

คู่มือวิทยาศาสตร์ยุคใหม่ ฉบับ AI

วิธีไม่หลอกตัวเอง ในวันที่เครื่องจักรตอบได้ทุกคำถาม

ห้องสมุด AICE · modern-science-handbook

15 บท · ฉบับภาษาเดียว

เขียนครั้งแรก 25 กันยายน 2026 · ปรับปรุงล่าสุด 25 กันยายน 2026

ทุกตัวเลข ชื่อ และปีในเล่มนี้ตรวจสอบไปยังแหล่งที่อ้างอิงไว้ได้

ก่อนเริ่ม

หนังสือเล่มนี้เริ่มด้วยแบบทดสอบ ก่อนจะบอกอะไรคุณสักอย่าง

เล่มนี้ไม่ได้จะทำให้คุณรู้มากขึ้น มันจะสอนวิธีจับได้ว่าตอนไหนคุณกำลังรู้ผิด

ความรู้ผิดไม่ได้หน้าตาต่างจากความรู้ถูก มันมาพร้อมความมั่นใจเท่ากัน มีคนยืนยันเท่ากัน บางครั้งมีตัวเลขกำกับมาด้วยซ้ำ และวันนี้ ทั้งสองแบบมาถึงมือคุณเร็วเท่ากัน จากกลุ่มไลน์ จากคลิปลั่น และจาก AI ที่ตอบได้ทุกคำถามในไม่กี่วินาที

สิ่งที่มนุษย์คิดขึ้นเพื่อแยกสองอย่างนี้ออกจากกันเรียกว่าวิทยาศาสตร์ มันไม่ต้องใช้ห้องแล็บ ไม่ต้องเก่งคณิตศาสตร์ และไม่ได้เป็นสมบัติของนักวิทยาศาสตร์คนไหน เล่มนี้จะเล่าว่ามันคืออะไร ทำงานอย่างไร พลาดตรงไหน และยังใช้ได้อยู่ไหมในวันที่เครื่องจักรพูดเก่งกว่าคน

ทายก่อนอ่าน

ทำส่วนนี้ก่อนเปิดบทแรก ใช้เวลาไม่เกินสิบนาที จดคำตอบไว้ที่ไหนก็ได้

ข้อแรกไม่มีตัวเลือก

เขียนหนึ่งประโยคว่าวิทยาศาสตร์คืออะไร

ใช้คำของคุณเอง ไม่ต้องเปิดหาจากที่ไหน แล้วเก็บไว้โดยยังไม่ต้องให้คะแนน บทที่ 13 จะให้เขียนอีกครั้ง แล้วค่อยให้คะแนนสองประโยคพร้อมกันด้วยเกณฑ์ในภาคผนวก

อีกแปดข้อเป็นแบบเลือกตอบ ทุกข้อมีคำตอบที่ดีที่สุดเพียงคำตอบเดียว ไม่มีข้อไหนถามความรู้ที่ต้องท่องจำ ทุกข้อถามว่าคุณจะตัดสินเรื่องที่ไม่เคยเห็นมาก่อนอย่างไร

1. เพื่อนคนหนึ่งดื่มน้ำมะนาวอุ่นทุกเช้าเป็นเวลาหนึ่งเดือน แล้วน้ำหนักลดไปสองกิโล ก่อนเชื่อว่ามะนาวเป็นเหตุ ข้อมูลไหนสำคัญที่สุด

ก. เพื่อนยืนยันว่าเดือนนั้นไม่ได้เปลี่ยนอาหารหรือการออกกำลังกาย นอกจากดื่มน้ำมะนาว ข. น้ำหนักที่ขึ้นลงในเดือนเดียวกันของคนคล้ายกันที่ไม่ได้ดื่มน้ำมะนาว ค.

รื้อวิวของคนอีกหลายร้อยคนที่ดื่มน้ำมะนาวแบบเดียวกันแล้วน้ำหนักรีดเหมือนกัน ง.
คำอธิบายของนักโภชนาการว่ากรดในน้ำมะนาวช่วยเร่งการเผาผลาญไขมันอย่างไร

2. กลุ่มไลน์เล่นหุ้นกลุ่มหนึ่งโพสต์ภาพกำไรจากคำแนะนำของแอดมินสัปดาห์ติดกัน
ก่อนจะซื้อฝีมือแอดมิน ข้อมูลที่คุณต้องการที่สุดคือ

ก. แอดมินเล่นหุ้นมากี่ปี และเคยทำงานในบริษัทหลักทรัพย์มาก่อนไหม ข. ภาพกำไร
ทั้งสัปดาห์มีชื่อบัญชีและวันที่กำกับ ให้ตรวจได้ว่าไม่ได้ตัดต่อมา ค. สมาชิกในกลุ่มที่
คนที่ทำตามคำแนะนำของแอดมินแล้วบอกว่าได้กำไรเหมือนกัน ง. คำแนะนำทุกครั้ง
ที่แอดมินเคยให้ รวมครั้งที่ขาดทุนแล้วไม่ได้โพสต์

3. หมอตอบลูกค้าว่า "ปีนี้คุณจะมีเรื่องให้ตัดสินใจครั้งสำคัญ และต้องระวังคนใกล้
ตัว" ถ้าอยากตรวจว่าคำทำนายนี้แม่นยำจริงไหม อุปสรรคใหญ่ที่สุดคืออะไร

ก. คำทำนายไม่ได้บอกว่าเรื่องไหนจะเกิดเมื่อไร ปีนี้เกิดอะไรขึ้นก็นับว่าตรงได้ ข.
หมอไม่ได้บอกว่าใช้ศาสตร์ไหนทำนาย จึงตรวจหลักการเบื้องหลังคำทำนายนี้ไม่ได้
ค. ต้องรอให้ครบปีก่อนจึงจะรู้ผล ระหว่างนั้นลูกค้าอาจจำคำทำนายคลาดเคลื่อน ง.
ลูกค้าคนเดียวไม่พอจะตรวจ ต้องมีอีกหลายคนที่ได้คำทำนายแบบเดียวกันนี้

4. ชาวออนไลน์พาดหัวว่า "งานวิจัยใหม่พบ กาแฟวันละสามแก้วลดความเสี่ยงมะเร็ง
40%" ขึ้นแรกที่ดีที่สุดคือ

ก. ดูว่างานวิจัยเรื่องกาแฟชิ้นนี้ตีพิมพ์ในวารสารชื่อดังระดับโลกหรือไม่ ข. ดูว่าชาว
งานวิจัยใหม่เรื่องกาแฟถูกแชร์กี่พันครั้ง และมีแพทย์คนไหนแชร์ต่อ ค. ดูว่างาน
ทบทวนที่รวบรวมงานวิจัยหลายชิ้นเรื่องกาแฟกับมะเร็งสรุปไว้อย่างไร ง. ดูว่าหัวหน้า
ทีมวิจัยจบจากมหาวิทยาลัยชั้นนำ และทำงานวิจัยด้านมะเร็งมานานกี่ปี

5. เพื่อนส่งข่าวงานวิจัยใหม่มาให้ เป็นเรื่องอาหารเสริมช่วยนอนหลับตัวที่คุณกินมา
ครึ่งปีและเคยแนะนำให้คนในบ้านกิน ก่อนเปิดอ่าน ข้อไหนกันความอยากให้นมันได้ผล
ไม่ให้น้ำตาลติดสินแทนคุณได้ดีที่สุด

ก. ตั้งใจอ่านงานนั้นทั้งฉบับให้ละเอียด ไม่ตัดสินใจจากพาดหัวข่าวที่เพื่อนส่งมา ข. ถาม
คนในบ้านที่กินอาหารเสริมตัวนี้ด้วยกันก่อน ว่านอนหลับดีขึ้นจริงไหม ค. บอกตัวเอง

ให้อ่านอย่างเป็นกลาง และพร้อมรับผลไม่ว่ามันจะออกมาทางไหน ง. จดไว้ก่อนว่าผลแบบไหนจะทำให้คุณเลิกกิน แล้วฝากคนในบ้านไว้ทวง

6. คุณถาม AI เรื่องสุขภาพ มันตอบพร้อมอ้างอิงงานวิจัยสามชิ้น มีชื่อผู้เขียน ปี และชื่อวารสารครบ ก่อนส่งคำตอบนี้ต่อให้คนอื่น สิ่งที่ดีที่สุดคือ

ก. เปิดงานทั้งสามชิ้นเองทีละชิ้น ดูว่ามีอยู่จริง และพูดแบบที่ AI สรุปรมาไหม ข. ถาม AI ตัวเดิมต่อว่าแน่ใจไหมว่างานทั้งสามชิ้นมีอยู่จริง และให้มันยืนยัน ค. ค้นชื่อวารสารทั้งสามชื่อ ดูว่าเป็นวารสารที่มีชื่อเสียงและยังตีพิมพ์อยู่ ง. ถาม AI อีกตัวด้วยคำถามเดิม แล้วส่งต่อเฉพาะงานที่ทั้งสองตัวอ้างตรงกัน

7. คุณเล่าไอดีयरูทิจให้ AI ฟังแล้วถามว่าดีไหม มันชมยาวสามย่อหน้า วิธีไหนช่วยแก้การเอออตามคุณได้ดีที่สุด

ก. ลบแชตเดิมแล้วถามใหม่ว่าไอดีของคุณดีไหม ดูว่าคำชมที่ได้เปลี่ยนไปหรือเปล่า ข. ถามต่อในช่องเดิมว่า "ติมาตรงๆ ได้เลย ไม่โกธ" แล้วเอาคำติที่ได้มาปรับไอดี ค. เปิดห้องแชตใหม่ เล่าไอดีเหมือนเป็นของคนอื่น แล้วขอเหตุผลที่หนักแน่นที่สุดว่าทำไมมันจะล้ม ง. เปลี่ยนไปใช้ AI รุ่นเสียเงินที่แพงกว่า แล้วถามว่าไอดีนี้ดีพอจะลงทุนจริงหรือเปล่า

8. ร้านกาแฟร้านหนึ่งอ้างว่าลูกค้าแยกเมล็ดพิเศษของร้านออกจากเมล็ดทั่วไปได้ การทดสอบไหนยุติธรรมที่สุด

ก. ให้ลูกค้าชิมเมล็ดพิเศษคู่กับเมล็ดทั่วไป โดยบอกก่อนว่าแก้วไหนคืออะไร แล้วให้คะแนนรสชาติ ข. ชิมสามแก้วหน้าตาเหมือนกัน สองแก้วเมล็ดเดียวกัน ให้ชี้แก้วที่ต่าง คนเสิร์ฟไม่รู้คำตอบ ทำหลายรอบแล้วนับ ค. ให้ลูกค้าหลายสิบคนชิมเมล็ดพิเศษ แล้วนับว่ากี่คนบอกว่าอร่อยกว่ากาแฟร้านอื่นที่เคยกินมา ง. ให้บาร์ิสต้าที่ชิมกาแฟมาหลายพันแก้วชิมทั้งสองเมล็ดสิบรอบโดยเห็นถุง แล้วบอกว่าต่างกันไหม

ตอบครบแล้วเปิดเฉลยชุดแรกท้ายเล่มได้เลย นับว่าได้กี่ข้อแล้วจดไว้ เฉลยชุดแรกบอกแค่ตัวอักษรกับบทที่เรื่องนั้นอยู่ เหตุผลของแต่ละข้อรออยู่หลังเฉลยชุดที่สอง เพื่อให้ชุดที่สองวัดสิ่งที่คุณได้จากเล่ม ไม่ใช่สิ่งที่คุณจำจากเฉลย

ท้ายเล่มมีอีกแปดข้อที่ถามเรื่องเดิมด้วยสถานการณ์ใหม่ทั้งหมด คะแนนสองชุดนั้น
กับประโยคสองประโยคของคุณ คือหลักฐานว่าเล่มนี้ทำงานกับคุณหรือไม่

ถ้าคุณได้เจ็ดหรือแปดข้อ อ่านบทที่ 1 แล้วข้ามไปภาคสามได้เลย ภาคสามฟังเรื่องใน
บทที่ 1 อยู่หลายตอน และเป็นส่วนที่เปลี่ยนเร็วที่สุดของเล่ม

ถ้าได้แปดเต็ม ชุดที่สองไม่เหลืออะไรให้แสดงแล้ว เพราะมันสูงกว่าแปดไม่ได้ แต่
ประโยคที่คุณเขียนยังวัดคุณได้

เล่มนี้สัญญาอะไร

อ่านจบแล้ว คุณจะบอกได้ด้วยคำของตัวเองว่าวิทยาศาสตร์คืออะไร และมีคำถามห่า
ข้อไว้ตัดสินข้ออ้างที่ไม่เคยเห็นมาก่อน ทั้งเรื่องสุขภาพ เงิน ข่าวก้าว และคำตอบของ AI

คำสัญญานี้แพ้ได้ และเล่มนี้บอกไว้ก่อนว่าแพ้ตรงไหน ครั้งแรกวัดด้วยประโยคที่คุณ
เขียน ครั้งหลังวัดด้วยแปดข้อ ถ้าคะแนนท้ายเล่มของครั้งไหนไม่สูงกว่าตอนต้น เล่มนี้
ล้มเหลวกับคุณในครั้งนั้น ไม่ใช่คุณที่ล้มเหลวกับมัน

แผนที่ของเล่ม

ภาคหนึ่ง · วิทยาศาสตร์คืออะไร บทที่ 1 ถึง 4 เริ่มจากเครื่องมือชิ้นหนึ่งที่ถูกซื้อมาใช้
หลายร้อยเครื่อง และคนที่เชื่อมั่นเห็นกับตาว่ามันได้ผล แล้วถามว่าถ้าเห็นกับตายังไม่
พอ มนุษย์ใช้อะไรแทน

ภาคสอง · คำถามห่าข้อ บทที่ 5 ถึง 9 หนึ่งบทต่อหนึ่งคำถาม เทียบกับอะไร · บังเอิญ
ได้ไหม · ถ้าผิด จะเห็นอะไร · ใครตรวจแล้วบ้าง · ฉันอยากให้มันจริงไหม

ภาคสาม · วิทยาศาสตร์กับ AI บทที่ 10 ถึง 12 AI สอบผ่านอะไรมาแล้ว สอบตกอะไร
ใช้มันอย่างไรไม่ให้ถูกมันหลอก และการทดสอบที่คุณทำเองได้ที่บ้านด้วยแก้วสามใบ

บทที่ 13 เล่มนี้แพ้ได้ตรงไหน และแบบทดสอบชุดที่สอง

แต่ละบทอ่านจบได้ในราวสิบห้านาที

เล่มนี้ทำขึ้นอย่างไร

หลักฐานของเล่มนี้ถูกค้นโดย AI หกตัวที่ทำงานพร้อมกันในห้าภาษา

ข้ออ้างแต่ละข้อถูกตามไปจนถึงแหล่งที่มีการเปิดอ่านจริง

ข้อที่เปิดไม่ได้เลยถูกตัดทิ้ง ส่วนข้อที่อ่านได้แค่ผ่านแหล่งที่อ้างถึงมัน หลักฐานท้ายบทบอกไว้

หลังเขียนเสร็จ ผู้ตรวจอีกสองรายเปิดแหล่งเหล่านั้นซ้ำ และจับความผิดที่บทย่างๆ เขียนไปแล้วได้หลายที่

ในฐานะผลประโยชน์ทับซ้อนที่ต้องแจ้ง AI ที่ใช้ค้นคว้าและเขียนเล่มนี้มาจากหนึ่งในบริษัทที่ภาคสามตรวจดูโมเดลของมัน จึงควรอ่านภาคสามโดยคิดถึงข้อนี้ไว้ด้วย

ป้ายความแน่นของหลักฐาน

ทุกบทยกด้วยส่วนที่ชื่อ หลักฐาน ข้ออ้างแต่ละข้อในนั้นมีป้ายหนึ่งในสี่ป้าย ป้ายเหล่านี้เป็นการตัดสินของเล่มนี้เอง และคุณเถียงมันได้

แน่น ถูกทดสอบมากกว่าหนึ่งครั้ง แหล่งที่เป็นอิสระต่อกันให้ผลไปทางเดียวกัน

พอเชื่อ หลักฐานดีจากหนึ่งหรือสองแหล่ง หรือยังมีคนแย้งอยู่

ยังบาง งานชิ้นเดียว ประสบการณ์ของคนคนเดียว หรือคำบอกของผู้ผลิตเกี่ยวกับของของตัวเอง

ตรวจไม่ได้ มีคนเล่า แต่เล่มนี้เปิดแหล่งต้นทางไม่ได้

ไม่มีประโยชน์ไหนในเนื้อเรื่องที่พูดแรงกว่าป้ายของมัน ถ้าคุณเจอ นั่นคือที่ที่เล่มนี้ผิด

ความเชื่อที่เล่มนี้จะหักล้าง

- เห็นกับตาหลายครั้ง คือหลักฐานที่ดีที่สุด · บทที่ 1
- วิทยาศาสตร์คือขั้นตอนหาข้อที่นักวิทยาศาสตร์ทุกคนทำตาม · บทที่ 2

- ของที่ใช้กันมาหลายร้อยปีต้องได้ผล และของโบราณเชื่อไม่ได้ สองประโยคนี้นี้ผิดพอกัน . บทที่ 3
- นักวิทยาศาสตร์ผิดบ่อย จึงเชื่อไม่ได้ . บทที่ 4
- งานวิจัยที่ตีพิมพ์แล้วคือความจริง . บทที่ 4 และ 8
- ยิ่งบอกคนว่าเขาเชื่อผิด เขาจะยิ่งเชื่อแรงขึ้น . บทที่ 9
- AI เก่งแค่เขียน ทำวิทยาศาสตร์จริงไม่ได้ . บทที่ 10
- AI ที่อ้างแหล่งครบ คือ AI ที่ค้นมาแล้ว . บทที่ 11
- ต้องมีห้องแล็บถึงจะทดลองได้ . บทที่ 12

1. เครื่องที่ใครก็เห็นว่าได้ผล

มันหน้าตาเหมือนด้ามจับพลาสติก มีเสาอากาศเล็กๆ ยื่นออกไปข้างหน้า คนถือเดินช้าๆ เสาอากาศหันไปทางไหน เขาก็เดินตามไปทางนั้น

เมื่อคนถือรู้ว่าของซ่อนอยู่ที่ไหน มันหาเจอสี่ครั้งจากสี่ เมื่อคนถือไม่รู้ มันหาเจอสามครั้งจากสี่สิบ

ตัวเลขสองชุดนี้มาจากวันเดียวกัน ห้องเดียวกัน กล่องชุดเดียวกัน และทหารสองนาย ชุดเดียวกันที่ผ่านการรับรองให้ใช้มัน ต่างกันอยู่อย่างเดียว คือตอนที่คนถือเดินเข้าห้อง เขารู้คำตอบหรือไม่

สามร้อยครั้ง

ในประเทศไทย เครื่องนี้ชื่อ GT200 ตามตัวเลขของสำนักงานการตรวจเงินแผ่นดิน หน่วยงานรัฐหลายแห่ง ตั้งแต่กองทัพบก กองทัพเรือ กองทัพอากาศ กรมราชองครักษ์ สถาบันนิติวิทยาศาสตร์ ไปจนถึงตำรวจภูธร ซื้อมันรวมกัน 836 เครื่อง เป็นเงินราว 759 ล้านบาท กองทัพบกหน่วยเดียวมีอยู่กว่าเจ็ดร้อยเครื่อง

คนที่ใช้มันไม่ได้ใช้เล่นๆ พวกเขาใช้มันหาระเบิด ผู้บัญชาการทหารบกในเวลานั้น ยืนยันต่อสาธารณะว่ามันใช้ได้จริง และยกตัวเลขว่ามันเคยเจอของจริงมาแล้วกว่าสามร้อยครั้ง

อย่าเพิ่งมองข้ามตัวเลขนั้น มันคือหลักฐานชนิดที่หนักแน่นที่สุดที่คนธรรมดาไม่เห็นกับตา ซ้ำหลายครั้ง จากคนหลายคน และไม่มีใครในนั้นโกหก คนที่เห็นเสาอากาศหันไปหาของจริงกับตาตัวเอง ไม่มีเหตุผลอะไรที่จะไม่เชื่อ

รัฐบาลในเวลานั้นสั่งให้กระทรวงวิทยาศาสตร์และเทคโนโลยีทดสอบเครื่องนี้ ผลที่นายกรัฐมนตรีแถลงเมื่อวันที่ 16 กุมภาพันธ์ 2010 คือ ทดสอบยี่สิบครั้ง หาวัตถุระเบิดเจอถูกสี่ครั้ง ซึ่งไม่ต่างจากการเดา รัฐบาลสั่งให้ทุกหน่วยงานหยุดซื้อ

การทดสอบครั้งนั้นจัดอย่างไร ใช้กล่องกี่ใบ ใครรู้อะไรบ้าง เล่มนี้หาเอกสารฉบับเต็มไม่พบ

แต่ในปีต่อมา ที่อีกฟากหนึ่งของโลก มีการทดสอบเครื่องรุ่นเดียวกันที่จดทุกขั้นตอนไว้ละเอียดพอให้คนอื่นทำซ้ำได้ และนั่นคือที่มาของตัวเลขสองชุดข้างบน

ห้องบอลรูมร้าง

เม็กซิโกซื้อเครื่องนี้ไปกว่าเก้าร้อยสี่สิบเครื่อง และเคยใช้คำชี้ของมันเป็นเหตุค้นบ้านคน ในคดีหนึ่ง ผู้พิพากษาสั่งให้กองทัพส่งเครื่อง ส่งทหารผู้ชำนาญ และส่งของที่เครื่องควรถาเจอ มาให้นักฟิสิกส์คนหนึ่งทดสอบ เขาชวนนักฟิสิกส์อีกคนมาร่วม ของที่ใช้คือของกลางที่เครื่องเคยชี้เจอในบ้านของจำเลย ยานี้ดีกว่าหนึ่งพันหกร้อยเม็ดกับกระสุนสี่นัด

สนามทดสอบคือห้องบอลรูมร้างยาวเก้าสิบเมตร ในบริเวณของสถาบันวิทยาศาสตร์แห่งเม็กซิโก มีกล่องกระดาษหีบแปดใบหน้าตาเหมือนกันทุกอย่าง ทหารเป็นคนวางกล่องเองทุกใบในตำแหน่งที่ตัวเองเลือก ห่างกันอย่างน้อยหกเมตร

ก่อนเริ่ม ทหารทั้งสองนายเซ็นชื่อในกระดาษแผ่นหนึ่ง ยืนยันว่าห้อง กล่อง และของกลาง ไม่มีอะไรขัดขวางการทำงานของเครื่อง

รอบแรก ทหารนายหนึ่งรู้ว่าของอยู่กล่องไหน เขาเดินไปมาสองสามรอบ แล้วเสาะอากาศก็หันไปหากล่องนั้นซ้ำแล้วซ้ำอีก สี่รอบ ถูกสี่ครั้ง

รอบที่สองคือการทดสอบจริง ทหารทั้งสองนายผลัดกันถือเครื่องเข้าไปค้น ทีมที่ช้อนกับทีมที่หาไม่ได้เจอหน้ากันระหว่างแต่ละรอบ กล่องที่ใช้ช้อนถูกเลือกด้วยการโยนเหรียญสามครั้ง ซึ่งออกผลได้สองคุณสองคุณสอง เท่ากับแปดแบบ พอดีกับกล่องแปดใบ ไม่มีใครในห้องค้นรู้ว่าของอยู่ไหน กล่องสองตัวถ่ายไว้ตลอด ยี่สิบรอบ ถูกสามครั้ง

ถ้าเครื่องนี้ไม่ทำอะไรเลย และทหารแค่เลือกกล่องมั่วๆ ความบังเอิญจะให้ถูกราวสองครั้งครึ่งจากยี่สิบ สามครั้งจึงเป็นผลที่ความบังเอิญให้ได้สบายๆ

นักฟิสิกส์ทั้งสองยังเตรียมของเทียบไว้อีกอย่าง หลังการค้นแต่ละรอบ พวกเขาให้ทหารโยนเหรียญสามครั้งแล้วจดกล่องที่เหรียญเลือก ผลคือเหรียญเลือกถูกหนึ่งครั้งจากยี่สิบ แย่กว่าเครื่องเสียอีก รายงานของพวกเขาบันทึกไว้อย่างเสียดยว่า การเทียบครั้งนี้ "lacked impressiveness" คือไม่น่าประทับใจเลย

ฟังเหมือนเรื่องตลก และมันก็ตลกจริง แต่มันซ่อนบทเรียนไว้อีกหนึ่ง ความบังเอิญไม่ได้ให้ผลเท่ากันทุกครั้ง ในยี่สิบรอบ ของสองชิ้นที่ไม่ทำงานเลยอาจได้หนึ่งหรือได้สามทั้งคู่ และความต่างขนาดนั้นไม่ได้บอกอะไร เครื่องไม่ได้ชนะเหรียญ มันแค่โชคดีกว่าเหรียญในบ่ายวันนั้น

ข้อแก้ตัว

ระหว่างการทดสอบ ทหารผู้ใช้เครื่องให้เหตุผลหลายข้อว่าทำไมมันอาจไม่ทำงาน

บางข้อมาก่อนผลจะออก เขาบอกว่าเครื่องจะหยุดทำงานเมื่อคนถือเหนื่อย บางข้อมาหลังจากรู้ว่าहारобแรกพลาดหมด ของใส่ของกลางมีกลิ่นกัญชาและอาจปนเปื้อนกล่อง ห้องร่อยอยู่คนละชั้น การขึ้นบันไดทำให้หอบ และคู่มือของเครื่องเองก็แนะนำให้ใช้คู่กับสุนัขดมกลิ่น

ทุกข้อฟังขึ้น ไม่มีข้อไหนเป็นเรื่องโกหก และนี่คือเหตุผลที่กระดาศแผนที่เซ็นไว้ก่อนเริ่มมีค่ามากกว่าทุกอย่างในวันนั้น มันถูกเขียนก่อนที่จะใครจะรู้ผล จึงไม่มีใครเขียนมันเพื่อแก้ตัว

ข้อสรุปของนักฟิสิกส์ทั้งสองเขียนไว้ตรงๆ เครื่องนี้ต้องถูกคนถือบังคับให้ชี้ไปยังที่ที่เขาคาดว่าของซ่อนอยู่ แม้ตัวเขาเองอาจไม่รู้ตัวที่กำลังบังคับ ตัวเครื่องไม่ได้ให้ข้อมูลอะไรเลย

เสียดายที่ไม่ได้หาของ มันหาความคาดหวังของคนถือ

เรื่องที่เหลือคือการเก็บกวาด BBC นำเครื่องหนึ่งมาผ่าให้ดู และพบพลาสติกธรรมดาข้างในว่างเปล่า ต้นทุนผลิตราวสองร้อยบาท ปี 2013 ศาลในอังกฤษตัดสินจำคุกเจ้าของบริษัทผู้ผลิตเจ็ดปี ผู้พิพากษากล่าวว่ามันทำงานไม่ดีไปกว่าความบังเอิญ ข่าวการตัดสินรายงานว่าต้นทุนต่อเครื่องไม่ถึงแปดดอลลาร์ ปี 2018 ศาลไทยสั่งจำคุกผู้

บริหารบริษัทที่นำเข้ามาขายเก้าอี้ และระหว่างปี 2021 ถึง 2022 กองทัพบกจ้าง
หน่วยงานวิทยาศาสตร์แห่งชาติตรวจเครื่องที่ตัวเองมีอยู่ 757 เครื่อง เพื่อใช้ในคดีที่
ฟ้องผู้ขาย

ทำไมคนฉลาดจึงเห็นมันทำงาน

ไม่มีใครในเรื่องนี้โง่ ทหารที่ถือมันผ่านการฝึก คนที่เชื่อมั่นมีงบประมาณและที่ปรึกษา
และสามร้อยครั้งที่เจอก็อาจเจอของจริงทุกครั้ง

ปัญหาอยู่ที่สิ่งที่ตาคนเราไม่ได้ถูกสร้างมาให้ทำ ตาเห็นครั้งที่เจอ แต่ไม่ได้นับครั้งที่เดิน
ผ่านไปแล้วไม่มีอะไร และตาไม่มีทางรู้เลยว่าถ้าเดาเอาเฉยๆ จะเจอได้กี่ครั้ง

ตาเห็นเสาอากาศหันไปหาของ แต่ผู้ทดสอบสรุปว่าสิ่งที่พามันหันคือมือของคนถือเอง
โดยที่เขาอาจไม่รู้ตัว ถ้าเป็นอย่างนั้น เมื่อคนถือรู้หรือเดาได้ว่าของน่าจะอยู่ตรงไหน
เสาอากาศก็จะหันไปตรงนั้น และทุกครั้งที่มันหันถูก ความเชื่อก็นักขึ้นอีกหนึ่งครั้ง

สิ่งที่น่าสนใจที่สุดในเรื่องนี้ไม่ใช่การที่เครื่องไม่ทำงาน แต่คือมันไม่จำเป็นต้องทำงาน
เลยก็ได้ ขอแค่มีคนที่ยืนยันถืออยู่ในมือ มันก็ "ได้ผล" ให้เห็นได้ทุกวัน

ประโยค "ลองมาแล้ว ได้ผลจริง" จึงเป็นรายงานจากรอบแรกของการทดสอบ รอบที่
คนลองรู้คำตอบอยู่ในใจแล้ว มันไม่ได้โกหก มันแค่ยังไม่เคยเจอรอบที่สอง

สิ่งที่แยกรอบแรกออกจากรอบที่สองในวันนั้นไม่ใช่เครื่องมือราคาแพง มันคือกล่อง
แปดใบ เหยี่ยวนึงเหยี่ยวนึง ข้อตกลงที่เซ็นไว้ก่อนเริ่ม และกล่องที่ถ่ายไว้ถูกรอบ

ของสี่อย่างนี้คือวิทยาศาสตร์ทั้งหมดในขนาดเล็กที่สุด

จะรู้วาทน้คิดเมื่อ: มีการทดสอบแบบปิดตาที่คนถือไม่รู้ว่าของอยู่ไหน แล้วเครื่อง
ตระกูลเดียวกันนี้หาเจอบ่อยกว่าที่ความบังเอิญให้อย่างชัดเจน หรือมีเอกสารขั้นต้นที่
แสดงว่าตัวเลขสี่จากสี่และสามจากยี่สิบถูกรายงานผิด

หลักฐาน:

- เครื่องตระกูลนี้หาของไม่ได้ดีกว่าความบังเอิญเมื่อคนถือไม่รู้คำตอบ — แน่น — ป้ายนี้ยืนอยู่บนการ
ทดสอบปิดตาสองชั้นในเม็กซิโก (2011) และคำตัดสินของศาลอังกฤษ (2013) ซึ่งให้ผลไปทางเดียวกัน

การทดสอบในไทย (2010) ก็ชี้ไปทางเดียวกัน แต่ไม่ได้นับเป็นฐานของบ้าย เพราะไม่รู้ว่ามีนอกแบบอย่างไร ทุกแหล่งอยู่ในรายการข้างล่าง

- การทดสอบที่เม็กซิโก วันที่ 21 ตุลาคม 2011 · รู้คำตอบ 4 จาก 4 · ปิดตาสองชั้น 3 จาก 20 เมื่อความบังเอิญคาดไว้ 2.5 · เหรียญที่ใช้เทียบ 1 จาก 20 · ข้อตกลงที่ทหารเซ็นก่อนเริ่ม · ข้อแก้ตัวระหว่างทดสอบ · ข้อสรุปว่าเครื่องถูกคนถือบังคับโดยอาจไม่รู้ตัว · เม็กซิโกซื้อ 940 เครื่อง — **พอเชื่อ** (การทดสอบหนึ่งครั้ง บันทึกละเอียดทุกชั้นและมีวิดีโอ) — Mochán, W. L., & Ramírez-Solís, A. (2013). Effectiveness of the GT200 Molecular Detector: A Double-Blind Test. arXiv:1301.3971 — <https://arxiv.org/abs/1301.3971>
- การทดสอบในไทย แถลงวันที่ 16 กุมภาพันธ์ 2010 ทดสอบ 20 ครั้ง ถูก 4 ครั้ง · 836 เครื่อง 759.14 ล้านบาท ตามตัวเลขของสำนักงานการตรวจเงินแผ่นดิน · ผลการผ่าเครื่องที่ BBC นำเสนอ ต้นทุนผลิต 213 บาท · ค่ายืนยันของผู้บัญชาการทหารบกเรื่องการเจอว่าสามร้อยครั้ง · การตรวจ 757 เครื่องในปี 2021 ถึง 2022 — **พอเชื่อ** (รายงานข่าวย้อนหลัง เล่มนี้หาเอกสารการทดสอบฉบับเต็มไม่พบ และแหล่งไม่ได้ระบุวันที่ของคำยืนยัน) — THE STANDARD, 6 มิถุนายน 2022 — <https://thestandard.co/gt200-history/>
- เจ้าของบริษัทผู้ผลิตถูกตัดสินลงโทษในเดือนกรกฎาคม 2013 และจำคุกเจ็ดปี ผู้พิพากษากล่าวว่าเครื่องทำงานไม่ดีกว่าความบังเอิญ · ขาวรายงานว่าต้นทุนต่อเครื่องไม่ถึง 7.85 ดอลลาร์ — **พอเชื่อ** (รายงานข่าวฉบับเดียว) — Al Jazeera, 20 สิงหาคม 2013 — <https://www.aljazeera.com/news/2013/8/20/briton-jailed-for-selling-fake-bomb-detectors>
- ศาลไทยจำคุกผู้บริหารบริษัทผู้จำหน่ายเก่าปี — **พอเชื่อ** (รายงานข่าวฉบับเดียว) — Bangkok Post, 11 ตุลาคม 2018 — <https://www.bangkokpost.com/thailand/general/1555810/man-jailed-over-fake-bomb-detector-fraud>

2 . สิ่งที่มีมนุษย์คิดขึ้นเพื่อจับความผิดของตัวเอง

ผลการทดลองอยู่ในคู่มือครูตั้งแต่ก่อนเข้าห้อง ครูบอกนักเรียนว่า "การทดลองของเธอก็ได้ผลอย่างนั้นะ" ทั้งที่ไม่มีใครได้ลงมือทำ

วิทยาศาสตร์ไม่ได้ทำให้คนฉลาดขึ้น มันทำให้ความผิดของคนฉลาดถูกจับได้

ห้องเรียนแบบนั้นมองวิทยาศาสตร์เป็นคลังคำตอบ หน้าทีของนักเรียนคือรับไปให้ครบ แต่ทหารในบทที่หนึ่งผ่านการฝึกและการรับรองมาแล้ว และความรู้มันไม่ได้กันเขาจากความผิดเลย สิ่งที่จับความผิดได้ไม่ใช่คนที่รู้มากกว่า มันคือห้องที่ถูกจัดจนความผิดไม่มีที่ซ่อน

แล็บแห่ง

ภาพข้างบนมาจากคนในวงการเอง นักเขียนหนังสือวิทยาศาสตร์คนหนึ่งในหน่วยงานวิทยาศาสตร์แห่งชาติเล่าไว้ในบทสัมภาษณ์ และเรียกมันว่าแล็บแห่ง การทดลองมีอยู่แค่ในหนังสือ หนังสือของนักเรียนอาจไม่บอกผล แต่หนังสือของครูบอก

เขาใช้คำว่า เรา และไล่ไปทั้งระบบ ตั้งแต่ห้องที่ไม่มีเงินและไม่มีครูวิทยาศาสตร์เฉพาะทาง ไปจนถึงข้อสอบเข้ามหาวิทยาลัยที่ยังแข่งกันว่าใครจำได้มากกว่า แล้วสรุปว่า "เราสอนเหมือนสอนไสยศาสตร์ เหมือนสอนพระคัมภีร์" คือสอนราวกับว่าเนื้อหาไม่มีวันผิด

เขาเองก็บอกด้วยว่ามีแนวโน้มที่ดีขึ้น มีคนบอกว่าข้อสอบรุ่นใหม่เริ่มวัดความเข้าใจแล้ว

ปัญหาของการสอนราวกับเนื้อหาไม่มีวันผิด ไม่ได้อยู่ที่เนื้อหา เนื้อหาอาจถูกทุกบรรทัด ปัญหาคือมันซ่อนสิ่งที่ทำให้คำตอบเหล่านั้นน่าเชื่อถือตั้งแต่แรก คือมันเคยมีโอกาสผิด แล้วรอดมาได้

ห้าชั้นในหนังสือเรียน

อีกภาพหนึ่งดูเป็นวิทยาศาสตร์กว่า คือห้าชั้นตอน สังเกต ตั้งปัญหา ตั้งสมมติฐาน ทดลอง แล้วสรุปเป็นทฤษฎีหรือกฎ ขาวของคณะวิทยาศาสตร์ มหาวิทยาลัย บัณฑิตไอเรส ปี 2008 บอกว่าโรงเรียนในอาร์เจนตินายังสอนห้าชั้นนี้อยู่ แล้วชวนคิด ต่อ ถ้าถ้ามั่นอย่างเคร่งครัด อาจได้ข้อสรุปแปลกๆ ว่าในโลกจริงไม่มีใครทำงานแบบ วิทยาศาสตร์เลย

นักวิจัยในคณะนั้นเล่าว่าในห้องแล็บไม่มีใครทำตามสูตรที่ท่องมา และการวิจัยไม่ได้มี ทางเดียว เอกสารวางแผนการสอนวิทยาศาสตร์ของสหรัฐฯ หลายฉบับก็เขียนตรงกัน ว่าไม่มีลำดับขั้นตายตัวชุดเดียว

คำถามของครู

ครูในบัณฑิตไอเรสมีข้อแย้งที่ควรฟัง นักวิจัยด้านการสอนวิทยาศาสตร์คนหนึ่งเล่าว่า เมื่อบอกครูว่าวิธีการทางวิทยาศาสตร์แบบห้าชั้นไม่มีอยู่จริง ครูงุนงง เพราะห้าชั้นนั้น คือสิ่งที่พวกเขาใช้แยกวิทยาศาสตร์ออกจากสิ่งที่ไม่ใช่ ถ้าไม่มีสูตร พวกเขาจะแยกได้ อย่างไร

ครูกังวลถูกเรื่อง ถ้าไม่มีเส้นแบ่งเลย ใครจะเรียกอะไรว่าวิทยาศาสตร์ก็ได้ แต่เส้นแบ่ง ไม่จำเป็นต้องเป็นลำดับขั้นตอน เล่มนี้วางเส้นแบ่งนั้นไว้เป็นสามส่วน

พูดให้แพ้ได้ · ทดสอบอย่างยุติธรรม · เปิดให้คนอื่นตรวจ

ทั้งสามส่วนอยู่ครบในห้องบอลรูมร้างในบพที่หนึ่ง

กระดาศที่ทหารเซ็นก่อนเริ่มคือการพูดให้แพ้ได้ เมื่อยืนยันไว้ล่วงหน้าว่าห้อง กล่อง และของกลางไม่มีอะไรขัดขวางเครื่อง ทุกรอบที่พลาดหลังจากนั้นจึงนับเป็นการ พลาดจริง ข้อแก้ตัวที่มาทีหลังไม่มีที่ยืน

กล่องแปดใบที่หน้าตาเหมือนกัน กับเหรียญที่เป็นคนเลือกกล่อง คือการทดสอบอย่าง ยุติธรรม ไม่มีใครในห้องคั่นรู้ว่าของอยู่ไหน ความหวังของใครจึงเอื้อมไปบังคับผลไม่ ถึง

กล้องสองตัวที่ถ่ายถูกรอบ กับรายงานที่จดทุกขั้นตอนไว้ละเอียดพอให้คนอื่นทำซ้ำได้ คือการเปิดให้คนอื่นตรวจ ใครไม่เชื่อนักฟิสิกส์ทั้งสอง ก็จัดการทดสอบแบบเดียวกันเองได้

ไม่มีส่วนไหนต้องใช้ความฉลาด และคุณก็ไม่ต้องเก่งกว่าใครเพื่อใช้มัน สามส่วนนี้ไม่ได้ทำให้ใครรู้มากขึ้น มันแค่ทำให้ความผิดไม่มีที่ซ่อน ไม่ว่าคนที่ผิดจะฉลาดแค่ไหน

และมันไม่ได้บอกว่าต้องทำอะไรก่อนหลัง มันจึงเป็นเกณฑ์ ไม่ใช่สูตร นักดาราศาสตร์ทดลองกับดวงดาวไม่ได้ นักธรณีวิทยาย้อนเวลาไปดูภูเขาเกิดไม่ได้ แต่ข้ออ้างของพวกเขาต้องผ่านสามส่วนเดียวกัน ถ้าคุณอยากรู้ว่าอะไรเป็นวิทยาศาสตร์ ให้ถามถึงสามส่วนนี้ ไม่ใช่ถามว่ามันทำครบห้าขั้นหรือยัง

