รายงานภาคนิพนธ์: ญาณวิทยาวิเคราะห์เชิงลึก
(Analytical Epistemology)

หัวข้อ: ในยุคปัญญาประดิษฐ์ เราสามารถกล่าวได้หรือไม่ว่า “ฉันรู้” เพียงเพราะ AI ให้คำตอบที่ถูกต้อง?

เสนอโดย: พระมหาเกรียงศักดิ์ สิริสกฺโก
อาจารย์ประจำรายวิชาญาณวิทยาวิเคราะห์

จัดทำโดย: พันธ์ ลิดดาศักดิ์ (Phan Liddasak)

รายงานฉบับนี้เป็นส่วนหนึ่งของการศึกษารายวิชาญาณวิทยาวิเคราะห์
ระดับปริญญาโท ภาคเรียนที่ 1 ปีการศึกษา 2569

บทคัดย่อ

การอุบัติขึ้นของแบบจำลองภาษาขนาดใหญ่ (Large Language Models - LLMs) ได้ท้าทายกระบวนทัศน์ทางญาณวิทยาอย่างถึงราก งานวิเคราะห์ชิ้นนี้มุ่งตอบคำถามทางปรัชญาที่ว่า เมื่อบุคคลได้รับคำตอบที่ถูกต้องจากปัญญาประดิษฐ์ บุคคลนั้นมีความชอบธรรมที่จะอ้างว่า "ฉันรู้" หรือไม่ ข้อเสนอหลักของรายงานฉบับนี้ยืนยันว่า การพึ่งพาคำตอบจาก AI ไม่เพียงพอต่อการบรรลุสภาวะของการ "รู้" ในความหมายที่สมบูรณ์ทางญาณวิทยา การวิเคราะห์นี้บูรณาการกรอบทฤษฎีที่หลากหลายเพื่อชำแหละปัญหา เริ่มตั้งแต่การวิพากษ์ผ่านทฤษฎีความรู้แบบดั้งเดิม (JTB) ที่ชี้ให้เห็นถึงความล้มเหลวในการให้เหตุผลเชิงประจักษ์ (Internalist Justification) เนื่องจากลักษณะความเป็นกล่องดำ (Black Box) ของอัลกอริทึม จากนั้นเชื่อมโยงสู่ปัญหาความโชคดีทางปริชาน (Epistemic Luck) ตามแนวคิดของ Gettier และ Pritchard เพื่อแสดงให้เห็นว่าความถูกต้องของ AI อาจเป็นเพียงความบังเอิญที่ปราศจากการติดตามความจริง (Truth-Tracking)

ยิ่งไปกว่านั้น งานชิ้นนี้ได้ยกระดับการวิเคราะห์สู่ญาณวิทยาเชิงคุณธรรม (Virtue Epistemology) ของ Ernest Sosa และ Linda Zagzebski เพื่อชี้ให้เห็นถึงภาวะ "ความเกียจคร้านทางปัญญา" (Epistemic Sloth) และการสูญเสียความเป็นผู้กระทำทางปัญญา (Epistemic Agency) นอกจากนี้ ยังได้นำเสนอมุมมองใหม่ผ่านญาณวิทยาสังคมของ Miranda Fricker และปรากฏการณ์วิทยาของ Heidegger เพื่ออธิบายสภาวะความแปลกแยกทางญาณวิทยา (Epistemic Alienation) บทสรุปของงานชี้ให้เห็นว่า ความรู้ที่แท้จริงเรียกร้องความรับผิดชอบและการลงมือปฏิบัติเชิงรุก ซึ่ง AI ไม่อาจทดแทนได้ แต่สามารถเป็นเพียงคู่สนทนาเชิงวิพากษ์เพื่อกระตุ้นปัญญาญาณของมนุษย์เท่านั้น

สารบัญ

บทที่ 1 บทนำ: รอยต่อระหว่างปัญญาประดิษฐ์และทฤษฎีความรู้
บทที่ 2 รื้อถอนโครงสร้าง JTB: ปัญหาของการให้เหตุผลในยุคกล่องดำทางอัลกอริทึม
บทที่ 3 ปัญหาเกตเทียร์ (Gettier) โชคทางญาณวิทยา และภาพหลอนของปัญญาประดิษฐ์
บทที่ 4 ภายในนิยม ปะทะ ภายนอกนิยม: ลัทธิความน่าเชื่อถือได้ (Reliabilism) เพียงพอหรือไม่?
บทที่ 5 ญาณวิทยาเชิงคุณธรรม (Virtue Epistemology): การพังทลายของความเป็นผู้กระทำทางปัญญา
บทที่ 6 ภาพลวงตาของจิตส่วนขยาย (The Extended Mind) และความแปลกแยกทางญาณวิทยา
บทที่ 7 ญาณวิทยาสังคมและความอยุติธรรมทางญาณวิทยา (Epistemic Injustice) ในยุค AI
บทที่ 8 ปรากฏการณ์วิทยาของการรู้ (Phenomenology of Knowing): การดำรงอยู่ของมนุษย์กับเครื่องมือ
บทที่ 9 การสังเคราะห์และทางออก: ปัญญาประดิษฐ์ในฐานะคู่สนทนาแบบโสกราตีส
บทที่ 10 บทสรุป
บรรณานุกรม

