

ปัญญาประดิษฐ์กับอคติทางสังคม-วัฒนธรรม

นิติ ภาวครพันธุ์

คณะรัฐศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

Niti Pawakapan

Faculty of Political Science, Chulalongkorn University

E-mail: npawakapan@gmail.com

ประเด็นที่ผมจะอภิปราย/ถกเถียงในที่นี้อาจไม่ใช่วิาทะหรือข้อโต้แย้งทางมานุษยวิทยาล่าสุด หรือได้รับความสนใจอย่างล้นหลามในหมู่นักวิชาการ/นักเรียนมานุษยวิทยา แต่ผมคิดว่ามีความสำคัญอย่างยิ่ง เพราะเป็นการตั้งคำถามเชิงสังคมต่อความคิด ทักษะคิดและพฤติกรรมของมนุษย์ อันเป็นผลพวงของแต่ละวัฒนธรรม นอกจากนี้ ยังเกี่ยวข้องกับหนึ่งในประเด็นใหญ่ที่ศูนย์มานุษยวิทยาสิรินธรเคยให้ความสนใจ กล่าวคือผมจะกล่าวถึงการศึกษาวิจัยและข้อถกเถียงเรื่องพาหนะไร้คนขับ (self-driving vehicles) ปัญญาประดิษฐ์ (Artificial Intelligence หรือเรียกสั้น ๆ ว่า AI – เอไอ ซึ่งจะใช้เรียกต่อไปเพื่อความกระชับ) และอคติของมนุษย์ที่เกิดขึ้นจากการอบรม/เรียนรู้ทางสังคมผ่านขบวนการทางวัฒนธรรม ซึ่งในด้านหนึ่ง แสดงให้เห็นว่าในการขับขีพาหนะ มนุษย์อาจกระทำการที่มีอคติแฝงอยู่ต่อคนบางกลุ่ม ซึ่งเกี่ยวข้องกับเรื่องสีผิวชาติพันธุ์ เพศ อายุ ฯลฯ แม้เป็นการกระทำที่อาจมิได้เกิดขึ้นอย่างจงใจหรือมีเจตนาก็ตาม แต่อาจก่อให้เกิดอันตรายต่อกลุ่มคนดังกล่าวได้ นอกจากนี้ ยังสะท้อนให้เห็นว่าสิ่งที่เรียกว่า “วัฒนธรรม” มีปัญหา และอาจเป็นอุปสรรคต่อพัฒนาการของมนุษย์ในการมีชีวิตอยู่ร่วมกัน ท่ามกลางความแตกต่าง/หลากหลาย อย่างปลอดภัย สันติสุขและเท่าเทียมกัน

ในการประชุมวิชาการประจำปีด้านมานุษยวิทยาและสังคมวิทยา ครั้งที่ 1 ในหัวข้อเรื่อง “เชื่อม ข้าม เผชิญหน้า ณ จุดตัด” เมื่อวันที่ 24-25 สิงหาคม 2561 ซึ่งมีศูนย์มานุษยวิทยาสิรินธร (องค์การมหาชน) เป็นเจ้าภาพ หนึ่งใน การอภิปรายกลุ่มย่อยที่ผมสนใจมากคือหัวข้อที่เกี่ยวกับปัญญาประดิษฐ์ ผู้ร่วมเสวนากล่าวถึงรายละเอียดต่าง ๆ ในหัวข้อนี้ ซึ่งบางเรื่องผมไม่มีความรู้เลย มีการพาดพิงถึงหลายประเด็น รวมถึง “กฎ 3 ข้อของหุ่นยนต์” (Three Laws

