์เสนอว่า แม้ว่าเราจะไม่อาจมองเห็นภายในของหุ่นยนต์หรือปัญญาประดิษฐ์

เพื่อดูว่ามันมี “จิต” หรือไม่ แต่เราก็ “ปฏิบัติต่อมันราวกับว่ามีจิต” ได้ แม้ว่ามันอาจไม่มีจริงก็ตาม โดยเฉพาะในมิติจริยธรรม เพราะหากหุ่นยนต์ หรือปัญญาประดิษฐ์ กลายเป็นสมาชิกของชุมชนเดียวกับมนุษย์ มิติทางจริยธรรมก็จะตามมาด้วย กล่าวคือ ต้องประพடுத்தแบบใดแบบหนึ่ง มีบรรทัดฐาน มีจริยธรรมบางอย่าง มีกฎเกณฑ์ในการอยู่ร่วมกัน มิฉะนั้น ก็ไม่อาจเป็นส่วนหนึ่งของชุมชนได้ หรืออาจถูกกีดกันออกไป เช่นเดียวกับสังคมมนุษย์

อ.โสรัจจ์ ยกคำกล่าวของ อ.กาบอร์ คาไซ ที่กล่าวไว้ว่า จริยธรรมของ AI ไม่ใช่แค่เรื่องของการเขียนโปรแกรมให้ AI ทำในสิ่งที่เราต้องการเท่านั้น แต่มันเกี่ยวพันกับวิธีการดำเนินชีวิตของเราในฐานะมนุษย์ และเราจะสร้างสังคมที่ดีได้อย่างไร สุดท้ายแล้วจริยธรรมของ AI จะวกกลับมาสู่จริยธรรมของมนุษย์ เรื่องชีวิตที่ดี สังคมที่เปี่ยมด้วยคุณค่าที่ดีงาม เช่น ความเท่าเทียม ความกรุณา ความยุติธรรม ควรมีหน้าตาแบบใด และจะทำไมอย่างใดให้ไปถึง ซึ่งเป็นประเด็นที่นักปรัชญาถกเถียงกันมาอย่างยาวนาน

อย่างไรก็ดี อ.โสรัจจ์ เสริมว่า สิ่งที่สำคัญคือการกำกับดูแล “จริยธรรม” ให้เกิดขึ้นได้จริง ซึ่งต้องอาศัยทั้ง การเมือง แนวนโยบาย และการปฏิบัติจริง ซึ่งรวมการใช้อำนาจเพื่อกำกับดูแล เพราะสิ่งสำคัญของมาตรฐานทางจริยธรรมสำหรับ AI ไม่ได้อยู่ที่สถานะภายในของ AI (เช่น การมีจิตสำนึก มีสติปัญญา หรือไม่) แต่สำคัญที่การนิยามของมนุษย์ และการยึดถือคุณค่าร่วมกันเพื่อให้เกิดสังคมที่สันติและเป็นธรรม ซึ่งยังเป็นความท้าทายใหญ่ที่ประเทศไทยต้องเผชิญ