บทที่ 1
บทนำ: รอยต่อระหว่างปัญญาประดิษฐ์และทฤษฎีความรู้

ตลอดหน้าประวัติศาสตร์อันยาวนานของปรัชญาตะวันตกและตะวันออก ญาณวิทยา (Epistemology) ทำหน้าที่เป็นด่านหน้าในการปกป้องและนิยามอาณาเขตของ 'ความรู้' ตั้งแต่การตั้งคำถามของโสเครตีสในกรุงเอเธนส์ การถกเถียงเรื่องเหตุผลนิยม (Rationalism) และประจักษ์นิยม (Empiricism) ในยุคฟื้นฟูศิลปวิทยาการ มาจนถึงการวิพากษ์โครงสร้างภาษาในศตวรรษที่ 20 มนุษย์ผูกขาดสิทธิแต่เพียงผู้เดียวในการเป็น 'ผู้สร้างและครอบครองความรู้' ทว่าเมื่อก้าวเข้าสู่ทศวรรษที่สามของศตวรรษที่ 21 การอุบัติขึ้นของปัญญาประดิษฐ์เชิงสร้างสรรค์ (Generative AI) ได้ท้าทายรากฐานนี้อย่างรุนแรงแบบที่ไม่เคยมีมาก่อนในประวัติศาสตร์

ปัญญาประดิษฐ์ในปัจจุบันไม่ได้ทำหน้าที่เป็นเพียงห้องสมุดเก็บข้อมูลเหมือนเสิร์ชเอนจินในยุคก่อน แต่มันมีความสามารถในการทำความเข้าใจบริบท สังเคราะห์ความรู้ใหม่ และนำเสนอคำตอบด้วยภาษาที่ลื่นไหลราวกับมนุษย์ เมื่อบุคคลหนึ่งเผชิญกับปัญหาที่ซับซ้อน ป้อนคำถามลงในหน้าต่างแชทของ AI และได้รับคำตอบที่ 'ถูกต้อง' กลับมาอย่างรวดเร็ว ปรากฏการณ์นี้ได้สร้างภาพลวงตาทางญาณวิทยาที่แนบเนียนที่สุด คำถามเชิงปรัชญาที่ทรงพลังจึงบังเกิดขึ้น: ผู้ใช้งานมีความชอบธรรมหรือไม่ที่จะประกาศว่า 'ฉันรู้' เพียงเพราะได้รับคำตอบที่ถูกต้องจากเครื่องจักร?

รายงานการศึกษาค้นคว้าขั้นสูงฉบับนี้ ถูกเขียนขึ้นด้วยจุดมุ่งหมายเพื่อรื้อถอนและชำแหละความซับซ้อนของคำถามดังกล่าว โดยจะไม่หยุดอยู่เพียงการวิเคราะห์ตามทฤษฎีความรู้แบบดั้งเดิม แต่จะดำดิ่งลงไปถึงรากฐานทางภววิทยา (Ontology) และญาณวิทยาเชิงสังคม (Social Epistemology) ข้อเสนอหลัก (Thesis) ของงานชิ้นนี้คือ การได้รับคำตอบที่ถูกต้องจาก AI เป็นเพียงการครอบครอง 'ข้อมูลที่มีความแม่นยำ' ทว่าห่างไกลอย่างยิ่งจากสภาวะของการ 'รู้' ในมิติของมนุษย์ การอ้างสิทธิ์ในความรู้ผ่าน AI โดยปราศจากการวิพากษ์คือการละทิ้งเจตจำนงเสรีทางปัญญา และนำไปสู่ความเสื่อมถอยของปัญญาญาณอย่างหลีกเลี่ยงไม่ได้

เพื่อพิสูจน์ข้อเสนอนี้ รายงานจะนำผู้อ่านเดินทางผ่าน 8 ทฤษฎีหลักทางญาณวิทยา เริ่มตั้งแต่ปัญหา JTB ของเพลโต ความท้าทายของ Edmund Gettier มุมมองด้านลัทธิความน่าเชื่อถือได้ของ Alvin Goldman ญาณวิทยาเชิงคุณธรรมของ Ernest Sosa ทฤษฎีจิตส่วนขยายของ Clark และ Chalmers ญาณวิทยาสังคมของ Miranda Fricker ไปจนถึงปรากฏการณ์วิทยาของ Martin Heidegger เพื่อวาดภาพที่สมบูรณ์ที่สุดของสถานะทางปัญญาของมนุษย์ในยุคที่เครื่องจักรกลายเป็นผู้ตอบคำถามของจักรวาล

บทที่ 2

รื้อถอนโครงสร้าง JTB: ปัญหาของการให้เหตุผลในยุคกล่องดำทางอัลกอริทึม

การสืบสวนทางญาณวิทยาใดๆ ย่อมต้องเริ่มต้นที่รากฐานอันมั่นคงที่สุด ซึ่งก็คือทฤษฎีความรู้ของเพลโตที่ปรากฏในบทสนทนา Theaetetus เพลโตได้วางกรอบว่า ความรู้คือ 'ความเชื่อที่เป็นจริงและมีเหตุผลรองรับ' (Justified True Belief - JTB) องค์ประกอบทั้งสามนี้เปรียบเสมือนเสาเข็มที่ค้ำยันปราสาทแห่งความรู้ หากเราป้อนคำถามให้ AI และ AI ตอบกลับด้วยข้อเท็จจริงที่พิสูจน์ได้ (เช่น ทฤษฎีสัมพัทธภาพของไอน์สไตน์) เงื่อนไขประการแรกคือ 'ความเป็นจริง' (Truth) ย่อมได้รับการเติมเต็ม และเมื่อผู้ใช้ยอมรับคำตอบนั้น เงื่อนไขประการที่สองคือ 'ความเชื่อ' (Belief) ก็เกิดขึ้น ทว่าวิกฤตทางปรัชญาซ่อนตัวอยู่ในเงื่อนไขที่สาม นั่นคือ 'การให้เหตุผล' (Justification)