of Robotics) ที่ Isaac Asimov นักเขียนและศาสตราจารย์ผู้โด่งดังเคยบัญญัติไว้เมื่อกว่าเจ็ดสิบปีที่แล้ว ผู้ร่วมเสวนาท่านหนึ่งให้ความเห็นว่ากฎทั้งสามนี้ล้าสมัยไปเสียแล้ว เหตุผลหนึ่งก็คือความรู้ ความเข้าใจและเทคโนโลยีด้านหุ่นยนต์ของเรามีความล้ำหน้าอย่างมาก ความเห็นของผู้ร่วมเสวนาคนนี้ดูไม่แตกต่างนักจากข้อถกเถียงที่พบในเว็บไซต์บางแห่ง ที่ระบุว่าเทคโนโลยีด้านหุ่นยนต์ได้พัฒนาอย่างไม่หยุดยั้งและมีความก้าวหน้าอย่างยิ่ง ควบคู่ไปกับการพัฒนาด้านเอไอและการเรียนรู้ของสมองกล (Machine Learning) กับซอฟต์แวร์ต่าง ๆ ที่จำเป็นต่อการดำเนินชีวิตประจำวัน การพัฒนาเหล่านี้นำไปสู่การประดิษฐ์หุ่นยนต์ประเภทต่าง ๆ ซึ่งทำให้ปฏิสัมพันธ์ระหว่างมนุษย์-หุ่นยนต์เพิ่มขึ้นอย่างมีนัยสำคัญ และเนื่องจากเกิดประดิษฐกรรมหุ่นยนต์สารพันประเภทและลักษณะเพื่อวัตถุประสงค์ต่าง ๆ เราคงต้องพิจารณาทบทวนกฎ 3 ข้อของหุ่นยนต์อีกครั้ง¹ ตัวอย่างเช่น หุ่นยนต์หรืออุปกรณ์ที่ประดิษฐ์เพื่อการทหารและการทำศึกสงคราม ซึ่งอาจทำให้เกิดปัญหากับกฎข้อที่ 1 กล่าวคือวัตถุประสงค์ทางการทหารมุ่งหมายที่จะปกป้องชีวิตของทหารและพลเรือน แต่ก็อาจก่อให้เกิดอันตรายต่อฝ่ายศัตรูในสนามรบได้ นอกจากนี้ ปัญหาใหญ่ที่เกี่ยวกับกฎสามข้อนี้อยู่ที่เป้าหมายของการทำวิจัยเรื่องเอไอมุ่งหมายที่จะพัฒนาให้มันสามารถคิดและกระทำอย่างสมเหตุสมผลและเหมือนมนุษย์ (think and act rationally and like a human) ทว่า ในการทำวิจัยเรื่องเอไอนั้น การศึกษาเกี่ยวกับการเลียนแบบพฤติกรรมมนุษย์ยังมีไม่เพียงพอ การพัฒนาในด้านพฤติกรรมที่สมเหตุสมผลจึงจำกัดอยู่ในวงแคบ ๆ เท่านั้น (Anderson, 2017)

ผมคิดว่าประเด็นนี้มีความสำคัญอย่างยิ่ง เพราะเราจะแน่ใจได้อย่างไรว่า “ความคิดและกระทำของมนุษย์มีความสมเหตุสมผล?” หากยังพิสูจน์ไม่ได้หรือไม่ชัดเจน/กระจ่างว่าความคิดและกระทำของมนุษย์สมเหตุสมผล เราจะสร้างให้เอไอมีความคิดและกระทำที่สมเหตุสมผลได้อย่างไร? การตั้งคำถามดังกล่าว ซึ่งนำไปสู่ข้อถกเถียงเกี่ยวกับความคิด ศักดิ์ศรีและพฤติกรรมของมนุษย์ เป็นผลมาจากรายงานบางส่วนที่เกี่ยวข้องกับการศึกษา/งานวิจัยเรื่องยานพาหนะขับเคลื่อนได้เอง (self-driving vehicles) อันเป็นความพยายามในการนำเอไอมาแทนที่มนุษย์ในการควบคุมยานพาหนะ

มีรายงานว่าความคิดเรื่องยานพาหนะขับเคลื่อนได้เองเริ่มปรากฏขึ้นในงาน Futurama หรือ New York World’s Fair ที่จัดขึ้นในปี 1939 โดยมี

General Motors Corporation (GM) เป็นสปอนเซอร์ นักออกแบบอุตสาหกรรมชาวอเมริกันนาม Norman Bel Geddes ได้แสดงประดิษฐ์กรรมต้นแบบของเขา เป็นยานพาหนะใช้พลังงานไฟฟ้าที่บังคับด้วยคลื่นวิทยุแม่เหล็กไฟฟ้า² ในปัจจุบัน ยานพาหนะขับเคลื่อนได้เองมีพัฒนาการที่ค่อนข้างซับซ้อน ความน่าสนใจประการหนึ่งคือการประยุกต์ใช้แนวคิดและวิธีการศึกษาเชิงมานุษยวิทยาในการทาวิจัยในด้านนี้ เช่น กรณีที่นิสสัน หนึ่งในบริษัทผลิตรถยนต์ที่ยิ่งใหญ่ของญี่ปุ่น ได้ว่าจ้างนักมานุษยวิทยานาม Melissa Cefkin เป็นนักวิจัยหลักของ Human Centered Systems ที่เป็นส่วนหนึ่งของ Nissan Research Center, Silicon Valley ในการศึกษา/วิจัยเกี่ยวกับพาหนะที่ขับเคลื่อนได้เองอย่างอิสระ³