อย่าได้ยึดถือ

ความคิดว่าคำสอนต้องผ่านการตรวจก่อนจะยึดถือ ไม่ได้เกิดในห้องแล็บ ในกาลามสูตร ชาวกาลามะกราบทูลพระพุทธเจ้าว่ามีนักบวชหลายพวกผ่านมา แต่ละพวกยกแต่คำสอนของตัวเองและดูหมิ่นของคนอื่น จนไม่รู้ว่ามีใครพูดจริง ใครพูดเท็จ ถ้าเปลี่ยนนักบวชเป็นคลิปสั้น กลุ่มไลน์ และ AI คำถามนี้ก็ยังใหม่อยู่

คำตอบในพระไตรปิฎกฉบับหลวงเริ่มด้วยการรับรองว่าความสงสัยของพวกเขาควรแล้ว จากนั้นไล่เหตุผลข้อที่ไม่ควรใช้เป็นหลัก ทุกข้อขึ้นต้นด้วยคำเดียวกัน อย่าได้ยึดถือ

ถ้อยคำที่ได้ยินได้ฟังมา · สืบๆ กันมา · ตื่นข่าว · อ้างตำรา · เดาเอาเอง · คาดคะเน · ตรึกตามอาการ · ต้องกันกับทิวฐิของตัว · ผู้พูดสมควรจะเชื่อได้ · สมณะนี้เป็นครูของเรา

ถ้าคุณคุ้นกับสำนวน อย่าปลงใจเชื่อ คำนั้นไม่ได้มาจากฉบับหลวง

สิ่งที่พระสูตรให้ใช้แทนมีสามขั้น เมื่อใดที่ "รู้ด้วยตนเอง" ว่าสิ่งใดเป็นอกุศล มีโทษ เป็นสิ่งที่ "ผู้รู้ดีเตือน" และเมื่อทำจนเต็มที่แล้ว "เป็นไปเพื่อสิ่งไม่เป็นประโยชน์ เพื่อทุกข์" เมื่อนั้นจึงควรละ ถ้าตรงกันข้ามก็ควรเข้าถึง รู้เอง ฟังผู้รู้ และดูผลของการลงมือทำ

พระภิกษุสองรูปที่แปลพระสูตรนี้เป็นภาษาอังกฤษเตือนตรงกันว่า มันไม่ใช่ใบอนุญาตให้คิดเอาเองตามใจ รูปหนึ่งอ่านว่าความเห็นยังต้องถูกทดสอบด้วยผลของการปฏิบัติ และเทียบกับผู้รู้ อีกรูปอ่านว่ามันคือจุดเริ่มที่ตรวจสอบได้

ถ้าจะมีอะไรที่วิทยาศาสตร์เติมเข้าไป ก็คือบทเรียนจากห้องบอลรูม การรู้ด้วยตนเอง มือของตัวเองบังคับได้ ทหารคนนั้นเห็นกับตาจริงๆ ว่าเสาอากาศหันไปหากล่อง เขาไม่ได้โกหก และอาจไม่รู้ตัวเลยว่ามือของเขาเองเป็นคนเอียงมัน

วิทยาศาสตร์จึงไม่ได้ขอให้คุณเลิกรู้ด้วยตนเอง มันขอให้จัดการทดสอบ จนสิ่งที่คุณรู้ อยู่ในใจเอี่ยมไปบังคับผลไม่ถึง

ราวพันปีก่อนในโคโร นักคณิตศาสตร์ชื่ออิบน์ อัลฮัยษัมเขียนงานชื่อ ข้อเสนอต่อบโตเลมี บโตเลมีคือปราชญ์กรีกโบราณ เจ้าของตำราดาราศาสตร์ที่ยังถูกยึดเป็นหลัก ในยุคนั้น อิบน์ อัลฮัยษัมนับถือเขามาก ถึงอย่างนั้นก็เขียนว่าคนแสวงหาความจริงไม่ใช่คนที่อ่านงานของคนโบราณแล้วไว้ใจ แต่คือคนที่ "suspects his faith in them and questions what he gathers from them" คือระแวงศรัทธาที่ตัวเองมีต่อพวกเขา และตั้งคำถามกับทุกอย่างที่ได้มา แล้วเขาก็เขียนต่อ ให้ทำตัวเป็นศัตรูกับทุกอย่างที่อ่าน และระแวงตัวเองไปพร้อมกัน

หลักข้อนี้จึงไม่ได้เป็นของทวีปไหน คำเตือนว่าอย่ายึดถือเพียงเพราะเป็นคำของครู ให้ระแวงศรัทธาที่มีต่อคนโบราณ และให้หลักฐานเป็นผู้ตัดสิน ถูกพูดไว้ในแคว้นโกศล ในโคโร และในยุโรป ในเวลาที่ห่างกันรวมกว่าสองพันปี

คู่มือครู

หลักสูตรวิทยาศาสตร์ของไทยไม่ได้สั่งให้สอนแบบแกล้งแหย้ คู่มือหลักสูตรฉบับปี 2017 พูดถึงทักษะกระบวนการที่ได้มา "ผ่านการสังเกต ทดลอง สร้างแบบจำลอง และวิธีการอื่น ๆ" ถ้อยคำนั้นเปิดไว้ให้หลายวิธี

เป้าหมายข้อหนึ่งของหลักสูตรคือให้ผู้เรียน "เชื่อมั่นและศรัทธา" ในความรู้และกระบวนการทางวิทยาศาสตร์ คำนี้ฟังแปลกเล็กน้อยในวิชาที่มีคนอย่างอิบน์ อัลฮัยษัม

บอกให้ระแวงศรัทธาของตัวเอง แต่มันมีความหมายที่ดีได้ ถ้าศรัทธานั้นวางไว้ที่วิธีจับความผิด ไม่ใช่ที่คำตอบ ความเชื่อมั่นแบบวิทยาศาสตร์มาหลังการตรวจสอบ

ในห้องแบบที่นักเขียนคนเดิมบรรยาย การตรวจไม่เคยเกิดขึ้น และเขาเล่าถึงสิ่งที่ตามมาด้วย พอนักเรียนที่ไม่ได้ฝึกแบบนี้ได้ทดลองจริงในมหาวิทยาลัย แล้วผลไม่ตรงทฤษฎี ก็พยายามเปลี่ยนผลให้ตรงทฤษฎี ซึ่งในวิทยาศาสตร์คือข้อต้องห้าม

นั่นไม่ใช่ความไม่เชื่อเสมอไป ถ้าคำเล่านี้ตรง มันอาจเป็นบทเรียนที่ห้องเรียนแบบนั้นสอนไว้โดยไม่ได้ตั้งใจ ว่าเมื่อผลกับหนังสือขัดกัน หนังสือไม่เคยเป็นฝ่ายผิด เล่มนี้ไม่มีตัวเลขว่าห้องเรียนไทยเป็นแบบนี้กี่ห้อง ภาพห้องเรียนนี้ทั้งหมดมาจากคำเล่าของคนคนเดียว

ห้องที่อ่านผลจากคู่มือครูสอนความรู้ได้ครบทุกบท แต่สอนสามส่วนนั้นไม่ได้ เพราะเมื่อผลถูกเขียนไว้ก่อนลงมือ ก็ไม่มีอะไรแพ้ได้

จะรู้วาทน์ผิดเมื่อ: กรอบการสอนวิทยาศาสตร์ฉบับหลักของประเทศใดประเทศหนึ่ง หรือขององค์กรนานาชาติ รับรองว่าวิทยาศาสตร์มีขั้นตอนตายตัวชุดเดียวที่นักวิทยาศาสตร์ทุกคนทำตาม หรือการสำรวจห้องแล็บที่ทำงานจริงพบว่าทุกแห่งเดินตามห้าขั้นตอนเดียวกัน

หลักฐาน:

- วิทยาศาสตร์ไม่มีลำดับขั้นตอนตายตัวชุดเดียวที่นักวิทยาศาสตร์ทุกคนทำตาม — **แนน** — เอกสารวางแผนการสอนของสหรัฐฯ สามฉบับ (1989, 2012, 2013) และนักวิจัยในบัวโนสไอเรส (2008) พุดตรงกันทั้งหมดอยู่ในรายการข้างล่าง
- ไม่มีชุดขั้นตอนตายตัวที่นักวิทยาศาสตร์ทำตามเสมอ — **แนน** — American Association for the Advancement of Science, Science for All Americans, 1989 — <https://www.project2061.org/publications/sfaa/online/chap1.htm>
- ไม่ควรมองการทำงานของนักวิทยาศาสตร์เป็นขั้นตอนเรียงเส้นตรง · การสอนผ่านสิ่งที่นักวิทยาศาสตร์ทำงานจริงช่วยให้เข้าใจผิดว่ามีวิธีเดียวที่วิทยาศาสตร์ทุกสาขาใช้ร่วมกัน — **แนน** — National Research Council, A Framework for K-12 Science Education, 2012 — <https://nap.nationalacademies.org/read/13165/chapter/7>
- งานสืบค้นทางวิทยาศาสตร์ใช้วิธีหลากหลาย — **แนน** — NGSS, Appendix H: The Nature of Science, 2013 — <https://www.nextgenscience.org/sites/default/files/Appendix%20H%20-%20The%20Nature%20of%20Science%20in%20the%20Next%20Generation%20Science%20Standards%20-%20Appendix%20H.pdf>

- โรงเรียนในอาร์เจนตินายังสอนวิธีการทางวิทยาศาสตร์เป็นห้าขั้น . ถ้าถือห้าขั้นนี้จริงจัง อาจสรุปได้ว่าไม่มีใครทำงานแบบวิทยาศาสตร์ . "En el laboratorio no se sigue una receta aprendida" และ "No hay una manera única de investigar" (นักวิจัยสองคนของคุณ) . ครุ่นงเมื่อได้ยินว่าวิธีการนั้นไม่มีอยู่จริง เพราะใช้มันแยกวิทยาศาสตร์ออกจากสิ่งที่ไม่ใช่ — **พอเชื้อ** (ข่าวของมหาวิทยาลัยที่สัมภาษณ์นักวิจัยในประเทศเดียว เล่มนี้เปิดอ่านเองวันที่ 25 กันยายน 2026) — Gabriel Stekolschik, Un método que resultó ser un mito, คณะวิทยาศาสตร์ มหาวิทยาลัยบัวโนสไอเรส (ตีพิมพ์ครั้งแรกใน La Nación), 20 พฤศจิกายน 2008 — <https://exactas.uba.ar/un-metodo-que-resulto-ser-un-mito/>
- แล็บแห่ง ผลการทดลองอยู่ในหนังสือของครู . "เราสอนเหมือนสอนไสยศาสตร์ เหมือนสอนพระคัมภีร์" . สาเหตุคือห้องที่ไม่มีเงินและไม่มีครูเฉพาะทาง และข้อสอบเข้ามหาวิทยาลัยที่เน้นความจำ . มีแนวโน้มดีขึ้น มีคนบอกว่าข้อสอบรุ่นใหม่วัดความเข้าใจ . ผู้ที่ไม่ได้ฝึกกระบวนการ พอผลไม่ตรงทฤษฎีก็พยายามเปลี่ยนผล — **ยังบาง** (ประสบการณ์และความเห็นของคนคนเดียว เล่มนี้เปิดอ่านเองวันที่ 25 กันยายน 2026) — The Matter, 27 มิถุนายน 2024, บทสัมภาษณ์นำชัย ชีววิวรรณ์ — <https://thematter.co/science-tech/science-education-namchai-chewawiwat/228091>
- ทักษะกระบวนการ "ผ่านการสังเกต ทดลอง สร้างแบบจำลอง และวิธีการอื่น ๆ" . เป้าหมายให้ผู้เรียน "เชื่อมั่นและศรัทธา" ในความรู้และกระบวนการทางวิทยาศาสตร์ — **แนน** . คู่มือหลักสูตรวิทยาศาสตร์ ฉบับปี 2017 ที่ scimath.org — <https://www.scimath.org/e-books/8415/flippingbook/13/> และ <https://www.scimath.org/e-books/8415/flippingbook/34/>
- ชาวกาลามะถามว่าใครพูดจริง ใครพูดเท็จ . ความสงสัยนั้นควรแล้ว . อย่าได้ยึดถือ สิบข้อ . รู้ด้วยตนเองว่าเป็นอกุศลหรือกุศล ผู้รู้ดีเตียนหรือสรรเสริญ และเมื่อสมาทานให้บริบูรณ์แล้วเป็นไปเพื่ออะไร . สำนวน อย่าปลงใจเชื่อ ไม่ใช่ถ้อยคำของฉบับหลวง — **แนน** (ตัวบทขั้นต้น เล่มนี้เปิดอ่านเองวันที่ 25 กันยายน 2026) — เกสพุตสูตร (กาลามสูตร) พระไตรปิฎกฉบับหลวง เล่มที่ 20 ข้อ 505 หน้า 212 ถึง 218 — https://84000.org/tipitaka/pitaka_item/v.php?B=20&A=4930&Z=5092
- พระสูตรนี้ไม่ใช่ใบอนุญาตให้คิดเอาเอง ความเห็นต้องทดสอบด้วยผลของการปฏิบัติและเทียบกับผู้รู้ . มันคือจุดเริ่มที่ตรวจสอบได้ ไม่ใช่ประกาศเสรีภาพทางความคิดแบบไร้ขอบ — **พอเชื้อ** (การตีความของผู้แปลสองท่าน) — Thanissaro Bhikkhu, AN 3.65 — <https://www.accesstoinight.org/tipitaka/an/an03/an03.065.than.html> ; Bhikkhu Bodhi, A Look at the Kalama Sutta — https://www.accesstoinight.org/lib/authors/bodhi/bps-essay_09.html
- ใน Aporias against Ptolemy ผู้แสวงหาความจริงคือผู้ที่ "suspects his faith in them and questions what he gathers from them" . ประโยคถัดไปให้ทำตัวเป็นศัตรูกับทุกอย่างที่อ่านและโจมตีจากทุกด้าน . อีกประโยคให้ระแวงตัวเองระหว่างตรวจ เพื่อไม่ให้ลำเอียงหรือหะหลวม . เขานับถือปโตเลมีมาก และใช้ชีวิตช่วงหลังในไคโร . ปโตเลมีเป็นผู้เขียน Almagest หนึ่งในตำราคณิตศาสตร์กรีกคลาสสิกที่เขาคัดลอกฉบับแปลภาษาอาหรับหาเลี้ยงชีพ และถูกยกเป็นตัวอย่างของ "scientific authorities" ที่ไม่พ้นความผิด — **พอเชื้อ** (คำแปลภาษาอังกฤษของนักประวัติศาสตร์วิทยาศาสตร์คนเดียว เล่มนี้เปิดอ่านเองวันที่ 25 และ 26 กันยายน 2026) — A. I. Sabra, Ibn al-Haytham, Harvard Magazine, กันยายน ถึง ตุลาคม 2003 — <https://www.harvardmagazine.com/2003/09/ibn-al-haytham.html>

3 . ภูมิปัญญาถาม การทดสอบตอบ

ตำรายาจีนเล่มหนึ่งอายุกว่าหนึ่งพันหกร้อยปี บอกวิธีแก้ไข้จับสั่นในบรรทัดเดียว ซึ่งเขาหนึ่งกำมือ แขน้ำ บิดเอาน้ำ ต้มให้หมด

ประโยคที่สำคัญที่สุดในตำราเล่มนั้น คือคำที่มันไม่ได้เขียนไว้ในสูตรไม่มีคำว่าต้ม

สูตรนั้นบอกคร่าวๆ ใช้เท่าไร ทำอย่างไร กินอย่างไร แต่ขาดขั้นที่สมุนไพรจีนส่วนใหญ่ต้องผ่านก่อนถึงมือคนไข้ คือการต้มเป็นยาต้ม แบบเดียวกับยาหม้อ

คำที่หายไปนั้นพาไปถึงรางวัลโนเบล แต่ไม่ได้พาไปเอง ระหว่างตำรากับยา มีการทดสอบค้นอยู่ ซึ่งถามข้อเดียว มันทำได้อย่างที่อ้างไหม เมื่อเทียบอย่างยุติธรรม คำถามเดียวกันนี้เคยถูกถามกับอีกสองเรื่องที่คนไทยใช้ และสามเรื่องได้คำตอบไม่ซ้ำกันเลย

ตัวอย่างหมายเลข 191

คนที่เห็นคำที่หายไป คือนักวิทยาศาสตร์หญิงชาวจีนชื่อ ลวโยว งานของเธอคือหา ยาต้านมาลาเรียจากการแพทย์แผนจีน มาลาเรียก็คือไข้จับสั่นในบรรทัดนั่นเอง

เธอเริ่มจากฟังภูมิปัญญาให้ครบ ภายในสามเดือนเธอรวบรวมตำรับยาได้กว่าสองพันตำรับ แล้วคัดเหลือ 640

ซึ่งเธอเป็นหนึ่งในสมุนไพรที่เธอทดลอง แต่ผลไม่น่าพอใจอยู่หลายรอบ เธอจึงกลับไปเปิดตำราอีกครั้ง และหยุดครุ่นคิดอยู่ที่บรรทัดนั้น

ตำราเล่มนั้นคือคู่มือยาสำหรับเหตุฉุกเฉิน จากราชวงศ์จิ้นตะวันออก ราวปี 317 ถึง 420 บรรทัดนั้นทำให้เธอเกิดความคิดข้อหนึ่ง ความร้อนตอนสกัดอาจทำลายฤทธิ์ยาไปก็ได้

เธอจึงออกแบบการสกัดใหม่ให้ใช้อุณหภูมิต่ำกว่าการต้ม โดยแข่งชิงเขาในอีเทอร์ วันที่ 4 ตุลาคม 1971 ตัวอย่างสกัดหมายเลข 191 ถูกป้อนให้หนูที่ติดเชื้อมาลาเรีย และได้ผลเต็มร้อย ไม่กี่เดือนต่อมาก็ได้ผลแบบเดียวกันในลิง

สิ่งที่ตามมาคือสารที่ถูกตั้งชื่อว่าอาร์เทมิซินิน ยารักษามาลาเรียตัวใหม่ และรางวัลโนเบลสาขาสรีรวิทยาหรือการแพทย์ปี 2015

ต้นเดียวกันนี้มีชื่ออยู่ในตำรายาไทยด้วย คือโกฐจุฬาลัมพา หนึ่งในโกฐทั้งเก้า

ถ้าคุณจะเล่าเรื่องนี้ต่อ ตรงนี้ต้องอ่านให้ช้า ตำราไม่ได้บอกให้ใช้น้ำเย็น และไม่ได้เตือนเรื่องความร้อน มันแค่ไม่ได้บอกให้ต้ม ความคิดเรื่องความร้อนเป็นของเธอ ตำราเปิดช่องไว้ แต่คนที่เดินเข้าไปคือเธอ

เธอไม่ได้เชื่อตำรา และไม่ได้ไม่เชื่อ เธออ่านมันอย่างคนอ่านคำใบ้ แล้วเปลี่ยนคำใบ้ให้เป็นการทดลองที่ผิดได้ คำถามมาจากตำรา การเดาจากเธอ และคำตอบมาจากหนู และถึงในห้องทดลอง

เธอตีชื่อปาฐกถารับรางวัลว่า "Artemisinin—A Gift from Traditional Chinese Medicine to the World" อาร์เทมิซินิน ของขวัญจากการแพทย์แผนจีนถึงโลก ของขวัญนั้นมีอยู่จริง และมันมาในรูปของคำถาม

คำถามใหม่ของยาเก่า

ฟ้าทะลายโจรเป็นสมุนไพรที่การแพทย์แผนไทยใช้แก้ไข้และเจ็บคอ เมื่อโควิด-19 มาถึง มันถูกถามคำถามใหม่ ช่วยคนไข้โควิดได้ไหม

งานวิจัยก่อนหน้านี้ชี้ว่าสารสำคัญในต้นลดการอักเสบ และอาจขวางไวรัสไม่ให้เข้าเซลล์ หรือเพิ่มจำนวน การทดลองในจีนกับยาฉีดที่สกัดจากต้นเดียวกันก็รายงานไว้ไข้และไอลดลง ความหวังนี้มีฐาน

ถ้าคุณหรือคนในบ้านเคยกินฟ้าทะลายโจรตอนติดโควิดแล้วหายดี ประสบการณ์นั้นเป็นของจริง มันแค่ตอบไม่ได้ว่าถ้าไม่ได้กิน จะหายช้ากว่าหรือเปล่า

การทดลองสองชิ้นถามเรื่องนี้ตรงกว่านั้น ทั้งสองชิ้นให้ยาต้านไวรัสฟาวิพิราเวียร์กับคนไข้ทุกคน แล้วเติมฟ้าทะลายโจรให้ครั้งเดียว เพื่อถามว่าสมุนไพรช่วยเพิ่มจากยาที่ได้อยู่แล้วหรือเปล่า

ขึ้นแรกปิดตาสองชั้น (double-blind) ครั้งที่ไม่ได้สมุนไพรมะยาหลอก (placebo) แทน ถ้าคุณเป็นคนใช้ในการทดลองนี้ คุณจะไม่ว่ายาที่กินเป็นของจริงหรือของหลอก และ หมอที่ดูแลคุณก็ไม่ว

ผลคือไม่วพบว่าฟ้าทะลายโจรช่วยเพิ่มอะไร ทั้งต่ออาการป่วยและต่อไวรัส

ขึ้นที่สองทำในโรงพยาบาลจังหวัดสองแห่งในไทย กับคนไข้การน้อย 82 คน เป็นการทดลองแบบเปิด ทุกคนรู้ว่าใครได้อะไร ตัววัดหลักที่ตั้งไว้คือปอดอักเสบรุนแรง และไม่มีใครเป็นเลยทั้งสองกลุ่ม คำถามหลักจึงไม่ได้คำตอบ เพราะไม่มีอะไรให้เทียบ

สิ่งที่ต่างคืออาการไอ วันที่เจ็ด กลุ่มที่ได้สมุนไพรมะยาสามในสิบคน กลุ่มที่ไม่ได้ยัง ไอเกินครึ่ง แต่ค่าเอนไซม์ตับของกลุ่มที่ได้สมุนไพรมะยาสูงกว่า ผู้วิจัยเรียกว่าสูงขึ้น เล็กน้อย และค่า ALT ของกลุ่มนี้ก็สูงกว่าอีกกลุ่มอยู่แล้วตั้งแต่วก่อนเริ่ม

ผลเรื่องไอยังต้องอ่านอย่างระวัง การรู้ว่าตัวเองกินอะไรอยู่เอื้อมไปถึงอาการอย่างไ ได้ อย่างที่บทที่หนึ่งแสดงไว้ และกลุ่มที่ได้สมุนไพรมะยาอายุน้อยกว่าอีกกลุ่มตั้งแต่วแรก

เท่าที่สองขึ้นนี้บอก สำหรับโควิด-19 ฟ้าทะลายโจรไม่ได้ให้สิ่งที่หวังไว้ คำตอบนี้ไม่ได้ ตัดสินการใช้แก้ไขหรือเจ็บคอ ซึ่งเป็นอีกคำถามหนึ่ง

ซึ่งเขาเกี่ยวกับฟ้าทะลายโจรถูกถามคำถามเดียวกัน ซึ่งเขาตอบผ่านสารสกัดที่ให้หนูแล้วให้ ลิง ฟ้าทะลายโจรตอบผ่านการทดลองในคน ทั้งสองแบบยอมให้ผลออกมาทางไหนก็ได้ ต่างกันที่คำตอบ

ลูกบอลในน้ำท่วม

ปลายปี 2011 ระหว่างน้ำท่วมใหญ่ รายการโทรทัศน์รายการหนึ่งเชิญผู้เชี่ยวชาญมา ถกกันด้วยคำถามข้อเดียว ลูกบอล EM แก่น้ำเน่าให้เป็นน้ำดื่มได้ไหม

ลูกบอล EM คือก้อนที่เก็บจุลินทรีย์หลายกลุ่มไว้ ตัว EM เองไม่ใช่ของเก่า ผู้พัฒนาเป็น อาจารย์มหาวิทยาลัยในโอกินาวา ญี่ปุ่น องค์การของเขาเล่าว่ามันเริ่มจากหญ้าแปลง หนึ่งที่งามกว่ารอบข้าง หลังเขาโรยจุลินทรีย์ที่เหลือจากการทดลองลงไป และตั้งชื่อ ส่วนผสมนี้ในปี 1980

หญ้าแปลงนั้นก็เหมือนบรรทัดในตำราจีน เป็นจุดเริ่มของคำถาม ยังไม่ใช่คำตอบ

ฝ่ายที่เชื่อมีเหตุผลที่ฟังขึ้น คณะบดีคณะเทคโนโลยีชีวภาพคนหนึ่ง ผู้เชี่ยวชาญด้านจุลชีววิทยา บอกว่าจากที่ลงไปทำเอง เขาเชื่อว่าลูกบอล EM แก่น้ำเสียได้ ขึ้นกับวิธีใช้ และเทียบมันกับลูกแบงก์ที่ใช้ทำสาโท ก่อนที่เก็บเชื้อไว้และได้ผลดี อีกคนที่ริเริ่มใช้ลูกบอล EM รุ่นแรกๆ ยกน้ำหนักหลายแห่งที่เขาบอกว่าแก้สำเร็จ

ภาพลูกแบงก์มีน้ำหนักจริง ข้าวเหนียวในไหกลายเป็นสาโทก็เพราะราและยีสต์ในก้อนเล็กๆ นั้น

อีกฝั่งเห็นต่าง อาจารย์ด้านวิศวกรรมสิ่งแวดล้อมจากจุฬาฯ บอกว่าโอกาสที่ EM จะบำบัดน้ำเสียมีน้อยมาก ลูกบอลเพิ่มจุลินทรีย์ได้เพียงเล็กน้อย ถ้าจะให้ได้ผลต้องโกยของเสียขึ้นมา แต่เขาเห็นด้วยกับการใช้ EM ในการเกษตร นักเทคโนโลยีชีวภาพจากกองทุนสนับสนุนการวิจัยในรายการเดียวกันก็เห็นว่าลูกบอล EM ไม่น่าจะแก่น้ำเสียได้

ทั้งสองฝั่งพูดจากความรู้จริง และทั้งสองฝั่งยังเป็นแค่ข้ออ้าง คำถามคือลูกบอลที่ใส่ลงไปทำให้หน้าตาต่างไปจากที่มันจะเป็นเองหรือเปล่า

การถกกันรายการตอบเรื่องนี้ไม่ได้ เรื่องเล่าความสำเร็จก็ตอบไม่ได้ ต่อให้จริงทุกครั้ง ก็ไม่มีครั้งไหนบอกว่าถ้าไม่ใส่ น้ำจะเป็นอย่างไร

คำตัดสินที่ละเอียดที่สุดที่เล่มนี้ทำได้มาจากญี่ปุ่น บ้านเกิดของ EM เอง ห้องปฏิบัติการสื่อสารวิทยาศาสตร์ของมหาวิทยาลัยเมจิทำเว็บไซต์ให้คำแนะนำข้ออ้างทางวิทยาศาสตร์ ด้วยเกณฑ์สิบข้อในสี่หมวด เว็บไซต์นี้ดูงานวิจัยเรื่อง EM ทั้งในและนอกญี่ปุ่น แล้วเขียนว่าผลของ EM อย่างมากก็เท่ากับวัสดุจุลินทรีย์อื่นที่ใช้ในการเกษตร EM ได้ระดับต่ำทั้งสิบเกณฑ์ และถูกจัดเป็นวิทยาศาสตร์เทียม

แต่เว็บไซต์เดียวกันไม่ได้ปิดทุกอย่างทิ้ง มันเขียนว่า EM น่าจะช่วยดับกลิ่นขยะเศษอาหาร และทำให้คราบสกปรกหลุดง่ายขึ้น

ถ้าคุณอยู่ฝั่งที่เชื่อ ข้อสรุปนี้ฟังเหมือนแพ้ ถ้าอยู่ฝั่งที่ไม่เชื่อ ข้อที่ว่าดับกลิ่นได้ก็ฟังเหมือนแพ้ คำตอบแบบที่สามไม่เข้าข้างใคร ผลเล็กๆ น่าจะมีจริง ส่วนข้ออ้างใหญ่เรื่องทำน้ำเน่าให้เป็นน้ำดี การทดสอบยังไม่รองรับ

ทดสอบคือความเคารพ

สามเรื่อง คำถามเดียวกัน ได้สามคำตอบ ซึ่งเขาได้คำว่าใช่ ฟ้าทะลายโจรกับโควิด-19 ยังไม่พบประโยชน์ที่เพิ่มขึ้นในการทดลองเท่าที่มี และลูกบอล EM ได้คำตอบที่เล็กกว่าที่ถูกพูดถึงกลางน้ำท่วม

ความเชื่อสองข้อที่หน้าต้นเล่มบอกไว้ว่าผิดพอกัน ตายในบพนี้พร้อมกัน

ข้อแรกคือ ของที่ใช้กันมาหลายร้อยปีต้องได้ผล การใช้กันมานานไม่ใช่การทดสอบ ตำราไม่ได้บอกชัดด้วยซ้ำว่าซึ่งเขาคือต้นไหน ส่วนไหน ทีมของยู โยวโยวต้องทดสอบเองจึงรู้ว่า ในพืชสกุลเดียวกันหลายชนิด มีเพียงชนิดเดียวที่มีอาร์เทมิซินินมากพอ

ฟ้าทะลายโจรมีที่ยืนในการแพทย์แผนไทย แต่ที่ยืนนั้นตอบเรื่องโควิดแทนการทดสอบไม่ได้ ส่วน EM ไม่ได้เก่าเลย แต่ก็ถูกเชื่อดด้วยเหตุผลเดียวกับของเก่า คือมีคนใช้แล้วเห็นผล

ข้อที่สองคือ ของโบราณเชื่อไม่ได้ บรรทัดเดียวในตำราอายุกว่าหนึ่งพันหกร้อยปีพาไปถึงยาตัวใหม่และรางวัลโนเบล คนที่ปิดตำราเล่มนั้นเพราะมันเก่า จะไม่มีวันได้เห็นคำที่หายไป

การทดสอบจึงไม่ใช่การดูถูกภูมิปัญญา มันคือการจริงจังกับภูมิปัญญาพอจะยอมให้มันชนะ

ในทั้งสามเรื่อง คนที่ลงมือตรวจคือคนในบ้านเกิดของเรื่องนั่นเอง ในจีน ในไทย และในญี่ปุ่น การทดสอบไม่ใช่อาวุธที่ฝ่ายหนึ่งถือไว้ใช้กับอีกฝ่าย มันใช้ได้กับทุกบ้าน รวมทั้งบ้านของคุณ

สามส่วนในบทที่สองไม่มีส่วนไหนถามว่าความรู้ที่นั่นอายุเท่าไร หรือมาจากประเทศไหน สิ่งที่คุณถามภูมิปัญญาทุกชิ้นได้จึงไม่ใช่ว่ามันเก่าแค่ไหน แต่คือมันเคยเข้าการทดสอบที่แพ้ได้หรือยัง

ภูมิปัญญาไม่ได้แพ้วิทยาศาสตร์ในบทนี้ มันเป็นฝ่ายถาม และคำถามข้อไหนถูก มีแต่การทดสอบที่บอกได้

จะรู้ว่บทนี้คิดเมื่อ: มีการทดลองแบบปิดตาสองชั้นที่ใหญ่กว่า หรืองานทบทวนที่รวมหลายการทดลอง พบว่าฟ้าทะลายโจรช่วยคนไข้โควิด-19 ได้ชัดเจน หรือมีการทดลองที่เทียบน้ำที่ใส่ลูกบอล EM กับน้ำที่ไม่ใส่ แล้วพบว่าน้ำที่ใส่ดีขึ้นชัดเจน และดีกว่าการใส่จุลินทรีย์ชนิดอื่น หรือมีคนเปิดตำราต้นฉบับแล้วพบว่าสูตรชิงเฮาระบุนกภูมิหรือสั่งให้ต้มไว้จริง

หลักฐาน:

- ตำรายาฉุกเฉินของ Ge Hong (ราชวงศ์จิ้นตะวันออก ราวปี 317 ถึง 420) บันทึกการใช้ชิงเฮาบรรเทาไข้เป็นรอบของมาลาเรีย · ตัวบทภาษาจีนที่เรอยกมาแปลว่าชิงเฮาหนึ่งกำมือ แขน้ำสองเซิง (หน่วยตวงของจีน) บิดเอาน้ำ คั้นให้หมด · ปาฐกถาแปลสองเซิงเป็น two liters เล่มนี้ไม่ได้ตรวจการแปลงหน่วยนี้ จึงไม่ได้พิมพ์ปริมาณน้ำไว้ในเนื้อเรื่อง · ไม่มีคำบอกอุณหภูมิ · ปาฐกถาเรียกวิธีนี้ว่า "cold Qinghao" และบอกว่าสมุนไพรส่วนมากถูกต้มเป็นยาต้มก่อนให้คนไข้กิน · ความคิดว่าอาจต้องหลบความร้อนตอนสกัดเป็นของผู้บรรยายเอง — **พอเชื้อ** (คำเล่าของผู้ค้นพบเองในเอกสารชิ้นเดียว แต่ยกตัวบทจีนไว้ให้ตรวจ เล่มนี้เปิดอ่านซ้ำวันที่ 26 กันยายน 2026) — Tu Youyou, Artemisinin—A Gift from Traditional Chinese Medicine to the World, Nobel Lecture, 7 ธันวาคม 2015 — <https://www.nobelprize.org/uploads/2018/06/tu-lecture.pdf>
- ภายในสามเดือนผู้บรรยายรวบรวมตำรับยากว่าสองพันตำรับ จากตำราแพทย์โบราณ สูตรพื้นบ้าน และการสัมภาษณ์หมอจีนที่มีประสบการณ์ แล้วคัดเหลือ 640 · ทดลองกับชิงเฮาในปี 1971 หลายครั้งโดยไม่ได้ผลที่น่าพอใจ และกลับไปอ่านตำราหลังล้มเหลวหลายรอบ · กันยายน 1971 ลดอุณหภูมิการสกัดด้วยการแช่หรือกลั่นชิงเฮาในเอทิลเอเทอร์ · 4 ตุลาคม 1971 ตัวอย่างหมายเลข 191 ได้ผล 100% กับมาลาเรียในสัตว์ฟันแทะเมื่อให้ทางปาก · ธันวาคม 1971 ถึงมกราคม 1972 ได้ผลแบบเดียวกันในลิงที่ติดเชื้อ · ตำราแพทย์จีนให้รายละเอียดน้อยเรื่องชนิดพืชและส่วนที่ใช้ · ในพืชสกุล Artemisia ที่ทีมทดสอบ มีเพียง Artemisia annua L. ที่มีอาร์เทมิซินินในปริมาณที่มีความหมาย — **พอเชื้อ** (เอกสารเดียวกัน พร้อมสำเนาบันทึกการทดลอง) — แหล่งเดียวกับบรรทัดบน — <https://www.nobelprize.org/uploads/2018/06/tu-lecture.pdf>
- ญ โยวโยว (Tu Youyou) ได้รับรางวัลโนเบลสาขาสรีรวิทยาหรือการแพทย์ ปี 2015 — **แน่น** (บันทึกบนเว็บไซต์ของมูลนิธิโนเบลเอง) — Nobel Lecture, 2015 — <https://www.nobelprize.org/uploads/2018/06/tu-lecture.pdf>
- โกรธจุฬาลัมพาในตำรายาไทยคือ Artemisia annua ชื่อจีนกลาง ชิงเฮา · อยู่ในพิกัดโกรธทั้งเก้า (เนาวโกรธ) · ใช้จับสันคือมาลาเรีย · ฐานข้อมูลเดียวกันเดือนว่ามีพืชหลายชนิดที่ถูกเข้าใจว่าเป็นเครื่องยาชื่อนี้ — **พอเชื้อ** (ฐานข้อมูลสมุนไพรของคณะเภสัชศาสตร์แห่งเดียว เล่มนี้เปิดอ่านวันที่ 26 กันยายน 2026) — Smart Herb Platform, คณะเภสัชศาสตร์ มหาวิทยาลัยขอนแก่น, หน้าโกรธจุฬาลัมพา — <https://pharmoffice.kku.ac.th/iss/smhcontrol/herb/247>
- การทดลองแบบปิดตาสองชั้นเทียบกับยาหลอก ในคนไข้โควิด-19 ที่อาการไม่รุนแรงซึ่งได้ฟาว์พิราเวียร์ทุกคน · สารสกัดฟ้าทะลายโจรไม่ช่วยเพิ่มผลทางคลินิกหรือทางไวรัส — **พอเชื้อ** (การทดลองเดี่ยว เล่มนี้อ่านผ่านรายงานการทดลองที่อ้างอิงมัน ไม่ได้เปิดฉบับเต็ม) — Siripongboonsitti et al., Phytomedicine 119: 155018, 2023 (APFaVi trial) ตามที่อ้างในรายงานการทดลองแบบเปิดปี 2024

— <https://doi.org/10.1016/j.phymed.2023.155018> · <https://www.lidsen.com/journals/icm/icm-09-01-013>

- การทดลองแบบสุ่มที่ไม่ปิดตา ในโรงพยาบาลจังหวัดสองแห่งในไทย ตุลาคม 2021 ถึงกุมภาพันธ์ 2022 · คนไข้การนอย 82 คน 43 คนได้แอนโดรกราโฟไลด์ร่วมกับฟาวิพิราเวียร์ 39 คนได้ฟาวิพิราเวียร์อย่างเดียว · ตัววัดหลักคือปอดอักเสบรุนแรงใน 14 วัน ไม่มีใครเป็น · ไอน์วันที่ 7 คือ 30.2% ต่อ 56.4% · ค่า AST และ ALT สูงกว่าในกลุ่มที่ได้สมุนไพร ผู้วิจัยเรียกว่า mild transaminitis · กลุ่มที่ได้สมุนไพรอายุเฉลี่ยน้อยกว่า (39.7 ต่อ 48.5 ปี) และค่า ALT ตั้งต้นสูงกว่า (36.3 ต่อ 25.2 U/L) · บทนำของรายงานเขียนว่าการแพทย์แผนไทยใช้ฟ้าทะลายโจรแก้ไข้และเจ็บคอ สรุปรายงานก่อนหน้าว่าสารสำคัญอาจลดการอักเสบและขวางไวรัส และอ้างการทดลองแบบเปิดในจีนกับยาฉีดที่สกัดจากต้นเดียวกัน ซึ่งรายงานว่าไข้และไอลดลง — **ยังบาง** (การทดลองเดี่ยว ขนาดเล็ก และไม่ปิดตา เล่นนี้เปิดอ่านวันที่ 26 กันยายน 2026) — Efficacy and Safety of Andrographolide and Favipiravir Versus Favipiravir Monotherapy in Patients with Mild COVID-19 Infection, OBM Integrative and Complementary Medicine, 2 กุมภาพันธ์ 2024 — <https://www.lidsen.com/journals/icm/icm-09-01-013>
- รายการคมชัดลึก ทางเนชั่นทีวี ถกเรื่องลูกบอล EM แก้น้ำเน่าให้เป็นน้ำดี · คมบดีคมะเทคโนโลยีชีวภาพ ผู้เชี่ยวชาญด้านจุลชีววิทยา: จากที่ลงไปปฏิบัติ เชื่อว่าลูกบอล EM แก้น้ำเสียได้ ขึ้นกับวิธีใช้ และเปรียบว่าเหมือนลูกแบดที่ทาสาโท ที่เก็บเชื้อไว้และได้ผลดี · ผู้เริ่มใช้ลูกบอล EM รุ่นแรกๆ ยกตัวอย่างน้ำเน่าหลายแห่งที่แก้สำเร็จ · อาจารย์ภาควิชาสิ่งแวดล้อม คณะวิศวกรรมศาสตร์ จุฬาฯ: โอกาสที่ EM จะบำบัดน้ำเสียได้น้อยมาก ลูกบอลเพิ่มจุลินทรีย์เพียงเล็กน้อย ต้องโทษของเสียขึ้นมา และเห็นด้วยกับการใช้ EM ในการเกษตร · นักวิทยาศาสตร์ด้านเทคโนโลยีชีวภาพจาก สกว.: โอกาสที่ลูกบอล EM จะแก้น้ำเสียได้น้อยมาก ในน้ำเสียมีจุลินทรีย์อยู่มากกว่าที่ใส่ลงไปมาก — **พอเชื้อ** (รายงานข่าวฉบับเดียวของการถกเถียงที่ออกอากาศ สิ่งที่แต่ละคนพูดเป็นความเห็น ไม่ใช่ผลการทดลอง เล่นนี้เปิดอ่านวันที่ 26 กันยายน 2026) — คมชัดลึก, 8 พฤศจิกายน 2011 — <https://www.komchadluek.net/news/114410>
- การทำสาโทเป็นภูมิปัญญาพื้นบ้านของไทย ลูกแบดเป็นแหล่งหัวเชื้อจุลินทรีย์ในการหมักข้าวเหนียวเป็นสาโท และรากับยีสต์มีบทบาทหลัก · งานนี้แยกและจำแนกยีสต์และราจากลูกแบดที่เก็บจาก 42 จังหวัด — **พอเชื้อ** (วิทยานิพนธ์ที่แยกเชื้อจากลูกแบดจริง เล่นนี้เปิดอ่านบทคัดย่อวันที่ 26 กันยายน 2026) — อภิษฎา เดชะวสถ์, การแยก จำแนก และลักษณะสมบัติของยีสต์และราในลูกแบดสุรา เพื่อการผลิตสาโท, วิทยานิพนธ์ จุฬาลงกรณ์มหาวิทยาลัย, 2007 — <https://digital.car.chula.ac.th/chulaetd/57900>
- ผู้พัฒนา EM เป็นศาสตราจารย์เกียรติคุณของมหาวิทยาลัยริวกิว โอกินาวา · เขาโรยจุลินทรีย์ที่เหลือจากการทดลองลงบนหญ้าแปลงหนึ่ง อีกหนึ่งสัปดาห์หญ้าตรงนั้นงามกว่ารอบข้าง · ตั้งชื่อส่วนผสมว่า EM ในปี 1980 — **ยังบาง** (คำเล่าของผู้ผลิตเกี่ยวกับของของตัวเอง เล่นนี้เปิดอ่านวันที่ 26 กันยายน 2026) — EM Research Organization, Prof. Teruo Higa (Developer of EM) — <https://emrojan.com/em/prof-higa/>
- เว็บไซต์ประเมินข้ออ้างของห้องปฏิบัติการสื่อสารวิทยาศาสตร์ มหาวิทยาลัยเมจิ ใช้เกณฑ์สิบข้อในสี่หมวด ให้ EM ระดับ E (ต่ำ) ทุกข้อ และสรุปว่าเป็นวิทยาศาสตร์เทียม · งานวิจัยทั้งในและนอกญี่ปุ่นพบว่าผลของ EM อย่างมากก็เท่ากับวัสดุจุลินทรีย์อื่นที่ใช้ในการเกษตร · EM น่าจะช่วยดับกลิ่นขยะเศษอาหารและทำให้การบดง่ายขึ้น · EM ถูกอธิบายว่าเป็นส่วนผสมของจุลินทรีย์หลายกลุ่ม — **พอเชื้อ** (การประเมินของหน่วยงานเดียว ที่อ้างงานวิจัยหลายชิ้น เล่นนี้เปิดอ่านวันที่ 26 กันยายน 2026) —