การให้เหตุผลทางญาณวิทยาแบ่งออกเป็นสองสำนักหลัก คือสำนักภายในนิยม (Internalism) และภายนอกนิยม (Externalism) หากพิจารณาตามหลักภายในนิยม บุคคลจะมีความชอบธรรมในการเชื่อสิ่งใดสิ่งหนึ่งได้ ก็ต่อเมื่อเขาสามารถเข้าถึง (Cognitively accessible) กระบวนการและเหตุผลที่นำมาซึ่งความเชื่อนั้นได้ด้วยตนเอง เมื่อเราสนทนากับมนุษย์ผู้เชี่ยวชาญ เราสามารถซักไซ้ไล่เลียงถึงที่มาของเหตุผล ขอดูงานวิจัย หรือถามถึงวิธีการอนุมานของพวกเขาได้ แต่สถาปัตยกรรมของปัญญาประดิษฐ์แบบโครงข่ายประสาทเทียมระดับลึก (Deep Neural Networks) นั้น มีลักษณะที่วงการวิทยาการคอมพิวเตอร์เรียกว่า 'กล่องดำ' (Black Box)

ในกล่องดำนี้ ข้อมูลจำนวนมหาศาลถูกประมวลผลผ่านตัวแปรนับล้านล้านพารามิเตอร์ (Trillions of parameters) เป็นการคำนวณน้ำหนักทางสถิติของเวกเตอร์คำที่ซับซ้อนเกินกว่าที่สติปัญญาของมนุษย์จะตามรอยได้ แม้แต่วิศวกรผู้พัฒนาระบบเองก็ไม่อาจอธิบายกระบวนการคิดเชิงตรรกะในแต่ละก้าวของ AI ได้อย่างชัดเจน เมื่อผู้ใช้งานยอมรับคำตอบจากกล่องดำนี้ การให้เหตุผลของเขาจึงมีลักษณะของการ 'จำนนต่ออำนาจมืด' (Blind Deference to Opaque Authority)

ผู้ใช้ไม่สามารถระบุได้ว่าคำตอบนั้นมาจากการอนุมานที่ถูกต้อง การรวบรวมหลักฐานเชิงประจักษ์ หรือเพียงแค่การเดาคำทางสถิติที่มีความน่าจะเป็นสูงสุดในฐานข้อมูล เมื่อผู้ใช้ถูกตัดขาดจากการเข้าถึงกระบวนการให้เหตุผล (Lack of internalist access) สถานะของความเชื่อนั้นจึงมีลักษณะเหมือน 'ศรัทธา' (Faith) มากกว่า 'ความรู้' (Knowledge) การกล่าวว่า 'ฉันรู้เพราะอัลกอริทึมบอกมา' จึงสอบตกในเงื่อนไขการให้เหตุผลเชิงวิพากษ์ของ JTB อย่างสิ้นเชิง

บทที่ 3

ปัญหาเกตเทียร์ (Gettier) โชคทางญาณวิทยา และภาพหลอนของปัญญาประดิษฐ์

สมมติว่ามีผู้เสนอข้อโต้แย้งว่า การให้เหตุผลจาก AI ไม่จำเป็นต้องโปร่งใส ขอแค่เราเชื่อมั่นว่ามันมีฐานข้อมูลที่ใหญ่และน่าเชื่อถือก็เพียงพอแล้ว ข้อโต้แย้งนี้จะถูกทำลายลงอย่างราบคาบเมื่อเผชิญหน้ากับ 'ปัญหาเกตเทียร์' (Gettier Problems) ในปี 1963 Edmund Gettier ได้ตีพิมพ์บทความที่ทำลายความสมบูรณ์ของทฤษฎี JTB โดยชี้ให้เห็นว่า คนเราสามารถมี 'ความเชื่อที่เป็นจริงและมีเหตุผลรองรับ' ได้โดยที่ไม่ได้มีความรู้จริงๆ หากความจริงนั้นเกิดจาก 'ความบังเอิญ' หรือ 'โชค'

นักปรัชญาร่วมสมัยอย่าง Duncan Pritchard ได้ขยายความเรื่อง 'โชคทางญาณวิทยา' (Epistemic Luck) โดยเฉพาะโชคประเภทแทรกแซง (Intervening Luck) ซึ่งเป็นหัวใจสำคัญในการประเมิน AI ปรากฏการณ์ที่พบได้บ่อยที่สุดใน AI คือ 'อาการประสาทหลอน' (AI Hallucination) ซึ่งเป็นภาวะที่ AI ไม่ทราบคำตอบ แต่กลับสร้างข้อความที่ดูสมเหตุสมผล

ขึ้นมาเองอย่างแนบเนียน ลองจินตนาการถึงสถานการณ์ที่นักศึกษาแพทย์ถาม AI เกี่ยวกับตัวยาตัวใหม่ AI เกิดอาการประสาทหลอนและสุ่มสร้างชื่อกลไกการออกฤทธิ์ทางชีวเคมีขึ้นมาใหม่ทั้งหมด