Cefkin เคยทำงานร่วมกับบริษัทธุรกิจใหญ่ ๆ หลายแห่งและเริ่มทำงานให้นิสสันในเดือนมีนาคม 2015 เธอสนใจว่ามนุษย์มีปฏิสัมพันธ์อย่างไรกับรถยนต์ และระบบขับเคลื่อนได้เองอย่างอิสระ (autonomous systems) จะตีความพฤติกรรมของมนุษย์อย่างไร ทีมงานของเธอจึงเริ่มด้วยการวิเคราะห์ปฏิสัมพันธ์ระหว่างคนขับรถยนต์บนถนน และระหว่างรถยนต์กับคนเดินเท้า (pedestrians) คนขี่จักรยาน (cyclists) และสภาพแวดล้อมบนถนน ด้วยความหวังว่าจะใช้ข้อมูลเหล่านี้สอนให้รถยนต์ขับเคลื่อนได้เองปฏิบัติเยี่ยงคนขับรถยนต์ที่ดี หนึ่งในบริบทที่เธอเน้นความสำคัญคือสี่แยก ซึ่งเธอสังเกตว่าเป็นบริเวณที่คนขับรถยนต์ คนขี่จักรยานและคนเดินเท้าจะสบสายตากัน (eye contact) เพื่อแสดงเจตจำนงของตนต่อผู้อื่น ซึ่งรถยนต์ขับเคลื่อนได้เองไม่สามารถทำได้ เธอจึงคิดว่ารถยนต์ขับเคลื่อนได้เองในอนาคตจะใช้แสงเพื่อสื่อสารกับคนเดินเท้า หรือต่อพฤติกรรมอื่น ๆ⁴ (Edelstein, 2016)

เราอาจคิดว่าความพยายามของนักมานุษยวิทยาอย่าง Cefkin พังสมเหตุสมผล นั่นคือเริ่มต้นด้วยการศึกษาปฏิสัมพันธ์ระหว่างผู้คนกลุ่มต่าง ๆ บนถนนและต่อสภาพแวดล้อม เพื่อทำความเข้าใจเกี่ยวกับพฤติกรรมของมนุษย์ในการขับขีรถยนต์ ทว่า พฤติกรรม อันเป็นผลพวงของความคิด ทศนคติและโลกทัศน์ของมนุษย์ ดูซับซ้อนและมีแนวโน้มที่จะขัดแย้งกัน ดูได้จากการศึกษาและงานวิจัยหลายชิ้น เช่น จากการสำรวจด้วยการใช้คำถามออนไลน์ที่ใช้ชื่อว่า “Moral Machine” ของสถาบัน MIT ในปี 2016 เพื่อทดสอบแนวการตัดสินใจเชิงจริยธรรม (ethical decisions) ของผู้ตอบคำถาม (ซึ่งมีลักษณะ

คล้ายกับการทดสอบที่เรียกว่า “Trolley Problem”⁵) ด้วยการสมมติว่าจะเกิดอุบัติเหตุรถชนขึ้น และให้ยูสเซอร์หรือผู้ใช้คอมพิวเตอร์ (user) ตอบคำถามเพื่อทดสอบปัจจัย 9 ประการที่เกี่ยวข้องกับอุบัติเหตุ เช่น ถ้าต้องเลือก ผู้ตอบคำถามเลือกที่จะชนผู้ขายหรือผู้หญิง คนหนุ่มมากหรือคนหนุ่มน้อย คนอายุน้อยหรือคนชรา คนเดินเท้าหรือคนที่ไม่ข้ามถนนตรงที่ให้ข้าม (jaywalkers) คนที่มีสถานภาพทางสังคมต่ำหรือสูง เป็นต้น