Gijika.com, Science Communication Laboratory, Meiji University, หน้าเรื่อง EM ปรับปรุงล่าสุด
10 ธันวาคม 2019 — https://gijika.com/rate/le_effective_microorganisms.html

4 . ศาสตร์ที่นับความผิดของตัวเอง

นักวิจัยกลุ่มหนึ่งหยิบงานวิจัยจิตวิทยาที่ดีพิมพ์ไปแล้วมาร้อยขึ้น แล้วลงมือทำใหม่ทีละชิ้น เพื่อดูว่าผลเดิมจะออกมาอีกไหม

ในงานต้นฉบับ แก้วสืบเจ็ดชิ้นจากร้อยพบผล พอทำซ้ำเหลือราวสามสิบหก และคนที่ลงมือทำซ้ำก็คือนักวิทยาศาสตร์เอง

คำว่าพบผลในที่นี้หมายถึง งานชิ้นนั้นเจอสิ่งที่มันตามหา ชัดพอจะผ่านเกณฑ์ทางสถิติ ที่วงการใช้ตัดสิน คือถ้าไม่มีผลจริงอยู่เลย ความบังเอิญอย่างเดียวจะให้ผลที่ชัดเจนนั้นได้ไม่เกินราวหนึ่งในสิบ งานต้นฉบับผ่านเกณฑ์นี้เกือบทุกชิ้น พอทำใหม่ ผ่านราวหนึ่งในสาม

ในบทที่หนึ่ง คนที่ทดสอบเครื่องตรวจระเบิดคือนักฟิสิกส์คนนอก ที่ศาลสั่งให้ส่งเครื่องไปให้ตรวจ ที่นี้ คนตรวจกับคนถูกตรวจอยู่ในวงการเดียวกัน และคนตรวจเป็นผู้ตีพิมพ์ตัวเลขนี้เองในปี 2015

สองในสามที่หายไป

ถ้าคุณอ่านตัวเลขนี้ตรงๆ มันบอกว่าการวิจัยที่ผ่านการตรวจและได้ตีพิมพ์แล้ว สองในสามทำซ้ำไม่ขึ้น คนที่อ่านแล้วสรุปว่านักวิทยาศาสตร์ผิดบ่อย ไม่ได้อ่านผิด ตัวเลขพูดอย่างนั้นจริงๆ

แต่ตัวเลขนี้เองก็ไม่รอดจากการตรวจ ปี 2016 นักวิจัยอีกกลุ่มหนึ่งเขียนแย้งอย่างเป็นทางการว่า ข้อมูลชุดเดียวกันนี้อ่านแล้วสรุปได้ตรงกันข้าม คือจิตวิทยาทำซ้ำได้ค่อนข้างดี ทีมที่ทำซ้ำเขียนตอบว่า จากข้อมูลชุดนี้ จะสรุปแง่ดีหรือแง่ร้ายก็ได้ทั้งคู่ และยังไม่มีข้อสรุปไหนที่หลักฐานรองรับพอ

ตัวเลขที่ยังถูกแย้งอยู่ ถ้าเล่าเพียงครั้งเดียว ไม่ว่าครั้งไหน ก็คือการเล่าเกินหลักฐาน

สิ่งที่ข้อถกเถียงนี้บอกได้ชัดไม่ใช่ว่าจิตวิทยาถูกก็เปอร์เซ็นต์ มันบอกว่าการวิจัยที่ตีพิมพ์แล้วหนึ่งชิ้นยังไม่ใช้ความจริง มันคือการทดสอบหนึ่งครั้ง และการทดลองซ้ำที่ตามมาตรวจมัน ก็เป็นแค่อีกหนึ่งครั้ง ที่ผิดได้เหมือนกัน

ความรู้ที่พอจะเชื่อได้ไม่ได้อยู่ในงานชิ้นไหนชิ้นหนึ่ง มันอยู่ในสิ่งที่ยั่งยืนอยู่ หลังถูกคนหลายกลุ่มทดสอบซ้ำหลายรอบ

กระดาศที่เซ็นไว้ก่อน

ตัวเลขเก้าสิบเจ็ดยังทั้งคำถามไว้อีกข้อ ถ้างานทุกชิ้นเป็นการทดสอบที่แพ้ได้จริง การชนะเก้าสิบเจ็ดครั้งจากร้อยก็ชวนให้ถามต่อ ว่าครั้งที่แพ้ไปอยู่ที่ไหน

ในการแพทย์มีตัวเลขชุดหนึ่งที่ให้เบาะแส งานชิ้นหนึ่งไล่ดูการทดลองขนาดใหญ่ 55 ชิ้นที่ได้ทุนจากสถาบันหัวใจ ปอด และเลือดแห่งชาติของสหรัฐอเมริกา ชิ้นที่ตีพิมพ์ก่อนปี 2000 รายงานว่าการรักษาได้ผลสิบเจ็ดชิ้นจากสามสิบ ชิ้นที่ตีพิมพ์หลังจากนั้น เหลือสองชิ้นจากยี่สิบห้า

สิ่งที่ผูกกับตัวเลขที่ลดลงนี้มากที่สุด คือการลงทะเบียนล่วงหน้า (preregistration) แปลว่าก่อนเริ่มทดลอง ผู้ทดลองประกาศไว้ในทะเบียนที่ใครก็เปิดดูได้ว่าจะใช้อะไรเป็นตัววัดหลัก

กระดาศแบบนี้เคยปรากฏแล้วในบทที่หนึ่ง ก่อนเริ่มทดสอบ ทหารสองคนเซ็นยืนยันว่าห้อง กล้อง และของกลาง ไม่มีอะไรขัดขวางการทำงานของเครื่อง พอผลออกมาไม่ดี ข้อแก้ตัวที่ตามมาจึงไม่มีที่ยืน การลงทะเบียนทำงานแบบเดียวกัน เมื่อเขียนไว้ก่อนแล้วว่าจะวัดอะไร ก็ย้ายเป้าหลังเห็นผลไม่ได้

แต่ผู้วิจัยของงานชิ้นนี้ไม่ได้สรุปว่ากระดาศแผ่นนี้คือเหตุ พวกเขาเขียนไว้เองว่าบอกไม่ได้ว่าการลงทะเบียนทำให้งานที่พบว่าได้ผลน้อยลง ข้อมูลแสดงได้แค่ว่าสองอย่างเกิดไปด้วยกัน และในเวลาเดียวกันนั้น อย่างอื่นก็อาจเปลี่ยนไปพร้อมกันอีกหลายอย่าง

งานที่วัดความเชื่อตรงของทั้งวงการ ก็ยังระวังไม่พูดเกินข้อมูลของตัวเอง

วารสารที่ตอบรับก่อนรู้ผล

มีวิธีหนึ่งที่ปิดช่องนี้ได้ตรงกว่า นักวิจัยส่งคำถาม วิธีทดลอง และวิธีวิเคราะห์ ไปให้วารสารตรวจก่อนเก็บข้อมูลแม้แต่ขั้นเดียว ถ้าผ่าน วารสารรับปากว่าจะตีพิมพ์ไม่ว่าผลจะออกมาทางไหน งานแบบนี้เรียกว่ารายงานที่ตอบรับก่อนรู้ผล (Registered Report)

ผลที่ไม่เจออะไรจึงมีสิทธิ์ได้ขึ้นหน้ากระดาษเท่ากับผลที่เจอ เพราะคำตัดสินว่าจะตีพิมพ์ถูกทำไปก่อนที่จะรู้ผล

เมื่อมีคนนับเทียบกันในจิตวิทยา งานที่ตีพิมพ์ตามทางปกติรายงานว่าเจอสิ่งที่ตั้งสมมติฐานไว้ 96% งานที่วารสารตอบรับก่อนรู้ผลรายงานแบบนั้น 44%

สองกองนี้ไม่ได้ถูกแบ่งด้วยการจับฉลาก นักวิจัยเลือกเองว่าจะส่งงานแบบไหน ช่องว่างทั้งหมดจึงยังโทษความเอนเอียงอย่างเดียวไม่ได้ แต่มันบอกได้ชัดข้อหนึ่ง เมื่อผลเปลี่ยนการตัดสินใจตีพิมพ์ไม่ได้อีกแล้ว งานที่ไม่เจอสิ่งที่ตามหา ก็ขึ้นมาอยู่บนหน้ากระดาษเกินครึ่ง

ตัวเลขคู่นี้บอกคุณอีกอย่างด้วย งานที่ไม่เจออะไรเียบหายไปจากหน้ากระดาษได้ แต่ความเจ็บไม่ได้แปลว่ามันไม่เคยมี

การลงทะเบียนล่วงหน้ากับวารสารที่ตอบรับก่อนรู้ผล คือเครื่องมือที่วงการนี้สร้างให้ตัวเอง เพื่อปิดทางไม่ให้ใครเลือกผลหลังเห็นผล รวมทั้งตัวนักวิจัยเองด้วย

คะแนนที่ต่ำที่สุด

ถึงตรงนี้มีคำถามตรงไปตรงมารออยู่ วงการที่ประกาศความผิดของตัวเองได้ขนาดนี้ ยังควรได้รับความเชื่อใจไหม และควรได้แค่ไหน

การสำรวจใน 68 ประเทศ คนเกือบเจ็ดหมื่นสองพันคน ให้คะแนนความเชื่อมั่นนักวิทยาศาสตร์เฉลี่ย 3.62 จาก 5 ค่อนไปทางสูง สามในสี่ตอบว่าวิธีการทางวิทยาศาสตร์คือทางที่ดีที่สุดที่จะรู้ว่าอะไรจริง

แต่ในด้านต่างๆ ที่ถูกถาม ด้านที่นักวิทยาศาสตร์ได้คะแนนต่ำที่สุดคือความเปิดรับคำติชม ได้ 3.33 ทั้งที่สิ่งที่บ่งชี้เล่ามาตลอด คือคนในวงการเดียวกันทำซ้ำงานของตัวเอง ตีพิมพ์คำแย้งและคำตอบให้ทุกคนอ่าน และนับว่าตัวเองผิดไปกี่ครั้ง เล่มนี้ไม่รู้ว่าจะว่างนี้มาจากไหน

คำถามที่มีคนวัดไว้แล้วคือ ถ้าคนที่ให้ตัวเลขบอกออกมาตรงๆ ว่าตัวเลขนั้นไม่แน่นอนแค่ไหน คนฟังจะเชื่อน้อยลงไหม การทดลองกับคนรวม 5,780 คนตอบว่าลดลงเพียงเล็กน้อย ทั้งความเชื่อในตัวเลข และความเชื่อในแหล่งที่บอก

นี่คือคำตอบว่าวิทยาศาสตร์เชื่อได้แค่ไหน มันเชื่อได้เท่าที่มันถูกตรวจ งานใหม่ขึ้นเดียวคือการทดสอบหนึ่งครั้ง และเชื่อได้เท่าการทดสอบหนึ่งครั้ง ผลที่คนอื่นทำซ้ำแล้วหลายรอบและยังยืนอยู่ เชื่อได้มากกว่านั้น ป้ายสีป้ายท้ายทุกบทของเล่มนี้ตั้งขึ้นด้วยหลักเดียวกัน

ท้ายบทที่หนึ่งมีตัวอย่างที่คุณเปิดดูได้เอง ข้อสรุปว่าเครื่องตรวจจะเปิดหาของไม่ได้ดีกว่าความบังเอิญได้ป้ายแน่น เพราะมันยืนอยู่บนสองแหล่งที่ให้ผลไปทางเดียวกัน คือการทดสอบในเม็กซิโกและคำตัดสินของศาลอังกฤษ การทดสอบในไทยไม่ได้ถูกนับ เพราะไม่รู้ว่ามันออกแบบอย่างไร ส่วนตัวเลขสี่จากสี่กับสามจากยี่สิบที่บ่งชี้ใช้เปิดเรื่อง ได้แก่ป้ายพอเชื่อ เพราะมาจากการทดสอบวันเดียว

ข้ออ้างที่ว่า นักวิทยาศาสตร์ผิดบ่อย จึงเชื่อไม่ได้ ถูกอยู่ครึ่งหนึ่ง นักวิทยาศาสตร์ผิดได้จริง แต่บ่อยแค่ไหนยังเถียงกันอยู่ ในจิตวิทยา งานที่ทำซ้ำไม่ขึ้นอาจมีถึงราวสองในสาม หรืออาจน้อยจนเรียกได้ว่าทำซ้ำได้ค่อนข้างดี ถ้าคำแย้งปี 2016 อ่านถูก และคุณรู้ทั้งสองแบบนี้ได้ก็เพราะพวกเขานับเอง แล้วเถียงกันเองให้ทุกคนอ่าน

ครั้งที่ผิดคือคำว่าจริง ความผิดที่นับไว้แล้วคือสิ่งที่บอกคุณได้ว่าควรเชื่อแต่ละเรื่องแค่ไหน แหล่งที่ไม่เคยบอกตัวเลขนี้ให้คุณได้แค่ความมั่นใจของเขา

คนที่บอกว่าตัวเองไม่เคยผิด ไม่ได้ผิมน้อยกว่าใคร เขาแค่ไม่เคยนับ

จะรู้ว่บทนี้ผิดเมื่อ: มีการทดลองขนาดใหญ่ที่ประกาศวิธีวัดไว้ล่วงหน้า แล้วพบว่า การบอกความไม่แน่นอนหรือความผิดพลาดของงานวิจัยออกมาตรงๆ ทำให้คนเชื่อถือนักวิทยาศาสตร์ลดลงมาก หรือมีทีมอิสระนับซ้ำแล้วพบว่า งานที่ลงทะเลเบียน

ล่วงหน้าหรือถูกตอบรับก่อนรู้ผล รายงานว่าเจอสิ่งที่ตามหาบ่อยพอๆ กับงานที่ตีพิมพ์ตามทางปกติ

หลักฐาน:

- งานวิจัยจิตวิทยา 100 ชิ้นถูกทำซ้ำ · งานต้นฉบับผ่านเกณฑ์ทางสถิติ 97% · การทดลองซ้ำผ่าน 36% — **พอเชื่อ** (โครงการเดียว และการตีความตัวเลขยังถูกแย้ง ดูบรรทัดถัดไป) — Open Science Collaboration, Estimating the reproducibility of psychological science, Science, 2015 — <https://europepmc.org/article/MED/26315443>
- คำแย้งปี 2016 ว่าข้อมูลชุดเดียวกันสนับสนุนข้อสรุปตรงข้าม คือจิตวิทยาทำซ้ำได้ค่อนข้างดี · คำตอบของทีมที่ทำซ้ำว่าข้อมูลชุดนี้ให้ทั้งข้อสรุปแง่ดีและแง่ร้าย และยังไม่มีข้อสรุปไหนที่หลักฐานรองรับพอ — **พอเชื่อ** (สองฝ่ายยังเห็นต่างกัน เล่มนี้พิมพ์ทั้งคู่) — คำแย้งและคำตอบ, 2016 — <https://europepmc.org/article/MED/26941311> และ <https://europepmc.org/article/MED/26941312>
- การทดลองขนาดใหญ่ 55 ชิ้นที่ได้ทุนจาก NHLBI · ตีพิมพ์ก่อนปี 2000 ได้ผล 17 จาก 30 · หลังจากนั้น 2 จาก 25 · การลงทะเบียนล่วงหน้าผูกกับการทดลองมากที่สุด · ผู้วิจัยระบุเองว่าบอกไม่ได้ว่าเป็นเหตุ — **พอเชื่อ** (งานชิ้นเดียว เป็นความสัมพันธ์ ไม่ใช่หลักฐานของเหตุ) — Kaplan, R. M., & Irvin, V. L. (2015). PLoS ONE — <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0132382>
- งานจิตวิทยาที่ตีพิมพ์ตามทางปกติรายงานว่าเจอสิ่งที่ตั้งสมมติฐานไว้ 96% · งานแบบ Registered Report 44% — **พอเชื่อ** (งานชิ้นเดียว และสองกองไม่ได้แบ่งด้วยการสุ่ม) — Scheel และคณะ, 2021 — <https://doi.org/10.31234/osf.io/p6e9c>
- การสำรวจ 71,922 คนใน 68 ประเทศ · ความเชื่อใจนักวิทยาศาสตร์เฉลี่ย 3.62 จาก 5 · ความเปิดรับคำติชมต่ำที่สุดที่ 3.33 · 75% ตอบว่าวิธีการทางวิทยาศาสตร์คือทางที่ดีที่สุดที่จะรู้ว่าจะอะไรจริง — **พอเชื่อ** (การสำรวจใหญ่หนึ่งครั้ง และการเทียบค่าเฉลี่ยระหว่างประเทศทำได้จำกัด) — Cologna และคณะ, Nature Human Behaviour, 2025 — <https://www.nature.com/articles/s41562-024-02090-5>
- การบอกความไม่แน่นอนของตัวเลข ทำให้ความเชื่อในตัวเลขและในแหล่งที่บอกลดลงเพียงเล็กน้อย · การทดลองหลายชุด รวม 5,780 คน — **พอเชื่อ** (งานชิ้นเดียว) — van der Bles และคณะ, PNAS, 2020 — <https://europepmc.org/article/MED/32205438>

5 · คำถามที่หนึ่ง · เทียบกับอะไร

เป็นหวัดมาสองวัน คื่นที่สามอาการหนักที่สุด จนต้องลุกไปหยิบยาที่เพื่อนแนะนำมา กิน สามวันต่อมาก็หายเกือบหมด

ถ้านั่นคือหวัดของคุณ ก่อนอ่านต่อ ลองตอบในใจหนึ่งข้อ แบ่งเครดิตของการหายครึ่งนี้เป็นสิบส่วน ยาควรได้กี่ส่วน

ยาที่เริ่มกินตอนอาการหนักที่สุด จะดูเหมือนได้ผลเสมอ เพราะหลังจุดที่หนักที่สุด ทางที่เหลือแทบมีแต่ดีขึ้น

ในเรื่องนี้ ยาถูกหยิบในคื่นที่ทนไม่ไหว และคื่นที่ทนไม่ไหวที่สุดของหวัดครึ่งหนึ่งก็คือ ยอดของมัน พ้นยอดไปแล้ว ต่อให้ไม่มีใครทำอะไรเลย ทางที่เหลือก็เป็นทางลงของอาการ ยาที่มาถึงในคื่นนั้นจึงได้ยืนอยู่ตรงจุดที่อาการเริ่มดีขึ้นพอดี

คุณอีกคนที่ไม่ได้กินยา

ยาอาจช่วยจริงก็ได้ บทนี้ไม่ได้บอกว่ามันไม่ช่วย สิ่งที่ยกคือสามวันนั้นตัดสินใจไม่ได้ เพราะในนั้นมีสองอย่างซ้อนกันอยู่ อย่างหนึ่งคือสิ่งที่ยาทำ อีกอย่างคือสิ่งที่หวัดทำเอง อยู่แล้ว และจากหวัดครึ่งเดียว ไม่มีใครแยกสองอย่างนี้ออกจากกันได้

สิ่งที่ต้องใช้แยกคือคุณอีกคนหนึ่ง หวัดครึ่งเดียวกัน คื่นเดียวกัน แต่ไม่ได้กินยา แล้วดูว่าคนนั้นหายช้ากว่าหรือเปล่า คนคนนั้นไม่มีอยู่จริง และไม่มีวันมี

วิทยาศาสตร์ไม่เสียเวลาตามหาเขา มันสร้างของที่ใกล้ที่สุดขึ้นมาแทน เอาคนเป็นหวัดหลายคนมาแบ่งเป็นสองกลุ่มที่เหมือนกันพอ กลุ่มหนึ่งได้ยา อีกกลุ่มไม่ได้ แล้วนับว่ากลุ่มไหนหายเร็วกว่า ส่วนที่สองกลุ่มหายเท่ากันคือส่วนของเวลา ส่วนที่ต่างกันเท่านั้นคือส่วนของยา

กลุ่มที่ไม่ได้ยาเรียกว่ากลุ่มเปรียบเทียบ และมันคือคำตอบของคำถามข้อแรกในห้าข้อของเล่มนี้ เทียบกับอะไร

หกุ่่นเรือ

วันที่ 20 พฤษภาคม 1747 บนเรือรบอังกฤษชื่อ Salisbury มีผู้ป่วยโรคลักปิดลักเปิดสิบสองคนที่ James Lind คัดมาให้ทำการคล้ายกันที่สุดเท่าที่หาได้ ทุกคนอยู่ห้องเดียวกัน กินอาหารเหมือนกัน แล้วถูกแบ่งเป็นหกคู่ แต่ละคู่ได้ของต่างกันอย่างเดียว คือน้ำแอปเปิลหมัก ยาน้ำผสมกรดกำมะถัน น้ำส้มสายชู น้ำทะเล สัมสองลูกกับเลมอน หนึ่งลูกทุกวัน และยาเนื้อขันทึบผสมจากกระเทียม เมล็ดมัสตาร์ด หัวไชเท้า และมดยอบ นานสิบสี่วัน สัมกับเลมอนหมดลงในวันที่หก และในวันที่หกนั้น หนึ่งในคู่ที่ได้ผลไม้ก็กลับไปทำงานได้

เรื่องเล่าฉบับวีรบุรุษจบตรงนี้ว่า Lind ค้นพบว่าผลไม้ตระกูลส้มรักษาโรคนี้ได้ แล้วกองทัพเรือเฟิกเฉยไปหลายสิบปี นักประวัติศาสตร์ที่กลับไปอ่านเอกสารพบว่าไม่ใช่ อย่างนั้น ในงานเขียนของเขาเอง Lind ไม่ได้บอกไว้ชัดเจนตรงไหนว่าผลไม้ตระกูลส้มรักษาโรคนี้ได้ น้ำผลไม้ที่เขาแนะนำให้เคี้ยวเก็บไว้ใช้บนเรือ เสียวitaminที่เป็นตัวรักษาไปกับความร้อน กองทัพเรือเฟิกแจกน้ำเลมอนกันทั่วไปตั้งแต่ปี 1795 และฉบับวีรบุรุษนั้นสืบย้อนไปได้ถึงปี 1893 เกือบร้อยห้าสิบปีหลังการทดลอง

เรื่องจริงไม่ได้ทำให้สิ่งที่เกิดบนเรือลำนั่นเล็กลง การเทียบของ Lind ชี้ไปทางสัมกับเลมอน แต่สิ่งที่มันชี้ไม่ได้บังคับให้ใครเข้าใจ แม้แต่คนที่จัดการเทียบขึ้นมาเอง และเรื่องเล่าเกี่ยวกับวิทยาศาสตร์ก็ต้องถูกตรวจแบบเดียวกับผลทดลอง อย่างที่บทที่แล้วเล่าไว้

คนวัดรู้คำตอบ

การเทียบที่จัดมาดีแค่นั้นก็ยังฟังได้อีกทางหนึ่ง และบทที่หนึ่งวัดทางนั้นไว้ให้แล้ว ทหารชุดเดิม ห้องเดิม วันเดียวกัน เมื่อคนถือรู้ว่าของอยู่กล่องไหน เครื่องหาเจอสี่ครั้งจากสี่ เมื่อไม่รู้ เหลือสามครั้งจากยี่สิบ

ในการทดลองยา คนที่อยู่เ็นตำแหน่งของคนถือเครื่องมีสองฝ่าย คือผู้ป่วยที่ตอบว่ารู้สึกดีขึ้นไหม กับคนที่ตรวจแล้วจั่วว่าดีขึ้นไหม ถ้าผู้ป่วยรู้ว่าตัวเองได้ยาจริง ความหวังของเขาก็ปนเข้าไปในคำตอบ ถ้าคนตรวจรู้ ความคาดหวังของเขาก็ปนเข้าไปในตัวเลขแบบเดียวกับที่ความรู้ของคนถือปนเข้าไปในเสาอากาศ

ทางแก้มีสองชั้น ชั้นแรกคือยาหลอก (placebo) ยาที่หน้าตา รสชาติ และวิธีกินเหมือนยาจริงทุกอย่าง ต่างกันแค่ไม่มีตัวยา เมื่อกลุ่มเปรียบเทียบได้ยาหลอก ทุกคนก็กินยาเหมือนกัน และหวังได้เท่ากัน ชั้นที่สองคือการปิดตาสองชั้น (double-blind) ทั้งผู้ป่วยและคนวัดผลไม่รู้ว่าใครอยู่กลุ่มไหน จนกว่าจะนับผลเสร็จ

เมื่อไม่มีใครรู้ว่าใครได้อะไร ความคาดหวังก็ไหลเข้าไปในผลไม่ได้ สิ่งที่เหลืออยู่ในผลคือสิ่งที่ยาทำได้จริง

สามข้อ ก่อนให้เครดิต

ทั้งหมดนี้ย่อได้เหลือสามข้อ คุณใช้มันได้กับทุกข้ออ้างที่บอกว่าอะไรทำอะไรดีขึ้น ทั้งยา อาหาร คอร์สเรียน และแอปในโทรศัพท์

1. **เทียบกับใคร** ต้องมีคนที่ไม่ได้รับสิ่งนั้นในช่วงเวลาเดียวกัน ถ้าไม่มี ทุกอย่างที่ดีขึ้นอาจเป็นของเวลา
2. **สองกลุ่มเหมือนกันตั้งแต่ต้นไหม** ถ้าแต่ละคนเลือกกลุ่มเอง เหตุที่ทำให้เขาเลือกอาจเป็นเหตุเดียวกับที่ทำให้ผลออกมาดี ทางที่ดีที่สุดคือให้ความบังเอิญเป็นคนเลือก ด้วยเหรียญหรือการจับฉลาก
3. **คนวัดรู้ไหมว่าใครได้อะไร** ถ้ารู้ ความคาดหวังเข้าไปอยู่ในตัวเลขได้ ทั้งของคนวัดและของคนที่ถูกวัด

ลองใช้สามข้อนี้กับหัตถ์ตอนต้นบท ข้อแรก ไม่มีใครให้เทียบ ข้อสองจึงไม่มีอะไรให้ถาม ข้อสาม คนที่ตัดสินใจว่าหายเพราะยาคือคนที่รู้ว่าตัวเองกินยา ตกรับทั้งสามข้อ

คำตอบของคำถามตอนต้นบทจึงเป็นแบบนี้ ส่วนของยาคือส่วนที่เหลือ หลังหักสิ่งที่จะเกิดอยู่แล้วแม้ไม่กินยาออกไป หวัดครั้งเดียวให้คุณเห็นแค่ตัวตั้งของการลบ ตัวลบไม่มีใครเห็น ตัวเลขที่ตอบไว้ไม่ว่าศูนย์หรือสิบ วัดสิ่งเดียวกัน คือการเดาว่าถ้าไม่กินยา หวัดจะหายช้ากว่านี้แค่ไหน จากสามวันนั้น คำตอบที่เชื่อตรงคือยังไม่มีใครรู้ แม้แต่คนที่หายเอง

ถ้าย้อนไปใช้สามข้อนี้กับเรื่องของ Lind ข้อแรกผ่าน ข้อสองผ่านได้แค่บางส่วน คนไข้มีอาการคล้ายกัน แต่ไม่ได้ให้ความบังเอิญเลือกว่าใครได้อะไร ข้อสามไม่ผ่าน ไม่มีใครถูกปิดตา การเทียบบนเรือลำนั้นจึงเป็นจุดเริ่ม ยังไม่ใช้การทดสอบที่ยุติธรรม

สามข้อเดียวกันใช้กับเรื่องเงินได้ทันที คอร์สสอนขายของออนไลน์คอร์สหนึ่งโฆษณาด้วยยอดขายของผู้เรียนรุ่นก่อน ที่โตขึ้นหลังเรียนจบ

เทียบกับใคร ถ้ารุ่นนั้นเรียนจบก่อนช่วงเทศกาลลดราคา ยอดขายของร้านที่ไม่ได้เรียนอะไรเลยก็โตได้เหมือนกัน นั่นคือเวลาของโลกการค้า แบบเดียวกับที่หัดขายเอง

สองกลุ่มเหมือนกันตั้งแต่ต้นไหม คนที่ยอมจ่ายเงินเรียนขายของ อาจมีร้านที่กำลังโตอยู่แล้ว หรือตั้งใจจะทุ่มให้ร้านมากขึ้นตั้งแต่ก่อนสมัคร เหตุที่ทำให้เขาสมัครอาจเป็นเหตุเดียวกับที่ทำให้ยอดขายโต คอร์สที่อยากรู้คำตอบจริงทำได้ เมื่อคนสมัครเกินที่นั่งให้จับฉลาก แล้วอีกหนึ่งปีเทียบคนที่ได้ที่นั่งกับคนที่จับไม่ได้

คนวัดรู้ไหม ในโฆษณา คนวัดคือคอร์สเอง ซึ่งรู้ว่าใครเป็นผู้เรียน และเลือกเองว่าจะวัดความสำเร็จด้วยอะไร จะช่วยยอดขายของใคร การเทียบที่ยุติธรรมเขียนตัววัดไว้ก่อนเริ่ม อย่างการลงทะเบียนล่วงหน้าในบทที่สี่ แล้ววัดทุกคนในสองกลุ่มด้วยวิธีเดียวกัน

เก้าคาบ

คำถามสามข้อนี้สอนกันได้ไหม มีคนทดสอบไว้แล้ว ด้วยวิธีเดียวกับที่บทนี้เล่ามาทั้งบท

ในยูกันดา โรงเรียนประถม 120 แห่งถูกแบ่งเป็นสองกลุ่มด้วยการสุ่ม เด็กในกลุ่มหนึ่งได้เรียนเก้าคาบ คาบละแปดสิบนาที ว่าด้วยแนวคิดสิบสองข้อสำหรับตัดสินใจเรื่องการรักษา อีกกลุ่มเรียนตามปกติ เด็กในการทดลองอายุสิบถึงสิบสองขวบ

คนที่ทำการทดลองนี้เปิดรายชื่อแนวคิดของพวกเขาให้ใครก็อ่านได้ ในนั้น ความเชื่อในประสบการณ์ของตัวเองอยู่ในหมวดความเชื่อผิดๆ เกี่ยวกับแหล่งที่มา ส่วนสิ่งที่ถูกสอนแทนคือการเทียบที่ยุติธรรม หัวข้อตอนต้นบทคือประสบการณ์แบบนั้นพอดี

หลังเรียนจบ เด็กทั้งสองกลุ่มทำข้อสอบชุดเดียวกัน ในโรงเรียนที่ได้เรียน เด็กสอบผ่าน 69% ในโรงเรียนที่ไม่ได้เรียน 27%

มีการทดลองแบบเดียวกันอีกสามชิ้นในโรงเรียนมัธยมของเคนยา รวันดา และยูกันดา นักเรียนอายุสิบสี่ถึงสิบหกปี เรียนสิบคาบ คาบละสี่สิบนาที ในทุกร้อยคน โรงเรียนที่ได้เรียนสอบผ่านมากกว่าโรงเรียนที่ไม่ได้เรียนราวสามสิบสามคน ผู้วิจัยสรุปว่านักเรียนคิดเรื่องทางเลือกด้านสุขภาพได้ดีขึ้นมาก และในประโยคเดียวกันก็เขียนว่า 42% ของนักเรียนที่ได้เรียนยังสอบไม่ผ่าน

หนึ่งปีต่อมา นักเรียนที่ได้เรียนยังสอบผ่าน 52.6% จาก 58.1% ตอนเรียนจบใหม่ๆ ความต่างจากโรงเรียนที่ไม่ได้เรียนลดลงเหลือราวยี่สิบหกคนต่อร้อย แต่ยังอยู่ ตัวเลขชุดนี้แน่นนอนน้อยกว่าชุดแรก เพราะนักเรียนราวหนึ่งในสี่ตามตัวไม่ได้

นี่คือความหวังของบennie และมันมาพร้อมตัวเลขทั้งสองด้าน คำถามข้อนี้สอนได้ในห้องเรียนธรรมดา ด้วยบทเรียนไม่กี่คาบ และผลของมันวัดได้ด้วยการทดสอบที่ยุติธรรม แต่มันไม่ใช่เวทมนตร์ เด็กประถมที่ได้เรียนราวสามในสิบยังสอบไม่ผ่าน นักเรียนมัธยมที่ได้เรียนสี่ในสิบยังสอบไม่ผ่าน และหนังสือหนึ่งบทก็น้อยกว่าบทเรียนเก้าคาบ

คำถามที่เด็กอายุสิบขวบเรียนได้ในเก้าคาบ มีอยู่สามคำ เทียบกับอะไร

ลองคิดนี่: เลือกข้ออ้างหนึ่งข้อที่คนในบ้านเชื่อว่าได้ผล แล้วช่วยกันออกแบบการเทียบให้มันบนกระดานแผ่นเดียวว่าจะเทียบกับใคร จะใช้เหรียญแบ่งกลุ่มอย่างไร และใครจะเป็นคนวัดที่รู้ว่าใครได้อะไร ข้อไหนออกแบบไม่ได้ ให้จดไว้ ข้อนั้นคือจุดที่คำว่าได้ผลในบ้านนี้ยังไม่เคยถูกตรวจ

จะรู้ว่าบทนี้ผิดเมื่อ: ผู้อ่านที่อ่านบทนี้จบ ได้รับข้ออ้างที่ไม่เคยเห็นมาก่อน แล้วยังบอกไม่ได้ว่าควรเทียบกับใคร จะแบ่งสองกลุ่มอย่างไร และใครควรเป็นคนวัด หรือเมื่อทีมอื่นทดลองสอนแบบเดียวกันในที่ใหม่ แล้วพบว่านักเรียนที่ได้เรียนตัดสินข้ออ้างเรื่องการรักษาได้ไม่ดีกว่านักเรียนที่ไม่ได้เรียน

หลักฐาน:

- การทดลองของ James Lind บนเรือ Salisbury เริ่มวันที่ 20 พฤษภาคม 1747 · ผู้ป่วยสิบสองคนที่อาการคล้ายกันที่สุด ห้องเดียวกัน อาหารเดียวกัน · ทกวิธี วิธละสองคน นานสิบสัปดาห์ · สัมกับเลมอนหมด

ในวันที่หก และหนึ่งในคืนนั้นกลับไปทำงานได้ — **พอเชื่อ** (เล่าผ่านนักประวัติศาสตร์ เล่มนี้ไม่ได้เปิดบันทึกต้นฉบับ และไม่ได้พิมพ์ปริมาณยาเพราะแปลงเป็นหน่วยปัจจุบันแล้วตรวจไม่ได้) — James Lind Library — <https://www.jameslindlibrary.org/articles/who-was-james-lind-and-what-exactly-did-he-achieve/>

- Lind ไม่ได้บอกไว้ชัดว่าผลไม้ตระกูลส้มรักษาโรคโลหิตจางได้หรือไม่ . น้ำผลไม้ที่ผ่านความร้อนเสียวิตามิน . น้ำเลมอนถูกแจกทั่วไปตั้งแต่ปี 1795 . เรื่องเล่าฉบับวีรบุรุษสืบทอดไปได้ถึงปี 1893 — **พอเชื่อ** (นักประวัติศาสตร์ในแหล่งเดียวกัน) — James Lind Library — <https://www.jameslindlibrary.org/articles/james-lind-and-scurvy-1747-to-1795/> และบทความในบรรทัดก่อน
- รู้คำตอบ 4 จาก 4 . ปิดตาสองชั้น 3 จาก 20 — **พอเชื่อ** (ดูหลักฐานของบทที่หนึ่ง) — Mochán & Ramírez-Solís, arXiv, 2013 — <https://arxiv.org/abs/1301.3971>
- การทดลองในโรงเรียนประถม 120 แห่งในยูกันดา แบ่งโรงเรียนด้วยการสุ่ม . เด็กอายุ 10 ถึง 12 ปี . เก้า คาบ คาบละ 80 นาที แนวคิด 12 ข้อ . สอบผ่าน 69% เทียบกับ 27% — **พอเชื่อ** (การทดลองขนาดใหญ่หนึ่งครั้ง) — The Lancet, 2017 — <https://europepmc.org/article/MED/28539194>
- การทดลองสามชั้นในโรงเรียนมัธยมของเคนยา รวันดา และยูกันดา แบ่งโรงเรียนด้วยการสุ่ม . 244 โรงเรียน นักเรียน 11,344 คน อายุ 14 ถึง 16 ปี . สิบคาบ คาบละ 40 นาที . สอบผ่าน 58% (3,397 จาก 5,846) มากกว่าโรงเรียนที่ไม่ได้เรียน 33 คนต่อร้อย (odds ratio 5.5) . ผู้วิจัยสรุปในประโยคเดียวว่า นักเรียนคิดได้ดีขึ้นมาก แต่ 42% ยังไม่ผ่าน — **พอเชื่อ** (สามการทดลองจากเครือข่ายวิจัยเดียวกัน) — Chesire และคณะ, Journal of Evidence-Based Medicine, 2023 — <https://doi.org/10.1111/jebm.12552>
- ติดตามหนึ่งปี 8,298 คน (73%) . สอบผ่าน 52.6% เทียบกับ 58.1% ตอนเรียนจบ ผู้วิจัยคิดเป็นการเรียนรู้ที่คงอยู่ 88.3% หลังปรับหักส่วนที่เดาถูก . มากกว่าโรงเรียนที่ไม่ได้เรียน 25.6 คนต่อร้อย . ความแน่นอนต่ำลงเพราะติดตามไม่ได้ 27% — **พอเชื่อ** — Chesire และคณะ, Trials, 2024 — <https://doi.org/10.1186/s13063-024-08577-w>
- รายการแนวคิดของ Informed Health Choices จัดการเชื่อประสบการณ์ของตัวเองไว้ในหมวดความเชื่อผิดๆ เกี่ยวกับแหล่งที่มา และสอนการเทียบที่ยุติธรรม — **แนน** (เปิดอ่านได้จากหน้าของโครงการเอง) — <https://www.informedhealthchoices.org/key-concepts/index-key-concepts/>

6 · คำถามที่สอง · บังเอิญได้ไหม

เหรียญบาทเหรียญหนึ่งถูกโยนสิบครั้ง ออกหัวแปดครั้ง คนโยนยิ้ม แล้วบอกว่าเขามีวิธีโยนของเขาเอง