แต่ด้วยความบังเอิญที่เหลือเชื่อ (A cosmic coincidence) กลไกทางชีวเคมีที่ AI สุ่มสร้างขึ้นมาอย่างมั่วซั่วนั้น ดันไปตรงกับผลวิจัยล่าสุดที่เพิ่งตีพิมพ์ในวารสารการแพทย์เช้านั้นพอดีเป๊ะ นักศึกษาแพทย์ผู้นั้นอ่านแล้วเชื่อ (Belief) และสิ่งนั้นเป็นความจริง (Truth) อีกทั้งเขายังมีเหตุผลให้ความเชื่อนั้นเพราะ AI ตัวนี้เคยให้คำตอบที่ถูกต้องมาตลอด (Justification) ตามหลักของเพลโต นักศึกษาคนนี้ควรจะ 'รู้' แต่ในทางปรัชญา สัญชาตญาณของเราต่อต้านอย่างรุนแรง เราตระหนักดีว่านักศึกษาคนนี้ไม่ได้รู้ เขาแค่ 'โชคดีอย่างเหลือเชื่อ' ประเด็นสำคัญของปัญหาเกตเทียร์ในยุค AI คือ คำตอบที่ถูกต้องอาจขาด 'การติดตามความจริง' (Truth-Tracking) กระบวนการสร้างคำตอบไม่ได้เชื่อมโยงเชิงสาเหตุ (Causal connection) กับข้อเท็จจริงในโลกแห่งความเป็นจริง แต่เป็นการเชื่อมโยงทางคณิตศาสตร์สถิติ ตราบใดที่เราไม่สามารถแยกแยะความแตกต่างระหว่าง 'คำตอบที่ถูกเพราะกระบวนการที่แท้จริง' กับ 'คำตอบที่ถูกเพราะความบังเอิญทางสถิติ' สถานะความรู้ของผู้ใช้ย่อมตกอยู่ในความสุ่มเสี่ยง และความสุ่มเสี่ยงคือศัตรูตัวฉกาจที่สุดของความรู้

บทที่ 4

ภายในนิยม ปะทะ ภายนอกนิยม: ลัทธิความน่าเชื่อถือได้ (Reliabilism) เพียงพอหรือไม่?

หากการเข้าถึงเหตุผลภายใน (Internalism) สร้างปัญหา บางสำนักปรัชญาอาจหันไปพึ่งพา ภายนอกนิยม (Externalism) โดยเฉพาะทฤษฎี 'ลัทธิความน่าเชื่อถือได้' (Reliabilism) ของ Alvin Goldman ทฤษฎีนี้เสนอว่า ความเชื่อจะเป็นความรู้ได้ก็ต่อเมื่อมันถูกผลิตขึ้นโดย 'กระบวนการก่อร่างความเชื่อที่เชื่อถือได้' (Reliable belief-forming process) โดยไม่จำเป็นที่ผู้เชื่อจะต้องเข้าใจกระบวนการนั้น เช่น เราใช้สายตาในการมองเห็นและเกิดความรู้ว่ามีต้นไม้อยู่ตรงหน้า เราไม่จำเป็นต้องเข้าใจกลไกการทำงานของเส้นประสาทตาตราบใดที่สายตาของเราเชื่อถือได้ (มีความแม่นยำสูง)

ผู้สนับสนุนเทคโนโลยีมักหยิบยกทฤษฎีนี้มาปกป้อง AI โดยให้เหตุผลว่า หากโมเดลภาษาขนาดใหญ่ (LLMs) มีสถิติการให้คำตอบที่ถูกต้องสูงถึง 98% การถามคำถามกับ AI ย่อมถือเป็นกระบวนการก่อร่างความเชื่อที่เชื่อถือได้ เทียบเท่ากับการรับรู้ผ่านประสาทสัมผัส หรือการทำงานของความทรงจำ ดังนั้น เมื่อ AI ให้คำตอบที่ถูก ผู้ใช้จึงมีความรู้โดยสมบูรณ์ ทว่า การเปรียบเทียบ AI กับประสาทสัมผัสทางชีววิทยาเป็นการอนุมานที่ผิดฝาผิดตัวอย่างร้ายแรง (False Analogy) ประสาทสัมผัสของมนุษย์มีวิวัฒนาการมาหลายล้านปีและมีความเสถียรในเชิงกายภาพ (Stable physical mechanism) ในขณะที่ AI เป็นอัลกอริทึมที่ขับเคลื่อนโดยบริษัทเอกชน ซึ่งมีการปรับปรุงโมเดล เปลี่ยนแปลงชุดข้อมูล (Training data) และแทรกแซงกติกา (Fine-tuning & RLHF) อยู่ตลอดเวลา กระบวนการที่แม่นยำในวันนี้อาจให้ผลลัพธ์ที่ผิดเพี้ยนในวันพรุ่งนี้เพียงเพราะบริษัทผู้ผลิตอัปเดตซอฟต์แวร์

นอกจากนี้ กระบวนการของ AI ยังเต็มไปด้วยอคติที่ซ่อนเร้น (Algorithmic Bias) การอนุญาตให้ระบบที่ไม่โปร่งใสและมีพลวัตแปรปรวนเช่นนี้ทำหน้าที่เป็น 'กระบวนการที่เชื่อถือได้ทางญาณวิทยา' จึงเป็นการสละอำนาจอธิปไตยทางปัญญาของมนุษย์ให้แก่องค์กรทุนนิยม การพึ่งพากระบวนการภายนอกที่ปราศจากความเข้าใจ จึงไม่อาจนับเป็นการบรรลุสภาวะของการรู้ทางปรัชญาได้

บทที่ 5
ญาณวิทยาเชิงคุณธรรม (Virtue Epistemology): การพังทลายของความเป็นผู้กระทำทางปัญญา