มียูสเซอร์จำนวน 40 ล้านคนจาก 233 ประเทศตอบคำถามแนวการตัดสินใจเชิงจริยธรรม จากข้อมูลทั้งหมดผู้ศึกษาได้สรุปว่าผู้ตอบคำถามเลือกที่จะรักษาชีวิตมนุษย์มากกว่าสัตว์ ของคนหนุ่มมากกว่าคนหนุ่มน้อย เด็กมากกว่าผู้ใหญ่ และเสนอว่าผู้วางแผนนโยบายควรนำปัจจัยเหล่านี้ไปใช้พิจารณาประกอบการออกกฎหมายสำหรับรถยนต์ขับเคลื่อนได้เอง แต่ก็ให้ความเห็นว่าข้อมูลที่ได้เหล่านี้แสดงถึงจริยศาสตร์เชิงบรรยาย (descriptive ethics) ของยูสเซอร์ที่ใช้ประกอบในการตัดสินใจ มิใช่การตัดสินใจในแง่ของจริยศาสตร์เชิงบรรทัดฐาน (normative ethics) ซึ่งเป็นเรื่องที่เกี่ยวข้องในขั้นตอนนี้ควรเป็นผู้ตัดสินใจชี้ขาด

ประเด็นสำคัญประการหนึ่งคือการตัดสินใจเชิงจริยธรรมของยูสเซอร์ผันแปรไปตามความแตกต่างของแต่ละประเทศ ซึ่งเกี่ยวข้องกับปัจจัยหลายประการ เช่น ด้านภูมิศาสตร์ ที่มีความแตกต่างระหว่างประเทศในเอเชียและยุโรป หรือด้านวัฒนธรรม ก็มีสังคมที่เน้นความเป็นปัจเจกบุคคลและสังคมส่วนรวม (Vincent, 2018) เช่น คนตอบคำถามที่มาจากวัฒนธรรมที่เน้นสังคมส่วนรวม อย่างประเทศจีนและญี่ปุ่น เลือกที่จะรักษาชีวิตคนที่มีอายุมากกว่ากว่าคนที่มีอายุน้อย ซึ่งอาจเกี่ยวข้องกับการที่สังคมทั้งสองเน้นการเคารพนับถือผู้อาวุโส ในขณะที่ผู้ที่มาจากประเทศยากจนที่สถาบันทางสังคม-การเมืองมีความอ่อนแอก็มีความอดกลั้น ยอมรับคนที่ไม่ข้ามถนนตรงที่ให้ข้ามได้มากกว่า ส่วนคนที่มาจากสังคมที่มีความไม่เท่าเทียมทางเศรษฐกิจสูงก็จะเห็นช่องว่างของการปฏิบัติต่อคนที่มีสถานภาพทางสังคมสูงและต่ำที่ต่างกันอย่างยิ่ง และมีการตั้งข้อสังเกตว่าผู้ตอบคำถามจำนวนมากจากวัฒนธรรมที่เน้นเรื่องปัจเจกบุคคล เช่น อังกฤษและสหรัฐอเมริกา เลือกที่จะรักษาชีวิตคนหนุ่มมาก Karen Hao (2018) ผู้เขียนบทความแสดงความเห็นว่านี่อาจเป็นเพราะสังคมทั้งสองเน้นความสำคัญของปัจเจกบุคคลก็ได้⁶

อย่างไรก็ตาม ทีมวิจัยของ MIT ได้ระบุว่าผู้ตอบคำถามส่วนใหญ่เป็น

ผู้เชี่ยวชาญด้านเทคโนโลยี (tech savvy) คอมพิวเตอร์ ดังนั้น การทดสอบนี้อาจมีความคลาดเคลื่อนบ้างก็ได้ (Hao, 2018) นอกจากนี้ ยังพบว่ามีความขัดแย้งกันเองเชิงจริยธรรม (ethical paradox) ในหมู่ผู้ตอบคำถาม ซึ่งกล่าวว่าต้องการรถยนต์ขับเคลื่อนได้เองที่ปกป้องคนเดินเท้าแม้ว่าจะทำให้ผู้โดยสารในรถเกิดอันตรายก็ตาม แต่ก็ระบุว่าตนจะไม่ซื้อรถยนต์ขับเคลื่อนได้เองที่ถูกรโปรแกรมให้ทำเช่นนั้น ข้อสรุปที่สำคัญอีกสองประการคือผู้ตอบคำถามส่วนใหญ่เลือกที่จะรักษาชีวิตมนุษย์มากกว่าสัตว์เลี้ยง และรักษาชีวิตคนทั้งกลุ่มมากกว่าปัจเจกบุคคล คำตอบเหล่านี้สอดคล้องกับรายงานและข้อเสนอแนะเรื่องรถยนต์ขับเคลื่อนได้เองที่เสนอต่อกรมการจราจรด้านจริยธรรมเกี่ยวกับระบบการขับเคลื่อนได้เองของประเทศเยอรมัน (German Ethics Commission on Automated and Connected Driving) ในปี 2017 (Maxmen, 2018)