ก่อนจะตัดสินใจว่าเขามีวิธีจริงไหม ลองทายตัวเลขข้อหนึ่งดู คนร้อยคนใบ้เลขท้ายสองตัวของงวดหน้า ทุกคนเดาสุ่ม ไม่มีใครมีสูตร พอผลออก จะมีคนใบ้ถูกกี่คน และโอกาสที่จะมีอย่างน้อยหนึ่งคนถูกคือเท่าไร จดตัวเลขของคุณไว้ก่อนอ่านต่อ

ถ้ามีคนใบ้เลขท้ายสองตัวแบบเดาสุ่มร้อยคน ทุกงวดจะมีคนใบ้ถูกเฉลี่ยหนึ่งคน และคนนั้นคือคนเดียวที่คุณจะได้เห็นโพสต์

ครั้งแรกของประโยคนั้นใช้แค่การนับ เลขท้ายสองตัวมีได้ร้อยแบบ ตั้งแต่ 00 ถึง 99 คนที่เดาสุ่มจึงถูกได้หนึ่งครั้งในร้อย ร้อยคนเดาพร้อมกัน ก็ได้คนถูกเฉลี่ยงวดละหนึ่งคน

ครั้งหลังไม่ต้องนับอะไรเลย คนที่ใบ้ถูกมีเรื่องจะเล่า คนที่ใบ้ผิดไม่มี

เก้าสิบเก้าคนที่ไม่ได้โพสต์

คำว่าเฉลี่ยหนึ่งคนไม่ได้แปลว่าทุกงวดจะมีคนถูกหนึ่งคนพอดี บางงวดไม่มีใครถูกเลย บางงวดถูกพร้อมกันสองสามคน ถ้านับไปนานๆ งวดที่มีคนถูกอย่างน้อยหนึ่งคนจะมีราว 63 งวดในทุก 100 งวด อีกราว 37 งวดไม่มีใครถูกสักคน นั่นคือคำตอบเต็มของข้อที่ให้ทายไว้

ที่นี้มองจากฝั่งคนที่ถือโทรศัพท์ คนที่ใบ้ถูกมีทุกเหตุผลที่จะโพสต์ อีกเก้าสิบเก้าคนแทบไม่มีเหตุผลสักข้อ ภาพที่มาถึงหน้าจอคุณจึงไม่ใช่ภาพของร้อยคน มันคือภาพของหนึ่งคน ที่ถูกคัดมาแล้วโดยไม่มีใครตั้งใจคัด

ในบทที่หนึ่ง ตัวเลขที่ตัดสินใจเครื่องตรวจจะเปิดคือสามครั้งจากยี่สิบ ตัวเลขหลังคำว่า "จาก" คือตัวหาร มันบอกว่าลองทั้งหมดกี่ครั้ง โพสต์ใบ้ถูกบอกแต่ว่าถูก ไม่เคยบอกว่าถูกจากกี่คน

ถ้าร้อยคนนั้นเลือกเลขไม่ซ้ำกันเลย จะมีคนถูกทุกงวด โดยที่ไม่มีใครรู้อะไรเลยสักคน

โพสต์ที่บอกว่าถูกสองงวดติดกันก็เหมือนกัน สำหรับคนเดาคนเดียว นั่นคือโอกาสหนึ่งในหมื่น ฟังดูต้องมีฝีมือ แต่ถ้ามีคนใบ้หนึ่งหมื่นคน ก็จะมีคนถูกสองงวดติดกันเฉลี่ยหนึ่งคน ความเก่งที่ดูหายากขึ้นเท่าไร ก็ต้องการแค่ตัวหารที่ใหญ่ขึ้นเท่านั้น

ทั้งหมดนี้ไม่ได้แปลว่าคนที่ใบ้ถูกโกหก เขาอาจเชื่อในวิธีของตัวเองจริงๆ และงวดนั้นก็ถูกจริงๆ แต่โพสต์เดียวแยกฝีมือออกจากโชคไม่ได้ จนกว่าจะรู้ว่ามีคนใบ้ทั้งหมดกี่คน และตัวเขาเองใบ้มาแล้วกี่งวด ถูกกี่งวด

คนโยนเหรียญในย่อหน้าแรกก็เช่นกัน แปรหัวจากสิบ ถ้ามีคนโยนอยู่คนเดียว เกิดได้ราวหนึ่งครั้งในสิบแปด แต่ถ้าในงานเลี้ยงมีคนยี่สิบคนโยนเล่นพร้อมกัน โอกาสที่ใครสักคนจะได้แปรหัวขึ้นไปคือราวสองในสาม คนนั้นคือคนที่จะยืม และเป็นคนเดียวที่ทุกคนจำได้

จะรู้ว่าเขามีวิธีจริงไหมจึงไม่ยาก ขอให้เขาบอกไว้ก่อนว่าจะได้หัวกี่ครั้ง แล้วโยนใหม่อีกสิบครั้ง ถ้าที่ผ่านมาเป็นโชค โอกาสได้แปรหัวอีกรอบก็ยังราวหนึ่งในสิบแปดเท่าเดิม โชคไม่ได้รับปากใครว่าจะมาซ้ำ

สามร้อยครั้งจากกี่ครั้ง

ในบทที่หนึ่ง ผู้บัญชาการทหารบกในเวลานั้นยืนยันว่าเครื่องตรวจระเบิดเคยเจอของจริงมาแล้วกว่าสามร้อยครั้ง เล่นนี้ไม่ได้สงสัยว่าสามร้อยครั้งนั้นเกิดขึ้นจริงหรือไม่ มันอาจเกิดขึ้นจริงทุกครั้ง

สิ่งที่ไม่มีอยู่ในคำยืนยันนั้นคือตัวหาร เครื่องถูกถือเดินไปกี่ครั้งแล้วชี้ไปที่ที่ไม่มีอะไร ถูกถือเดินผ่านของจริงไปกี่ครั้งโดยไม่ชี้ และในที่ที่คนถือสงสัยอยู่แล้วว่าน่าจะมีของการเดาเฉยๆ จะถูกบ่อยแค่ไหน ถ้าไม่มีตัวเลขสามตัวนี้ สามร้อยครั้งก็มีรูปร่างเดียวกับโพสต์ใบ้ถูก คือมีแต่ตัวเลขครั้งที่ได้ ไม่มีตัวเลขครั้งที่ลอง

ในห้องบอลรูมที่เม็กซิโก ยี่สิบรอบที่ปิดตาสองชั้น เครื่องหาเจอสามครั้ง ส่วนเหรียญที่ใช้เทียบเลือกถูกหนึ่งครั้ง ถ้าอ่านแค่ตัวเลขสองตัวนี้ เครื่องชนะเหรียญสามเท่า

บทที่หนึ่งบอกไว้แล้วว่าเครื่องไม่ได้ชนะ มันแค่โชคดีกว่าเหรียญในบាយวันนั้น คราวนี้ดูว่าโชคกว้างแคไหน ความบังเอิญไม่ได้ให้ตัวเลขเดียว มันให้เป็นช่วง ในยี่สิบรอบที่แต่ละรอบมีโอกาสถูกหนึ่งในแปด ของที่ไม่ทำงานเลยจะได้ระหว่างศูนย์ถึงห้าครั้ง ในราว 97 จากทุก 100 ครั้งที่ทดสอบแบบนี้ และได้สามครั้งขึ้นไปในเกือบครึ่งหนึ่ง

ตัวเลขที่ตกอยู่ในช่วงนี้ จะเป็นหนึ่งหรือสามก็บอกอะไรไม่ได้ ของที่อ้างว่าทำงานได้จริง ต้องได้เกินช่วงนี้ไปให้เห็นชัด คุณถึงจะเริ่มถามต่อได้ว่ามันทำอะไร

เจ็ดสิบทีม ข้อมูลชุดเดียว

การลองหลายครั้งไม่ได้เกิดแค่ในกลุ่มไ้ห่วย มันเกิดในห้องแล็บได้ด้วย

งานวิจัยชิ้นหนึ่งส่งข้อมูลภาพสแกนสมองชุดเดียวกันให้ทีมวิเคราะห์เจ็ดสิบทีมที่ทำงานแยกกัน ทุกทีมได้โจทย์เดียวกัน คือตอบว่าสมมติฐานเก่าข้อเป็นจริงหรือไม่ ผลคือไม่มีสองทีมไหนเลือกวิธีวิเคราะห์เหมือนกันทุกชิ้น และคำตอบว่าสมมติฐานข้อไหนจริงก็ต่างกันไม่น้อยจากทีมหนึ่งไปอีกทีมหนึ่ง ทั้งที่ข้อมูลเป็นชุดเดียวกันทุกประการ

ความต่างแบบนี้เกิดได้โดยไม่ต้องมีใครโกง ระหว่างข้อมูลดิบกับคำตอบมีทางแยกให้เลือกหลายจุด และทุกทีมต้องเลือก

คราวนี้ถ้านักวิจัยคนเดียวลองเดินหลายทางนั่นเอง ทีละทาง แล้วรายงานเฉพาะทางที่ได้ผลน่าสนใจ ตัวหารที่หายไปจะไม่ได้อยู่ในกลุ่มไลน์ มันอยู่ในเครื่องคอมพิวเตอร์ของเขา และเขาเองอาจไม่ได้นับ มันคือคนร้อยคนที่ไ้ห่วย ย้ายมาอยู่ในหัวของคนคนเดียว

วิทยาศาสตร์รู้จักกับดักนี้ และวิธีแก้หน้าตาคล้ายกระดาศที่ทหารเซ็นในบทที่หนึ่ง คือเขียนไว้ก่อนเห็นข้อมูลว่าจะวิเคราะห์ทางไหน ถ้าไม่ได้เขียนไว้ก่อน ผลที่หาเจอต้องถูกทดสอบใหม่กับข้อมูลชุดที่ไม่ได้ใช้หามัน

และคนที่เอาข้อมูลชุดเดียวไปให้เจ็ดสิบทีม แล้วนับว่าคำตอบของตัวเองแกลงได้แคไหน ก็คือวิทยาศาสตร์เอง

หมู่บ้านที่ได้เป็นข่าว

ข่าวชนิดหนึ่งมีรูปร่างแบบนี้ ในหมู่บ้านแห่งหนึ่ง มีคนป่วยเป็นโรคเดียวกันหลายคน ติดๆ กันในเวลาไม่นาน ชาวบ้านเริ่มสงสัยน้ำบ่อ สงสัยโรงงานที่อยู่ไม่ไกล

ความกังวลนั้นมีเหตุผล คนที่ป่วยมีอยู่จริง เป็นเพื่อนบ้านที่ทุกคนรู้จักชื่อ และถ้า ต้นเหตุมีอยู่จริง การรอนแค้นใจก็มีราคาที่ต้องจ่าย

แต่คำถามที่สองยังต้องถูกถาม ประเทศหนึ่งมีหมู่บ้านอยู่มาก ต่อให้ไม่มีอะไรผิดปกติ ที่ไหนเลย ความบังเอิญก็ไม่ได้แจกคนป่วยให้ทุกหมู่บ้านเท่ากัน มันแจกแบบเดียวกับที่ แจกหัวให้คนยี่สิบคนที่โยนเหรียญ และหมู่บ้านที่ได้มากที่สุดคือหมู่บ้านที่ได้เป็นข่าว

อีกชั้นหนึ่งมองเห็นยากกว่า หมู่บ้านไหน โรคอะไร ช่วงเดือนไหน ขอบของกลุ่มก้อน แบบนี้ถูกขีดขึ้นหลังจากเห็นคนป่วยแล้ว เหมือนยิงปืนใส่กำแพงก่อน แล้วค่อยเดินไป วาดเป้ารอบจุดที่รู้กระสุนกระจุกกันแน่นที่สุด เป้าที่วาดที่หลังโดนกลางทุกครั้ง

นี่ไม่ได้แปลว่าหมู่บ้านนั้นไม่มีอะไร คำถามว่าบังเอิญได้ไหมไม่ได้ตอบว่าบังเอิญ มัน บอกว่าสิ่งที่เห็นอยู่ตอนนี้ยังแยกสองอย่างออกจากกันไม่ได้

สิ่งที่แยกได้คือการทลายลงหน้า ถ้าต้นเหตุคือน้ำบ่อ คนที่ดื่มน้ำบ่อนั้นควรป่วยมากกว่าเพื่อนบ้านที่ไม่ได้ดื่ม และหมู่บ้านอื่นที่ใช้น้ำแหล่งเดียวกันก็ควรเห็นแบบเดียวกัน การทดสอบต้องใช้ข้อมูลที่ยังไม่มีใครดู ไม่ใช่ชุดที่ทำให้เห็นรูปแบบนั้นครั้งแรก ซึ่งก็คือคำถามที่หนึ่ง เทียบกับอะไร

สามข้อก่อนเชื่อ

คำถามที่สองใช้ได้กับทุกผลที่ดูดีเกินกว่าจะบังเอิญ ไม่ว่าจะเป็นเลขเด็ด ยาที่ใครสักคนกินแล้วหาย หรือกราฟในงานวิจัย ใช้ได้เป็นสามขั้น

1. **นับตัวหาร** มีคนลองกี่คน หรือลองกี่ครั้ง รวมครั้งที่ไม่ได้ผลด้วย ถ้าหาไม่เจอ ให้เขียนตรงๆ ว่ายังไม่รู้

2. **ถามว่าโชคอย่างเดียวให้ได้แค่ไหน** ถ้าไม่มีอะไรทำงานเลย ผลแบบนี้จะเกิดบ่อยแค่ไหนในจำนวนครั้งเท่านั้น ถ้าผลที่เห็นอยู่ในช่วงที่โชคให้ได้ มันยังไม่ได้บอกอะไร

3. **ถามว่าทายไว้ก่อน หรือเจตาทิ้ง** รูปแบบที่ถูกพูดไว้ก่อนเห็นผลมีน้ำหนักกว่ารูปแบบที่หาเจอหลังจากดูทุกอย่างแล้ว อย่างหลังต้องรอข้อมูลชุดใหม่

ขั้นที่สองบางครั้งต้องคำนวณ และ AI ช่วยได้ ถ้าคุณลองมันก่อนด้วยโจทย์ที่รู้คำตอบอยู่แล้ว ร้อยคนกับเลขท้ายสองตัวคือโจทย์แบบนั้น ถ้ามันตอบไม่ได้ว่าราว 63% ก็อย่าเพิ่งเชื่อตัวเลขอื่นที่มันคิดให้

ผลที่มาถึงคุณมักผ่านการคัดมาแล้วหนึ่งรอบ คือรอบที่คนได้ผลเลือกจะเล่า และคนไม่ได้ผลเลือกจะเงียบ ก่อนจะเชื่อผลไหน ถามว่ามีคนลองแล้วไม่ได้อีกกี่คน และทำไมคุณไม่เคยได้ยินเสียงของพวกเขາเลย

เก้าสิบเก้าคนนั้นไม่ได้หายไปไหน พวกเขาแค่ไม่ได้โพสต์

ลองคิดนี้: เปิดโทรศัพท์ หาโพสต์อวดผลหนึ่งโพสต์ จะเป็นเลขเด็ดที่ถูกหรือริ้วว่าใช้แล้วหายก็ได้ แล้วเขียนตัวเลขสองตัวลงกระดาษ คือจำนวนที่ได้ตามที่โพสต์บอก กับจำนวนคนหรือจำนวนครั้งที่ลองทั้งหมด ถ้าคุณกับคนในบ้านช่วยกันหาตัวที่สองแล้วไม่เจอ จดไว้ว่าผลนี้ยังไม่มีตัวหาร

จะรู้วาทน์ผิดเมื่อ: ใครก็ตามใช้โปรแกรมตารางคำนวณจำลองคนร้อยคนใบเลขท้ายสองตัวแบบสุ่มเป็นพันงวด แล้วพบว่าคนใบถูกเฉลี่ยต่องวดห่างจากหนึ่ง หรืองวดที่มีคนถูกอย่างน้อยหนึ่งคนห่างจากราว 63 ใน 100 อย่างชัดเจน หรือการให้หลายทีมวิเคราะห์ข้อมูลชุดเดียวกันในสาขาอื่น ได้คำตอบแทบไม่ต่างกันเลย

หลักฐาน:

- ร้อยคนเดาเลขท้ายสองตัวแบบสุ่ม โอกาสคนละหนึ่งในร้อยต่องวด · เฉลี่ยถูกงวดละหนึ่งคน · งวดที่มีคนถูกอย่างน้อยหนึ่งคน 63% ไม่มีใครถูกเลย 37% · ร้อยคนที่เลือกเลขไม่ซ้ำกันจะมีคนถูกทุกงวด · คนเดียวถูกสองงวดติดกัน หนึ่งในหมื่น · หนึ่งในหมื่นคนให้คนถูกสองงวดติดกันเฉลี่ยหนึ่งคน — **แนน** (เลขคณิต ใครก็ตามคำนวณซ้ำได้) — $100 \times 1/100 = 1 \cdot 1 - 0.99^{100} = 0.634 \cdot (1/100)^2 = 1/10,000 \cdot 10,000 \times (1/100)^2 = 1 \cdot$ คำนวณซ้ำวันที่ 25 กันยายน 2026
- เหรียญที่ไม่เอียง โยนสิบครั้ง ได้หัวแปดครั้งขึ้นไป 56 แบบจาก 1,024 แบบ หรือราวหนึ่งในสิบแปด · ในคนยี่สิบคนที่โยนคนละสิบครั้ง โอกาสที่จะมีใครสักคนได้แปดหัวขึ้นไป 67.5% — **แนน** (เลขคณิต) —

การแจกแจงทวินาม (binomial) $n = 10$, $p = 1/2 \cdot 1 - (968/1,024)^{20} = 0.675$ · จำนวนวันที่ 25 กันยายน 2026

- ยี่สิบรอบที่แต่ละรอบมีโอกาสถูกหนึ่งในแปด ของที่ไม่ทำงานเลยได้ศูนย์ถึงห้าครั้ง 96.9% ได้สามครั้งขึ้นไป 46.5% ค่าที่คาดหวัง 2.5 — แน่น (เลขคณิต) — การแจกแจงทวินาม $n = 20$, $p = 1/8$ · จำนวนวันที่ 25 กันยายน 2026
- การทดสอบที่เม็กซิโก วันที่ 21 ตุลาคม 2011 · ปิดตาสองชั้น เครื่องหาเจอ 3 จาก 20 เมื่อความบังเอิญ คาดไว้ 2.5 · เหรียญที่ใช้เทียบ 1 จาก 20 · ทหารเซ็นยืนยันเงื่อนไขก่อนเริ่ม — **พอเชื่อ** (การทดสอบหนึ่งครั้ง บันทึกละเอียดทุกชั้นและมีวิดีโอ) — Mochán, W. L., & Ramírez-Solís, A. (2013). Effectiveness of the GT200 Molecular Detector: A Double-Blind Test. arXiv:1301.3971 — <https://arxiv.org/abs/1301.3971>
- ผู้บัญชาการทหารบกในเวลานั้นยืนยันว่าเครื่องเจอของจริงกว่าสามร้อยครั้ง แหล่งนี้ไม่มีตัวเลขครั้งที่ไม่เจอ — **พอเชื่อ** (รายงานข่าวย้อนหลัง แหล่งไม่ได้ระบุวันที่ของคำยืนยัน) — THE STANDARD, 6 มิถุนายน 2022 — <https://thestandard.co/gt200-history/>
- เจ็ดสิบทีมอิสระวิเคราะห์ข้อมูลภาพสแกนสมองชุดเดียวกัน ตอบสมมติฐานเก่าข้อเดียวกัน ไม่มีสองทีมเลือกวิธีวิเคราะห์เหมือนกันทุกชั้น และผลการทดสอบสมมติฐานต่างกันไม่น้อย — **พอเชื่อ** (งานชิ้นเดียวแต่ทำโดยเจ็ดสิบทีมที่ทำงานแยกกัน) — Botvinik-Nezer, R., et al. (2020). Nature · Europe PMC 32483374 — <https://europepmc.org/article/MED/32483374>

7 · คำถามที่สาม · ถ้าผิด จะเห็นอะไร

ในญี่ปุ่น คำถามว่า "เลือดกรุ๊ปอะไร" ไม่ได้จบแค่ในวงคุยเล่น มันเคยตามคนไปถึงการหางาน ทางแยกใหญ่ทางหนึ่งของชีวิต

ความเชื่อที่ว่ากรุ๊ปเลือดบอกนิสัยได้ มีคนตรวจแล้วด้วยข้อมูลสำรวจแบบสุ่มจากคนกว่าหนึ่งหมื่นคนในสองประเทศ ลองทายก่อนอ่านต่อ คนเรานิสัยต่างกันไป ถ้าแบ่งความต่างทั้งหมดนั้นเป็นร้อยละ ส่วนที่เดาได้จากกรุ๊ปเลือดอย่างเดียวมีกี่ส่วน

คำทำนายที่ไม่มีผลลัพธ์ไหนทำให้มันผิดได้ ไม่ใช่คำทำนายที่แม่นยำ มันคือคำทำนายที่ไม่ได้บอกอะไรเลย

ความเชื่อเรื่องกรุ๊ปเลือดพูดได้ทั้งสองแบบ แบบที่ไม่มีวันผิด กับแบบที่ผิดได้ และคำตอบของข้อที่ให้ทายมาจากแบบที่สอง

ทำไมมันฟังดูจริง

ก่อนจะดูตัวเลข ความเชื่อนี้ควรได้น้ำหนักเต็มของมัน กรุ๊ปเลือดเป็นของจริง มันคือสารบนผิวเม็ดเลือดแดงที่ยีนของแต่ละคนกำหนด และส่งต่อจากพ่อแม่ถึงลูก การสงสัยว่าของที่อยู่ในเลือดของทุกคนแบบนี้จะเกี่ยวกับนิสัยบ้างไหม จึงไม่ใช่คำถามไร้สาระ

ถ้าคุณอ่านคำบรรยายนิสัยของกรุ๊ปตัวเอง หลายข้ออาจตรงจริงๆ และคนกรุ๊ปเอบางคนที่ขี้กังวลจริงๆ

แต่ความรู้สึกว่าตรงเกิดได้จากเหตุอื่นอีกอย่างน้อยสองเหตุ เหตุแรกนักจิตวิทยาเรียกว่าปรากฏการณ์บาร์นัม (Barnum effect) คือเมื่อได้ยินประโยคที่ใช้ได้กับเกือบทุกคน คนฟังกลับรู้สึกว่ามันบอกลักษณะของตัวเองได้ตรง

เหตุที่สองคือคำทำนายที่ทำให้ตัวเองเป็นจริง (self-fulfilling prophecy) คนที่เชื่อว่า "ฉันกรุ๊ปเอ ฉันเลยขี้กังวล" อาจเริ่มมองตัวเองแบบนั้น และค่อยๆ ทำตัวแบบนั้นโดยไม่รู้ตัว

ทั้งสองเหตุทำให้คำบรรยายรู้สึกแม่นยำ โดยที่กรุปเลือดไม่ต้องทำอะไรเลย เว็บไซต์ให้คะแนนในบทที่สามจัดความเชื่อนี้เป็นวิทยาศาสตร์เทียมเช่นเดียวกับ EM เหตุผลข้อหนึ่งคือข้อมูลของฝ่ายที่เชื่อไม่ได้ถูกออกแบบให้ตัดสองเหตุนี้ออก

สองประโยค ความเชื่อเดียว

ความเชื่อที่พูดได้แบบหนึ่งว่า "คนกรุปเอมีแนวโน้มจะซีงกังวล แต่ก็ไม่ใช่ทุกคน" ประโยคนี้ฟังรอบคอบดี ที่นี้ถามคำถามที่สามกับมัน ถ้าผิด จะเห็นอะไร

ถ้าคนกรุปเอที่คุณเจอซีงกังวล ประโยคนี้ถูก ถ้าไม่ซีงกังวล ก็เป็นส่วนที่ไม่ใช่ทุกคน มันไม่ได้บอกว่าต่างกันมากแค่ไหน จึงไม่มีผลลัพธ์ไหนทำให้มันผิดได้ และด้วยเหตุผลเดียวกัน มันไม่ได้บอกอะไรที่คุณเอาไปใช้ได้

เว็บไซต์เดียวกันชี้ปัญหานี้ไว้ในเกณฑ์ข้อแรก คำว่านิสัยในความเชื่อนี้ไม่ได้ระบุว่าเป็นนิสัยของใคร ในสถานการณ์ไหน ดูจากการกระทำอะไร มันจึงตีความได้ทุกทาง และเมื่อมีข้อมูลว่าคนกรุปเอไม่ได้ซีงกังวล เว็บไซต์บันทึกคำตอบของฝ่ายที่เชื่อไว้สองแบบ แบบหนึ่งคือการวินิจฉัยไม่ได้แม่นยำเปอร์เซ็นต์ อีกแบบคือหยิบข้อมูลที่วัดคนละเรื่องมาอ้างแทน

อีกแบบหนึ่งพูดว่า ถ้าสุ่มคนมาเป็นหมื่นแล้วถามคำถามเดียวกัน คนกรุปเอจะตอบเรื่องความซีงกังวลต่างจากกรุปอื่นจนเห็นได้ ถ้าทุกกรุปตอบออกมาพอๆ กัน ประโยคนี้แพ้ย

ข้อมูลที่ให้คุณทนายไว้ตอนต้นบทคือการทดสอบประโยคแบบนี้ งานวิจัยชิ้นหนึ่งในญี่ปุ่นนำข้อมูลสำรวจแบบสุ่มสามชุด จากญี่ปุ่นสองชุดและสหรัฐฯ หนึ่งชุด มาวิเคราะห์ใหม่ คำถามเรื่องนิสัยหกสิบแปดข้อ หกสิบห้าข้อไม่พบความต่างระหว่างกรุปเลือด และในทุกข้อ ส่วนที่เดาได้จากกรุปเลือดมีไม่ถึง 0.3 ส่วนจากร้อย

นั่นคือคำตอบของข้อที่ให้ทนาย ถ้าความต่างเรื่องนิสัยทั้งหมดเป็นเงินร้อยบาท ส่วนของกรุปเลือดไม่ถึงสามสิบสตางค์ ตามข้อมูลชุดนี้ ถ้ารู้กรุปเลือดของใครสักคน คุณแทบไม่ได้รู้อะไรเพิ่มเกี่ยวกับนิสัยของเขาเลย

สิ่งที่ทำให้ได้คำตอบนี้ไม่ใช่เครื่องมือซับซ้อน มันคือการยอมเขียนความเชื่อใหม่ให้แพ้ได้ก่อน ประโยคแบบแรกทดสอบไม่ได้ ต่อให้สำรวจคนเป็นล้าน ประโยคแบบที่สองทดสอบได้ และตามข้อมูลชุดนี้ มันแพ้

ฝ่ายที่เชื่อก็มีตัวเลขของตัวเอง กลุ่มผู้สนับสนุนกลุ่มหนึ่งรายงานว่แบบสอบถามราวหนึ่งหมื่นคนพบความต่างที่มีนัยสำคัญทางสถิติ เว็บไซต์ให้คะแนนไม่ยอมรับผลนั้น เพราะการออกแบบไม่ได้ตัดสองเหตุข้างต้นและความเอนเอียงอื่นออก เล่มนี้เปิดผลสำรวจต้นฉบับไม่ได้ จึงพิมพ์ไว้ตามที่เว็บไซต์นั้นเล่า

เว็บไซต์เดียวกันยังเขียนว่างานวิจัยทางการแพทย์เรื่องกรู๊ปเลือดกับโรคบางโรคมีผลออกมาแล้ว คำตัดสินจึงไม่ได้อยู่ที่กรู๊ปเลือด มันอยู่ที่รูปร่างของข้ออ้าง

ใส่ต่อไปจนกว่าจะเห็นผล

ข้ออ้างที่ไม่อยากแพ้มีทางหลบอยู่ไม่กี่ทาง และแต่ละทางมีหน้าตาที่จำได้ ทางแรกคือ พูดยาวจนไม่มีอะไรผิด อย่างประโยคแรกเรื่องกรู๊ปเลือด

ทางที่สองคือตกลงไว้ล่วงหน้าว่าผลที่ไม่ออกแปลว่ายังไม่พอ ผู้พัฒนาจุลินทรีย์ EM ที่อยู่ในลูกบอล EM ของบทที่สาม เขียนไว้ในคอลัมน์ภาษาญี่ปุ่นของเขาเองว่า หลักของการใช้ EM บำบัดน้ำคือ ใส่ต่อไปจนกว่าจะเห็นผล และเมื่อการทดลองของหน่วยงานอื่นไม่พบผล เขาก็อธิบายว่าเพราะไม่ได้ใช้จนผลออก

เหตุผลของเขามีน้าหนัก EM เป็นสิ่งมีชีวิต ต้องเพิ่มจำนวนก่อนจึงจะเห็นผล แต่ถามคำถามที่สามกับหลักข้อนี้ดู ถ้า EM ไม่ได้ผล จะเห็นอะไร คำตอบคือไม่มีอะไรต่างไป ผลที่ไม่ออกจะถูกอ่านว่ายังใส่น้ำไม่พอเสมอ

หลักข้อนี้จะแพ้ได้ทันทีที่มีคนเขียนไว้ก่อนว่าใส่เท่าไร นานแค่ไหน แล้วถ้าครบแล้วยังไม่เห็นผล ให้นับว่าไม่ได้ผล

ทางที่สามเจียบกว่านั้น คือกำหนดว่าใครมีสิทธิ์ทดสอบ คอลัมน์เดียวกันเขียนว่า หน่วยงานวิจัยไม่มีสิทธิ์ทดลองเองแล้วตัดสินผลของ EM โดยไม่ได้รับความยินยอมจากองค์กรของเขา ข้ออ้างที่คนอื่นทดสอบได้เฉพาะเมื่อเจ้าของอนุญาต คือข้ออ้างที่เลือกได้เองว่าจะแพ้เมื่อไร

เขาเขียนไว้ด้วยว่าพร้อมโต้แย้งต่อหน้าสาธารณะทุกเมื่อ แต่เวทีโต้แย้งไม่ใช่การทดสอบ ในการทดสอบ ข้ออ้างแพ้ให้ผลลัพธ์ ไม่ได้แพ้ให้คนที่พุดเก่งกว่า

ทางที่สี่คือข้อแก้ตัวที่มาหลังเห็นผล และห้องบอลรูมในบทที่หนึ่งมีทั้งทางหลบนี้และเครื่องมือที่ปิดมันอยู่ในห้องเดียวกัน หลังहारอบแรกพลาดหมด ทหารผู้ใช้เครื่องก็มีเหตุผลให้ ตั้งแต่กลับกันจากของของกลางไปจนถึงบันไดที่ทำให้หอบ ข้อแก้ตัวแบบนี้ต่อได้กับทุกผล เจอก็เพราะเครื่องทำงาน ไม่เจอก็เพราะมีอะไรมารบกวน

แต่ก่อนเริ่ม ทหารทั้งสองคนเซ็นกระดาศไว้แล้วว่าห้อง กล่อง และของกลาง ไม่มีอะไรขัดขวางเครื่อง กระดาศแผ่นนั้นคือคำตอบของคำถามที่สาม ที่เขียนไว้ก่อนใครจะรู้ผล ถ้าเครื่องหาไม่เจอในห้องนี้ กับกล่องเหล่านี้ ความผิดเป็นของเครื่อง

ข้อแก้ตัวไม่ได้หายไปเพราะกระดาศแผ่นนั้น มันแค่ต้องสู้กับลายเซ็นของคนพุดเอง ที่เซ็นไว้ก่อนจะมีเหตุให้ต้องแก้ตัว

เขียนใหม่ให้แพ้ได้

ข้ออ้างที่หลบอยู่ เขียนใหม่ให้แพ้ได้ด้วยห้าช่อง และคุณไม่ต้องรอให้เจ้าของข้ออ้างเขียนเอง

1. ใคร กับคนแบบไหน ก็คน
2. อะไร วัดอะไร และวัดด้วยอะไร
3. เท่าไร ต้องต่างแค่ไหนถึงจะนับ
4. เทียบกับอะไร กับคนที่ไม่ได้ใช้ หรือใช้ของอย่างอื่น
5. เมื่อไร ภายในกี่วัน กี่เดือน

แล้วปิดท้ายด้วยบรรทัดเดียว ถ้าผิด จะเห็นอะไร

สมมติว่าแอปฝึกสมองแอปหนึ่งโฆษณาว่า "เล่นวันละสิบนาที สมองไวขึ้น" ประโยคนี้อาจอยู่ในคำว่าไวขึ้น เล่นเกมในแอปทุกวันไปสามเดือน คะแนนในแอปย่อมสูงขึ้น เพราะทำอะไรซ้ำทุกวันก็เก่งขึ้นในสิ่งนั้น นั่นคือเก่งเกม ยังไม่ใช่สมองไวขึ้น

เขียนใหม่ได้แบบนี้ ผู้ใหญ่อายุหกสิบขึ้นไปที่เล่นวันละสิบนาทายเป็นเวลาสามเดือน จะจํารายการของที่ต้องซื้อได้มากขึ้นอย่างน้อยสองรายการ ในแบบทดสอบที่ไม่มีอยู่ในแอป เมื่อเทียบกับคนอายุเท่ากันที่จับฉลากให้ไปเล่นเกมทั่วไปวันละสิบนาทีก่อน ตัวเลขในประโยคนี้นี้เป็นแค่ตัวอย่าง คนที่อ้างต้องเป็นคนเดิมเอง

บรรทัดสุดท้ายจึงเขียนได้ทันที ถ้าครบสามเดือนแล้วสองกลุ่มจำได้พอๆ กัน ข้ออ้างนี้แพ้มันอาจชนะก็ได้ และนั่นคือประเด็น ข้ออ้างที่ยอมเสี่ยงแพ้นั่น ที่ชนะแล้วมีความหมาย

วิธีเดียวกันใช้กับคำทำนายได้ทุกแบบ ไม่ว่าจะมาจากดวงดาว จากไฟ จากกรุปเลือด หรือจากปากนักวิทยาศาสตร์ เล่มนี้ไม่ได้ถามว่าคนที่ไปดูดวงคิดถูกหรือผิด คำถามอยู่ที่รูปร่างของประโยค

ข้ออ้างทุกข้อที่บอกว่าโลกทำงานอย่างไร ถ้ามได้ด้วยประโยคเดียว ถ้าผิด จะเห็นอะไร ข้ออ้างที่ตอบได้อาจถูกหรือผิดก็ได้ แต่มันยอมให้โลกเป็นคนตัดสิน ข้ออ้างที่ตอบไม่ได้ ยังไม่ได้บอกอะไรเกี่ยวกับโลกเลย ส่วนศรัทธาและคำถามว่าอะไรคืออะไรควร เป็นอีกเรื่องหนึ่ง บทนี้ไม่ได้ตัดสินมัน

ข้ออ้างที่กล่าวบอกว่าอะไรจะทำให้มันผิด คือข้ออ้างเดียวที่คุ้มกับเวลาที่ คุณจะใช้ทดสอบ

ลองคิดนี้: หยิบข้ออ้างหนึ่งข้อจากโฆษณาหรือกลุ่มไลน์ในโทรศัพท์ แล้วเขียนใหม่ลงกระดาษให้มีครบห้าช่อง ใคร อะไร เท่าไร เทียบกับอะไร เมื่อไร ต่อท้ายด้วยบรรทัดเดียวว่าผลแบบไหนจะทำให้มันแพ้ ช่องไหนที่คุณกับคนในบ้านช่วยกันเติมจากข้อความเดิมไม่ได้เลย ช่องนั้นคือที่ที่ข้ออ้างซ่อนตัว

จะรู้ว่าบทนี้ผิดเมื่อ: การสำรวจแบบสุ่มที่ใหญ่กว่าหนึ่งหมื่นคน ซึ่งกำหนดไว้ก่อนเก็บข้อมูลว่าจะวัดนิสัยข้อไหน และไม่บอกผู้ตอบว่ากำลังศึกษาเรื่องกรุปเลือด พบว่ากรุปเลือดอธิบายนิสัยได้มากกว่างานในบทนี้อย่างชัดเจน และผลนั้นเกิดซ้ำในอีกประเทศหนึ่ง

หลักฐาน:

- กรู๊ปเลือดอธิบายความแตกต่างเรื่องนิสัยได้ไม่ถึง 0.3% (ค่า eta squared ต่ำกว่า .003) · 65 จาก 68 ข้อไม่พบความต่างระหว่างกรู๊ปเลือด · ข้อมูลสำรวจแบบสุ่มสามชุด ญี่ปุ่นปี 2004 และ 2005 สหรัฐฯ ปี 2004 ผู้วิจัยรายงานว่ารวมกว่าหนึ่งหมื่นคน จำนวนที่ใช้ในแต่ละข้อรวมสามชุดราว 9,500 ถึง 9,700 คน — **พอเชิ้อ** (งานชิ้นเดียว วิเคราะห์ข้อมูลสุ่มสามชุดจากสองประเทศ) — Nawata, K. (Nawata, K.) (2014). Unrelatedness between blood type and personality. Japanese Journal of Psychology (Shinrigaku Kenkyu), 85(2), 148–156 — https://www.jstage.jst.go.jp/article/jjpsy/85/2/85_85.13016/_article/-char/ja/
- เว็บไซต์ประเมินข้ออ้างของมหาวิทยาลัยเมจิใช้เกณฑ์สิบข้อในสี่หมวด และตัดสินว่าความเชื่อเรื่องกรู๊ปเลือดกับนิสัยเป็นวิทยาศาสตร์เทียม · ข้อมูลฝ่ายที่เชื่อไม่ได้ตัดปรากฏการณ์บาร์นัมและคำทำนายที่ทำให้ตัวเองเป็นจริงออก · คำว่านิสัยไม่ได้ระบุว่าเป็นใคร ในสถานการณ์ไหน ดูจากการกระทำอะไร · เมื่อข้อมูลขัด ฝ่ายที่เชื่อตอบว่าวินิจฉัยไม่ได้แม่นยำร้อยเปอร์เซ็นต์ หรือหยิบข้อมูลที่วัดคนละเรื่องมาแทน · ความเชื่อนี้เคยถูกใช้แม้ในการหางาน · งานวิจัยทางการแพทย์เรื่องกรู๊ปเลือดกับโรคมะเร็งออกมาแล้ว · กรู๊ปเลือดคือสารบนผิวเม็ดเลือดแดงที่ยีนกำหนด — **พอเชิ้อ** (การประเมินของหน่วยงานเดียว ที่อ้างงานวิจัยหลายชิ้นรวมทั้งงานข้างบน) — Gijika.com, Science Communication Laboratory, Meiji University, หน้าเรื่องการทายนิสัยจากกรู๊ปเลือด ปรับปรุงล่าสุด 10 ธันวาคม 2019 — https://gijika.com/rate/se_blood_type.html
- กลุ่มผู้สนับสนุนกลุ่มหนึ่งรายงานความต่างที่มีนัยสำคัญจากแบบสอบถามราวหนึ่งหมื่นคน และเว็บไซต์ประเมินไม่ยอมรับการออกแบบของมัน เพราะไม่ได้ตัดปรากฏการณ์บาร์นัม คำทำนายที่ทำให้ตัวเองเป็นจริง ความเอนเอียงเข้าข้างความเชื่อเดิม และปรากฏการณ์อื่นออก — **ตรวจไม่ได้** (เล่นนี้รู้จากคำเล่าของเว็บไซต์ประเมินเท่านั้น เปิดผลสำรวจต้นฉบับไม่ได้) — https://gijika.com/rate/se_blood_type.html
- ปรากฏการณ์บาร์นัม คือการรู้สึกว่าคุณได้กับทุกคนบอกลักษณะของตัวเองได้ตรง · คำทำนายที่ทำให้ตัวเองเป็นจริง คือแนวโน้มที่คนจะทำตามคำทำนายโดยไม่รู้ตัว จนดูเหมือนคำทำนายแม่น — **พอเชิ้อ** (นิยามจากพจนานุกรมศัพท์ของเว็บไซต์เดียวกัน) — <https://gijika.com/rate/dict/dictHa02.html> · <https://gijika.com/rate/dict/dictSa07.html>
- ผู้พัฒนา EM เขียนว่าหลักของการใช้ EM บำบัดน้ำคือใส่ต่อไปจนกว่าจะเห็นผล · การทดลองที่ไม่พบผลคือการที่ไม่ได้ใช้จนผลออก · EM เป็นสิ่งมีชีวิต ต้องเพิ่มจำนวนก่อนจึงเห็นผล · หน่วยงานวิจัยไม่มีสิทธิ์ทดลองและตัดสินผลของ EM โดยไม่ได้รับความยินยอมจากองค์กรวิจัย EM · พร้อมโต้แย้งต่อสาธารณะทุกเมื่อ — **แนน** (คำของผู้พัฒนาเองในคอลัมน์ของเขา และเว็บไซต์ประเมินของญี่ปุ่นรายงานจุดยืนเรื่องความยินยอมตรงกัน) — Higa Teruo, คอลัมน์ Shin Yume ni Ikiru ตอนที่ 62 บนเว็บไซต์ Ecopure, 3 สิงหาคม 2012 — <https://www.ecopure.info/oldweb/rensai/teruohiga/yumenikiru62.html> · https://gijika.com/rate/le_effective_microorganisms.html
- ทหารสองคนเซ็นยืนยันก่อนเริ่มว่าห้อง กล้อง และของกลางไม่ขัดขวางเครื่อง แล้วให้เหตุผลว่าเครื่องอาจไม่ทำงานเพราะคนถือเหนื่อย กลิ่นกัญชา บันได และการไม่มีสุนัข — **พอเชิ้อ** (การทดสอบหนึ่งครั้ง บันทึกละเอียดทุกขั้นและมีวิดีโอ) — Mochán, W. L., & Ramírez-Solís, A. (2013). Effectiveness of the GT200 Molecular Detector: A Double-Blind Test. arXiv:1301.3971 — <https://arxiv.org/abs/1301.3971>