ในบรรดากรอบวิเคราะห์ทั้งหมด ญาณวิทยาเชิงคุณธรรม (Virtue Epistemology) คือเครื่องมือที่แหลมคมที่สุดในการชำแหละปัญหา AI ญาณวิทยาสาขานี้เปลี่ยนจุดสนใจจากการประเมิน 'ความเชื่อ' (The Belief) มาสู่การประเมิน 'ผู้มีความเชื่อ' (The Believer หรือ Epistemic Agent) Ernest Sosa นำเสนอแนวคิด 'ความเชื่อที่สัมฤทธิ์ผล' (Apt Belief) โดยเปรียบเปรยกระบวนการหาความรู้เสมือนนักธนูที่ยิงธนูสู่เป้า การยิงที่ดีต้องประกอบด้วย

ความแม่นยำ (Accuracy - เข้าเป้า), ความเชี่ยวชาญ (Adroitness - ท่วงท่าและทักษะที่ดี), และ ความสัมฤทธิ์ผล (Aptness - เข้าเป้า 'เพราะ' ทักษะ ไม่ใช่เพราะบังเอิญลมพัดช่วย) เมื่อพิจารณาในบริบทของการถาม AI คำตอบที่ถูกต้องคือ 'ความแม่นยำ' (Accuracy) อัลกอริทึมที่ทรงพลังของ AI คือ 'ความเชี่ยวชาญ' (Adroitness) แต่เมื่อถามถึง 'ความสัมฤทธิ์ผล' (Aptness) เรากลับพบรอยโหว่ขนาดใหญ่ ความจริงที่เกิดขึ้นนั้นไม่ได้เป็นผลมาจาก 'ทักษะทางปัญญาของผู้ใช้' เลยแม้แต่น้อย มันเป็นผลงานของเครื่องจักรล้วนๆ หากเปรียบกับนักธนู ผู้ใช้ก็เหมือนคนที่ใช้ปืนเลเซอร์นำวิถีอัตโนมัติกดปุ่มเพียงครั้งเดียวแล้วเข้าเป้าตรงกลาง เราย่อมไม่สามารถยกย่องคนผู้นั้นว่าเป็น 'ผู้เชี่ยวชาญการยิงธนู' ในทำนองเดียวกัน การอ้างว่า 'ฉันรู้' ผ่านความสำเร็จของ AI จึงเป็นการสวมรอยขโมยผลงานทางปัญญามาเป็นของตนเอง ลึกลงไปกว่านั้น Linda Zagzebski ได้เน้นย้ำถึง 'คุณธรรมแห่งจิตใจ' (Virtues of the Mind) เช่น ความรักในความจริง ความกล้าหาญทางปัญญา ความรอบคอบ และการเปิดกว้าง กระบวนการแสวงหาความรู้ของมนุษย์คือสนามฝึกฝนคุณธรรมเหล่านี้ เมื่อมนุษย์เผชิญความสงสัย เขาต้องค้นคว้า ต้องอดทนกับความไม่รู้ และประเมินหลักฐานด้วยตนเอง แต่เทคโนโลยี AI ได้ทำลายกระบวนการเหล่านี้ลงอย่างราบคาบ

การได้รับคำตอบสำเร็จรูปในเสี้ยววินาทีก่อให้เกิดสภาวะที่เรียกว่า 'ความเกียจคร้านทางญาณวิทยา' (Epistemic Sloth) มนุษย์ลดทอนสถานะตนเองจากการเป็น 'ผู้แสวงหาและสร้างความรู้' (Active Knower) ลงมาเป็นเพียง 'ภาชนะรองรับข้อมูล' (Passive Information Receptacle) การขาดการดิ้นรนทางปัญญาทำให้กล้ามเนื้อทางความคิดของมนุษย์ฝ่อลง การอ้างความรู้ในสภาวะที่จิตใจปราศจากคุณธรรมทางญาณวิทยา จึงเป็นเพียงการท่องจำข้อความบนหน้าจอ ปราศจากซึ่งความเข้าใจเชิงลึก (Understanding) และปัญญาญาณ (Wisdom)

บทที่ 6

ภาพลวงตาของจิตส่วนขยาย (The Extended Mind) และความแปลกแยกทางญาณวิทยา

นักคิดสายวิพากษ์ทางเทคโนโลยีอาจหยิบยกทฤษฎี 'จิตส่วนขยาย' (The Extended Mind) ของ Andy Clark และ David Chalmers มาโต้แย้ง ทฤษฎีนี้เสนอสมมติฐานที่ท้าทายว่า จิตใจและกระบวนการทางปัญญาของมนุษย์ไม่ได้ถูกจำกัดอยู่เพียงแค่ภายในสมอง (Skull-bound) แต่สามารถขยายออกไปเชื่อมต่อกับสภาพแวดล้อมและวัตถุภายนอกได้ คลาร์กและชาลเมอร์สยกตัวอย่าง 'อ็อตโต' (Otto) ผู้ป่วยโรคอัลไซเมอร์ที่พึ่งพาสมุดบันทึกส่วนตัวในการดำรงชีวิต สมุดบันทึกนั้นทำหน้าที่เสมือนเซลล์สมองเทียมของเขา เมื่ออ็อตโตต้องการทราบที่อยู่ของพิพิธภัณฑ์ เขาเปิดสมุดและพบคำตอบ ทฤษฎีนี้ยืนยันว่า อ็อตโต 'รู้' ตำแหน่งของพิพิธภัณฑ์อย่างแท้จริง ผ่านจิตใจส่วนขยายของเขา