การศึกษาที่น่าสนใจอีกชิ้นหนึ่งคืองานวิจัยของ Georgia Institute of Technology ที่ค้นพบว่าระบบตรวจจับ (detection systems) ที่ใช้ในรถยนต์ขับเคลื่อนได้เอง เช่น เซ็นเซอร์และกล้อง จะตรวจจับผู้ที่มีผิวโทนสีจางหรือขาว (lighter skin tones) ได้ดีกว่าผู้ที่มีผิวสีเข้ม/ดำ ด้วยเหตุนี้ จึงเป็นไปได้มากกว่ารถยนต์อาจตรวจจับผู้ที่มีผิวสีเข้ม/ดำได้ไม่ทันและชนคนกลุ่มนี้ ทีมวิจัยพบว่าระบบดังกล่าวมีความคลาดเคลื่อนค่อนข้างสูงเมื่อเกี่ยวข้องกับประชากรบางกลุ่ม กล่าวคือไร้ความไม่แม่นยำต่อผู้ที่มีผิวสีเข้ม/ดำ แม้ว่าทีมวิจัยจะทดลองด้วยการเปลี่ยนให้เป็นเวลากลางวันหรือกั้นระบบตรวจจับภาพผลที่ออกมาก็เหมือนกัน รายงานวิจัยสรุปว่านี่เป็นหลักฐานที่แสดงว่าปัญหาที่แท้จริงของระบบคืออคติที่แฝงอยู่ (Cuthbertson, 2019)

ทั้งหมดนี้เป็นเพียงบางส่วนของ การค้นคว้า/งานวิจัยเกี่ยวกับรถยนต์ขับเคลื่อนได้เอง และแม้ว่านี่จะเป็นเพียงจุดเริ่มต้นของการศึกษาอันยาวนานในอนาคต แต่คงช่วยให้เราได้ตระหนัก/สำนึกว่าผู้ขับขี่รถยนต์ ไม่ว่าจะเป็นคนชาติใด เพศอะไร อาศัยอยู่ที่ไหน อายุเท่าไร นับถือศาสนาอะไร ประกอบอาชีพอะไร ก็อาจมีอคติในการตัดสินใจหรือเลือกที่จะกระทำการใด (และไม่กระทำการใด) เป็นอคติที่เป็นผลพวงของการเรียนรู้ผ่านคุณค่า ทักษะคติ โลกทัศน์ ความเชื่อของสังคมที่ตนเกิด เติบโตขึ้นมาและอาศัยอยู่

อาจไม่ผิดนักหากผมจะเรียกว่านี่คืออคติเชิงสังคม-วัฒนธรรมที่เราไม่ได้ตระหนักถึงหรือสำนึกว่าเรามี ดังนั้น ในฐานะนักเรียนมานุษยวิทยา ผมคิดว่า

เราควรพิจารณาเรื่องนี้อย่างจริงจัง และตั้งคำถามต่อวัฒนธรรมที่เป็นส่วนหนึ่งของเรา โดยเฉพาะอย่างยิ่ง เมื่อเราต้องตัดสินใจเชิงจริยธรรม ว่าวัฒนธรรมของเรามีอคติอะไรบ้าง? ส่งผลต่อการกระทำของเราอย่างไร? และผลลัพธ์ของการกระทำคืออะไร?

เชิงอรรถ

1 อันได้แก่

กฎข้อที่ 1 หุ่นยนต์มีอำนาจการตัดสินใจที่เป็นอันตรายต่อมนุษย์ หรือหนึ่งเฉยปล่อยให้มนุษย์ตกอยู่ในอันตราย

กฎข้อที่ 2 หุ่นยนต์ต้องเชื่อฟังคำสั่งที่ได้รับจากมนุษย์ ยกเว้นว่าคำสั่งนั้นขัดแย้งกับกฎข้อแรก