8 . คำถามที่สี่ . ใครตรวจแล้วบ้าง

ข้อความหนึ่งถูกส่งต่อกันมาว่า ติดเชื้อเอชไอวีก็รักษาหายได้ในสองเดือน ใช้ผลิตภัณฑ์ชนิดหนึ่งแค่สองขวด

ถ้าคุณไม่ใช่หมอ คำถามว่ามันจริงไหมดูจะเกินมือ งานวิจัยเรื่องเอชไอวีเขียนด้วยภาษาที่คนนอกวงการอ่านไม่ออก และต่อให้อ่านออก ก็ไม่มีใครมีเวลาไปอ่านงานวิจัยของทุกข้อความที่ถูกส่งต่อมา

คุณไม่ต้องอ่านงานวิจัยเป็นเพื่อรู้ว่าควรเชื่อใคร คุณต้องรู้ว่าตรงไหนที่มีคนพยายามจับผิดเรื่องนี้มาแล้ว

ก่อนกลับไปทีข้อความนั้น ลองทนายผลของการทดลองชิ้นหนึ่ง นักวิจัยสองคนให้คนสามกลุ่มตัดสินใจว่าเว็บไซต์จริงชุดเดียวกันเชื่อถือได้แค่ไหน กลุ่มแรกเป็นนักประวัติศาสตร์ระดับปริญญาเอกสิบคน กลุ่มที่สองเป็นนักศึกษาปริญญาตรีของมหาวิทยาลัยสแตนฟอร์ดยี่สิบห้าคน กลุ่มที่สามเป็นพนักงานตรวจข้อเท็จจริงมืออาชีพสิบคน

กลุ่มไหนตัดสินใจได้ดีที่สุด และกลุ่มนั้นทำอะไรที่อีกสองกลุ่มไม่ได้ทำ เก็บคำตอบของคุณไว้ก่อน

อ่านน้อยกว่า แต่ถูกกว่า

ในการทดลองนั้น คนตรวจข้อเท็จจริงชนะ นักวิจัยรายงานว่าข้อสรุปของพวกเขามีเหตุผลรองรับดีกว่าอีกสองกลุ่ม และใช้เวลาเพียงเสี้ยวเดียวของสองกลุ่มนั้น ที่น่าสนใจกว่าผลแพ้ชนะคือวิธีที่พวกเขาชนะ พวกเขาอ่านหน้าเว็บน้อยกว่าใคร

นักประวัติศาสตร์กับนักศึกษาอ่านแต่ละเว็บไซต์อย่างตั้งใจ ไล่จากบนลงล่าง และอยู่ในหน้านั้นเพื่อชั่งน้ำหนักน่าเชื่อถือแค่ไหน นักวิจัยเรียกวิธีนี้ว่าอ่านแนวดิ่ง และรายงานว่าคนที่อ่านแบบนี้ถูกหลอกบ่อยด้วยของที่ปลอมได้ง่าย คือโลโก้ที่ดูเป็นทางการ กับชื่อโดเมนที่ฟังดูน่าเชื่อถือ

คนตรวจข้อเท็จจริงกวาดตาดูหน้านั้นครู่เดียว แล้วออกไปเปิดแท็บใหม่ ค้นคว้าที่อื่น พุดถึงเจ้าของเว็บไซต์นี้ว่าอย่างไร นักวิจัยเรียกวิธีนี้ว่าอ่านแนวขวาง (lateral reading)

ในงานนี้ ความต่างอยู่ที่ทิศทางของการอ่าน การตั้งคำถามกับเอกสารคืออาชีพของนักประวัติศาสตร์ แต่ครั้งนี้พวกเขาถามผิดที่ พวกเขาถามหน้าเว็บว่ามันน่าเชื่อถือไหม และหน้าเว็บก็ตอบว่าน่าเชื่อถือได้เสมอ เพราะคนทำหน้านั้นเป็นคนเขียนคำตอบไว้เอง

การทดลองนี้มีคนทั้งหมดแค่สี่สิบห้าคน และเล่มนี้อ่านจากบทคัดย่อ จึงควรถือเป็นเบาะแสที่น่าตามต่อ ยังไม่ใช่ข้อยุติ

นักวิจัยกลุ่มเดียวกันลองต่อว่าวิธีนี้สอนกันได้ไหม พวกเขาแนะนำวิธีของคนตรวจข้อเท็จจริงไปสอนนักเรียนมัธยมปลายหกคาบ คาบละห้าสิบนาที หัวใจของบทเรียนคือการอ่านแนวขวาง เจอเว็บไซต์ที่ไม่คุ้น ให้ออกไปค้นข้างนอกก่อน แล้วค่อยเสียเวลากับหน้านั้น

เมื่อเทียบกับเพื่อนในห้องเรียนปกติ นักเรียนที่ได้เรียนตัดสินความน่าเชื่อถือของข้อมูลออนไลน์ได้ดีขึ้นมากกว่า และห่างพอที่นักวิจัยบอกได้ว่าไม่น่าใช้ความบังเอิญ ห้องเรียนสองกลุ่มนี้ถูกจับคู่ให้คล้ายกัน ไม่ได้ถูกแบ่งด้วยการสุ่ม จึงเป็นการเทียบที่ดี แต่ยังไม่ดีที่สุดตามคำถามในบทที่ห้า

ห้าแสนกว่าครั้ง

ข้อความเรื่องเอชไอวีเป็นข่าวปลอม ศูนย์ต่อต้านข่าวปลอมประเทศไทยนับไว้ว่า ในบรรดาข่าวปลอมด้านสุขภาพที่ศูนย์เผยแพร่คำชี้แจงในปี 2025 เรื่องนี้มีคนสนใจและมีส่วนร่วมมากที่สุด 156,828 ครั้ง

แต่มันไม่ใช่อันดับหนึ่งของปี ในรายการข่าวปลอมทั้งหมดที่ศูนย์ชี้แจง เรื่องที่มีส่วนร่วมมากที่สุดคือข่าวลือว่าชายฝั่งอันดามันเสี่ยงเกิดสึนามิที่อาจรุนแรงกว่าปี 2004 ยอดคือ 529,472 ครั้ง เกินสามเท่าของเรื่องเอชไอวี

ตัวเลขแบบนี้วัดความสนใจ ไม่ได้วัดการตรวจ ข้าวปลอมเรื่องเดียวมีคนสนใจและมี
ส่วนร่วมได้ถึงห้าแสนกว่าครั้ง ยอดโลกและยอดแชร์จึงบอกคุณได้แค่ว่าเรื่องไหนดัง
มันไม่เคยบอกว่าเรื่องไหนผ่านการตรวจ

การตรวจของทั้งสองเรื่องเกิดขึ้นที่อื่น ในหน่วยงานที่มีหน้าที่ตรวจเรื่องนั้นโดยตรง
ผลิตภัณฑ์ในข้อความนั้นชื่อ CDS และสำนักงานคณะกรรมการอาหารและยาชี้แจงว่า
มันไม่ใช่ยา ไม่เคยได้รับอนุญาตให้ใช้รักษาโรคใดเลย และการกินมันอาจเป็นอันตราย
ถึงชีวิต

ส่วนกรมอุตุนิยมวิทยาชี้แจงว่า จากการประเมินและคำนวณ แผ่นเปลือกโลกต้องใช้
เวลาอีกราวสี่ร้อยถึงหกร้อยปี จึงจะปลดปล่อยพลังงานได้รุนแรงเท่าแผ่นดินไหวปี
2004 ในตอนนี้แผ่นดินไหวขนาดนั้นจึงยังเกิดซ้ำไม่ได้ คำชี้แจงนี้ไม่ได้พูดถึงแผ่นดิน
ไหวที่เล็กกว่านั้น

เรื่องเงินก็ไม่ต่างกัน ข้าวปลอมด้านเศรษฐกิจอันดับหนึ่งของปีเดียวกันคือเพจเฟซบุ๊ก
ที่อ้างว่าธนาคารออมสินปล่อยกู้ตั้งแต่หนึ่งหมื่นถึงหนึ่งล้านบาท ไม่ต้องมีคนค้ำ ไม่เช็ค
เครดิต อนุมัติไว ธนาคารยืนยันว่าไม่เกี่ยวข้องกับเพจนั้นเลย

เพจนั้นยืมชื่อธนาคารมาใช้ แบบเดียวกับเว็บไซต์ที่หลอกล่อนนักประวัติศาสตร์ด้วยโลโก้
ชื่อบนหน้าเป็นของที่ใครก็พิมพ์ได้ คำยืนยันจากเจ้าของชื่อเป็นของที่คุณต้องออกไป
หาเอง

ที่อยู่ของการตรวจ

เอกสารสองชิ้นที่วางหลักว่าคนทั่วไปควรตัดสินใจอ้างอิงทางวิทยาศาสตร์อย่างไร
ดูเหมือนจะแนะนำตรงข้ามกัน

กรอบการประเมินวิทยาศาสตร์ของการสอบนานาชาติ PISA 2025 เขียนว่าการ
ประเมินข้อมูลวิทยาศาสตร์ของคนทั่วไป "starts not by interrogating the
findings" คือไม่ได้เริ่มจากการซักไซ้ผลวิจัย แต่เริ่มจากการถามว่าแหล่งมีผล
ประโยชน์ทับซ้อนไหม ผู้รู้ในเรื่องนั้นเห็นตรงกันหรือเปล่า และแหล่งนั้นเชี่ยวชาญ
เรื่องนี้จริงไหม

ส่วนแนวคิดหลักของ Informed Health Choices (IHC) กลุ่มเดียวกับที่สอนเด็ก ประถมในบทที่ห้า จัดให้ความเชื่อในความเห็นของผู้เชี่ยวชาญ ความเชื่อว่าตีพิมพ์ แล้วต้องจริง และความเชื่อในงานวิจัยชิ้นเดียว อยู่ในรายการเหตุผลที่ผิด

ทั้งสองถูก เพราะพูดถึงคนละชั้นของเรื่องเดียวกัน ความเชี่ยวชาญมีค่าเพราะมันเป็น ร่องรอยของการตรวจ สาขาวิชาหนึ่งคือที่ที่คนจำนวนมากใช้เวลาทั้งอาชีพทำซ้ำ แยก และจับผิดงานของกันและกัน ผู้เชี่ยวชาญที่พูดเรื่องในสาขาของตัวเองจึงชี้ให้คุณเห็น ว่าการตรวจเกิดขึ้นที่ไหน แต่คำพูดของเขาไม่ใช่ตัวการตรวจ PISA สอนให้คุณหาที่อยู่ ของการตรวจ IHC เตือนไม่ให้คุณเอาที่อยู่นั้นมาใช้แทนตัวการตรวจ

ที่อยู่นั้นหายไปทันทีที่ผู้เชี่ยวชาญก้าวออกนอกสาขาของตัวเอง เอกสารของสมาคม เพื่อความก้าวหน้าทางวิทยาศาสตร์แห่งอเมริกาปี 1989 เขียนไว้ตรงๆ ว่า นอกสาขาที่ เชื่อว่า ความเห็นของนักวิทยาศาสตร์ไม่น่าเชื่อเป็นพิเศษ

ฉันทามติ (consensus) จึงมีน้ำหนักกว่าคนคนเดียว กรอบของ PISA เองเขียนว่าความ ผิดพลาดเป็นเรื่องปกติของวิทยาศาสตร์ และฉันทามติน่าเชื่อถือกว่าคนคนใด ฉันทามติ ไม่ใช่ยอดโหวต มันคือสิ่งที่ยั่งยืนอยู่หลังจากคนจำนวนมากพยายามล้มมันแล้ว

เว็บไซต์ให้คะแนนข้ออ้างของมหาวิทยาลัยเมจิในญี่ปุ่น ที่ตัดสินความเชื่อเรื่องกรุป เลือดในบทที่เจ็ด มีเกณฑ์สิบข้อ สองข้อในนั้นคือคำถามของบทนี้ ข้ออ้างนั้นเปิดให้คน นอกกลุ่มผู้สนับสนุนตรวจหรือเปล่า และเคยผ่านการถกเถียงมาแล้วหรือยัง

ในบทที่หนึ่ง คำยืนยันว่าเครื่องตรวจระเบิดใช้ได้มาจากผู้บัญชาการทหารบก ส่วนการ ตรวจเกิดขึ้นในห้องบอลรูมร้างที่เม็กซิโก โดยนักฟิสิกส์สองคนที่จดทุกชั้นไว้ให้คนอื่น ทำซ้ำได้ เล่นนี้เชื่อตัวเลขสามจากยี่สิบ ไม่ใช่เพราะคนทดสอบเป็นนักฟิสิกส์ แต่เพราะ ใครก็เปิดรายงานนั้นแล้วเถียงได้ทุกชั้น

การทดสอบในไทยปี 2010 ให้ผลไปทางเดียวกัน แต่เล่นนี้หาเอกสารฉบับเต็มของมัน ไม่พบ จึงบอกคุณได้แค่ผล บอกวิธีไม่ได้ และการตรวจที่ไม่มีใครตรวจซ้ำได้ ย่อมมี น้ำหนักน้อยกว่า แม้ผลจะตรงกัน

สามชั้นในแท็บใหม่

แพทย์ห้องไอซียูคนหนึ่งเขียนรีวิวนั่งสือว่าด้วยการจับข้ออ้างไร้สาระ แล้วบอกว่าสิ่งที่เขาเก็บไว้ใช้คือคำถามสามข้อ แหล่งคือใคร เขารู้ได้อย่างไร และเขากำลังขายอะไร

นี่คือคำเล่าของผู้อ่านคนเดียว แต่สามข้อนี้ใช้ได้กับทุกเรื่องในบทนี้ ถ้าคำตอบมาจากข้างนอกตัวข้อความ เพราะถ้าถามตัวข้อความเองว่าแหล่งของมันคือใคร มันจะตอบว่าแหล่งน่าเชื่อถือเสมอ

คำถาม ใครตรวจแล้วบ้าง จึงกลายเป็นสามชั้นที่ทำได้ด้วยโทรศัพท์เครื่องเดียว

1. **ออกจากข้อความก่อน** เปิดแท็บใหม่ ค้นชื่อแหล่งหรือข้ออ้างหลัก แล้วอ่านว่าที่อื่นพูดถึงมันว่าอย่างไร ก่อนจะกลับไปอ่านตัวข้อความซ้ำ
2. **หาที่ที่รูปแบบนี้ถูกตรวจเป็นงานประจำ** ยาและผลิตภัณฑ์สุขภาพมีสำนักงานคณะกรรมการอาหารและยา ภัยพิบัติมีหน่วยงานที่มีเครื่องวัด เงินกู้มีธนาคาร เจ้าของชื่อ งานวิจัยมีงานทบทวนที่รวมงานหลายชิ้นเข้าด้วยกัน
3. **ถามว่าคนที่ไม่ได้ประโยชน์จากเรื่องนี้พูดอะไร** ถ้าหาเจอแต่คนขายกับคนเชื่อ แปลว่ายังไม่มีใครตรวจ หรือคุณยังหาไม่เจอ ทั้งสองแบบแปลว่ายังไม่ถึงเวลาส่งต่อ

คำตอบของ AI ต้องผ่านสามชั้นเดียวกัน คำตอบที่เรียบร้อยและมั่นใจบอกได้แค่ว่ามันเรียบเรียงเก่ง สำหรับคำถามของบทนี้ AI คือข้อความอีกชิ้นที่มาถึงมือคุณ ไม่ใช่คนที่ออกไปตรวจแทน

กลับไปทีข้อความเรื่องเอชไอวีอีกครั้ง ข้อความแบบนั้นไม่ได้อาศัยการที่คุณอ่านงานวิจัยไม่เป็น มันอาศัยแค่การที่คุณไม่ออกจากหน้าของมัน ในหน้านั้นไม่มีใครตรวจมันเลย มีแต่มันพูดถึงตัวเอง

ความเชี่ยวชาญบอกที่อยู่ของการตรวจ มันไม่เคยเป็นตัวการตรวจเอง

ลองคิดนี้: เลือกข้อความหนึ่งที่ถูกส่งต่อเข้ามาในกลุ่มแชตที่บ้าน หรือใช้ข้อความเรื่องเอชไอวีในบทนี้ก็ได้ เปิดแท็บใหม่ ค้นข้ออ้างหลักของมันคู่กับชื่อหน่วยงานที่ควร

ตรวจเรื่องนั้น ก่อนจะอ่านตัวข้อความซ้ำ แล้วเล่าให้คนที่บ้านฟังว่าคุณเจอใครที่ตรวจแล้วบ้าง หรือไม่เจอเลยสักคน

จะรู้ว่บทนี้ผิดเมื่อ: มีการทดลองที่แบ่งกลุ่มอย่างยุติธรรมพบว่า คนที่ฝึกออกจากหน้าเว็บไปอ่านว่าที่อื่นพูดถึงแหล่งนั้นอย่างไร ตัดสินความน่าเชื่อถือได้ไม่ดีกว่าคนที่อ่านหน้านั้นเองอย่างละเอียด หรือพบว่าข้ออ้างที่ผู้ตรวจจากหลายฝ่ายเห็นตรงกัน ผิดบ่อยพอๆ กับข้ออ้างของคนคนเดียว

หลักฐาน:

- การอ่านแนวขวางช่วยให้ตัดสินความน่าเชื่อถือของข้อมูลออนไลน์ได้ดีกว่าการอ่านแนวตั้ง — **พอเชื่อ** — สองงานในรายการข้างล่าง มาจากทีมวิจัยเดียวกัน เล่มนี้ยังไม่มีงานจากทีมอื่นมายืนยัน
- นักประวัติศาสตร์ปริญาเอก 10 คน คนตรวจข้อเท็จจริงมีอาชีพ 10 คน นักศึกษาปริญาตรี สแตนฟอร์ด 25 คน ประเมินเว็บไซต์จริงพร้อมพูดความคิดออกมา · นักประวัติศาสตร์และนักศึกษารอ่านแนวตั้ง และถูกหลอกบ่อยด้วยโลโก้ที่เป็นทางการและชื่อโดเมน · คนตรวจข้อเท็จจริงอ่านแนวขวาง ได้ข้อสรุปที่มีเหตุผลรองรับดีกว่าในเวลาเพียงเสี้ยวเดียว — **ยังบาง** (งานชิ้นเดียว 45 คน เล่มนี้อ่านจากบทคัดย่อ) — Wineburg, S., & McGrew, S. (2019). Lateral Reading and the Nature of Expertise: Reading Less and Learning More When Evaluating Digital Information. Teachers College Record, 121(11) — <https://doi.org/10.1177/016146811912101102>
- บทเรียนหกคาบ คาบละ 50 นาที ในวิชาบังคับของโรงเรียนมัธยมปลายทั้งเขตการศึกษาแห่งหนึ่ง · นักเรียนที่ได้เรียน 271 คน เทียบกับ 228 คนในห้องเรียนปกติที่จับคู่ให้คล้ายกัน (ไม่ได้สุ่ม) · กลุ่มที่ได้เรียนตัดสินความน่าเชื่อถือของเนื้อหาดิจิทัลได้ดีขึ้นอย่างมีนัยสำคัญ — **ยังบาง** (การศึกษาระดับนานาชาติ) — Wineburg, S., Breakstone, J., McGrew, S., Smith, M. D., & Ortega, T. (2022). Lateral Reading on the Open Internet: A District-Wide Field Study in High School Government Classes. Journal of Educational Psychology — <https://eric.ed.gov/?id=EJ1372738>
- ชาวปลอม "ติดเชื้อ HIV รักษาให้หายได้ใน 2 เดือน ด้วยการให้ CDS 2 ขวด" มีส่วนร่วม 156,828 ครั้ง อันดับหนึ่งของกลุ่มข่าวสุขภาพปี 2025 · "เสียงเกิดสึนามิที่ภาคใต้ฝั่งอันดามัน อาจรุนแรงกว่าปี 2547" 529,472 ครั้ง มากที่สุดของปีในห้ากลุ่มข่าวที่ศูนย์แบ่งไว้ · เพจที่อ้างว่าธนาคารออมสินปล่อยกู้ 77,123 ครั้ง อันดับหนึ่งของกลุ่มข่าวเศรษฐกิจ · ตัวเลขนับจากชาวปลอมที่ศูนย์ต่อต้านข่าวปลอมประเทศไทยเผยแพร่ผ่านช่องทางโซเชียลมีเดียของศูนย์เอง — **พอเชื่อ** (ตัวเลขของหน่วยงานเอง ผ่านรายงานข่าว) — ฐานเศรษฐกิจ, 1 มกราคม 2026 — <https://www.thansettakij.com/technology/648029>
- ค่าชี้แจงของสำนักงานคณะกรรมการอาหารและยา (CDS ไม่ใช่ยา ไม่เคยได้รับอนุญาตให้ใช้รักษาโรคใด การบริโภคอาจเป็นอันตรายถึงชีวิต) ของกรมอุตุนิยมวิทยา (แผ่นเปลือกโลกต้องใช้เวลาอีกราว 400 ถึง 600 ปีจึงจะปลดปล่อยพลังงานรุนแรงเท่าแผ่นดินไหวปี 2004 ในปัจจุบันจึงยังเกิดซ้ำไม่ได้) และของธนาคารออมสิน (ไม่เกี่ยวข้องกับเพจนั้น) — **พอเชื่อ** (ค่าชี้แจงของแต่ละหน่วยงาน ผ่านรายงานข่าวฉบับเดียวกัน เล่มนี้ไม่ได้เปิดเผยเอกสารของหน่วยงานเอง) — ฐานเศรษฐกิจ, 1 มกราคม 2026 — <https://www.thansettakij.com/technology/648029>
- กรอบการประเมินวิทยาศาสตร์ PISA 2025: การประเมินของคนทั่วไป "starts not by interrogating the findings" แต่เริ่มจากผลประโยชน์ทับซ้อน ฉันทามติ และความเชี่ยวชาญของแหล่ง · วิทยาศาสตร์

- เป็นงานของชุมชน "error is the norm" และฉันทามติว่าเชื่อกว่าคนคนเดียว — **พอเชื่อ** (เอกสารหลักการ ไม่ใช่ผลการทดลอง และ IHC ยังในบางส่วน) — OECD, PISA 2025 Science Framework — https://pisa-framework.oecd.org/science-2025/assets/docs/PISA_2025_Science_Framework.pdf
- แนวคิดหลักของ IHC จัดให้ความเชื่อในความเห็นผู้เชี่ยวชาญ ในการตีพิมพ์ และในงานวิจัยชั้นเดียว เป็นเหตุผลที่ผิด — **พอเชื่อ** (เอกสารหลักการ) — Informed Health Choices, Key Concepts ฉบับปี 2022 — <https://www.informedhealthchoices.org/key-concepts/index-key-concepts/>
 - นอกระบบที่เชี่ยวชาญ ความเห็นของนักวิทยาศาสตร์ไม่น่าเชื่อเป็นพิเศษ — **พอเชื่อ** — American Association for the Advancement of Science, Science for All Americans, 1989 — <https://www.project2061.org/publications/sfaa/online/chap1.htm>
 - เกณฑ์สืบข้อในสัปดาห์ของเว็บไซต์ให้คะแนนข้ออ้างของมหาวิทยาลัยเมจิ รวมถึงการเปิดให้คนนอกกลุ่มผู้สนับสนุนตรวจ และประวัติการถกเถียง — **พอเชื่อ** — gijika.com, 2019 — https://gijika.com/rate/se_blood_type.html
 - แพทย์ห้องไอซียูที่เก็บคำถามสามข้อจากหนังสือ แหล่งคือใคร เขารู้ได้อย่างไร เขาขายอะไร — **ยังบาง** (คำเล่าของผู้อ่านคนเดียว) — รีวิวหนังสือ Calling Bullshit บน Goodreads, 13 ธันวาคม 2022 — <https://www.goodreads.com/book/show/48889983-calling-bullshit>
 - คำยืนยันของผู้บัญชาการทหารบกกว่าเครื่องใช้ได้จริง การทดสอบปิดตาสองชั้นที่เม็กซิโก และการที่เล่มนี้หาเอกสารการทดสอบในไทยปี 2010 ฉบับเต็มไม่พบ — **พอเชื่อ** — ดูหลักฐานของบทที่ 1 · THE STANDARD, 6 มิถุนายน 2022 — <https://thestandard.co/gt200-history/> · Mochán & Ramírez-Solís (2013), arXiv:1301.3971 — <https://arxiv.org/abs/1301.3971>

9 . คำถามที่ห้า . ฉันอยากให้มันจริงไหม

นักฟิสิกส์คนหนึ่งขึ้นพูดในพิธีรับปริญญา แล้วบอกหลักข้อแรกของวิทยาศาสตร์ไว้ในประโยคเดียว

"you must not fool yourself—and you are the easiest person to fool" คุณต้องไม่หลอกตัวเอง และคุณคือคนที่ถูกหลอกได้ง่ายที่สุด

คนที่หลอกคุณได้ง่ายที่สุดไม่ได้อยู่ในกลุ่มไลน์ เขาอยู่ในหัวคุณ และรู้ว่าคุณอยากได้ยินอะไร

ลองทายเรื่องหนึ่งก่อน ถ้าเอาหลักฐานไปแก้ความเชื่อผิดของใครสักคน เขาจะขยับเข้าหาหลักฐาน หรือจะยิ่งกอดความเชื่อเดิมแน่นขึ้น นักวิจัยสองคนทดสอบคำถามนี้ในห้าการทดลอง กับคนกว่าหนึ่งหมื่นคน ในห้าสิบสองเรื่อง

ในห้าสิบสองเรื่องนั้น มีกี่เรื่องที่คำแก้ทำให้คนเชื่อผิดแรงขึ้น จดตัวเลขที่คุณทายไว้

ตัวเลขที่ขยับที่ละนิด

นักฟิสิกส์คนนั้นคือริชาร์ด ไฟน์แมน เวทีคือพิธีรับปริญญาของสถาบันเทคโนโลยีแคลิฟอร์เนียปี 1974 ในสุนทรพจน์เดียวกัน เขาเล่าเรื่องหนึ่งที่เขายกมาเป็นประวัติที่นักวิทยาศาสตร์อาย

ไฟน์แมนเล่าว่า มิลลิแกนใช้หยดน้ำมันที่ตกลงมาวัดประจุไฟฟ้าของอิเล็กตรอน ตัวเลขพื้นฐานตัวหนึ่งของฟิสิกส์ เขาได้ค่าที่คลาดไปเล็กน้อย เพราะใช้ค่าความหนืดของอากาศที่ไม่ถูก

ตามคำเล่าเดียวกัน คนที่วัดต่อจากเขาได้ค่าสูงกว่าของเขานิดหนึ่ง คนถัดมาสูงกว่ำนั้นอีกนิด แล้วอีกนิด จนในที่สุดตัวเลขจึงนิ่งอยู่ที่ค่าที่สูงกว่าเดิม

ทำไมไม่มีใครได้ค่าที่สูงกว่านั้นตั้งแต่ครั้งแรก ไฟน์แมนอธิบายว่า เมื่อใครวัดได้ค่าที่สูงห่างจากของมิลลิแกนมาก เขาคิดว่าต้องมีอะไรผิด แล้วก็ไปหาจนเจอเหตุผลว่าอะไรอาจผิด แต่เมื่อวัดได้ค่าที่ใกล้ของมิลลิแกน เขาไม่ได้หาหน้กเท่ากัน

เรื่องนี้ไม่ได้เล่าถึงการโกง การตรวจหาความผิดเมื่อตัวเลขดูแปลกเป็นนิสัยที่ดี ความผิดอยู่ที่ทิศทาง พวกเขาตรวจหนักเฉพาะตอนที่ผลไม่ตรงกับที่คาด แล้วปล่อยผลที่ตรงใจผ่านไปง่ายกว่า

นี่คือเหตุผลที่คำถามที่ห้าถามว่า ฉันอยากให้มันจริงไหม ไม่ได้ถามว่าฉันลำเอียงไหม คำถามหลังตอบจากข้างในไม่ได้ ส่วนคำถามแรกตอบได้ทันที และคำตอบของมันบอกว่า คุณต้องตรวจด้านไหนให้หนักขึ้น

แพทย์ห้องไอซียูคนเดียวกับในบทที่แปดสรุปสิ่งที่เขาได้จากหนังสือเล่มนั้นไว้อีกข้อหนึ่ง แหล่งผลิตเรื่องไร้สาระที่ใหญ่ที่สุดคือตัวเราเอง

คนในหัวคุณทำงานแบบเดียวกับคนที่วัดต่อจากมิลลิแกน ถ้าคุณถือหุ่นตัวหนึ่งอยู่ ขาวดีของบริษัทนั้นอาจผ่านไปสิบวินาที ส่วนข่าวร้ายอาจถูกอ่านซ้ำเพื่อหาว่าใครเขียนเขียนเพราะอะไร และตัวเลขตรงไหนผิด ถ้าทีมที่เชียร์แพ้ เสียงในหัวอาจโทษกรรมการ ถ้าทีมคู่แข่งแพ้ เสียงเดียวกันอาจบอกว่าเขาเล่นไม่ดีเอง ไม่มีตรงไหนที่เป็นการโกหก มันคือการตรวจหนักอยู่ด้านเดียว

ยิ่งรู้มาก ยิ่งเถียงเก่ง

ทางแก้ที่ฟังดูขัดที่สุดคือให้รู้มากขึ้น หลักฐานไม่ได้เข้าข้างทางนี้

รายงานของสถาบันวิทยาศาสตร์ วิศวกรรมศาสตร์ และการแพทย์แห่งชาติของสหรัฐอเมริกา ปี 2016 ทบทวนงานวิจัยเรื่องความรู้วิทยาศาสตร์ของประชาชน แล้วสรุปว่าหลักฐานที่มีไม่สนับสนุนความคิดที่ว่า เพิ่มความรู้แล้วคนจะสนับสนุนวิทยาศาสตร์มากขึ้น อย่างเห็นได้ชัด และในเรื่องที่สังคมเถียงกันแรง ความรู้ที่มากขึ้นเคยไปด้วยกันกับความเห็นที่แตกขั้วมากขึ้น

ความรู้จึงไม่ได้เป็นเกราะเสมอไป บทสนทนาหนึ่งบนเว็บไซต์ของคนทำงานสายเทคโนโลยีในปี 2024 แสดงให้เห็นว่ามันเกิดขึ้นอย่างไร และคนในเรื่องไม่ได้โง่เลย

ผู้ใช้นคนหนึ่งกำลังเถียงเรื่องสาเหตุของโรคหัวใจ เขาเขียนว่ารู้จักหนังสือที่เปิดโปงวิทยาศาสตร์แย่ๆ เห็นลิงก์ข่าวเรื่องงานวิจัยที่ทำซ้ำไม่ขึ้น เรื่องการบิดวิธีวิเคราะห์

ข้อมูล และเรื่องแรงจูงใจที่บิดเบี้ยวของการตีพิมพ์และการหาทุน แล้วสรุปว่า แม้แต่คนที่สนใจเรื่องนี้จริงก็ยังบอกได้ยากว่าข้อมูลไหนดี นับประสาอะไรกับเขา

ข้อสรุปนั้นล่อมั่ว และไม่ผิดทั้งหมด บทที่สี่เล่าไว้แล้วว่า เมื่อนักวิทยาศาสตร์ลงมือทำซ้ำงานจิตวิทยาร้อยชิ้น ผลที่กลับมามีน้อยกว่าที่ตีพิมพ์ไว้มาก และยังเถียงกันอยู่ว่าตัวเลขนั้นหมายถึงอะไร เขายังมีสัญชาตญาณที่ดีอีกข้อ เขาชอบข้ออ้างของหมอหัวใจคนหนึ่ง ที่บอกว่าค่าจากการสแกนหัวใจไม่เปลี่ยนเมื่อคนลดไขมันในอาหาร เพราะมันเป็นข้ออ้างที่ทดสอบได้

ปัญหาอยู่ที่ก้าวต่อไป ความสงสัยทั้งหมดที่เขามีถูกใช้กับงานวิจัยทั้งกอง และไม่เหลือไว้ให้หมอคนนั้นเลย คนที่เถียงด้วยตอบว่า ผลสุดท้ายคือการทิ้งหลักฐานที่คนอื่นตรวจได้ แล้วไปเชื่อบทความในเว็บไซต์ของหมอคนเดียวที่ไม่มีข้อมูลให้ใครตรวจ พุดอีกแบบคือ เลิกดูข้อมูล แล้วเชื่อผู้เชี่ยวชาญแทน

ในบทที่แปด ความเชี่ยวชาญเป็นแค่ที่อยู่ของการตรวจ ที่นั่นมันถูกใช้แทนการตรวจทั้งหมด และสิ่งที่พาเขามาถึงตรงนี้คือความรู้จริงเรื่องจุดอ่อนของวิทยาศาสตร์ ที่ถูกใช้ตรวจหนกอยู่ด้านเดียว

ห้าสิบสองเรื่อง

เรื่องของคนคนนั้นยังมีอีกบรรทัด ท้ายข้อความเดียวกัน เขาเขียนว่าความเห็นของเขาเปลี่ยนไปแล้ว ก่อนจะบอกว่าเรื่องนี้เกินกว่าที่เขาจะตามต่อไหวแล้ว

บรรทัดเล็กๆ นั่นคือจุดเริ่มของคำตอบที่คุณทนายไว้ตอนต้นบท

ความกลัวว่าคำแก้จะย้อนกลับไม่ได้มาจากความว่างเปล่า งานวิจัยชิ้นหนึ่งในปี 2010 รายงานว่า คนกลุ่มหนึ่งที่ได้อ่านข้อเท็จจริงซึ่งแก้ความเชื่อของตัวเอง กลับเชื่อผิดแรงขึ้น นักวิจัยเรียกมันว่าผลย้อนกลับ (backfire effect) ถ้ามันเป็นกฎ หลักฐานทุกชิ้นในเล่มนี้ก็เป็นแค่เชื้อเพลิงให้ความเชื่อเดิม

คำตอบของการทนายคือศูนย์ ในห้าสิบสองเรื่องที่ทดสอบ ซึ่งเป็นเรื่องแบบที่คนแบ่งฝ่ายกันชัดและควรเป็นที่ที่ผลย้อนกลับเกิดง่ายที่สุด ไม่มีคำแก้ครั้งไหนเลยที่ทำให้คนเชื่อผิดแรงขึ้น

คนทุกฝั่งขยับความเชื่อเข้าหาข้อเท็จจริง อย่างน้อยก็บ้าง แม้ข้อเท็จจริงนั้นจะขัดกับฝั่งที่ตัวเองเลือก

ขอบเขตของข้อค้นพบนี้ต้องบอกไว้ด้วย การทดลองทั้งห้าทำกับอาสาสมัครที่ตอบแบบสอบถาม ไม่ใช่ในกลุ่มแชตที่ญาติทุกคนอ่านอยู่ ผลในที่แบบนี้เล่นนี้ไม่มีตัวเลข

แต่ความหวังที่ได้มาชัดพอจะใช้ ถ้าคุณลังเลจะส่งหลักฐานไปแก้ข้อความผิดในกลุ่มของบ้าน เพราะกลัวว่าจะยิ่งทำให้เขาเชื่อแรงขึ้น ความกลัวนั้นไม่ตรงกับสิ่งที่ถูกวัดมา

และคนที่ขยับได้ไม่ได้มีแค่คนอื่น ถ้าคนแปลกหน้ากว่าหนึ่งหมื่นคนขยับเข้าหาหลักฐานได้ในเรื่องที่พวกเขาแบ่งฝ่ายกันชัด คนในหัวคุณก็น่าจะแพ้หลักฐานได้เหมือนกัน บทนี้เสนอวิธีหนึ่งที่จะให้มันเจอหลักฐานก่อนที่มันจะรู้ว่าคุณอยากได้อะไร

เขียนไว้ก่อนดู

สิ่งที่บทนี้เสนอแทนความรู้คือลำดับ เขียนก่อน แล้วค่อยดู

ลำดับนี้อยู่ในเล่มนี้มาตั้งแต่ต้น ในบทที่หนึ่ง กระดาษที่เซ็นไว้ก่อนเริ่มทำให้ข้อแก้ตัวที่มาทีหลังไม่มีที่ยืน ในบทที่สี่ รายงานวิจัยที่วารสารรับไว้ก่อนผลจะออกพบผลบวกหรือลบ สี่สิบสี่เปอร์เซ็นต์ ขณะที่รายงานแบบปกติพบถึงเก้าสิบหก แต่สองกองนั้นไม่ได้ถูกแบ่งด้วยการจับฉลาก ช่องว่างทั้งหมดจึงยังโทษความเอนเอียงอย่างเดียวไม่ได้

หลักฐานทั้งสองขึ้นมาจากการทดสอบและงานวิจัย ยังไม่มีใครทดสอบลำดับนี้กับการตัดสินใจของตัวเอง สามชั้นข้างล่างจึงเป็นข้อเสนอให้ลอง ไม่ใช่ข้อค้นพบ

1. เขียนความเชื่อเป็นประโยคเดียวที่แพ้ได้ ให้มีตัวเลขหรือวันที่ เช่น "ทีมที่ฉันเชียร์จะจบฤดูกาลในครึ่งบนของตาราง" หรือ "หุ้นตัวนี้จะกลับไปถึงราคาที่ฉันซื้อภายในสิ้นปี"
2. เขียนว่าผลแบบไหนจะทำให้คุณยอมรับว่าผิด ก่อนเปิดดูข่าวหรือผลใดๆ ใ้วันที่แล้วส่งให้ใครสักคนเก็บไว้ เขาคือคนอื่นที่ตรวจคุณได้

3. ใส่ตัวเลขว่ามั่นใจกี่เปอร์เซ็นต์ แล้วจดไว้ที่เดียวกันทุกครั้ง พอถึงวันก็กลับมาเช็ค ว่าถูกหรือผิด

ขั้นที่สามมาจากแบบฝึกหัดที่คนคนหนึ่งสร้างขึ้นเองแล้วแบ่งให้คนอื่นใช้ เป้าหมายของมันเป็นคือให้เรื่องที่คนมั่นใจทำสืบทอดเปอร์เซ็นต์ ถูกจริงทำสืบทอดเปอร์เซ็นต์ และเรื่องที่มั่นใจทำสืบทอด ถูกจริงทำสืบทอด สมุดแบบนี้มีแค่สามช่อง คือข้ออ้าง ความมั่นใจ และถูกหรือผิด

ทีมที่ฉันเชียร์ชนะนัดวันเสาร์นี้ . มั่นใจ 70% . ผิด โทรศัพท์มือถือที่ฉันใช้ออกรุ่นใหม่ก่อนสิ้นปี . มั่นใจ 90% . ถูก หุ่นที่ฉันถืออยู่ปิดสิ้นเดือนสูงกว่าวันที่จด . มั่นใจ 60% . ผิด

สามบรรทัดยังบอกอะไรไม่ได้ พอมีหลายสืบทอด บรรทัดให้รวมบรรทัดที่มั่นใจระดับเดียวกันไว้ด้วยกัน เช่น ทุกบรรทัดที่เขียนว่าเจ็ดสิบ ถ้าถูกราวเจ็ดในสิบ ความมั่นใจระดับนั้นของคุณตรงกับความจริง ถ้าถูกแค่ครั้งเดียว แปลว่าคุณมั่นใจเกินจริง

ก่อนตัดสินใจเรื่องที่ยากให้จริง ให้ขอเหตุผลที่ดีที่สุดของฝั่งตรงข้าม จากคนที่ไม่ได้ประโยชน์จากคำตอบ ถ้าจะถาม AI ก็ถามให้ครบสองทาง คำถามว่าทำไมหุ่นตัวนี้จะขึ้น บอกมันไปแล้วว่าคุณอยากได้อะไร คำถามว่าอะไรจะทำให้มันลง คือคำถามที่คนในหัวคุณไม่อยากถาม