หากประยุกต์ทรรศนะนี้ ผู้สนับสนุน AI จะอ้างว่า สมาร์ทโฟนและหน้าต่างแชทของ ChatGPT ก็คือสมุดบันทึกของอ็อตโตในศตวรรษที่ 21 มันคือจิตใจส่วนขยายของมนุษย์ ดังนั้น เมื่อ AI ประมวลผลและให้คำตอบ มนุษย์ก็มีความชอบธรรมที่จะอ้างว่า 'ฉันรู้' ทว่า การเปรียบเปรยนี้เต็มไปด้วยความลักลั่นและข้อบกพร่องเชิงตรรกะ (Logical Fallacy)

เงื่อนไขสำคัญที่ทำให้สมุดบันทึกของอ็อตโตเป็นจิตส่วนขยายได้คือ หลักการของ 'ความสอดคล้องและการกลั่นกรองในอดีต' (Past Endorsement) ข้อมูลในสมุดนั้นถูกอ็อตโตคิดไตร่ตรอง และจดบันทึกด้วยตัวเขาเองในอดีต มันคือความทรงจำที่ถูกแช่แข็งไว้ แต่ปัญญาประดิษฐ์ไม่ได้ทำหน้าที่เป็นคลังเก็บความทรงจำที่นิ่งเฉย (Static storage) AI เป็นสิ่งมีชีวิตทางคณิตศาสตร์ที่มีพลวัต (Dynamic Entity) ข้อมูลที่มันนำเสนอไม่ใช่สิ่งที่มนุษย์เคยไตร่ตรองมาก่อน แต่มันถูก 'สร้างขึ้นใหม่' (Generated) จากชุดข้อมูลที่มนุษย์ไม่ได้ครอบครอง

ความสัมพันธ์ระหว่างมนุษย์กับ AI จึงไม่ใช่การเชื่อมต่อกับความทรงจำของตนเอง แต่เป็น 'ปฏิสัมพันธ์ระหว่างตัวแทน' (Agent-to-Agent Interaction) หรือการปรึกษากับสติปัญญาอื่น (Alien Intelligence) การรับข้อมูลจากบุคคลภายนอกไม่สามารถนับเป็นการ

กระทำของจิตส่วนขยายได้ สิ่งที่เกิดขึ้นจึงไม่ใช่การขยายตัวของจิต (Mind Extension) แต่เป็นสภาวะ 'ความแปลกแยกทางญาณวิทยา' (Epistemic Alienation) ที่มนุษย์หลงคิดว่ากระบวนการคิดของเครื่องจักรคือกระบวนการคิดของตนเอง

บทที่ 7

ญาณวิทยาสังคมและความอยุติธรรมทางญาณวิทยา (Epistemic Injustice) ในยุค AI

ปรัชญาความรู้ไม่ได้ดำรงอยู่ในสุญญากาศ แต่ผูกพันอย่างแยกไม่ออกกับมิติทางสังคม ญาณวิทยาสังคม (Social Epistemology) ชี้ให้เห็นว่า การอ้างว่า 'ฉันรู้' เป็นการกระทำทางสังคม (Speech Act) ที่มีผลผูกพันทางจริยธรรม เมื่อบุคคลหนึ่งประกาศว่าตนรู้สิ่งใดสิ่งหนึ่ง เขาไม่ได้แค่รายงานสถานะทางจิตใจ แต่เขากำลังให้ 'คำมั่นสัญญาทางญาณวิทยา' (Epistemic Promise) ต่อชุมชนของเขา ว่าเขาได้ตรวจสอบข้อมูลนั้นแล้ว มีวิจารณญาณเพียงพอ และพร้อมที่จะรับผิดชอบต่อผลที่ตามมาหากข้อมูลนั้นผิดพลาด

การรับคำตอบสำเร็จรูปจาก AI และนำไปใช้ในสังคมโดยอ้างว่าเป็นความรู้ของตน เป็นการละเมิดพันธสัญญาทางสังคมนี้อย่างร้ายแรง หากสถาปนิกใช้คำนวณโครงสร้างจาก AI แล้วตึกถล่ม การแก้ตัวว่า 'ผมไม่รู้เรื่อง AI มันบอกมาแบบนี้' คือการยอมรับอย่างโจ่งแจ้งว่าเขาไม่เคย 'รู้' เลยตั้งแต่แรก ความรู้เรียกร้องหน้าที่ (Epistemic Duty) ที่จะตั้งคำถาม ปกป้อง และรับผิดชอบ ซึ่งเป็นสิ่งที่ผู้ใช้แบบมืดบอดไม่สามารถให้ได้

ยิ่งไปกว่านั้น การพึ่งพา AI อย่างแพร่หลายยังนำไปสู่สิ่งที่ Miranda Fricker เรียกว่า 'ความอยุติธรรมทางญาณวิทยา' (Epistemic Injustice) โดยเฉพาะ 'ความอยุติธรรมทางประจักษ์พยาน' (Testimonial Injustice) ในอนาคตอันใกล้ สังคมอาจเริ่มให้น้ำหนักและเครดิตกับคำตอบที่มาจากอัลกอริทึม (Algorithmic Authority) มากกว่าประสบการณ์และเหตุผลที่มาจากมนุษย์ด้วยกันเอง เสียงของคนชายขอบ ภูมิปัญญาท้องถิ่น หรือประสบการณ์ทางอารมณ์ที่ไม่สามารถแปลงเป็นข้อมูลเชิงสถิติได้ จะถูกลดทอนคุณค่าและถูกเบียดขับออกจากปริมณฑลของความรู้