กฎข้อที่ 3 หุ่นยนต์ต้องปกป้องสถานะการมีตัวตนของตนไว้ トラบเท่าที่การกระทำนั้นมีได้ขัดแย้งกับกฎข้อแรกหรือกฎข้อที่สอง – ดู Mark Robert Anderson, “After 75 years, Isaac Asimov’s Three Laws of Robotics need updating”, in *The Conversation*. Academic rigour, journalistic flair, March 17, 2017 <<https://theconversation.com/after-75-years-isaac-asimovs-three-laws-of-robotics-need-updating-74501>> (สืบค้นข้อมูลในอินเทอร์เน็ต วันที่ 15 กันยายน 2563), คำแปลของผม

2 Bonnie Gringer, “History of the Autonomous Cars” <<https://www.titlemax.com/resources/history-of-the-autonomous-car/>> (สืบค้นข้อมูลในอินเทอร์เน็ต วันที่ 21 กันยายน 2563)

3 Center for Automotive Research, “Dr. Melissa Cefkin” <<https://www.cargroup.org/person/dr-melissa-cefkin/>> และดู Molly Shade, “Melissa Cefkin / A Profile”, Profile Series, EPIC, Advancing the Value of Ethnography in Industry, November 23, 2015 <<https://www.epicpeople.org/melissa-cefkin-a-profile/>> (สืบค้นข้อมูลในอินเทอร์เน็ต วันที่ 20 กันยายน 2563) Melissa Cefkin คงไม่ใช่ นักมานุษยวิทยาคนแรกหรือคนเดียวที่ทำงานในภาคเอกชน เพราะในสหรัฐอเมริกา นักมานุษยวิทยาได้เข้าไปมีส่วนร่วมกับหรือทำงานให้ภาคธุรกิจมานานหลายทศวรรษแล้ว นำไปสู่การเรียนการสอนวิชามานุษยวิทยาธุรกิจ (Business Anthropology) ในมหาวิทยาลัยหลายแห่ง

4 Stephen Edelstein, “An anthropologist will help Nissan design self-driving cars”, digitaltrends, August 11, 2016 <<https://www.digitaltrends.com>>

com/cars/nissan-hired-an-anthropologist-to-help-design-self-driving-cars/> (สืบค้นข้อมูลในอินเทอร์เน็ต วันที่ 22 กันยายน 2563)

- 5 เป็นข้อเสนองานของ Philippa Foot จากงานเขียนชื่อ “Abortion and the Doctrine of Double Effect” ตีพิมพ์ในปี 1967 เป็นส่วนหนึ่งของวิวาทะเรื่องการฆ่าแกง และถูกพัฒนาขยายความโดยนักปรัชญานาม Judith Jarvis Thomson ในเวลาต่อมา “Trolley problem” เป็นการตั้งคำถามต่อจริยธรรมของมนุษย์ และเป็นตัวอย่างของแนวคิดเชิงปรัชญาที่เรียกว่า “consequentialism” ซึ่งเสนอคำนิยามว่าจริยธรรม (morality) คือผลลัพธ์ของการกระทำและผลลัพธ์นั้นแหละที่สำคัญ จึงเกิดคำถามต่อเนื่องว่าแล้วมนุษย์ตัดสินใจกระทำการด้วยอะไร อย่างไร? Josh Clark ผู้เขียนบทความนี้ยกตัวอย่างว่าสมมติว่าคุณกำลังเดินเล่นตามทางรกรางอยู่ ทันใดนั้นคุณได้ยินเสียงคนร้องอะอะขอความช่วยเหลือ มีรถรางที่แล่นมาจากด้านหลังของคุณ เกิดเบรคแตก หยุดไม่ได้ มีผู้โดยสารบนรถรางคันนั้น 5 คน บังเอิญว่าคุณยืนอยู่ใกล้ที่สับรางรถรางและมีกองทรายอยู่ข้าง ๆ ถ้าคุณสับราง รถรางก็จะเปลี่ยนทิศทางแล่นไปชนกองทรายกองนั้น ผู้โดยสารทั้งห้าบนรถรางจะปลอดภัย ทว่ามีผู้ชายคนหนึ่งยืนไม่รู้ชื่อไหนชื่อไหนอยู่ที่กองทราย ถ้าคุณสับราง รถรางจะพุ่งชนชายคนนั้นและเขาอาจเสียชีวิตได้ คุณจะตัดสินใจเลือกทำอะไร? จะปกป้องชายห้าคนหรือคนเดียว? (ดู Josh Clark, “How the Trolley Problem Works”, howstuffworks, <<https://people.howstuffworks.com/trolley-problem.htm>> (สืบค้นข้อมูลในอินเทอร์เน็ต วันที่ 27 กันยายน 2563) ยังมีรายละเอียดอื่น ๆ อีกที่ผมไม่อาจอธิบายในที่นี้ได้หมด เพราะจะทำให้บทความนี้มีความยาวเกินไป
- 6 Karen Hao, “Should a self-driving car kill the baby or the grandma? Depends on where you’re from. The infamous “trolley problem” was put to millions of people in a global study, revealing how much ethics diverge across cultures.”, MIT Technology Review, Tech policy / AI Ethics, October 24, 2018 <<https://www.technologyreview.com/2018/10/24/139313/a-global-ethics-study-aims-to-help-ai-solve-the-self-driving-trolley-problem/>> (สืบค้นข้อมูลในอินเทอร์เน็ต วันที่ 24 กันยายน 2563) และดู James Vincent, “Global preferences for who to save in self-driving car crashes revealed” ซึ่งอ้างอิงถึงบททดสอบเดียวกัน