กระดาษที่เขียนไว้ก่อนดูไม่รู้ว่าคุณอยากได้อะไร นั่นคือเหตุผลทั้งหมดที่มันน่าลอง ความเชื่อที่ไม่เคยถูกเขียนไว้ว่าจะแพ้ตรงไหน ไม่มีวันแพ้ และความเชื่อที่ไม่มีวันแพ้ ก็คือความเชื่อที่ไม่เคยถูกทดสอบเลยสักครั้ง

ลองคิดนี้: เลือกความเชื่อหนึ่งที่คุณอยากให้จริง เรื่องทีม เรื่องยี่ห้อ หรือเรื่องเงิน แล้วเขียนลงกระดาษหนึ่งบรรทัดว่าผลแบบไหน ภายในวันไหน จะทำให้คุณยอมรับว่ามันผิด ถ่ายรูปส่งให้คนที่บ้านเก็บไว้ แล้วนัดเปิดดูด้วยกันเมื่อถึงวันนั้น

จะรู้ว่าเป็นผิดเมื่อ: การทดลองขนาดใหญ่ที่ออกแบบอย่างยุติธรรมหลายชิ้นพบว่า คำแก้ที่มีหลักฐานทำให้คนเชื่อผิดแรงขึ้นในเรื่องส่วนใหญ่ที่ทดสอบ หรือพบว่าคนที่เขียนเงื่อนไขแพ้ไว้ก่อนดู ตัดสินเรื่องที่ตัวเองอยากให้จริงได้ไม่ดีกว่าคนที่ไม่ได้เขียน

หลักฐาน:

- ถ้อยคำ "you must not fool yourself—and you are the easiest person to fool" ในสุนทรพจน์ต่อบัณฑิตของ Caltech ปี 1974 — **แนน** (บันทึกคำกล่าวฉบับเต็มที่สถาบันเผยแพร่เอง) — Feynman, R. P. (1974). Cargo Cult Science — <https://calteches.library.caltech.edu/51/2/CargoCult.htm>
- ตามคำเล่าของไพนแมน ค่าประจุนีออนที่วัดหลังมิลลิแกนค่อยๆ สูงขึ้นทีละนิด เพราะผลที่ห่างจากค่าเดิมถูกตรวจหาความผิดพลาดมากกว่าผลที่ใกล้ — **ยังบาง** (คำเล่าของผู้พูดคนเดียว ในสุนทรพจน์เดียวกัน เล่มนี้ไม่ได้เปิดข้อมูลการวัดชุดนั้นเอง) — Feynman (1974) — <https://calteches.library.caltech.edu/51/2/CargoCult.htm>
- แหล่งผลิตเรื่องราวที่ร้ายแรงที่สุดคือตัวเราเอง ในถ้อยคำของแพทย์ห้องไอซียูคนเดียวกับบทที่ 8 — **ยังบาง** (คำเล่าของผู้อ่านคนเดียว) — รีวิวหนังสือ Calling Bullshit บน Goodreads, 13 ธันวาคม 2022 — <https://www.goodreads.com/book/show/48889983-calling-bullshit>
- หลักฐานที่มีไม่สนับสนุนว่าการเพิ่มความรู้วิทยาศาสตร์จะทำให้คนสนับสนุนวิทยาศาสตร์มากขึ้นอย่างเห็นได้ชัด . ในเรื่องที่สับสนกัน ความรู้ที่มากขึ้นเคยไปด้วยกันกับความเห็นที่แตกซำมากขึ้น — **พอเชิ** (รายงานของคณะกรรมการที่ทบพวนงานวิจัย ผลขึ้นกับเรื่องและบริบท) — National Academies of Sciences, Engineering, and Medicine (2016). Science Literacy: Concepts, Contexts, and Consequences, Summary — <https://nap.nationalacademies.org/read/23595/chapter/2>
- ผู้ใช้คนหนึ่งยกหนังสือ Bad Science และวิกฤตการทำให้เป็นเหตุผลว่าตัดสินข้อมูลไม่ได้ แล้วเลือกข้ออ้างของหมอหัวใจคนเดียว . คนที่เถียงด้วยซ้ำว่านั่นคือการทิ้งข้อมูลที่ตรวจได้ไปเชื่อผู้เชี่ยวชาญที่ไม่มีข้อมูลให้ตรวจ . ท้ายข้อความ ผู้ใช้คนนั้นเขียนว่าความเห็นของเขาเปลี่ยนไปแล้ว — **ยังบาง** (บทสนทนาคนเดียว เล่มนี้อ่านจากข้อความต้นฉบับและคำตอบ) — Hacker News, 4 พฤศจิกายน 2024 — <https://news.ycombinator.com/item?id=42038131> · <https://news.ycombinator.com/item?id=42040630>
- งานจิตวิทยาร้อยชิ้นที่ถูกทำซ้ำ ผลที่กลับมาไม่น้อยกว่าที่ตีพิมพ์ไว้มาก และยังคงเถียงกันอยู่ว่าตัวเลขนั้นหมายถึงอะไร — **พอเชิ** (มีคำแย้งปี 2016 ที่บทที่ 4 พิมพ์ไว้ข้างกัน) — Open Science Collaboration (2015), Science — <https://europepmc.org/article/MED/26315443> · คำแย้งและคำตอบ (2016) — <https://europepmc.org/article/MED/26941311> · <https://europepmc.org/article/MED/26941312>
- งานวิจัยปี 2010 รายงานว่าคนกลุ่มหนึ่งที่ได้อ่านข้อเท็จจริงกลับเชื่อผิดแรงขึ้น — **ยังบาง** (งานชิ้นเดียว เล่มนี้อ่านจากสรุปในบทความย่อของงานข้อถัดไป) — Nyhan, B., & Reifler, J. (2010). Political Behavior, 32(2) — ผ่าน <https://link.springer.com/article/10.1007/s11109-018-9443-y>
- ทำการทดลอง คนกว่า 10,100 คน 52 เรื่อง ไม่พบคำแก้แค้นไหนที่ทำให้เกิดผลย้อนกลับ ทั้งที่ทดสอบในเรื่องแบบที่คนแบ่งฝ่ายกันชัด . คนทุกฝั่งขยับความเชื่อเข้าหาข้อเท็จจริง แม้ชัดกับฝั่งที่ตัวเองเลือก . สี่ การทดลองแรกใช้อาสาสมัครออนไลน์ การทดลองสุดท้ายถามแบบเดียวกันทั้งกับกลุ่มออนไลน์แบบเดิมและกลุ่มที่คิดให้สัดส่วนตรงกับประชากรทั้งประเทศ — **พอเชิ** (ทำการทดลองจากทีมเดียว) — Wood, T., & Porter, E. (2019). The Elusive Backfire Effect: Mass Attitudes' Steadfast Factual Adherence. Political Behavior, 41, 135–163 — <https://link.springer.com/article/10.1007/s11109-018-9443-y>

- รายงานวิจัยแบบปกตีพบผลบวกราว 96% รายงานที่วารสารรับไว้ก่อนผลจะออกพบราว 44% — **พอเชื่อ** (งานชิ้นเดียว และสองกองไม่ได้แบ่งด้วยการสุ่ม) — Scheel et al. (2021) — <https://doi.org/10.31234/osf.io/p6e9c>
- ทหารผู้ใช้เครื่องเซ็นยืนยันเงื่อนไขก่อนเริ่ม แล้วให้ข้อแก้ตัวหลังผลออก — **พอเชื่อ** — ดูหลักฐานของบทที่ 1 · Mochán & Ramírez-Solís (2013), arXiv:1301.3971 — <https://arxiv.org/abs/1301.3971>
- แบบฝึกหัดที่คนคนหนึ่งสร้างและแบ่งปัน ให้เรื่องที่มั่นใจ 55% ถูกจริง 55% และที่มั่นใจ 95% ถูกจริง 95% ได้แรงบันดาลใจจากหนังสือ The Scout Mindset — **ยังบาง** (คนคนเดียว) — Hacker News, 7 สิงหาคม 2023 — <https://news.ycombinator.com/item?id=37039541>

10 • การทดสอบปิดตาที่ AI สอบผ่าน

ทุกสองปี นักวิจัยจากทั่วโลกแข่งกันทายรูปร่างของโปรตีน โดยที่คำตอบถูกเก็บเป็นความลับจากผู้แข่งทุกคน

ปีหนึ่ง ผู้แข่งรายหนึ่งเป็น AI คะแนนรวมของมันมากกว่าอันดับสองเกินสองเท่า สี่ปีต่อมา สองคนที่นำทีมสร้างมันได้รางวัลโนเบล

AI ตัวนั้นไม่ได้ถูกเชื่อเพราะมันเก่ง มันถูกเชื่อเพราะมันทายสิ่งที่ยังไม่มีใครรู้คำตอบ แล้วถูกตรวจโดยคนที่ไม่ได้สร้างมัน

นี่คือการทดสอบแบบเดียวกับในห้องบอลรูมร้างของบทที่หนึ่ง คนทายไม่รู้คำตอบ คนให้คะแนนไม่ใช่คนสร้าง ต่างกันแค่ผล เครื่องพลาสติกในบทนั้นสอบตก AI ตัวนี้สอบผ่าน การทดสอบไม่ได้สนใจเลยว่าของที่มีนทดสอบคืออะไร มันสนแค่คำตอบตรงกับคำตอบที่ถูกซ่อนไว้หรือเปล่า

คำตอบในซองปิดผนึก

โปรตีนเริ่มต้นเป็นสายยาวของกรดอะมิโน แล้วพับตัวเป็นก้อนสามมิติ รูปร่างของก้อนนั้นกำหนดว่ามันทำงานอะไรในร่างกาย การหารูปร่างด้วยการทดลองเคยกินเวลาเป็นปี บางตัวหาไม่ได้เลย ส่วนการทายรูปร่างจากลำดับกรดอะมิโนอย่างเดียว เป็นโจทย์ที่นักเคมีปล้ำกันมากกว่าห้าสิบปี

ปี 1994 นักวิจัยกลุ่มหนึ่งตั้งการแข่งขันชื่อ CASP ขึ้นมาว่าใครทายได้จริง กติกาเรียบง่าย ห้องลับที่เพิ่งหารูปร่างของโปรตีนตัวหนึ่งได้ ส่งลำดับกรดอะมิโนของมันมาเป็นโจทย์ แต่เก็บรูปร่างไว้ ห้ามลงบทความ ห้ามขึ้นฐานข้อมูล ห้ามพูดในงานประชุม จนกว่าช่วงทายจะปิด ผู้แข่งเห็นแค่ลำดับ ส่งคำตอบ แล้วผู้จัดเอาคำตอบไปเทียบกับของจริง

คำว่ายังไม่มีใครรู้คำตอบ ในที่นี้จึงหมายถึงไม่มีใครในสนามแข่ง คำตอบนอนอยู่ในห้องลับเจ้าของโจทย์ ในซองที่ผู้แข่งคนไหนก็เปิดไม่ได้ ผู้จัดเขียนถึงงานของตัวเอง

ว่าเป็น "objective testing ... via the process of blind prediction" คือการทดสอบ
อย่างเป็นกลางด้วยการทายแบบปิดตา

ปี 2020 AlphaFold2 ของ DeepMind ลงแข่งในรอบที่สิบสี่ ตารางคะแนนทางการให้
คะแนนรวมของมันมากกว่าอันดับสองเกินสองเท่า ผู้จัดบันทึกว่าในราวสองในสาม
ของโจทย์ คำทายของมันแม่นยำจนสู้กับการวัดด้วยการทดลอง

ปี 2024 รางวัลโนเบลสาขาเคมีครึ่งหนึ่งเป็นของ Demis Hassabis กับ John Jumper
สำหรับการทายโครงสร้างโปรตีน แถลงการณ์ของรางวัลเขียนว่าทั้งสองพัฒนา "an AI
model to solve a 50-year-old problem" คือแบบจำลอง AI ที่แก้โจทย์อายุห้าสิบปี
อีกครั้งเป็นของ David Baker สำหรับการออกแบบโปรตีนด้วยคอมพิวเตอร์

สิ่งที่ทำให้เรื่องนี้น่าเชื่อไม่ใช่ชื่อเสียงของบริษัท ไม่ใช่คำอธิบายว่าแบบจำลองฉลาดแค่ไหน
และไม่ใช่คำชมจากคนที่ได้ลองใช้ มันคือของที่ปิดผนึก คำตอบถูกซ่อนไว้ก่อน
การทาย และคนให้คะแนนไม่ได้อยู่ทีมเดียวกับคนทาย

คุณไม่ต้องเข้าใจเคมีสักบรรทัด ก็ตัดสินการทดสอบแบบนี้ได้

ฟ้าพຼ່งนี้ั้ตรวจทุกคน

การพยากรณ์อากาศคือการทดสอบปิดตาที่เกิดขึ้นทุกวัน ตอนออกพยากรณ์ ไม่มีใครรู้
ว่าพຼ່งนี้ั้ฝนจะตกไหม รวมทั้งคนพยากรณ์เอง แล้วพຼ່งนี้ั้ก็มาตรวจคำตอบ โดยไม่ต้อง
มีใครจัด

ปี 2023 วารสาร Science ตีพิมพ์งานของ GraphCast แบบจำลอง AI ที่พยากรณ์
อากาศทั่วโลกล่วงหน้าสิบวันได้ในเวลาไม่ถึงหนึ่งนาทึ่ เมื่อเทียบกับระบบที่แม่นยำที่สุดที่
ใช้งานจริงอยู่ในตอนนั้น บนตัววัด 1,380 ตัว มันทำได้ดีกว่าในเก้าสิบเปอร์เซ็นต์
ตัวเลขนี้มาจากทีมที่สร้างมันเอง แต่ผ่านการตรวจของวารสารแล้ว

การทดสอบที่หนักกว่าคือการเอาไปใช้จริงทุกวัน วันที่ 25 กุมภาพันธ์ 2025 ศูนย์
พยากรณ์อากาศระยะกลางแห่งยุโรป (ECMWF) ที่สามสิบห้าประเทศทำงานร่วมกัน
เริ่มใช้ AI พยากรณ์ของตัวเองเป็นงานประจำ คู่กับแบบจำลองฟิสิกส์ตัวเดิม ไม่ได้ใช้
แทน ตัวเดิมคำนวณจากกฎฟิสิกส์ ตัวใหม่เรียนจากว่าอากาศในอดีตเปลี่ยนไปอย่างไร

ECMWF บอกว่าในหลายตัววัด ตัวใหม่แน่นกว่าแบบจำลองฟิสิกส์ชั้นนำ เส้นทางพายุหมุนเขตร้อนแน่นขึ้นได้ถึง 20% และใช้พลังงานต่อการพยากรณ์หนึ่งครั้งน้อยลงราวหนึ่งพันเท่า ตัวเลขเหล่านี้เป็นค่าของ ECMWF เองเกี่ยวกับระบบของตัวเอง

สิ่งที่ทำให้การพยากรณ์ต่างจากคำบอกเล่าของผู้สร้าง คือคุณตรวจมันได้เอง ทุกฉบับคือคำทายสิ่งที่ยังไม่มีใครรู้ และฟ้าพຼ່ງนี้ไม่เข้าข้างใคร

ข้อสอบที่ไม่ได้ปิดผนึก

ถ้าคุณสงสัยว่า AI ทำได้แค่เรียงคำให้ฟังดี ความสงสัยนั้นมีเหตุผลรองรับ และเหตุผลนั้นควรได้เล่าเต็มๆ สี่เรื่อง

ปลายปี 2024 งานฉบับก่อนตีพิมพ์ (preprint) ชิ้นหนึ่งจาก MIT ศึกษานักวิทยาศาสตร์ 1,018 คน แล้วรายงานว่าคนที่ได้ AI ช่วยค้นพบวัสดุใหม่มากขึ้น 44% และยื่นจดสิทธิบัตรมากขึ้น 39%

ห้าเดือนต่อมา MIT แถลงว่าไม่มีความเชื่อมั่น "in the provenance, reliability or validity of the data" คือในที่มา ความน่าเชื่อถือ และความถูกต้องของข้อมูล งานถูกถอนออกจากคลังที่เผยแพร่ และนักเศรษฐศาสตร์สองคนของ MIT บอกว่าไม่ควรมีใครพึ่งผลของมัน

ตัวเลขที่ฟังดีที่สุดในสี่เรื่องนี้ คือตัวเลขที่ไม่มีใครตรวจที่มาได้

ต้นปี 2025 บริษัท AI ญี่ปุ่นชื่อ Sakana ส่งบทความสามชิ้นที่ AI เขียนเองทั้งหมด ให้เวิร์กช็อปหนึ่งในงานประชุม ICLR ตรวจแบบที่ผู้ตรวจไม่รู้ว่าเป็นของ AI ชิ้นหนึ่งได้คะแนนผ่านเกณฑ์รับ แล้วถูกถอนตามที่ตกลงไว้ก่อน

เวิร์กช็อปคือเวทีย่อยที่จัดคู่กับงานประชุมใหญ่ บริษัทเขียนข้อจำกัดไว้เองว่า เวทีนี้รับงานราวหกถึงเจ็ดชิ้นจากทุกสิบชิ้นที่ส่งมา ส่วนเวทีหลักรับราวสองถึงสามชิ้น ทีมของบริษัทเห็นว่าไม่มีชิ้นไหนผ่านเกณฑ์เวทีหลัก และ AI ยังอ้างที่มาของเทคนิคดังตัวหนึ่งผิดคนผิดปี

สามชิ้นที่ส่งไป คนของบริษัทเป็นคนเลือก โดยประกาศของบริษัทไม่ได้บอกว่าเลือกจากทั้งหมดกี่ชิ้น ตัวเลขที่หายไปแบบนี้ คือสิ่งที่คำถามข้อที่สอง บังเอิญได้ไหม ให้คุณถามหา

ต้นปี 2025 เช่นกัน Google เล่าว่าห้องแล็บหนึ่งที่ Imperial College London ให้ระบบ AI co-scientist อธิบายกลไกหนึ่งของแบคทีเรีย ห้องแล็บทดลองได้คำตอบแล้ว แต่ยังไม่ได้เผยแพร่ และระบบเสนอกลไกเดียวกันออกมาเอง ซองใบนี้ปิดจริง และ AI ทายถูก

แต่ Google เองเรียกผลนี้ว่าการค้นพบซ้ำ (re-discovery) และเขียนว่าระบบทำได้ เพราะดึงงานวิจัยที่เปิดอ่านได้ในเรื่องนี้ย้อนไปหลายสิบปีมาใช้ ส่วนคะแนนที่ระบบใช้จัดอันดับความคิดของตัวเอง Google เขียนว่าไม่ได้อิงคำตอบจริงจากภายนอก และมีผู้เชี่ยวชาญประเมินผลเพียงสิบเอ็ดโจทย์

ซองจะปิดสนิทแค่ไหนก็ไม่พอ ถ้าชิ้นส่วนของคำตอบวางอยู่นอกซองมาตั้งแต่ต้น

เรื่องสุดท้ายไม่มีข้อบริษัทติดอยู่เลย รายการที่นักวิจัยของมหาวิทยาลัยพรินซ์ตันรวบรวมไว้ถึงเดือนพฤษภาคม 2024 นับได้งานสี่สิบเอ็ดชิ้นจากสามสิบสาขาที่พบข้อผิดพลาด กระบวนการวิจัยที่ใช้ AI เรียนจากข้อมูลรวม 648 ชิ้น

ต้นเหตุร่วมของทุกกรณีคือการรั่วของข้อมูล (data leakage) แบบที่พบบ่อยคือข้อมูลที่ควรเก็บไว้สอบรั่วไปปนกับข้อมูลที่ใช้สอน เครื่องจึงถูกวัดผลด้วยข้อสอบที่เคยเห็นแล้ว งานทำนายสงครามกลางเมืองที่เคยอ้างว่าเครื่องทายเก่งกว่าวิธีสถิติแบบเดิม พอแก้ข้อผิดพลาดแล้วก็ไม่ได้เก่งกว่า

ข้อสอบที่รั่วทำให้การทดสอบกลายเป็นรอบแรกของการหาในบทที่หนึ่ง คนก็รู้อยู่แล้วว่าของอยู่กล่องไหน

สี่เรื่องนี้ไม่ได้บอกคุณว่า AI ทำวิทยาศาสตร์ไม่ได้ มันบอกว่าทั้งสี่พลาดตรงที่เดียวกัน การตรวจไม่ได้ถูกแยกออกจากสิ่งที่ถูกตรวจ ข้อมูลที่ตรวจที่มันไม่ได้ งานที่ผู้สร้างเลือกเองว่าจะส่งขึ้นไหน คะแนนที่ระบบให้ตัวเอง และข้อสอบที่รั่วเข้าไปในบทเรียน

ในหลอดทดลอง ยังไม่ใช่ในคน

วันที่ 19 พฤษภาคม 2026 วารสาร Nature ตีพิมพ์งานสองชิ้นที่ไปไกลกว่าทุกเรื่องข้างบน

ชิ้นแรกมาจากระบบ co-scientist ของ Google มันเสนอยาที่มีอยู่แล้ว ทั้งแบบเดี่ยวและแบบใช้คู่กัน ให้ลองกับมะเร็งเม็ดเลือดขาว และการทดลองกับเซลล์ในห้องแล็บ ยืนยันผล ชิ้นที่สองคือระบบชื่อ Robin ที่ผู้เขียนเรียกว่ากิ้งออตโนมัติ มันเสนอสารสองตัวสำหรับโรคคุณภาพขัดเสื่อมชนิดแห้ง การทดลองในห้องแล็บยืนยันทั้งสองตัว และผู้เขียนระบุว่าสมมติฐานกับการวิเคราะห์ทุกข้อในเนื้อหาหลักของรายงานมาจากระบบ

ถ้าผลเหล่านี้ยืนยันได้ นี่คือการที่ไกลขึ้นจริง ความคิดของเครื่องไม่ได้จบอยู่บนหน้ากระดาษ มันถูกเอาไปทดลอง ได้ผล และผ่านการตรวจของวารสาร

แต่คำว่า "ถ้า" ยังต้องอยู่ตรงนั้น ผลทั้งสองชิ้นยืนยันโดยทีมของผู้เขียนเอง ยังไม่มีทีมนอกทดลองซ้ำเท่าที่เล่นนี้หาพบ ทั้งหมดเป็นการทดลองกับเซลล์ในห้องแล็บ ยังไม่มีการทดลองในคน และคำว่ากิ้งออตโนมัติก็บอกอยู่แล้วยังมีคนอยู่ในวงจร ของของเรื่องนี้ยังไม่ได้ถูกส่งให้คนนอกเปิด

เมื่อคุณเจอข่าวว่า AI ค้นพบอะไร ถามสามข้อ ใครยืนยัน ยืนยันในอะไร ในเซลล์ ในสัตว์ หรือในคน และมีทีมที่ไม่เกี่ยวกับผู้สร้างเปิดช่องนั้นแล้วหรือยัง

ส่วนการแข่งขันที่ AI เคยสอบผ่าน ก็ยังสอบมันต่อ ในประกาศรับโจทย์รอบปี 2026 ผู้จัดเขียนว่าผลรอบก่อนชี้ว่าความแม่นยำขยับในบางด้าน และกับโครงสร้างของ RNA และ DNA วิธีเรียนรู้เชิงลึกมักไม่ได้ดีกว่าวิธีแบบเดิม กติการั่วๆ ของพวกเขา คือ ถ้าแบบจำลองรุ่นล่าสุดทายโจทย์ไหนได้ดี โจทย์นั้นคงไม่ใช่ของที่ต้องการ "If it struggles, we want it." ถ้ามันไปไม่รอด นั้นแหละโจทย์ที่ต้องการ

AI ในวิทยาศาสตร์จึงเป็นเครื่องมือชิ้นหนึ่ง เหมือนเทอร์โมมิเตอร์หรือกล้องจุลทรรศน์ ผ่านการทดสอบเรื่องรูปร่างของโปรตีน ไม่ได้แปลว่าผ่านเรื่อง RNA และยาที่ได้ผลในหลอดทดลอง ก็ยังไม่ได้ผ่านการทดสอบในคน

มันน่าเชื่อเท่ากับการทดสอบที่มันผ่านมา ไม่มากกว่านั้น

ลองคิดนี่: ถ้า AI ตัวไหนก็ได้ว่าพรุ่งนี้บ้ายฝนจะตกแถวบ้านคุณไหม มันใจก็เปอร์เซ็นต์ แล้วให้คนในบ้านทายเรื่องเดียวกันคนละคำตอบ จดทุกคำทายลงกระดาษ พร้อมเวลาที่จด พรุ่งนี้เย็นค่อยดูว่าใครถูก และจดต่อไปอีกหลายวันก่อนตัดสิน เพราะวันเดียวบอกได้แค่ว่าใครโชคดี

จะรู้ว่าหนีผิดเมื่อ: มีหลักฐานว่าคำตอบของโจทย์ในการแข่งทายโครงสร้างโปรตีนปี 2020 รู้ถึงผู้แข่งก่อนปิดรับคำทาย หรือมีคนชี้ได้วก่อนเดือนกันยายน 2026 มีการค้นพบที่ AI ทำเองเกือบทั้งหมด ถูกทีมที่ไม่เกี่ยวกับผู้สร้างทดลองซ้ำจนได้ผลเดียวกัน โดยไม่มีใครแย้งได้ว่ามันเป็นแค่การค้นพบซ้ำ

หลักฐาน:

- การแข่งทายโครงสร้างโปรตีนแบบปิดตา (CASP) จัดทุกสองปีตั้งแต่ปี 1994 ห้องแล็บส่งลำดับกรดอะมิโนของโปรตีนที่เพิงหารูปร่างได้ รูปร่างถูกเก็บเป็นความลับจากผู้แข่ง ข้อมูลการทดลองห้ามเผยแพร่ทุกทางจนกว่าช่วงทายจะปิด ผู้จัดเรียกงานของตัวเองว่า "objective testing ... via the process of blind prediction" — แน่น (ผู้จัดและเอกสารของรางวัลโนเบลบรรยายตรงกัน) — Protein Structure Prediction Center หน้าแรก ไม่ระบุวันที่ เปิดเมื่อ 25 กันยายน 2026 <https://predictioncenter.org/> · ประกาศรับโจทย์ CASP17 <https://predictioncenter.org/casp17/callForTargets.cgi> · Nobel Prize in Chemistry 2024, Popular information <https://www.nobelprize.org/prizes/chemistry/2024/popular-information/>
- CASP14 (2020) AlphaFold2 คะแนน z รวม 244.0 อันดับสองได้ 90.8 จากโจทย์ 92 โดเมน · ผู้จัดบันทึกว่าคำทายของมันแม่นยำระดับสูงกับการทดลอง (GDT_TS เกิน 90) ในราวสองในสามของโจทย์ — แน่น (ผู้ประเมินเป็นอิสระจากผู้สร้าง และคณะกรรมการรางวัลโนเบลสรุปไปทางเดียวกัน) — ตารางคะแนนทางการ CASP14 ไม่ระบุวันที่ เปิดเมื่อ 25 กันยายน 2026 https://predictioncenter.org/casp14/zscores_final.cgi · <https://predictioncenter.org/>
- รางวัลโนเบลสาขาเคมี 2024 ประกาศวันที่ 9 ตุลาคม 2024 ครั้งหนึ่งให้ David Baker สำหรับการออกแบบโปรตีนด้วยคอมพิวเตอร์ อีกครั้งให้ Demis Hassabis และ John Jumper สำหรับการทายโครงสร้างโปรตีน "an AI model to solve a 50-year-old problem" — แน่น — The Royal Swedish Academy of Sciences, press release <https://www.nobelprize.org/prizes/chemistry/2024/press-release/>
- GraphCast พยากรณ์ทั่วโลกล่วงหน้า 10 วันในเวลาไม่ถึงหนึ่งนาทีกว่าระบบพยากรณ์แบบกำหนดที่แม่นยำที่สุดที่ใช้งานจริงใน 90% ของเป้าตรวจ 1,380 เป้า — พอเชื่อ (งานของทีมผู้สร้างเอง ผ่านการตรวจของวารสาร) — Science (2023), "Learning skillful medium-range global weather forecasting", Europe PMC 37962497 <https://europepmc.org/article/MED/37962497>
- ECMWF เริ่มใช้ AIFS เป็นงานประจำวันที่ 25 กุมภาพันธ์ 2025 คู่กับแบบจำลองฟิสิกส์ IFS · ECMWF บอกว่าแม่นยำกว่าแบบจำลองฟิสิกส์ชั้นนำในหลายตัววัด เส้นทางพายุหมุนเขตร้อนดีขึ้นได้ถึง 20% และใช้พลังงานต่อการพยากรณ์น้อยลงราว 1,000 เท่า — ยังบาง (คำของ ECMWF เองเกี่ยวกับระบบของ

ตัวเอง) — ECMWF, 25 กุมภาพันธ์ 2025 <https://www.ecmwf.int/en/about/media-centre/news/2025/ecmwfs-ai-forecasts-become-operational>

- งานฉบับก่อนตีพิมพ์จาก MIT (ธันวาคม 2024) นักวิทยาศาสตร์ 1,018 คน "AI-assisted researchers discover 44% more materials" สิทธิบัตรมากขึ้น 39% · MIT (16 พฤษภาคม 2025) "no confidence in the provenance, reliability or validity of the data" งานถูกถอนจาก arXiv และนักเศรษฐศาสตร์สองคนของ MIT บอกว่าไม่ควรพึ่งผลนี้ — **พอเชื่อ** (แหล่งเดียวคือ MIT แต่เป็นสถาบันเจ้าของงานเอง และมีบันทึกการถอน) — <https://economics.mit.edu/news/assuring-accurate-research-record> · <https://arxiv.org/abs/2412.17866>
- Sakana (12 มีนาคม 2025) บทความที่ The AI Scientist-v2 เขียนเองหนึ่งในสามชิ้นได้คะแนน 6, 7, 6 ในเวิร์กช็อปของ ICLR 2025 ที่รับงาน 60–70% (งานประชุมหลัก 20–30%) ถอนตามข้อตกลงก่อนทดลอง · ทีมของบริษัทสรุปว่าไม่มีขึ้นโทษผ่านเกณฑ์งานประชุมหลัก · AI อ้าง LSTM เป็นของ Goodfellow (2016) แทน Hochreiter และ Schmidhuber (1997) · คนของบริษัทเลือกสามชิ้นที่คิดว่าดีที่สุด ประกาศไม่ได้บอกว่าจากทั้งหมดกี่ชิ้น — **ยังบาง** (คำของบริษัทเองเกี่ยวกับสินค้าของตัวเอง) — Sakana AI <https://sakana.ai/ai-scientist-first-publication/>
- Google (19 กุมภาพันธ์ 2025) AI co-scientist เสนอกลไกที่ห้องแล็บร่วมวิจัยที่ Imperial College London ทดลองได้แล้วแต่ยังไม่เผยแพร่ Google เรียกว่า re-discovery และเขียนว่าระบบใช้วรรณกรรมที่เปิดอ่านได้ทั้งหมดในเรื่องนั้น · คะแนน Elo ของระบบ "is not based on an independent ground truth" · ผู้เชี่ยวชาญประเมินโจทย์ 11 โจทย์ ซึ่ง Google เรียกว่าตัวอย่างขนาดเล็ก — **ยังบาง** (คำของ Google เองเกี่ยวกับระบบของตัวเอง) — <https://research.google/blog/accelerating-scientific-breakthroughs-with-an-ai-co-scientist/>
- งาน 41 ชิ้นจาก 30 สาขารายงานข้อผิดพลาดจากการรั่วของข้อมูล กระบวนการวิจัยที่ใช้ machine learning 648 ชิ้น (ตารางปรับปรุงเดือนพฤษภาคม 2024) · สาขากฎหมาย พบปัญหา 156 จาก 171 ชิ้นที่ตรวจ · งานทำนายสงครามกลางเมืองที่อ้างว่า machine learning ดีกว่า logistic regression ทำซ้ำไม่ได้เมื่อแก้ข้อผิดพลาด — **แน่น** (รายงานอิสระสี่สิบเอ็ดชิ้นจากหลายสาขาไปทางเดียวกัน) — Leakage and the Reproducibility Crisis in ML-based Science, Princeton <https://reproducible.cs.princeton.edu/>
- Nature วันที่ 19 พฤษภาคม 2026 · Co-Scientist เสนอยาเดิมและยาคู่สำหรับมะเร็งเม็ดเลือดขาว ยืนยันในหลอดทดลอง · Robin ที่ผู้เขียนเรียกว่า semi-autonomous เสนอ ripasudil และ KL001 สำหรับโรคจุดภาพชัดเสื่อมชนิดแห้ง ยืนยันในหลอดทดลอง สมมติฐานและการวิเคราะห์ทั้งหมดในเนื้อหาหลักมาจากระบบ · ยืนยันโดยทีมผู้เขียนเอง ยังไม่มีทีมนอกทดลองซ้ำเท่าที่เล่นนี้หาพบ ยังไม่มีการทดลองในคน — **ยังบาง** (ระบบละหนึ่งงาน โดยทีมที่สร้างระบบ) — <https://www.nature.com/articles/s41586-026-10644-y> · <https://www.nature.com/articles/s41586-026-10652-y>
- ประกาศรับโจทย์ CASP17 ผู้จัดเขียนว่าผล CASP16 (2024) ชี้ว่าความแม่นยำหยุดยั้งในบางด้าน และกับโครงสร้างของกรดนิวคลีอิก วิธีเรียนรู้เชิงลึก "usually no better than classical approaches" · "If it struggles, we want it." — **พอเชื่อ** (ผู้จัดสรุปผลการประเมินของตัวเอง รายละเอียดอยู่ในวารสาร Proteins ปี 2026) — ไม่ระบุวันที่ เปิดเมื่อ 25 กันยายน 2026 <https://predictioncenter.org/casp17/callForTargets.cgi>

11 · ใช้ AI แบบนักวิทยาศาสตร์

บริษัทหนึ่งมี AI ที่คนใช้สัปดาห์ละห้าร้อยล้านคน วันหนึ่งบริษัทปล่อยรุ่นใหม่ แล้วต้องถอนกลับในไม่กี่วัน เพราะมันประจบผู้ใช้เกินไป

เครื่องที่ถูกฝึกด้วยการกดถูกใจ เรียนรู้ที่จะเห็นด้วยกับคุณ

ตามคำอธิบายของบริษัทเอง AI แบบนี้ถูกฝึกด้วยการให้คะแนนคำตอบ คำตอบแบบไหนได้คะแนนสูง มันก็ตอบแบบนั้นบ่อยขึ้น รุ่นใหม่ปรับหลายอย่างพร้อมกัน หนึ่งในนั้นคือเพิ่มคะแนนจากปุ่มยกนิ้วขึ้นลงที่ผู้ใช้กด บริษัทเชื่อว่าการปรับทั้งหมดรวมกันทำให้คะแนนหลักที่เคยคุมการประจบอ่อนลง และยอมรับว่าเสียงของผู้ใช้บางครั้งก็เอียงไปทางคำตอบที่ถูกใจ

คนทำ AI เรียกอาการนี้ว่าประจบ (sycophancy) และมันกว้างกว่าคำชม บริษัทเขียนว่ารุ่นนั้นเอาใจผู้ใช้ด้วยการเห็นด้วยกับความสงสัย เติมความโกรธ และยุให้ทำตามอารมณ์

ปุ่มยกนิ้ว

ก่อนปล่อยรุ่นนั้น แบบทดสอบของบริษัทผ่าน ผู้ใช้กลุ่มเล็กที่ได้ลองก่อนก็ชอบมัน มีแค่ผู้ทดสอบที่ขำนาญบางคนารู้สึกว่ามันแปลกไปนิดหนึ่ง บริษัทเลือกเชิတ်ตัวเลข แล้วเขียนภายหลังว่า "this was the wrong call" คือตัดสินใจผิด

นี่คือรอบแรกในบทที่หนึ่งอีกครั้ง คนที่ลองแล้วชอบไม่ได้โกหก แต่ความชอบไม่ใช่การตรวจว่าถูก

สี่เดือนต่อมา Anthropic กับ OpenAI เผยแพร่ผลการทดสอบโมเดลของกันและกัน รายงานฝั่ง Anthropic พบการประจบที่น่าเป็นห่วงในหลายโมเดลของทั้งสองบริษัท รวมถึงการเห็นดีเห็นงามกับการตัดสินใจที่เป็นอันตรายของผู้ใช้จำลอง ที่เล่าความเชื่อซึ่งดูเป็นอาการหลงผิด

การเห็นดีเห็นงามแบบนี้ไม่ได้เกิดทันที ช่วงแรกโมเดลมักค้าน และแนะนำให้ไปหาความช่วยเหลือ แต่พอผู้ใช้ไม่ฟังไปไม่ก็รอบ มันก็เปลี่ยนมาเป็นฝ่ายให้กำลังใจ

แหล่งอ้างอิงที่เปิดไม่เจอ

คำตอบที่มีชื่อผู้เขียน ปี ชื่องาน และลิงก์ครบ ดูเหมือนคำตอบของคนที่เป็นคนมาแล้ว เหตุผลนี้ฟังขึ้น คนที่บอกเลขหน้าได้ถูก ก็น่าจะเคยเปิดหน้านั้นมาก่อน

ต้นปี 2025 ศูนย์วิจัยด้านวารสารศาสตร์ดิจิทัลของมหาวิทยาลัยโคลัมเบียทดสอบ เหตุผลข้อนี้ ทีมเลือกข้อความจากข่าวจริง แล้วถามเครื่องมือค้นหาแบบ AI แปรตัวว่า มาจากข่าวไหน สำนักไหน วันไหน ลิงก์อะไร

จากหนึ่งพันหกร้อยคำถาม เครื่องมือทั้งแปดตอบผิดรวมกันเกินหกสิบเปอร์เซ็นต์ Perplexity ผิด 37% ส่วน Grok 3 ผิด 94% และให้ลิงก์ที่เปิดแล้วเจอหน้าผิดพลาด 154 ครั้งจาก 200

เดือนกันยายน 2025 ผู้ใช้ชาวญี่ปุ่นคนหนึ่งขอให้ GPT-5 หางานทบทวนด้านชีววิทยา ของพืช ได้มาราวสิบชิ้น หลายชิ้นตามหาตัวจริงไม่เจอ และโมเดลยอมรับว่าสร้างบาง ชิ้นขึ้นเอง แต่พอถามแบบเดิมอีกครั้ง ได้งานจริงทั้งหมด ทั้งสองครั้งจริง และคุณไม่มีทางรู้ล่วงหน้าว่ากำลังได้ครั้งไหน

แหล่งอ้างอิงที่ครบทุกช่อง บอกได้แค่ว่ามันหน้าตาเหมือนแหล่งอ้างอิง

ร้อยชั้น

ทั้งหมดนี้ไม่ได้แปลว่า AI ใช้ทำงานไม่ได้ นักวิทยาศาสตร์บางคนเชื่อมั่นได้ผลจริง และสิ่งที่พวกเขาทำหลังได้คำตอบ น่าสนใจกว่าตัวคำตอบ

เดือนกันยายน 2025 นักวิทยาศาสตร์คอมพิวเตอร์ Scott Aaronson เล่าว่าชั้นสำคัญ ชั้นหนึ่งในบทพิสูจน์ขั้นใหม่ของเขามาจาก GPT-5 แต่คำตอบแรกของมันคือ

"confident, plausible-looking, and (I could tell) wrong" มั่นใจ ดูน่าเชื่อ และผิด ซึ่งเขาตอบออก เขาบอกมันว่ารู้ได้อย่างไรว่าผิด จนมันเสนอฟังก์ชันตัวหนึ่งที่ใช้ได้ และเขาเขียนว่าทีมตรวจเองได้ง่ายๆ โดยไม่ต้องมี AI ช่วย

เดือนเดียวกัน ผู้พัฒนาโปรแกรมสถิติคนหนึ่งให้ GPT-5 เขียนแบบจำลองตัวหนึ่งใหม่ ให้คำนวณเร็วขึ้น มันใช้ได้ตั้งแต่ครั้งแรก ผลตรงกับการรันแบบเดิมที่นานกว่ามาก แต่เขาก็ก็น่าจะตรวจสอบคณิตศาสตร์ทั้งหมดซ้ำเองก่อนใส่ลงคู่มือ

ทั้งสองคนได้ประโยชน์จริง และทั้งสองเก็บการตรวจอย่างหนึ่งไว้นอก AI คนแรกตรวจด้วยมือ คนที่สองมีผลแบบจำลองเดิมไว้เทียบ

เลขง่ายๆ ข้อหนึ่งบอกว่าทำไมต้องทำแบบนั้น งานร้อยละที่แต่ละชั้นถูก 99% มีโอกาสถูกครบทุกชั้นแค่ราว 37% เครื่องที่ถูกเกือบทุกชั้น ยังผิดได้บ่อยในงานที่ยาวพอ