เมื่อสังคมนิยามความรู้ว่าคือ 'ผลลัพธ์ที่ตรงกับที่ AI ตอบ' มนุษยชาติจะเผชิญกับ 'ความอยุติธรรมเชิงความหมาย' (Hermeneutical Injustice) ที่กรอบความคิดและภาษาที่ใช้ทำความเข้าใจโลก ถูกกำหนดโดยบรรษัทเทคโนโลยีที่ควบคุมโมเดล AI การอ้างสิทธิ์ในการรู้ผ่านกรอบของ AI จึงไม่ใช่แค่เรื่องปัญหาส่วนบุคคล แต่เป็นปัญหาสิทธิมนุษยชนและการเมืองแห่งความรู้ (Politics of Knowledge) ในศตวรรษที่ 21

บทที่ 8
ปรากฏการณ์วิทยาของการรู้ (Phenomenology of Knowing): การดำรงอยู่ของมนุษย์กับเครื่องมือ

ก้าวข้ามจากญาณวิทยาสายวิเคราะห์ (Analytic Epistemology) สู่สายทวีปยุโรป (Continental Philosophy) ปรากฏการณ์วิทยา (Phenomenology) ของ Martin Heidegger มอบเครื่องมืออันลึกซึ้งในการพิจารณาความสัมพันธ์ระหว่างมนุษย์กับเทคโนโลยี ไฮเดกเกอร์แบ่งแยกสถานะของสิ่งต่างๆ รอบตัวมนุษย์เป็นสองลักษณะ คือ 'ความพร้อมใช้งาน' (Ready-to-hand หรือ Zuhandenheit) เช่น ค้อนที่เราใช้ตอกตะปูโดยไม่ต้องคิดถึงตัวค้อน และ 'การปรากฏตรงหน้า' (Present-at-hand หรือ Vorhandenheit) เมื่อค้อนนั้นหัก เราจึงหยุดและพิจารณามันในฐานะวัตถุ

ในอุดมคติ AI ควรทำหน้าที่เป็นเพียงเครื่องมือ 'ความพร้อมใช้งาน' ที่ช่วยให้เราสำรวจโลก แต่ในความเป็นจริง ความทรงพลังของมันได้ทำให้มันกลายเป็น 'ศูนย์กลางของโลก' การปฏิสัมพันธ์กับ AI ทำลายกระบวนการที่มีชีวิตชีวาของการก่อร่างความรู้ (The Living Experience of Knowing) สำหรับมนุษย์ 'การเรียนรู้' (Learning) ไม่ใช่แค่การบรรลุถึงผลลัพธ์ แต่รวมถึงความสับสน ความหงุดหงิด การค้นพบ การอิ่มเอมใจเมื่อไขปริศนาได้ (Aha! Moment) ประสบการณ์เชิงปรากฏการณ์วิทยาเหล่านี้คือสิ่งที่หล่อหลอมตัวตน (Dasein) ของมนุษย์

เมื่อ AI เสนอคำตอบให้ในคลิกเดียว มันได้ตัดตอนประสบการณ์อันรุ่มรวยเหล่านี้ทิ้งไป เหลือเพียงจุดเริ่มต้นและจุดจบ ปราศจากการเดินทางตรงกลาง ความรู้จึงถูกลดทอนลงจาก

'กระบวนการแห่งการดำรงอยู่' (An Existential Process) ไปสู่ 'สินค้าบริโภค' (A Consumable Commodity) เมื่อเป็นเช่นนี้ การกล่าวว่า 'ฉันรู้' จากการอ่านคำตอบของ AI จึงเปรียบเสมือนการกล่าวว่า 'ฉันปีนถึงยอดเขาเอเวอเรสต์แล้ว' เพียงเพราะนั่งเฮลิคอปเตอร์ไปลงที่ยอดเขา ความหมายของการไปถึงเป้าหมายได้ถูกทำลายลงโดยสิ้นเชิงด้วยเครื่องมือที่รวดเร็วเกินไป

บทที่ 9

การสังเคราะห์และทางออก: ปัญญาประดิษฐ์ในฐานะคู่สนทนาแบบโสกราตีส

การวิพากษ์อย่างถอนรากถอนโคนในบทก่อนหน้า ไม่ได้มีเจตนาที่จะนำไปสู่ลัทธิต่อต้านเทคโนโลยี (Neo-Luddism) หรือการปฏิเสธการใช้ปัญญาประดิษฐ์ในวงการศึกษา การหลีกหนีจากเทคโนโลยีไม่ใช่ทางออก ทว่ามนุษยชาติจำเป็นต้อง 'สังเคราะห์' (Synthesis) และนิยามกระบวนทัศน์ทางญาณวิทยาใหม่ในการอยู่ร่วมกับ AI ทางออกของวิกฤตนี้คือการเปลี่ยนสถานะของ AI จาก 'ศาสดาพยากรณ์ผู้ให้คำตอบ' (Oracle of Truth) มาสู่ 'คู่สนทนาเชิงโสกราตีส' (Socratic Conversationalist)