บรรณานุกรม

ภาษาอังกฤษ

Anderson, M.R. 2017. “After 75 years, Isaac Asimov’s Three Laws of Robotics need updating”. in The Conversation. Academic rigour,

journalistic flair, March 17, 2017 <<https://theconversation.com/after-75-years-isaac-asimovs-three-laws-of-robotics-need-updating-74501>> (สืบค้นข้อมูลในอินเทอร์เน็ต วันที่ 15 กันยายน 2563).

Center for Automotive Research. 2015. Dr. Melissa Cefkin.<<https://www.cargroup.org/person/dr-melissa-cefkin/>> และดู Molly Shade, “Melissa Cefkin / A Profile. Profile Series, EPIC, Advancing the Value of Ethnography in Industry, November 23, 2015 Available at <https://www.epicpeople.org/melissa-cefkin-a-profile/> (สืบค้นข้อมูลในอินเทอร์เน็ต วันที่ 20 กันยายน 2563).

Cuthbertson, A. 2019. “Self-driving cars more likely to drive into Black people, study claims”. New study suggests autonomous vehicles might be racist. The Independent, Indy/Life, 6 March 2019 Available at <https://www.independent.co.uk/life-style/gadgets-and-tech/news/self-driving-car-crash-racial-bias-black-people-study-a8810031.html> (สืบค้นข้อมูลในอินเทอร์เน็ต วันที่ 27 กันยายน 2563).

Edelstein, S. 2016. “An anthropologist will help Nissan design self-driving cars”. Digitaltrends, August 11, 2016 Available at <https://www.digitaltrends.com/cars/nissan-hired-an-anthropologist-to-help-design-self-driving-cars/> (สืบค้นข้อมูลในอินเทอร์เน็ต วันที่ 22 กันยายน 2563).

Gringer, B. “History of the autonomous cars”. Available at <https://www.titlemax.com/resources/history-of-the-autonomous-car/> (สืบค้นข้อมูลในอินเทอร์เน็ต วันที่ 21 กันยายน 2563).

Hao, K. 2018. Should a self-driving car kill the baby or the grandma? Depends on where you’re from. The infamous “trolley problem” was put to millions of people in a global study, revealing how much ethics diverge across cultures. MIT Technology Review, Tech policy / AI Ethics, October 24, 2018 Available at <https://www.technologyreview.com/2018/10/24/139313/a-global-ethics-study-aims-to-help-ai-solve-the-self-driving-trolley-problem/> (สืบค้นข้อมูลในอินเทอร์เน็ต วันที่ 24 กันยายน 2563).

Maxmen, A. 2018. “Self-driving car dilemmas reveal that moral choices are not universal”. Survey maps global variations in ethics for programming autonomous vehicles. Nature, News, 24 October 2018 Available at <https://www.nature.com/articles/d41586-018-07135-0> (สืบค้นข้อมูลในอินเทอร์เน็ต วันที่ 27 กันยายน 2563).

Vincent, J. 2018. “Global preferences for who to save in self-driving car crashes revealed”. The Verge, October. 24, 2018 Available at <https://www.theverge.com/2018/10/24/18013392/self-driving-car-ethics-dilemma-mit-study-moral-machine-results> (สืบค้นข้อมูลในอินเทอร์เน็ต วันที่ 24 กันยายน 2563).