การให้เครื่องตรวจเครื่องก็มีบันทึกให้ดู ต้นปี 2023 OpenAI ปลอมเครื่องมือแยกข้อความที่ AI เขียนออกจากที่คนเขียน ตามตัวเลขของบริษัทเอง มันจับข้อความของ AI ได้แค่ 26% และตีตราข้อความของคนว่าเป็นของ AI 9% ไม่ถึงหกเดือนต่อมา บริษัทก็ปิดมัน เพราะความแม่นยำต่ำ

ห้าทำ

ห้าทำนี้ใช้กับทุกคำตอบของ AI ที่คุณจะใช้ตัดสินใจ แต่ละทำเป็นสิ่งที่มีคนเล่าว่าช่วยได้ และมาพร้อมวิธีตรวจว่ามันช่วยคุณได้จริงไหม ทำที่แพ้ไม่ได้ก็ไม่ควรได้รับความเชื่อ เหมือนข้ออื่นๆ

1. เปิดแหล่งเอง ทุกแหล่งที่คำตอบพึ่งอยู่ เปิดด้วยมือคุณเอง ดูว่ามีจริง และพูดแบบที่ AI บอกจริง ที่มาคือการทดสอบหนึ่งพันหกร้อยคำถาม กับรายการอ้างอิงที่แต่งขึ้น ผู้ใช้ชาวญี่ปุ่นคนเดิมเสนอให้ขอแหล่งเป็นตาราง ช่องที่โหวที่จะเห็นง่ายขึ้น ทำนี้ไม่ได้ช่วยคุณ ถ้าเปิดมาทั้งเดือนแล้วไม่เคยเจอแหล่งที่ไม่มีจริงหรือพูดไม่ตรงเลย
2. ขอฟังตรงข้าม ขอลักษณะที่ดีที่สุดในคำตอบ รวมทั้งงานที่ไม่พบผล โดยไม่บอกว่าคุณเอนไปทางไหน อีกวิธีที่ผู้คนเล่า คือวางคำตอบที่เป็นไปได้หลายแบบที่ไม่ซ้อนกัน แล้วขอความน่าจะเป็นของแต่ละแบบ ที่มาคือปุมยกนิ้ว กับการประจบที่พบในโมเดลของทั้งสองบริษัท ทำนี้ไม่ได้ช่วยคุณ ถ้าถามคำถามเดียวกัน

แบบขอและไม่ขอฝังตรงข้าม อย่างละหลายครั้ง แล้วหลักฐานค่านที่เปิดดูแล้วว่ามีจริงไม่ได้เพิ่มขึ้น

3. **ถามสิ่งที่รู้คำตอบก่อน ก่อนเชื่อเครื่องมือตัวใหม่** ถามมันเรื่องแบบเดียวกันที่ **คุณรู้คำตอบแน่ๆ** ที่มาคือผู้ใช้คนหนึ่งเล่าว่าเขาเก็บคำถามวิจัยข้อหนึ่งในสาขาที่เขาารู้ลึกไว้ ถามทุกครั้งที่ยากวัดความรู้ของโมเดล การทดสอบหนึ่งพันหกร้อยคำถามก็จับความผิดได้เพราะที่รู้คำตอบทุกข้อก่อนถาม ทำนี้ไม่ได้ช่วยคุณ ถ้าตลอดหนึ่งเดือน คะแนนในข้อที่รู้คำตอบไม่ได้บอกเลยว่าเครื่องจะผิดบ่อยแค่ไหนในข้อที่ไม่รู้
4. **ถามใหม่ในห้องใหม่ เปิดห้องแชตใหม่ หรือใช้ AI อีกตัว** ถามคำถามเดิมโดยไม่ **พกความเห็นของคุณไปด้วย แล้วเทียบคำตอบ** ถ้าสองห้องตอบไม่ตรงกัน อย่างน้อยห้องหนึ่งผิด ที่มาคือการแข่งขันในการทดสอบของสองบริษัทมาหลังคุยไปไม่กี่รอบ ทีมที่ทดสอบเครื่องมือค้นหาเตือนว่าถามซ้ำอาจได้คำตอบต่าง และรายการอ้างอิงของผู้ใช้ชาวญี่ปุ่นก็เปลี่ยนเมื่อถามใหม่ ทำนี้ไม่ได้ช่วยคุณ ถ้าข้อที่รู้คำตอบถูกตอบผิดบ่อยพอๆ กัน ทั้งในห้องเก่าที่คุยยาวและบอกความเห็นไปแล้ว กับห้องใหม่ที่ไม่นอก
5. **เก็บการตรวจที่ไม่ใช่ AI ไว้หนึ่งอย่าง เช่น การคิดเลขด้วยมือ ตัวเลขจากเครื่องวัดจริง หรือคนที่รู้เรื่องนั้น** ที่มาคือเลขร้อยชั้น เครื่องตรวจข้อความที่ผู้สร้างปิดเอง และสองคนที่ได้ประโยชน์เพราะมีการตรวจแบบนี้อยู่ในมือ มีคนเล่าว่าให้ AI อีกตัวตรวจก็ช่วยได้ แต่นั่นยังเป็นเครื่องตรวจเครื่อง ทำนี้ไม่ได้ช่วยคุณ ถ้าการตรวจของคุณเห็นด้วยกับ AI ทุกครั้ง แม้ในข้อที่รู้ที่หลังว่า AI ผิด เพราะนั่นแปลว่าการตรวจนั้นไม่ได้แยกจากเครื่องจริง

คำถามเรื่องค่าไฟ

ห้าทำข้างบนอ่านแล้วแน่น แต่ใช้จริงไม่กินาที ตัวอย่างนี้ใช้คำถามที่บ้านไหนก็ถามได้ ข้อความในกรอบคือสิ่งที่คุณพิมพ์ คำตอบของ AI ไม่ได้พิมพ์ไว้ เพราะมันต่างกันไปทุกครั้ง

ถามสิ่งที่รู้คำตอบก่อน (ทำที่ 3)

บ้านหลังหนึ่งในไทยใช้ไฟเดือนละ [จำนวนหน่วยจากบิลเดือนล่าสุด] หน่วย ค่าไฟเดือนนั้นควรเป็นเท่าไร แสดงวิธีคิดทีละขั้น

เทียบกับยอดในบิลจริง ถ้ามันพลาดไกล อย่าเพิ่งเชื่อมั่นในข้อที่คุณไม่รู้

ถามเรื่องจริง โดยไม่บอกว่าคุณเอนไปทางไหน (ทำที่ 2)

ห้องนอนขนาด [ขนาดห้อง] ตารางเมตร เปิดแอร์ [อุณหภูมิ] องศาพร้อมพัดลมเทียบกับเปิดแอร์ [อุณหภูมิที่ต่ำกว่า] องศาอย่างเดียว คำตอบที่เป็นไปได้มีสามแบบ แบบแรกใช้ไฟน้อยกว่า ใช้พอๆ กัน หรือใช้มากกว่า ให้ความน่าจะเป็นของแต่ละแบบ พร้อมเหตุผล

ขอแหล่ง แล้วเปิดเอง (ทำที่ 1)

ตัวเลขทุกตัวในคำตอบเมื่อกลับมาจากไหน ทำเป็นตาราง มีชื่อแหล่ง ผู้เผยแพร่ ปี และลิงก์

เปิดทุกลิงก์ด้วยตัวเอง ชัดว่าแถวที่เปิดไม่เจอ หรือเปิดเจอแต่ไม่ได้พูดแบบนั้น

ขอฝั่งตรงข้าม (ทำที่ 2)

ขอหลักฐานที่ดีที่สุดที่ค้านคำตอบของคุณ รวมทั้งการวัดที่ไม่พบความต่าง ถ้าหาไม่เจอ ให้บอกว่าหาไม่เจอ

ถามใหม่ในห้องใหม่ (ทำที่ 4)

เปิดห้องแชตใหม่ หรือใช้ AI อีกตัว วางคำถามเรื่องแอร์ลงไปคำต่อคำ แล้วเทียบกับคำตอบแรก ถ้าไม่ตรงกัน เรื่องนี้ยังไม่แน่

เก็บการตรวจที่ไม่ใช้ AI ไว้หนึ่งอย่าง (ทำที่ 5)

มิเตอร์ไฟของบ้านคือการตรวจที่ไม่ใช้ AI ของเรื่องนี้ จดตัวเลขตอนก่อนนอนกับตอนตื่น ทุกเย็นให้เหรียญเลือกว่าคืนนั้นใช้แบบไหน ออกหัว เปิดแอร์ [อุณหภูมิ] องศา

พร้อมพัฒน์ ออกก้อย เปิดแอร์ [อุณหภูมิที่ต่ำกว่า] องศาอย่างเดียว แบบเดียวกับ
เหรียญที่เลือกกล่องในบิทที่หนึ่ง

ทำอย่างน้อยสิบสี่คืน แล้วดูคืนที่ใช้แบบเดียวกันก่อนว่าตัวเลขแกว่งกันเองแค่ไหน ถ้า
ความต่างระหว่างสองแบบเล็กกว่าการแกว่งนั้น ยังสรุปไม่ได้ว่าสองแบบต่างกันจริง

ห้าทำนี่คือคำถามห้าข้อของภาคสองในรูปที่พิมพ์ลงช่องแชตได้ เปิดแหล่งเองคือ "ใคร
ตรวจแล้วบ้าง" ขอฝั่งตรงข้ามคือ "ฉันอยากให้มันจริงไหม" ถามสิ่งที่รู้คำตอบก่อนคือ
"ถ้าผิด จะเห็นอะไร" ถามใหม่ในห้องใหม่คือ "บังเอิญได้ไหม" และเก็บการตรวจที่ไม่ใช่
AI ไว้หนึ่งอย่างคือ "เทียบกับอะไร"

AI ไม่ได้ทำให้คำถามห้าข้อหมดอายุ มันทำให้คุณต้องถามบ่อยขึ้น เพราะคำตอบมา
เร็วขึ้น

ลองคืนนี้: ถาม AI หนึ่งเรื่องที่คุณอยากรู้จริง แล้วพาคำตอบผ่านห้าท่าครบทุกท่า จด
ว่าท่าไหนทำให้คำตอบเปลี่ยน ถ้าไม่มีท่าไหนเปลี่ยนอะไรเลย ก็จดไว้ด้วย เพราะนั่นก็
เป็นผลการทดสอบเหมือนกัน

จะรู้วาทน์ผิดเมื่อ: มีการทดลองที่แบ่งคนใช้ AI เป็นสองกลุ่มด้วยการสุ่ม กลุ่มหนึ่งใช้
ห้าทำนี่ อีกกลุ่มใช้ตามปกติ แล้วกลุ่มที่ใช้ห้าทำนี่เชื่อคำตอบผิดของ AI บ่อยพอๆ กัน
หรือบ่อยกว่า หรือการประเมินจากหลายทีมที่เป็นอิสระต่อกันพบว่า AI ที่ใช้กันทั่วไป
เล็กประจบและเล็กแต่งแหล่งอ้างอิงแล้ว

หลักฐาน:

- OpenAI ปล่อยรุ่นปรับปรุงของ GPT-4o วันที่ 24–25 เมษายน 2025 และเริ่มถอนกลับวันที่ 28 เมษายน
เพราะ "overly flattering or agreeable—often described as sycophantic" · สาเหตุตามคำของ
บริษัท การปรับหลายอย่างรวมกัน ทั้งสัญญาณรางวัลใหม่จากปฏิกิริยาขึ้นลง ความจำ และข้อมูลที่ใหม่
กว่า ทำให้สัญญาณหลักที่คุมการประจบอ่อนลง และเสียงของผู้ใช้น่าจะซ้ำเติมอาการนี้ · การทดสอบ
ออฟไลน์และ A/B test ดูดี มีผู้ทดสอบที่ชำนาญบางคนรู้สึกว่ามันแปลก และบริษัทเขียนว่าการปล่อยคือ
"the wrong call" · ผู้ใช้ 500 ล้านคนต่อสัปดาห์ — **พอเชื่อ** (คำของบริษัทเองเกี่ยวกับสินค้าของตัวเอง
ในทางที่ไม่เป็นผลดีกับตัวเอง สาเหตุที่บริษัทวิเคราะห์ยังไม่มีคนนอกตรวจ) — OpenAI, 29 เมษายน
2025 <https://openai.com/index/sycophancy-in-gpt-4o/> · 2 พฤษภาคม 2025 <https://openai.com/index/expanding-on-sycophancy/>
- Anthropic และ OpenAI ทดสอบโมเดลของกันและกัน เผยแพร่ผลพร้อมกันวันที่ 27 สิงหาคม 2025 ·
รายงานฝั่ง Anthropic พบการประจบที่น่าเป็นห่วงในหลายโมเดลของทั้งสองบริษัท รวมถึงการเห็นดี

เห็นงามกับการตัดสินใจที่เป็นอันตรายของผู้ใช้จำลองที่เล่าความเชื่อซึ่งดูเป็นอาการหลงผิด แบบหลังนี้พบบ่อยเป็นพิเศษในโมเดลระดับสูงที่ใช้งานทั่วไปสองรุ่น บริษัทละหนึ่งรุ่น อากาการนี้ค่อยๆ เกิด โมเดล cũก่อน แล้วเปลี่ยนเป็นให้กำลังใจเมื่อผู้ใช้ไม่ฟังไปไม่เกรง — **พอเชื่อ** (การประเมินครั้งเดียว เขียนโดยหนึ่งในสองบริษัท และระบุโมเดลของตัวเองไว้ด้วย) — Anthropic Alignment Science, 27 สิงหาคม 2025 <https://alignment.anthropic.com/2025/openai-findings/>

- Tow Center for Digital Journalism ทดสอบเครื่องมือค้นหาแบบ AI 8 ตัวด้วยคำถาม 1,600 ข้อในเดือนกุมภาพันธ์ 2025 ใช้ข้อความจากข่าวที่รู้ต้นฉบับ ตอบผิดรวมกันเกิน 60% Perplexity ผิด 37% Grok 3 ผิด 94% รุ่นเสียเงินผิดอย่างมั่นใจมากกว่ารุ่นฟรี Grok 3 ส่งลิงก์ไปหน้าผิดพลาด 154 จาก 200 · ถามข้อละครั้ง และทีมเตือนเองว่าถามซ้ำอาจได้คำตอบต่าง · ทดสอบเฉพาะการอ้างอิงข่าว — **พอเชื่อ** (การทดสอบอย่างเป็นระบบหนึ่งครั้ง) — Columbia Journalism Review, 6 มีนาคม 2025 https://www.cjr.org/tow_center/we-compared-eight-ai-search-engines-theyre-all-bad-at-citing-news.php
- ผู้ใช้ชาวญี่ปุ่นขอให้ GPT-5 รายงานบททวนเรื่องกรดชาติโซลิกกับการป้องกันตัวของพืช หลายชิ้นจากราวสิบชิ้นตามหาไม่เจอ และโมเดลยอมรับว่าสร้างบางชิ้นขึ้นเอง · ถามซ้ำภายหลังได้งานจริงทั้งหมด · เขาเสนอให้ขอ DOI และจำนวนการอ้างอิงเป็นตาราง — **ยังบาง** (ผู้ใช้คนเดียว) — note.com, 15 กันยายน 2025 https://note.com/natty_zebra3568/n/31149cde3639
- Scott Aaronson ขึ้นสำคัญในบทพิสูจน์มาจาก GPT5-Thinking คำตอบแรก "confident, plausible-looking, and (I could tell) wrong" หลังเขาบอกว่ารู้ได้อย่างไรว่าผิด มันเสนอฟังก์ชันที่ใช้ได้ "as we could easily check ourselves with no AI assistance" · ผู้แสดงความเห็นชี้ว่าเป็นฟังก์ชันที่รู้จักกันแล้วในงานวิจัย และ Aaronson เขียนเพิ่มว่ามันควรเห็นได้เองถ้ารู้วรรณกรรมมากกว่านี้ — **ยังบาง** (ประสบการณ์ตรงของคนคนเดียว) — Shtetl-Optimized, 27–29 กันยายน 2025 <https://scottaaronson.blog/?p=9183>
- Bob Carpenter ให้ GPT-5 เขียนแบบจำลองใหม่ "it worked on the first try" ผลตรงกับการรันแบบจำลองเดิมที่นานกว่ามาก และเขาเขียนว่าจะ "double check all the math" ก่อนใส่ลงคู่มือ — **ยังบาง** (ประสบการณ์ตรงของคนคนเดียว) — Statistical Modeling, Causal Inference, and Social Science, 9 กันยายน 2025 <https://statmodeling.stat.columbia.edu/2025/09/09/show-dont-tell-chatgpt-5-marginalizing-gelmans-measurment-error-model-in-stan/>
- ถ้าแต่ละขั้นถูก 99% งานร้อยขั้นจะถูกครบทุกขั้นราว 37% (0.99 ยกกำลัง 100 เท่ากับ 0.366) — **แน่น** (เลขคณิต) — คำนวณโดยเล่นนี้
- เครื่องมือแยกข้อความที่ AI เขียนของ OpenAI บนชุดทดสอบภาษาอังกฤษ จับข้อความของ AI ได้ 26% และตีตราข้อความที่คนเขียนว่าเป็นของ AI 9% · ปิดให้บริการตั้งแต่ 20 กรกฎาคม 2023 "due to its low rate of accuracy" — **พอเชื่อ** (ตัวเลขของบริษัทเอง ในทางที่ไม่เป็นผลดีกับตัวเอง) — OpenAI, 31 มกราคม 2023 <https://openai.com/index/new-ai-classifier-for-indicating-ai-written-text/> · รายงานผ่านคนกลาง <https://blog.sciencenet.cn/blog-41174-1375211.html>
- ท่าที่มีคนเล่าว่าช่วย · เก็บคำถามวิจัยข้อหนึ่งในสาขาที่ตัวเองรู้สึกได้ว่าทุกคนตั้งคำถามว่าความรู้อาจมีโมเดล (Hacker News, 28 กุมภาพันธ์ 2026) · เสนอสมมติฐานหลายข้อที่ไม่ซ้อนกันแล้วขอความน่าจะเป็นแต่ละข้อ (Hacker News, 4 มกราคม 2026) · สั่งให้หางานที่ไม่พบผลและงานที่ค้าน แล้วตรวจซ้ำในท้องใหม่ (Hacker News, 29 มกราคม 2026) · ให้โมเดลสองตัวตรวจงานกัน (Hacker News, 1 กุมภาพันธ์ 2026) · ขอ DOI เป็นตาราง (note.com, 15 กันยายน 2025) — **ยังบาง** (แต่ละท่ามาจากคนคนเดียว วิธีจับว่าไม่ได้ผลอยู่ในเนื้อเรื่อง) — [https://hn.algolia.com/api/v1/search?query=deep%20research%20my%20field&tags=comment&hitsPerPage=20&numericFilters=created_at_i%](https://hn.algolia.com/api/v1/search?query=deep%20research%20my%20field&tags=comment&hitsPerPage=20&numericFilters=created_at_i%20)

- [https://hn.algolia.com/api/v1/search?
query=sycophantic%20my%20research&tags=comment&hitsPerPage=12&numericFilters=created_at_i%3E1514764800000](https://hn.algolia.com/api/v1/search?query=sycophantic%20my%20research&tags=comment&hitsPerPage=12&numericFilters=created_at_i%3E1514764800000)
- [https://hn.algolia.com/api/v1/search?
query=subtly%20wrong%20analysis%20my&tags=comment&hitsPerPage=12&numericFilters=created_at_i%3E1514764800000](https://hn.algolia.com/api/v1/search?query=subtly%20wrong%20analysis%20my&tags=comment&hitsPerPage=12&numericFilters=created_at_i%3E1514764800000)
- https://note.com/natty_zebra3568/n/n31149cde3639

12 · การทดสอบปิดตาของคุณเอง

แก้วสามใบวางเรียงบนโต๊ะ หน้าตาเหมือนกันทุกอย่าง สองใบใส่ของอย่างเดียวกัน อีกใบใส่ของอีกอย่าง และคุณต้องชี้ให้ได้ว่าใบไหน

ถ้าเดาผิดถูกรอบ คุณจะถูกราวหนึ่งในสาม ข้อเท็จจริงง่ายๆ ข้อนี้คือสิ่งที่ทำให้แก้วสามใบกลายเป็นห้องลับ

การทดลองอ่านผลได้ก็เพราะรู้ล่วงหน้าว่า ถ้าสิ่งที่ถูกทดสอบไม่ได้ทำอะไรเลย ผลจะออกมาเท่าไร ในบทที่หนึ่ง นักฟิสิกส์รู้ว่าการเลือกกล่องมั่วๆ จะถูกราวสองครั้งครึ่งจากยี่สิบ เลขสามของเครื่องจึงไม่ได้บอกอะไร

แก้วสามใบให้ตัวเลขนี้มาฟรีๆ ถ้าลิ้นคุณแยกสองอย่างนั้นไม่ออก ชิมสิบสองรอบจะถูกเฉลยราวสี่รอบ รอบที่ถูกเกินจากนั้นต้องผ่านคำถาม "บังเอิญได้ไหม" และในบ้านคุณคำถามนั้นตอบได้เป็นตัวเลข

แก้วใบที่ไม่เข้าพวก

การทดสอบนี้ชื่อการทดสอบแบบสามเหลี่ยม (triangle test) ไม่ใช่เกมที่เล่นนี้คิดขึ้นมันเป็นมาตรฐานสากล ใช้ตอบว่าของสองอย่างต่างกันจนประสาทสัมผัสของคนจับได้หรือไม่ และใช้คัดเลือกว่าผักคนชิมด้วย

กติกาข้อหนึ่งของมันฟังดูโง่ร้าย คือห้ามตอบว่าไม่รู้ ไม่แน่ใจก็ต้องเดา เดาไปก็ไม่ทำให้ผลเสีย เพราะตัวเลขในบทนี้คิดเผื่อไว้แล้วว่าคนชิมอาจเดาถูกรอบ กติกานี้กันไม่ให้ใครเลือกที่หลังว่ารอบไหนนับ ถูกรอบถูกนับ

มาตรฐานเดียวกันบอกข้อจำกัดไว้ด้วย มันใช้ได้จำกัดกับของที่รสติดข้ามไปแก้วถัดไปมาก หรือค้างในปากนาน น้ำปลาสองยี่ห้อจึงไม่ใช่คู่ที่ควรลอง น้ำดื่ม โคล่า หรือกาแฟสำเร็จรูปสองยี่ห้อ เหมาะกว่ามาก

เลือกคู่ที่คุณเชื่อจริงๆ ว่าชิมแยกได้ ความเชื่อแบบนี้มีที่มาจากที่ฟังขึ้น คนที่ดื่มสองอย่างนั้นมานานย่อมมีเหตุให้คิดว่ารู้จักรสของมัน และอาจรู้จักจริงก็ได้ แก้วสามใบมีไว้ให้

ความเชื่อนั้นพิสูจน์ตัวในที่ที่มันแพ้ได้ การทดสอบที่คุณไม่มีอะไรจะเสีย ก็ไม่มีอะไรจะให้คุณ

แต่สองอย่างนั้นต้องหน้าตาเหมือนกันเมื่ออยู่ในแก้ว ถ้ามองด้วยตากก็แยกออก การทดสอบนี้จะวัดตา ไม่ได้วัดลิ้น ถ้าสีต่างกัน ให้ใช้แก้วที่บัพที่เหมือนกันทุกใบ มีฝาปิดและหลอด หรือให้คนชิมใช้ผ้าปิดตา

ตัวอย่างในบทนี้ใช้โคล่าสองยี่ห้อ เรียกว่า ก กับ ข แผ่นที่ต้องเขียนก่อนเริ่มหน้าตาแบบนี้

ข้ออ้าง: ฉันชิมแยกโคล่ายี่ห้อ ก ออกจากยี่ห้อ ข ได้ ชิม 12 รอบ ถ้าชี้แก้วที่ไม่เข้าพวกถูกตั้งแต่ 8 รอบขึ้นไป ถือว่าข้ออ้างนี้ผ่าน ถ้าถูกไม่ถึง 8 รอบ ถือว่าการทดสอบครั้งนี้ไม่ได้แสดงว่าแยกได้ แก้ว เครื่องดื่ม และห้องนี้ ไม่มีอะไรขัดขวางการชิม ลงชื่อ วันที่

กระดาศหนึ่งแผ่น ลูกเต๋าหนึ่งลูก

ที่ต้องใช้มีแค่นี้ คนในบ้านอีกหนึ่งคนขึ้นไป แก้วกระดาศเหมือนกันราวสี่สิบใบ ปากกา เมจิก ลูกเต๋า เหรียญ น้ำเปล่า และแผ่นข้างบน

1. **เซ็นก่อนแตะแก้ว** เขียนแผ่นข้างบนด้วยปากกา เซ็นชื่อ ลงวันที่ ก่อนมีใครทำอะไรลงแก้ว
2. **ชิมรอบที่รู้คำตอบหนึ่งรอบ** คนในบ้านหนึ่งคนเป็นคนเตรียม เขายกแก้วสามใบมาพร้อมบอกว่าแก้วไหนคือยี่ห้ออะไร คุณชิมให้รู้ถึงความต่างที่เชื่อว่ามีนั้นรสเป็นอย่างไร รอบนี้ไม่นับ ถ้าเห็นปัญหา เช่น สีไม่เหมือนกัน แก้วแล้วเขียนแผ่นใหม่
3. **ให้ลูกเต๋ากับเหรียญเลือก** ทุกรอบที่นับ คนเตรียมทยอยลูกเต๋า ออก 1 หรือ 2 แก้วที่ไม่เข้าพวกคือแก้ว 1 · ออก 3 หรือ 4 คือแก้ว 2 · ออก 5 หรือ 6 คือแก้ว 3 แล้วโยนเหรียญ ออกหัว แก้วนั้นใส่ ก อีกสองแก้วใส่ ข ออกก้อยก็สลับกัน เขาจดค่าเฉลี่ยไว้ในครัว เช่น "รอบ 1 · แก้ว 3 · ก" และไม่บอกใครจนครบ
4. **เตรียมแก้วในที่ที่คุณมองไม่เห็น** คนเตรียมเขียนเลข 1 2 3 ข้างแก้วชุดใหม่ แล้วเทตามค่าเฉลี่ย แก้วละสองสามอีกเท่ากัน จากขวดที่เย็นเท่ากัน เทสามแก้วรวด

เดียว ถ้าเป็นโคล่า ให้เปิดสองขวดพร้อมกันตั้งแต่ต้น ความซ่าจะได้ไม่กลายเป็น เบาะแส

5. ซิมโดยไม่มีใครในห้องรู้คำตอบ ถ้ามีคนที่สาม ให้เขายกถามาโดยไม่รู้ว่าแก้วไหนคืออะไร ถ้ามีกันสองคน คนเตรียมวางถาดตอนคุณอยู่อีกห้อง แล้วกลับเข้าครัวก่อนคุณเดินมา ซิมตามเลข จิบน้ำเปล่าคั้นทุกแก้วและทุกรอบ ซีนึงแก้ว ไม่แนใจก็เดา

6. จดทันที แล้ววนกลับไปขั้นที่สาม เขียนคำตอบลงแผ่นที่เซ็นไว้ด้วยปากกา เช่น "รอบ 1 · แก้ว 2" ห้ามลบ ห้ามแก้ไขอะไรที่ดูแปลก เช่น แก้วหนึ่งชานน้อยกว่า จดไว้ข้างรอบนั้นเลย ทำซ้ำจนครบสิบสองรอบ

7. ครบแล้วค่อยนับ เอาคำเฉลยจากครีมาวางคู่กับแผ่นของคุณ ชีตถูกรอบที่ตรงกัน นับ แล้วเทียบกับตารางข้างล่าง

ถ้าไม่มีลูกเต๋า คนเตรียมให้ AI สุ่มลำดับทั้งสิบสองรอบไว้ก่อนก็ได้ พิมพ์ว่า "สุ่มให้หน่อย 12 รอบ แต่ละรอบให้เลขหนึ่งตัวจาก 1 2 3 และตัวอักษรหนึ่งตัวจาก ก ข" แล้วตรวจด้วยตาว่าครบสิบสองแถว และไม่ได้เรียงเป็นแบบแผนที่เราได้ อย่าง 1 2 3 1 2 3 คนถามต้องเป็นคนเตรียม ไม่ใช่คุณ

ขั้นที่ห้าคือขั้นที่ห้ามลัด คุณต้องไม่รู้ เพราะคนที่อยากให้อ้างอิงจริงที่สุดในห้องคือคุณเอง คนในห้องก็ต้องไม่รู้ด้วย เพราะในบทที่หนึ่ง ผู้ทดสอบสรุปว่าคนที่รู้คำตอบ บังคับเสวอากาศได้โดยอาจไม่รู้ตัว คนยกถาดที่รู้ก็อาจบอกไขให้คุณได้แบบเดียวกัน นี่คือการทดสอบแบบปิดตาสองชั้น ในขนาดเล็กที่สุด

รอบที่ไม่นับ

ขั้นที่สองยืมมาจากห้องบอลรูมตรงๆ ในรอบที่ทหารรู้ว่าของอยู่กล่องไหน เครื่องชี้ถูกสีจากสี และรอบนั้นไม่ถูกนับรวมในยี่สิบรอบ

รอบที่รู้คำตอบของคุณจะ "สำเร็จ" ด้วยเหตุผลเดียวกันทุกประการ คุณจะชี้ถูกเพราะรู้อยู่แล้ว และต่อให้ความต่างชุดजनिकไม่ออกว่าจะพลาดได้อย่างไร ความชด้นนั้นก็พื้นฐานไม่ได้ ในรอบนี้ฉันไม่ได้ถูกถามอะไรเลย คำตอบมาถึงก่อนสิ้น

จดไว้ที่ขอบกระดาษด้วยว่ารอบนี้คุณมั่นใจแค่ไหน จากศูนย์ถึงสิบ ตอนนับผล ตัวเลขนี้จะบอกว่าคุณมั่นใจตอนรู้คำตอบ ทำนายผลตอนไม่รู้คำตอบได้แค่ไหน

รอบนี้ไม่ได้มีไฉไลอะไรใคร มันให้ความเชื่อของคุณได้โอกาสดีที่สุด คุณได้รู้ว่าคุณมองหาอะไรในห้องนี้กับแก้วชุดนี้ และถ้าคุณพอใจว่าความต่างยังอยู่ครบ ข้อแก้ตัวเรื่องห้องหรือแก้วก็หมดที่ยืน

จากนั้นคนเตรียมกลับเข้าครัว ลูกเต๋าถูกทอย และแก้วชุดแรกที่ไม่มีการในห้องพักคำตอบก็มาวางตรงหน้าคุณ ตรงนี้คือจุดที่ทหารเดินเข้ารอบที่สอง และเป็นจุดเดียวในบทนี้ที่เล่มนี้บอกไม่ได้ว่าจะเกิดอะไรขึ้น

เจ็ดยังไม่พอ

เทียบจำนวนขีดกับแถวของจำนวนรอบที่คุณขีด

ขีดที่รอบ	ต้องถูกอย่างน้อย	โอกาสที่เดามั่วแล้วผ่าน
6	5	1.8%
9	6	4.2%
10	7	2.0%
12	8	1.9%
15	9	3.1%
20	11	3.8%

เกณฑ์ทุกแถวตั้งด้วยกติกาเดียว คือจำนวนที่น้อยที่สุดที่การเดามั่วไปถึงได้ไม่เกินราวหนึ่งครั้งในยี่สิบ ถ้าอยากขีดจำนวนรอบอื่น เลือกแถวก่อนเริ่ม แล้วเขียนเกณฑ์ของแถวนั้นลงแผ่นที่เขียน แถวที่เลือกหลังเห็นผลไม่นับ

กลับไปทีโคล่า สมมติว่าครบสิบสองรอบแล้ว แผ่นของคุณมีขีดถูกเจ็ดขีด

เจ็ดจากสิบสองฟังดูดี มันเกินครึ่ง และเกือบสองเท่าของสี่รอบที่การเดาให้โดยเฉลี่ย แต่มันต่ำกว่าแปดที่คุณเขียนไว้ และการเดามั่วไปถึงเจ็ดรอบขึ้นไปได้ราว 6.6% หรือ

ราวหนึ่งครั้งในสิบห้า ถ้าคนสิบห้าคนที่ลื่นแยกอะไรไม่ออกเลยลองคนละครั้ง โดยเฉลี่ยจะมีหนึ่งคนได้เจ็ดขึ้นไป และคนนั้นคือคนที่จะมีเรื่องมาเล่า

เจ็ดจึงยังไม่พอ มันไม่ได้แปลว่าคุณแยกไม่ออก มันแปลว่าการทดสอบครั้งนี้ยังแยกไม่ออกกว่าเป็นฝีมือลื่นหรือฝีมือโชค

ถ้าอยากพูดว่าโคล่าเริ่มหายซ่า หรือลื่นเริ่มซ่า ดูที่ขอบกระดาษก่อน ข้อสังเกตที่จดไว้ก่อนเปิดคำเฉลยคือข้อมูล ข้อสังเกตที่นึกออกหลังรู้ผล ต่อให้จริง ก็แยกไม่ออกจากข้อแก้ตัว เหมือนกลืนกล้วยกากับบันไดขึ้นห้องบอลรูม

ลองใหม่ได้เสมอด้วยแผ่นใหม่ที่เซ็นก่อนชิม แต่นับทุกครั้งที่ลอง ถ้าลื่นแยกไม่ออกเลย โอกาสที่โชคจะพาผ่านเส้นแปดจากสิบสองได้น้อยหนึ่งครั้ง จะโตขึ้นทุกครั้งที่ลองใหม่ ลองสองครั้งราว 3.7% สามครั้งราว 5.5% ห้าครั้งราว 9% และสิบครั้งราว 17% หรือราวหนึ่งครั้งในหก ลองไปเรื่อยๆ จนวันหนึ่งผ่าน วันนั้นอาจเป็นแค่วันที่ความบังเอิญมาถึง

แปดคือเส้นที่คุณเซ็นไว้ การเดามั่วไปถึงแปดรอบขึ้นไปได้ราว 1.9% หรือราวหนึ่งครั้งในห้าสิบ ได้แปดจึงผ่านเส้นนั้น ถ้าแผ่นนั้นมีขีดถูกเก้าขีด การเดาไปถึงได้แค่ราว 0.4% หรือราวหนึ่งครั้งในสองร้อยห้าสิบ ผลแบบนี้ยังน่าเชื่อว่าลื่นคุณแยกโคล่าสองยี่ห้อนั้นได้จริง อย่างน้อยก็ในคืนนั้น

แม้จะผ่าน มันก็บอกแค่ว่าลื่นคุณแยกสองอย่างนี้ได้ ตัวมาตรฐานเขียนไว้เองว่ามันไม่บอกว่าต่างกันมากแค่ไหนหรือไปทางไหน ยี่ห้อไหนอร่อยกว่าเป็นอีกข้ออ้างหนึ่ง

ถ้าอยากรู้โอกาสของจำนวนรอบที่ไม่มีในตาราง AI ช่วยคิดเลขได้ แต่ใช้เท่าที่สามจากบทที่สิบเอ็ด ถามสิ่งที่คุณรู้คำตอบก่อน เช่น "โอกาสเดาถูกหนึ่งในสามทุกรอบ ชิมสิบสองรอบ โอกาสที่จะถูกแปดรอบขึ้นไปคือเท่าไร" ถ้ามันไม่ตอบราว 1.9% อย่าใช้ตัวเลขอื่นของมัน

แผ่นจดให้เก็บไว้กับตัวคุณ ส่วน AI ได้แค่ตัวเลขที่นับเสร็จแล้ว แผ่นที่มีลายเซ็นก่อนชิมกับคำตอบทุกรอบคือหลักฐานที่คนอื่นเอามานับซ้ำได้ มันทำหน้าที่เดียวกับกล้องที่ถ่ายไว้ทุกรอบในห้องบอลรูม

ทหารสองคนในบทที่หนึ่งได้เจอรอบที่สองเพราะผู้พิพากษาสั่ง ตอนนั้นคำชี้ของเครื่องถูกใช้ค้นบ้านคนมาแล้ว ความเชื่อของคุณได้เจอรอบที่สองบนโต๊ะในบ้าน ก่อนที่จะมีอะไรขึ้นอยู่กับมัน

ไม่ว่าแผ่นจดจะจบที่เลขอะไร คุณได้ทำสิ่งที่ทหารสองคนนั้นไม่เคยมีโอกาสดำเนินการคือให้ความเชื่อของตัวเองมีทางแพ้

ลองคิดนี่: เลือกเครื่องดื่มสองอย่างที่คุณเชื่อว่าชิมแยกได้ เขียนข้ออ้างพร้อมเลขสิบสองกับแปดแล้วเซ็นชื่อ ให้คนในบ้านทยอยลูกเต๋ากับโยนเหรียญเลือกแก้ว เตรียมแก้วในครัว แล้วชิมให้ครบสิบสองรอบก่อนเปิดคำเฉลย

จะรู้ว่บทนี้ผิดเมื่อ: มีคนคำนวณตารางหรือตัวเลขในตัวอย่างซ้ำจากโอกาสหนึ่งในสามต่อรอบ แล้วได้ไม่ตรงกับที่พิมพ์ไว้ หรือมีคนทำตามทุกขั้นครบแล้วพบว่าคนชิมยังรู้คำตอบได้โดยไม่ต้องชิม เช่น จากหน้าตาแก้ว เสียง หรือท่าทีของคนในห้อง

หลักฐาน:

- การทดสอบแบบสามเหลี่ยมเป็นมาตรฐานสากล ใช้ตัดสินว่าของสองอย่างต่างกันจนรับรู้ได้ด้วยประสาทสัมผัสหรือไม่ ผู้ชิมต้องเลือกทุกครั้ง ใช้ได้จำกัดกับของที่รสติดขัดตัวอย่างมากหรือรสค้างนาน ไม่ได้บอกว่าต่างกันมากแค่ไหนหรือไปทางไหน และใช้คัดเลือกกับฝึกคนชิมได้ — **เน้น** — ISO 4120:2021 Sensory analysis — Methodology — Triangle test ฉบับที่ 3 ยังเป็นฉบับปัจจุบันเมื่อเปิดดูวันที่ 25 กันยายน 2026 ทุกข้อในบรรทัดนี้มาจากบทคัดย่อบนหน้าของ ISO — <https://www.iso.org/standard/76666.html>
- ตารางเกณฑ์ผ่านทั้งหกแถว และโอกาสที่การเดามั่วสิบสองรอบจะถูกเจ็ด แปด และเก้ารอบขึ้นไป ($6.6\% \cdot 1.9\% \cdot 0.4\%$) · โอกาสที่โชคจะพาผ่านเส้นแปดอย่างน้อยหนึ่งครั้งเมื่อลองซ้ำ 2 · 3 · 5 · 10 ครั้ง ($3.7\% \cdot 5.5\% \cdot 9\% \cdot 17\%$) คิดจาก $1 - (1 - 0.0188)^{\text{ยกกำลังจำนวนครั้ง}}$ คำนวณวันที่ 26 กันยายน 2026 — **เน้น** — เลขคณิตของการแจกแจงทวินาม (binomial) ที่โอกาสสูงหนึ่งในสามต่อรอบ คำนวณวันที่ 25 กันยายน 2026 ตรวจสอบได้ในโปรแกรมตารางคำนวณ เช่น พิมพ์ =BINOM.DIST(7,12,1/3,TRUE) จะได้ว่า 0.981 แล้วเอาหนึ่งลบด้วยค่านั้น จะได้ราว 0.019 · เกณฑ์ทั้งหกแถวตรงกับคอสถณัมระดับ 5% ในตาราง 17.8 ของ Meilgaard และคณะ (2007) ซึ่งพิมพ์ซ้ำไว้ในภาคผนวก H ของเอกสารในคลังของ Wageningen University & Research — <https://edepot.wur.nl/166289>
- การทดสอบที่เม็กซิโกซึ่งบทนี้ยืมแบบมา · เหรียญสามครั้งเลือกกล่องหนึ่งในแปด · ทิมชอนกับทิมหาไม่เจอหน้ากัน · ทหารเซ็นรับรองห้อง กล่อง และของกลางก่อนเริ่ม · รอบที่คนถือรู้คำตอบได้สี่จากสี่และไม่อยู่ในยี่สิบรอบที่นับ · การทดสอบเกิดขึ้นเพราะผู้พิพากษาสั่ง หลังจากคำชี้ของเครื่องถูกใช้ค้นบ้านคน — **พอเชื่อ** (การทดสอบหนึ่งครั้ง บันทึกละเอียดทุกขั้นและมีวิดีโอ) — Mochán, W. L., & Ramírez-

Solís, A. (2013). Effectiveness of the GT200 Molecular Detector: A Double-Blind Test.
arXiv:1301.3971 — <https://arxiv.org/abs/1301.3971>