โสกราตีสไม่เคยให้คำตอบสำเร็จรูปแก่ลูกศิษย์ แต่ใช้วิธีวิภาษวิธี (Dialectic) และการตั้งคำถาม (Maieutics) เพื่อสูตินารีเวชกรรมทางปัญญา ดึงความจริงที่ซ่อนอยู่ในตัวผู้เรียนออกมา ในทำนองเดียวกัน มนุษย์ต้องปรับเปลี่ยนวิธีการใช้ AI อย่างถอนรากถอนโคน เราต้องใช้ AI เพื่อกระตุ้นความสงสัย ใช้เพื่อทดสอบสมมติฐานที่ตนเองตั้งขึ้น ใช้เพื่อจำลองมุมมองที่แตกต่าง (Devil’s Advocate) หรือใช้เป็นตัวเร่งปฏิกิริยาทางปัญญา (Epistemic Catalyst) เมื่อผู้ใช้งานได้รับข้อมูลจาก AI แทนที่จะรับมาเป็นความเชื่อทันที เขาต้องนำมันเข้าสู่กระบวนการสอบทาน (Cross-examination) นำไปเปรียบเทียบกับเอกสารปฐมภูมิ วิเคราะห์อคติที่อาจแฝงอยู่ และบูรณาการข้อมูลนั้นเข้ากับโครงสร้างความรู้และโลกทัศน์เดิมของตน การกระทำเช่นนี้คือการทวงคืน 'ความเป็นผู้กระทำทางปัญญา' (Epistemic Agency) และเมื่อกระบวนการนี้สิ้นสุดลง การครอบครองความรู้นั้นก็จะเกิดขึ้นด้วยความยากลำบาก ทว่ามี

ความชอบธรรมอย่างสมบูรณ์ และเมื่อนั้น มนุษย์จึงจะสามารถประกาศคำว่า 'ฉันรู้' ได้อย่างภาคภูมิ โดยไม่ขัดต่อหลักคุณธรรมทางญาณวิทยา

บทที่ 10 บทสรุป

คำถามที่ว่า 'เราสามารถกล่าวได้หรือไม่ว่า ฉันรู้ เพียงเพราะ AI ให้คำตอบที่ถูกต้อง?' ได้นำพาเราเดินทางผ่านปริมณฑลของปรัชญาญาณวิทยาที่ลึกซึ้งที่สุด จากการศึกษาและวิเคราะห์อย่างรอบด้าน รายงานขั้นสูงฉบับนี้ขอสรุปด้วยท่าทีที่หนักแน่นและไร้ข้อกังขาว่า 'เราไม่สามารถกล่าวเช่นนั้นได้'

ความถูกต้อง (Accuracy) ของข้อมูล ไม่ใช่สิ่งเดียวกับ ความรู้ (Knowledge) ปัญญาประดิษฐ์สามารถผลิตคำตอบที่เป็นความจริงได้อย่างไร้ขีดจำกัด แต่ตราบใดที่กระบวนการนั้นปราศจากความโปร่งใสในการให้เหตุผล (Justification) ปราศจากการติดตามความจริงเชิงสาเหตุเพื่อต้านทานโชคทางญาณวิทยา (Epistemic Luck) และปราศจากซึ่งคุณธรรมแห่งความวิริยะอุตสาหะทางปัญญาของผู้ใช้งาน การรับคำตอบจากเครื่องจักรย่อมไม่อาจยกระดับขึ้นเป็นความรู้ได้

มนุษย์ไม่สามารถพึ่งพาสมมติฐานจิตส่วนขยาย (The Extended Mind) เพื่อสวมรอยเอาความฉลาดของอัลกอริทึมมาเป็นของตนเองได้ เพราะจิตใจที่แท้จริงเรียกร้องเจตจำนงเสรีและความรับผิดชอบต่อสังคม ความเสี่ยงที่ใหญ่หลวงที่สุดของยุคสมัยนี้ ไม่ใช่การที่ AI จะฉลาดล้ำหน้าจนควบคุมมนุษย์ (Singularity) แต่คือการที่มนุษย์พึ่งพา AI จนสูญเสียความสามารถในการคิด การไตร่ตรอง และการรับผิดชอบต่อความจริง (Epistemic Devolution) ท้ายที่สุดแล้ว ปัญญาประดิษฐ์เป็นเพียงเครื่องมือสร้างแผนที่ทางข้อมูล (Data Mapping) แต่ทักษะในการอ่านแผนที่ การตัดสินใจเลือกเส้นทาง และการก้าวเดินด้วยเท้าของตนเอง ยังคงเป็นศิลปะ ศักดิ์ศรี และภาระหน้าที่อันศักดิ์สิทธิ์ที่มนุษย์ต้องกระทำด้วยตนเอง

บรรณานุกรม

Clark, A., & Chalmers, D. (1998). The Extended Mind. Analysis, 58(1), 7-19.

Floridi, L. (2011). The Philosophy of Information. Oxford University Press.

Fricker, M. (2007). Epistemic Injustice: Power and the Ethics of Knowing. Oxford University Press.

Gettier, E. L. (1963). Is Justified True Belief Knowledge? Analysis, 23(6), 121-123.

Goldman, A. I. (1979). What Is Justified Belief? In G. S. Pappas (Ed.), Justification and Knowledge (pp. 1-23). D. Reidel.

Heidegger, M. (1962). Being and Time (J. Macquarrie & E. Robinson, Trans.). Harper & Row.

Plato. (1921). Theaetetus (H. N. Fowler, Trans.). Loeb Classical Library.

Pritchard, D. (2005). Epistemic Luck. Oxford University Press.

Sosa, E. (2007). A Virtue Epistemology: Apt Belief and Reflective Knowledge, Volume I. Oxford University Press.

Zagzebski, L. T. (1996). Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge. Cambridge University Press.

พันธ์ ลิดดาศักดิ์. (2569). ข้อเสนอเชิงวิพากษ์ขั้นสูงว่าด้วยญาณวิทยากับเทคโนโลยี AI. เอกสารประกอบการศึกษารายวิชาญาณวิทยาวิเคราะห์ มหาวิทยาลัยมหาจุฬาลงกรณราชวิทยาลัย.