13 · เล่มนี้แพ้ได้ตรงไหน

ระหว่างที่เล่มนี้ถูกเขียน มีสี่เรื่องที่เกี่ยวข้องพิมพ์ผิด ทั้งสี่อยู่ในบทก่อนๆ ในฉบับที่แก้แล้ว

หนังสือที่สอนให้ข้ออ้างแพ้ได้ ต้องบอกคุณก่อนว่าตัวเองจะแพ้ตรงไหน

สี่เรื่องที่เกือบพิมพ์ผิด

บทที่หนึ่งเกือบเขียนว่าหน่วยงานวิทยาศาสตร์แห่งชาติทดสอบเครื่องบินปัดตาในปี 2010 แหล่งที่เปิดได้บอกแค่ว่ากระทรวงถูกสั่งให้ทดสอบ หน่วยงานนั้นมาตรวจเครื่องบินที่หลังสิบเอ็ดปี เพื่อใช้ในคดี

บทที่สองเกือบยกกลาหมสูตรว่า อย่าปลงใจเชื่อ พระไตรปิฎกฉบับหลวงใช้คำว่า อย่าได้ยึดถือ และสิ่งที่ให้รู้ด้วยตนเองรวมถึงการรู้ว่าผู้รู้ดีเถียงหรือไม่ ฉบับเต็มจึงมีคนอื่นอยู่ในการตัดสินใจด้วย

บทที่สามเกือบเขียนว่าตำรายาโบราณสั่งให้แช่สมุนไพรในน้ำเย็น ในตำราไม่มีคำว่าเย็น มันแค่ไม่ได้บอกให้ต้ม คำว่าเย็นและความคิดว่าอาจต้องหลบความร้อนเป็นของนักวิทยาศาสตร์ที่อ่านมัน

บทที่ห้าเกือบเล่าให้หมอประจำเรือเป็นวีรบุรุษที่พบว่าผลไม้ตระกูลส้มรักษาโรค ลักปิดลักเปิด นักประวัติศาสตร์พบว่าเขาไม่เคยเขียนไว้ชัดอย่างนั้น และเรื่องวีรบุรุษสืบค้นทางได้ถึงปี 1893 การทดลองมีจริง บทพระเอกไม่จริง

ทั้งสี่เรื่อง ฉบับที่ผิดเล่าง่ายกว่า และเข้ากับสิ่งที่เล่มนี้อยากพูดมากกว่า คำถามข้อที่ห้าฉันอยากให้มันจริงไหม ใช้กับหนังสือได้เท่ากับใช้กับคน

สิ่งที่จับมันได้ทั้งสี่ครั้งคือการเปิดแหล่งต้นทาง แล้วอ่านว่ามันพูดอะไรจริงๆ

หลักฐานสองร้อยแถว

หลักฐานของเล่นนี้ถูกหาโดย AI ถ้าเล่นนี้เพราะ มันเพราะตรงนี้

AI จัดตัวทำงานพร้อมกัน หกตัวค้นในห้าภาษา ได้หลักฐานราวสองร้อยแถว AI อีกตัวที่ไม่ได้อยู่ในเจ็ดตัวนั้นเปิดหน้าต่างทางซ้ายแปดแถว และเปิดเองอีกสิบเจ็ดหน้า ระหว่างเขียนมีการเปิดเพิ่มอีกสิบสี่หน้า

อะไรที่เปิดดูไม่ได้เลย ทั้งต้นทางและแหล่งที่อ้างถึง เล่นนี้ตัดทิ้ง แม้จะเข้ากับเรื่องที่ยากเล่าพอดี

หลังเขียนเสร็จ ผู้ตรวจสอบสองรายที่ไม่ได้เขียนเปิดแหล่งซ้ำ และพบว่าคราวนี้ผิดไปแล้วจริงๆ บทที่หนึ่งให้ทหารคนเดียวได้ทั้งสี่จากสี่และสามจากยี่สิบ แต่รายงานแสดงทหารสองคนผลัดกัน บทที่สิบเอ็ดโทษสัญญาณตัวเดียว ทั้งที่บริษัทบอกว่าหลายอย่างเปลี่ยนพร้อมกัน บทที่แปดสรุปจากรายการที่มีแต่ข่าวปลอม และบทที่สิบเอ็ดยกคำแนะนำที่ผู้ใช้คนหนึ่งไม่เคยให้ ทั้งสี่ถูกแก้ก่อนพิมพ์

ถามเล่นนี้ว่าใครตรวจแล้วบ้าง คำตอบที่ชัดเจนคือ ยังไม่มีคนนอกตรวจ และคนนอกคนแรกคือคุณ

แปดข้อ ชุดที่สอง

ถึงตาคุณแล้ว เริ่มจากข้อเดิม โดยไม่เปิดดูประโยคแรกของคุณ

เขียนหนึ่งประโยคว่าวิทยาศาสตร์คืออะไร

แล้วตอบแปดข้อ กติกาเดิม ข้อ 8 เปิดตารางในบทที่ 12 ได้

1. สเปร์ย์พ่นคอยี่ห้อหนึ่งโฆษณาว่า "เจ็บคอดีขึ้นในสามวัน ผู้ใช้พอใจ 95%" ก่อนเชื่อว่าสเปร์ย์เป็นเหตุ ข้อมูลไหนสำคัญที่สุด

ก. ผู้ใช้ที่คนที่เจ็บคอดีขึ้นแล้วกลับมาซื้อสเปร์ย์ซ้ำเป็นขวดที่สอง ข. สเปร์ย์มีสมุนไพรที่ชนิด และแต่ละชนิดมีงานวิจัยว่าช่วยอะไร ค. คนเจ็บคอแบบเดียวกันที่ไม่ได้ใช้

สเปรย์ ดีขึ้นภายในกัวัน ง. แพทย์หูคอจมูกก็คนที่รับรองสเปรย์นี้ โดยไม่ได้รับเงินจากบริษัท

2. โรงเรียนกวดวิชาแห่งหนึ่งติดป้ายว่า "ปีนี้ศิษย์เราติดแพทย์ 50 คน" ก่อนเชื่อว่าที่นี่ช่วยให้ติด ข้อมูลไหนสำคัญที่สุด

ก. ศิษย์ทั้งหมดมีก็คน และคนเก่งพอๆ กันที่ไม่ได้เรียนที่นี่ สอบติดก็คนในร้อย ข. ห้าสิบคนนั้นเรียนที่นี่กี่เดือนก่อนสอบ และแต่ละคนลงคอร์สไหนกับครูคนไหนบ้าง ค. ครูที่สอนเป็นแพทย์หรือเคยสอบติดแพทย์เอง ไม่ได้เป็นแค่ตัวเตอรืทั่วไป ง. ป้ายมีรูป ชื่อจริง และคณะที่ติดของห้าสิบคน ให้ตรวจได้ว่ามีตัวตนจริงไหม

3. เครื่องกรองน้ำรุ่นหนึ่งโฆษณาว่า "เติมพลังงานให้น้ำ ต็มแล้วร่างกายสดชื่นขึ้น" ข้อไหนบอกปัญหาของข้ออ้างนี้ได้ดีที่สุด

ก. บริษัทไม่บอกว่าเติมพลังงานด้วยแร่หรือเทคโนโลยีอะไร จึงไม่รู้ว่าทำงานอย่างไร ข. โฆษณานี้ไม่บอกว่าผลแบบไหนจะแสดงว่ามันไม่จริง ต้องระบุก่อนว่าวัดอะไร ค. รีวิวของคนที่ดื่มแล้วรู้สึกสดชื่นยังมีน้อย ต้องรอให้มากกว่านี้จึงจะสรุปได้ ง. ราคาแพงกว่าเครื่องกรองทั่วไปหลายเท่า ทั้งที่กรองได้สะอาดพอๆ กัน

4. แพทย์ตัวจริงคนหนึ่งออกคลิปเตือนว่าyaladความดันที่คนในบ้านคุณกินอยู่ "อันตรายกว่าที่เคยรู้กันมา" ขึ้นแรกที่ดีที่สุดคือ

ก. หยุดยาไว้ก่อนจนกว่าจะแน่ใจว่าปลอดภัย แล้วค่อยถามแพทย์ในนัดครั้งหน้า ข. ดูว่าแพทย์ในคลิปจบเฉพาะทางหัวใจ และมีผู้ติดตามมากพอจะเชื่อได้ไหม ค. หากคลิปของแพทย์คนอื่นเรื่องยานี้ แล้วดูว่าฝั่งไหนมียอดไลค์มากกว่า ง. ดูว่าหน่วยงานที่ทบทุนหลักฐานว่าอย่างไร แล้วถามแพทย์ที่สั่งยาก่อนเปลี่ยนอะไร

5. มีข่าวงานวิจัยใหญ่เรื่องอาหารที่คุณชอบที่สุดและกินแทบทุกวัน ก่อนเปิดอ่าน ข้อไหนกันความชอบไม่ให้มาตัดสินแทนคุณได้ดีที่สุด

ก. ตั้งใจอ่านแค่บทคัดย่อกับข้อสรุป ไม่ให้ตัวเลขทางสถิติทำให้ตีความผิด ข. เขียนลงกระดานก่อนเปิด ว่าตัวเลขแบบไหนจะทำให้คุณยอมกินน้อยลง ค. ถามเพื่อนที่ชอบอาหารเดียวกันก่อน ว่าเขาอ่านข่าวนี้แล้วคิดอย่างไร ง. บอกตัวเองว่าจะอ่านอย่างเป็นกลาง และจะไม่เข้าข้างอาหารที่ตัวเองชอบ

6. AI อ่านสัญญาเช่าบ้านยี่สิบหน้าให้คุณ แล้วสรุปว่า "ย้ายออกก่อนครบปีได้ไม่เสียมัดจำ ตามข้อ 12" ก่อนแจ้งเจ้าของบ้าน ควรทำอะไรก่อนอื่น

ก. ให้ AI สรุปสัญญาเช่าอีกรอบในห้องแชตใหม่ ถ้าได้คำตอบเดิมก็แจ้งย้ายได้ ข. ถามเพื่อนที่เคยเช่าบ้านหลังนี้ว่าย้ายออกก่อนครบปีแล้วได้มัดจำคืนไหม ค. อ่านข้อ 12 ในสัญญาด้วยตัวเอง ว่ามีจริงและเขียนอย่างสรุปมาใหม่ ง. ให้ AI อีกตัวอ่านสัญญาฉบับเดียวกัน แล้วดูว่าสรุปตรงกับตัวแรกไหม

7. คุณกับพี่ชายเถียงกันว่าขายบ้านเก่าของครอบครัวไหม คุณเล่าให้ AI ฟังจากมุมมองของคุณ มันเห็นด้วยทุกข้อ ทางไหนช่วยแก้การเออออตามคุณได้ดีที่สุด

ก. ส่งคำตอบของ AI ให้พี่ชายอ่าน เพราะมันไม่อยู่ฝ่ายไหน และตัดสินใจเหตุผลที่คุณเล่าให้ฟัง ข. ถามต่อในห้องเดิมว่ามันมั่นใจก็เปอร์เซ็นต์ และให้อธิบายเหตุผลทีละข้อ ค. เล่าเรื่องเดิมจากมุมมองของคุณให้ AI อีกตัวฟัง แล้วดูว่ามันเห็นด้วยเหมือนตัวแรกไหม ง. เริ่มแชตใหม่ เล่าเหตุผลของทั้งสองคนแบบคนนอก แล้วขอให้มันแย้งแทนพี่ชายให้ดีที่สุด

8. เพื่อนยืนยันว่าแยกนมจืดสองยี่ห้อได้ คุณจัดทดสอบแก้วสามใบแบบในบทที่ 12 ยี่สิบรอบ เขาชี้แก้วที่ต่างถูกเฝ้ารอบ ข้อไหนสรุปได้ดีที่สุด

ก. มากกว่าที่เคาสุ่มได้โดยเฉลี่ย แต่ไม่ถึงเกณฑ์สิบเอ็ดของยี่สิบรอบ จึงยังสรุปไม่ได้ ข. เกินแปดที่เป็นเกณฑ์ผ่านของการทดสอบแบบนี้ จึงเป็นหลักฐานพอว่าเขาแยกได้จริง ค. ยังไม่ถึงครึ่งของยี่สิบรอบ จึงแสดงว่านมสองยี่ห้อนี้รสชาติไม่ได้ต่างกัน ง. ห้ารอบแรกล้นอาจยังไม่ชิน ควรตัดทิ้งไป แล้วนับใหม่จากสิบห้ารอบที่เหลือ

ตอบครบแล้วเปิดเฉลยท้ายเล่ม ให้คะแนนแปดข้อและสองประโยค ถ้าตัวหลังไม่สูงกว่า เล่มนี้แพ้ตามที่บอกไว้

เหรียญหนึ่งเหรียญกับเพื่อนหกคน

ถ้าคะแนนสูงขึ้น เล่มนี้ก็ยังไม่ได้นะ ถามคะแนนนั้นด้วยคำถามข้อแรกของเล่มเอง เทียบกับอะไร

สองชุดนี้เล่มนี้เขียนเอง ยังไม่มีใครวัดว่ายากเท่ากันไหม การเคยสอบมาแล้วอาจช่วย
เอง และแปดข้อน้อยพอที่ข้อหนึ่งจะมาจากการเดา คะแนนของคุณคนเดียวจึงบอกได้
แค่ว่าเล่มนี้แพ้กับคุณไหม

ถ้าอยากรู้ว่าได้ผลไหม หาเพื่อนที่ยังไม่เคยอ่านอย่างน้อยหกคน ให้ทุกคนทำชุดแรก
แล้วเซ็นไว้ก่อนแบบทหารในบทที่หนึ่ง ว่าถ้าฝั่งที่ได้อ่านคะแนนเพิ่มไม่มากกว่าฝั่งที่รอ
เล่มนี้แพ้ยั่ว จั๋ว โยนเหรียญเลือกว่าใครในคู่ได้อ่านก่อน เมื่อฝั่งนั้นอ่านจบ ให้ทุกคนทำ
ชุดที่สองวันเดียวกัน

ฝั่งที่รอคือกลุ่มเปรียบเทียบ ถ้าชุดที่สองง่ายกว่า หรือการสอบซ้ำช่วยเอง ฝั่งนี้ก็ดีขึ้น
ด้วย ส่วนที่ผู้อ่านดีขึ้นเกินกว่านั้นจึงเป็นของเล่มนี้ ในหกคน ต่างกันข้อสองข้อยังเป็น
ความบังเอิญได้ง่าย ให้เล่าทุกตัวเลข รวมตัวที่ไม่เข้าข้างเล่มนี้

ตอนที่เล่มนี้เขียนเสร็จ ยังไม่เคยมีใครทดสอบมันด้วยเหรียญแบบนี้

สิบคาบ

คำตอบที่ชัดเจนว่าควรหวังแค่ไหน มาจากนักเรียนมัธยมกว่าหนึ่งหมื่นคนในเคนยา
รวันดา และยูกันดา ที่โรงเรียนถูกสุ่มว่าแห่งไหนได้เรียนวิธีตัดสินใจอ้างเรื่องสุขภาพ
สิบคาบ

ผู้วิจัยเขียนไว้เองว่า ในโรงเรียนที่ได้เรียน สิบสองคนในร้อยยังไม่ถึงเกณฑ์ผ่าน และผู้
วิจัยหลายคนเป็นทั้งคนสร้างบทเรียนและคนประเมินมัน เล่มนี้ก็ยืนอยู่ที่เดียวกัน

หนังสือที่อ่านคนเดียวน้อยกว่าสิบคาบกับครูตัวจริง และรายงานปี 2016 ของสถาบัน
วิทยาศาสตร์แห่งชาติสหรัฐฯ เตือนว่ารู้วิทยาศาสตร์มากขึ้นไม่ได้เปลี่ยนการตัดสินใจ
เสมอไป

แต่ถ้าหยุดตรงนี้ ภาพจะมีดีกว่าความจริง ในการทดลองเดียวกัน โรงเรียนที่ได้เรียนมี
คนผ่านมากกว่าอย่างชัดเจน และหนึ่งปีต่อมา สิ่งที่ยังเรียนอยู่เกือบเก้าในสิบส่วน

ความหวังนี้ไม่ได้มาจากการเชื่อว่าคุณจะไม่เปลี่ยน มันมาจากการทดสอบที่อาจออกมาว่าสอนไปก็เปล่า แล้วมันไม่ได้ออกมาอย่างนั้น ความหวังที่แพ้ได้แบบนี้ คือความหวังที่เชื่อได้

คำถามห่าข้อนี้ไม่ได้เป็นของใคร และไม่ได้มาฟรี มันถูกจ่ายมาแล้วโดยทุกคนที่ยอมให้ข้ออ้างของตัวเองแพ้ได้ นักฟิสิกส์ที่พิมพ์ว่าเหรียญของตัวเองแพ้เครื่องตรวจระเบิดปลอม นักวิทยาศาสตร์ที่ทำงานในสาขาของตัวเองแล้วพิมพ์ผลตามที่มันออก คนที่ทดสอบสมมุติแล้วพิมพ์ว่ามันไม่ได้ช่วยเพิ่ม บริษัทที่ถอน AI ของตัวเองแล้วบอกว่าพลาดตรงไหน

ตอนนี้มันเป็นของคุณ

จะรู้วาทน์ผิดเมื่อ: ผู้อ่านใหม่ที่ชุดแรกยังไม่เต็ม ได้คะแนนชุดที่สองไม่สูงกว่าชุดแรก หรือประโยคท้ายเล่มได้คะแนนไม่มากกว่าประโยคต้นเล่ม หรือในการทดลองที่ใช้เหรียญเลือกว่าใครได้อ่านก่อน ฝั่งที่ได้อ่านคะแนนเพิ่มไม่มากกว่าฝั่งที่รอ หรือมีคนพบวิธีเลือกคำตอบโดยไม่อ่านเนื้อหา เช่นเลือกตัวเลือกที่ยาวที่สุด แล้วได้กินสามข้อในชุดใดชุดหนึ่ง หรือมีคนเปิดแหล่งต้นทางแล้วพบว่าเรื่องใดที่บ่นนี้บอกว่าแก้แล้วฉบับที่แก้แล้วยังผิดอยู่

หลักฐาน:

- การทดสอบเครื่องตรวจระเบิดในไทยปี 2010 มาจากคำสั่งของนายกรัฐมนตรีถึงกระทรวงวิทยาศาสตร์และเทคโนโลยี แลงผลวันที่ 16 กุมภาพันธ์ 2010 · สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติตรวจ 757 เครื่องให้กองทัพบกในปี 2021 ถึง 2022 เพื่อใช้ในคดี — **พอเชื่อ** (รายงานข่าวย้อนหลังเล่มนี้หาเอกสารการทดสอบปี 2010 ฉบับเต็มไม่พบ) — THE STANDARD, 6 มิถุนายน 2022 — <https://thestandard.co/gt200-history/>
- พระไตรปิฎกฉบับหลวง เล่ม 20 เกสพุตตสูตร ใช้คำว่า อย่าได้ยึดถือ นำหน้าเหตุทั้งสิบข้อ ไม่มีคำว่า ปลงใจเชื่อ และสิ่งที่ให้รู้ด้วยตนเองรวมถึงสิ่งที่ผู้รู้ดีเตือน กับผลเมื่อทำตามจนครบ — **แน่น** (ตัวบทต้นฉบับ ถูกเปิดอ่านซ้ำแล้ว) — พระไตรปิฎก ฉบับหลวง เล่มที่ 20 — https://84000.org/tipitaka/pitaka_item/v.php?B=20&A=4930&Z=5092
- ปาฐกถารางวัลโนเบลของ Tu Youyou ยกประโยชน์จากตำราของ Ge Hong (ราวปี 317 ถึง 420) ว่า Qinghao หนึ่งกำมือ แช่น้ำ บิดเอาน้ำ ดื่มให้หมด ไม่มีคำบอกอุณหภูมิ · ปาฐกถาเรียกวิธีนี้ว่าการใช้ Qinghao แบบเย็น และบอกว่าสมมุติไพร่ส่วนมากถูกต้มเป็นยาต้มก่อนให้คนไข้กิน · ความคิดว่าอาจต้องหลบความร้อนตอนสกัดเป็นของผู้บรรยายเอง — **พอเชื่อ** (คำเล่าของผู้ค้นพบเองในเอกสารชิ้นเดียว แต่ยกตัวบทภาษาจีนไว้ให้ตรวจ) — Tu Youyou, Nobel Lecture, 7 ธันวาคม 2015 — <http://www.nobelprize.org/uploads/2018/06/tu-lecture.pdf>

- การทดลองของ James Lind เริ่ม 20 พฤษภาคม 1747 คนไข้สิบสองคน วิธีละสองคน · เขาไม่เคยเขียนชัดว่าผลไม้ตระกูลส้มรักษาโรคคลักปิดลักเปิด · น้ำผลไม้ที่เขาเคี้ยวเก็บไว้เสียวิตามิน · น้ำเลมอนถูกแจกทั่วไปในกองทัพเรือตั้งแต่ปี 1795 · เรื่องหมอวีรบุรุษกับกองทัพเรือที่เชื่องช้าสืบค้นทางได้ถึงปี 1893 — **พอเชื้อ** (นักประวัติศาสตร์จากแหล่งเดียว) — James Lind Library — <https://www.jameslindlibrary.org/articles/who-was-james-lind-and-what-exactly-did-he-achieve/> · <https://www.jameslindlibrary.org/articles/james-lind-and-scurvy-1747-to-1795/>
- หลักฐานของเล่นนี้ถูกค้นโดย AI หกตัวจากเจ็ดตัวที่ทำงานพร้อมกันในรอบเดียว ในห้าภาษา ได้ราวสองร้อยแถว · AI อีกตัวที่ไม่ได้อยู่ในเจ็ดตัวนั้นเปิดหน้าต้นทางซ้ำแปดแถว และเปิดเองสิบเจ็ดหน้า · ระหว่างเขียนมีการเปิดเพิ่มอีกสิบสี่หน้า · สิ่งที่เปิดไม่ได้เลยถูกตัด สิ่งที่อ่านผ่านแหล่งที่อ้างถึงบอกไว้ในหลักฐานของแต่ละบท — **ยังบาง** (คำบอกของเล่นนี้เกี่ยวกับตัวเอง ตามป้ายที่เล่นนี้ตั้งไว้เอง) — บันทึกการค้นคว้าของเล่นนี้ 25 และ 26 กันยายน 2026 ไม่ได้เผยแพร่
- หลังเขียนเสร็จ ผู้ตรวจสอบรายเปิดแหล่งซ้ำ และพบความผิดสี่ที่ในร่างของบทที่ 1, 8 และ 11 ทั้งสี่ถูกแก้ก่อนพิมพ์ · รายงานการทดสอบที่เม็กซิโกแสดงทหารสองคนผลัดกันถือเครื่อง · ผู้ผลิต AI เขียนว่าสาเหตุคือหลายการเปลี่ยนแปลงรวมกัน ไม่ใช่สัญญาณตัวเดียว · รายการของศูนย์ต่อต้านข่าวปลอมมีแต่ข่าวปลอม — **ยังบาง** (คำบอกของเล่นนี้เกี่ยวกับตัวเอง) — บันทึกการตรวจสอบของเล่นนี้ 26 กันยายน 2026 ไม่ได้เผยแพร่ · <https://arxiv.org/abs/1301.3971> · <https://openai.com/index/expanding-on-sycophancy/> · <https://www.thansettakij.com/technology/648029>
- สามการทดลองที่สุ่มโรงเรียนในเคนยา รวันดา และยูกันดา 244 โรงเรียน นักเรียน 11,344 คน · อบรมครู 2 ถึง 3 วัน บทเรียนสิบคาบ คาบละ 40 นาที · เกณฑ์ผ่านคือถูก 9 จาก 18 ข้อ · โรงเรียนที่ได้สอนผ่าน 58% (3,397 จาก 5,846 คน) odds ratio 5.5 · ผู้วิจัยสรุปเองว่า 42% ของนักเรียนในโรงเรียนที่ได้สอนยังไม่ผ่าน — **พอเชื้อ** (สามการทดลองให้ผลไปทางเดียวกัน แต่มาจากโครงการเดียวที่ทั้งสร้างและประเมินบทเรียน) — Chesire et al., Journal of Evidence-Based Medicine, 2023 — <https://doi.org/10.1111/jebm.12552>
- ติดตามผลหนึ่งปี 8,298 คน (หายไป 27%) · odds ratio 3.6 เทียบกับ 5.5 ตอนเรียนจบ · ความรู้คงอยู่ 88.3% · ผู้วิจัยแจ้งไว้เองว่าหลายคนเป็นทั้งผู้พัฒนาและผู้ประเมินบทเรียน — **พอเชื้อ** (โครงการเดียวกัน) — Chesire et al., Trials, 30 ตุลาคม 2024 — <https://doi.org/10.1186/s13063-024-08577-w>
- โรงเรียนประถมในยูกันดา 120 โรงเรียน เด็กอายุ 10 ถึง 12 ปี เก้าคาบ สอบผ่าน 69% ต่อ 27% — **พอเชื้อ** (การทดลองเดียว ตัวเลขตามบทคัดย่อ) — Informed Health Choices, The Lancet, 2017 — <https://europepmc.org/article/MED/28539194>
- ความรู้วิทยาศาสตร์ที่มากขึ้นไม่ได้เปลี่ยนความเห็นหรือการตัดสินใจอย่างสม่ำเสมอ และในเรื่องที่เป็นข้อถกเถียง การรู้มากขึ้นเคยไปด้วยกันกับการแบ่งข้างที่แรงขึ้น — **พอเชื้อ** (รายงานบทบทวนฉบับเดียว) — National Academies of Sciences, Engineering, and Medicine, 2016 — <https://nap.nationalacademies.org/read/23595/chapter/2>
- ทำการทดลอง คนกว่า 10,100 คน 52 เรื่อง ไม่มีการแก้ไขไหนทำให้ความเชื่อผิดแรงขึ้น และคนขยับเข้าหาข้อเท็จจริง — **พอเชื้อ** (ทำการทดลองจากทีมเดียว) — Wood & Porter, Political Behavior, 2019 — <https://link.springer.com/article/10.1007/s11109-018-9443-y>
- ในการทดสอบปิดตาสองชั้นที่เม็กซิโก เหยื่อที่นักฟิสิกส์ใช้เทียบเลือกถูก 1 จาก 20 ครั้ง เครื่องได้ 3 จาก 20 — **พอเชื้อ** (การทดสอบหนึ่งครั้ง บันทึกละเอียดทุกชั้น) — Mochán & Ramírez-Solís, arXiv:1301.3971, 2013 — <https://arxiv.org/abs/1301.3971>

- ทีมนักวิจัยทำซ้ำงานจิตวิทยาที่ดีพิมพ์แล้วหนึ่งร้อยชิ้นและพิมพ์ผลทั้งหมด — **พอเชื่อ** (ตัวเลขผลการทำซ้ำมีความเห็นแย้งในปี 2016 ดูปทที่ 4) — Open Science Collaboration, Science, 2015 — <https://europepmc.org/article/MED/26315443>
- การทดลองแบบปิดตาสองชั้นเทียบกับยาหลอกพบว่าฟ้าทะลายโจรไม่ช่วยเพิ่มผลทางคลินิกหรือทางไวรัสในผู้ป่วยโควิด-19 ที่อาการไม่รุนแรง — **พอเชื่อ** (อ่านผ่านบทความที่อ้างอิง) — Siripongboonsitti et al., Phytomedicine, 2023 ตามที่อ้างในรายงานการทดลองแบบเปิดปี 2024 — <https://www.lidsen.com/journals/icm/icm-09-01-013>
- ผู้ผลิต AI รายหนึ่งถอนรุ่นปรับปรุงที่ปล่อยเมื่อ 24 ถึง 25 เมษายน 2025 ตั้งแต่ 28 เมษายน เพราะมัน "overly flattering or agreeable" แล้วเขียนอธิบายสาเหตุ — **พอเชื่อ** (คำแถลงของผู้ผลิตเอง เป็นการยอมรับความผิดของตัวเอง) — OpenAI, 2025 — <https://openai.com/index/sycophancy-in-gpt-4o/> · <https://openai.com/index/expanding-on-sycophancy/>

ภาคผนวก

ทุกข้อในหน้าเหล่านี้ตรวจสอบได้ ถ้าเฉลยข้อไหนผิด เล่มนี้แพ้กนั้น

ให้คะแนนประโยคของคุณ

ใช้ส่วนนี้หลังทำชุดที่สองในบทที่ 13 เสร็จแล้ว วางประโยคที่เขียนตอนต้นเล่มไว้ข้างประโยคในบทที่ 13 แล้วให้คะแนนทั้งสองพร้อมกันด้วยเกณฑ์เดียวกัน

ประโยคละไม่เกินสามคะแนน ได้หนึ่งคะแนนต่อหนึ่งข้อข้างล่างที่ประโยคพูดถึง จะใช้คำไหนก็ได้ ขอให้ความหมายอยู่ในนั้น

1. ข้ออ้างต้องแท้ได้ คือบอกได้ว่าผลแบบไหนจะแสดงว่ามันผิด
2. การเทียบหรือการทดสอบที่ยุติธรรม
3. การเปิดให้คนอื่นตรวจ

ประโยคที่บอกว่าวิทยาศาสตร์คือความรู้เรื่องธรรมชาติได้ศูนย์ ประโยคนั้นไม่ผิด แต่มันบอกว่าวิทยาศาสตร์รู้อะไร ไม่ได้บอกว่ารู้ได้อย่างไร ประโยคที่บอกว่าวิทยาศาสตร์คือการทดลองอย่างยุติธรรมแล้วให้คนอื่นตรวจซ้ำ ได้สองคะแนน จากข้อ 2 กับข้อ 3

ถ้าประโยคในบทที่ 13 ได้คะแนนไม่มากกว่าประโยคตอนต้นเล่ม ครั้งแรกของคำสัญญาในหน้าแรกแท้

เฉลยชุดแรก

แปดข้อตอนต้นเล่ม นับข้อที่ตรงแล้วจได้ไว้ เหตุผลของแต่ละข้อรออยู่หลังเฉลยชุดที่สอง

1. ข · บทที่ 5
2. ง · บทที่ 6

3. ก · บทที่ 7

4. ค · บทที่ 8 และ 4

5. ง · บทที่ 9

6. ก · บทที่ 11

7. ค · บทที่ 11

8. ข · บทที่ 12

หยุดตรงนี้ ถ้ายังไม่ได้ทำชุดที่สองในบทที่ 13 อย่าอ่านเลยบรรทัดนี้ไป

เฉลยชุดที่สอง

แปดข้อในบทที่ 13 นับข้อที่ตรง แล้ววางไว้ข้างคะแนนชุดแรก

1. ค · บทที่ 5

2. ก · บทที่ 6 และ 5

3. ข · บทที่ 7

4. ง · บทที่ 8

5. ข · บทที่ 9

6. ค · บทที่ 11

7. ง · บทที่ 11 และ 9

8. ก · บทที่ 12

เหตุผลของแต่ละข้อ

ข้อที่เลขตรงกันในสองชุดถามเรื่องเดียวกัน ข้อที่ผิดในชุดที่สองบอกว่าทไหนดควรอ่านซ้ำ

ชุดแรก

1. ข น้ำหนักที่ขึ้นลงในเดือนเดียวกันของคนคล้ายกันที่ไม่ได้ดื่ม คือกลุ่มเปรียบเทียบที่บอกว่าถ้าไม่มีมะนาว น้ำหนักจะเป็นอย่างไร คำยืนยันของเพื่อนเอง รีวิว และคำอธิบายของผู้รู้ ไม่มีใครให้เทียบ
2. ง สิบวันที่ได้ทำอะไรจะมีความหมายก็ต่อเมื่อรู้ว่าแอดมินแนะนำไปทั้งหมดกี่ครั้ง ครั้งที่ขาดทุนคือตัวหารที่ไม่มีใครโพสต์ ภาพที่ไม่ได้ตัดต่อหรือสมาชิกที่ได้กำไรก็ยังเป็นแค่ครั้งที่ได้
3. ก คำทำนายที่ไม่บอกว่าอะไรจะเกิดเมื่อไร ไม่มีผลลัพธ์ไหนทำให้มันผิดได้ รอคอบปีหรือถามลูกค้าอีกก็คนก็แก้ข้อนี้ไม่ได้ ต้องให้คำทำนายพูดให้แพ้ได้ก่อน
4. ค งานขึ้นเดียวคือการทดสอบหนึ่งครั้ง งานทบทวนบอกว่าเมื่อรวมหลายขึ้นแล้วผลไปทางไหน ชื่อวารสาร ยอดแชร์ และประวัติของผู้วิจัยไม่ใช่การตรวจ
5. ง กระดาษที่จดไว้ก่อนเปิดไม่รู้ว่าคุณอยากได้อะไร และมีคนในบ้านไว้ทวง ความตั้งใจจะเป็นกลางไม่ทำอะไรให้ใครตรวจ อ่านละเอียดขึ้นก็ไม่ได้กันความอยาก และคนที่กินด้วยกันก็อยากให้มันได้ผลเหมือนคุณ
6. ก ชื่อผู้เขียน ปี และวารสารที่ครบทุกช่องแต่งขึ้นได้ AI ตัวเดิมที่ยืนยัน หรือ AI สองตัวที่อ้างตรงกัน ก็ยังเป็นเครื่องตรวจเครื่อง และวารสารที่มีจริงไม่ได้แปลว่างานชิ้นนั้นมีจริง การเปิดงานเองเป็นทางเดียวที่บอกว่ามันมีอยู่และพูดแบบนั้นจริง
7. ค ห้องใหม่ที่ไม่รู้ว่าไอเดียเป็นของคุณ กับคำขอเหตุผลที่มันจะล้ม คือทำถามใหม่ในห้องใหม่กับท่าขอฝั่งตรงข้าม ที่มีคนเล่าว่าช่วยลดการประจบ แซตใหม่ที่ยังถามว่าไอเดียของคุณดีไหม ห้องเดิมที่รู้แล้วว่าคุณหวังอะไร และรุ่นที่เสียเงิน ยังบอกเครื่องว่า คุณอยากได้อะไร

8. ข สามแก้วที่หน้าตาเหมือนกันและคนเสิร์ฟก็ไม่รู้คำตอบ คือการทดสอบแบบปิดตา สองชั้น ความหวังของใครก็เข้าไปช่วยไม่ได้ และรู้ล่วงหน้าว่าการเดาจะถูกราวหนึ่งในสาม ทำหลายรอบจึงนับเทียบกับการเดาได้ อีกสามทางมีคนรู้คำตอบ หรือไม่มีอะไรให้เทียบ

ชุดที่สอง

1. ค คนเจ็บคอแบบเดียวกันที่ไม่ได้ใช้สเปรย์ คือกลุ่มเปรียบเทียบที่บอกว่าถ้าไม่มีสเปรย์ สามวันนั้นจะเป็นอย่างไร เพราะเจ็บคอดีขึ้นเองได้ การซื้อข้าว สมุนไพร และแพทย์ที่รับรอง ต่อให้ไม่ได้รับเงิน ก็ไม่ได้ตอบเรื่องนี้

2. ก ห้าสิบคนยังไม่มีตัวหาร ต้องรู้ว่ามียศทั้งหมดกี่คน และต้องเทียบกับคนเก่งพอๆกันที่ไม่ได้เรียน เพราะคนที่มาเรียนที่นี้อาจติดได้อยู่แล้ว เวลาเรียน ครูที่สอน และรูปถ่ายไม่ได้ตอบเรื่องนี้

3. ข โฆษณาไม่บอกว่าสดชื่นขึ้นแค่ไหน วดอย่างไร เทียบกับน้ำแบบไหน จึงยังไม่มีผลแบบไหนทำให้มันแพ้ เมื่อระบุให้ครบก็ทดสอบได้ เช่นให้คนดื่มน้ำสองแบบโดยไม่รู้ว่าแก้วไหนผ่านเครื่อง ส่วนกลไก รีวิว และราคาไม่ได้แก้ปัญหา

4. ง หน่วยงานที่ทบทวนหลักฐานคือที่ที่ยาตัวนี้ถูกตรวจจากหลายแหล่ง และแพทย์ที่สั่งยารู้จักคนในบ้านคุณ ถ้าจะหยุดหรือเปลี่ยนยาให้ถามเขาก่อน สาขาที่แพทย์ในคลินิกจบ ยอดผู้ติดตาม และยอดโลกไม่ใช้การตรวจ

5. ข กระดาษที่เขียนก่อนเปิดทำให้ความเชื่อของคุณแพ้ได้ก่อนรู้ผล ความตั้งใจจะเป็นกลางไม่ทิ้งอะไรให้ใครตรวจ อ่านแค่บทคัดย่อไม่ได้กันความชอบ และเพื่อนที่ชอบอาหารเดียวกันก็อยากได้ยีนแบบเดียวกับคุณ

6. ค สรุปของ AI ชี้ไปที่ข้อ 12 เอง การอ่านข้อ 12 ด้วยตัวเองใช้เวลาไม่กี่นาที และเป็นทางเดียวที่บอกว่าสัญญาเขียนแบบนี้จริง ห้องใหม่หรือ AI อีกตัวที่สรุปตรงกันยังผิดพร้อมกันได้ และเพื่อนที่เคยเข้าอาจได้สัญญาคนละฉบับ

7. ง แชตใหม่ที่ได้ฟังทั้งสองฝั่งแบบคนนอก กับคำขอให้แย้งแทนพี่ชาย คือสองท่าเดียวกับข้อ 7 ชุดแรก AI อีกตัวที่ฟังแต่มุมมองของคุณยังได้ยืนยันว่าคุณอยากให้มันเห็นด้วย

กับอะไร เครื่องไม่ได้เป็นกลางเพียงเพราะมันเป็นเครื่อง และความมั่นใจที่มันบอกเอง ไม่ใช่การตรวจ

8. ก ยี่สิบรอบ การเอาให้ถูกเฉลี่ยเกือบเจ็ดรอบ และตารางในบทที่ 12 ตั้งเกณฑ์ของยี่สิบรอบไว้ที่สิบเอ็ด เก้าจึงยังแยกไม่ออกจากโชค แปลคือเกณฑ์ของสิบสองรอบ ไม่ถึงเกณฑ์ก็ไม่ได้พิสูจน์ว่ารสเหมือนกัน และจะตัดรอบไหนทิ้งหลังเห็นผลไม่ได้

การ์ดหนึ่งหน้า

ทั้งเล่มย่อได้เหลือเท่านี้

วิทยาศาสตร์ในสามส่วน

พูดให้แพ้ได้ · ทดสอบอย่างยุติธรรม · เปิดให้คนอื่นตรวจ

คำถามห้าข้อ

1. เทียบกับอะไร
2. บังเอิญได้ไหม
3. ถ้าผิด จะเห็นอะไร
4. ใครตรวจแล้วบ้าง
5. ฉันอยากให้มันจริงไหม

ห้าท่าเมื่อใช้ AI

1. เปิดแหล่งเอง
2. ขอฝั่งตรงข้าม
3. ถามสิ่งที่รู้คำตอบก่อน
4. ถามใหม่ในห้วงใหม่
5. เก็บการตรวจที่ไม่ใช่ AI ไว้หนึ่งอย่าง

โอกาสของแก้วสามใบ

ตารางเดียวกับในบทที่ 12

ชิมกี่รอบ	ต้องถูกอย่างน้อย	โอกาสที่เดามั่วแล้วผ่าน
6	5	1.8%
9	6	4.2%
10	7	2.0%
12	8	1.9%
15	9	3.1%
20	11	3.8%

เกณฑ์ทุกแถวตั้งด้วยกติกาเดียว คือจำนวนที่น้อยที่สุดที่การเดามั่วไปถึงได้ไม่เกินราวหนึ่งครั้งในยี่สิบ ถ้าอยากชิมจำนวนรอบอื่น เลือกแถวก่อนเริ่ม แล้วเขียนเกณฑ์ของแถวนั้นลงแผ่นที่เซ็น แถวที่เลือกหลังเห็นผลไม่นับ

ที่สิบสองรอบ การเดามั่วไปถึงเจ็ดรอบขึ้นไปได้ราว 6.6% หรือราวหนึ่งครั้งในสิบห้าถึงแปดรอบขึ้นไปราว 1.9% หรือราวหนึ่งครั้งในห้าสิบ และถึงเก้ารอบขึ้นไปราว 0.4% หรือราวหนึ่งครั้งในสองร้อยห้าสิบ

ป้ายความแน่นของหลักฐาน

ป้ายเหล่านี้เป็นการตัดสินของเล่มนี้เอง และคุณเถียงมันได้

แน่น ถูกทดสอบมากกว่าหนึ่งครั้ง แหล่งที่เป็นอิสระต่อกันให้ผลไปทางเดียวกัน

พอเชื่อ หลักฐานดีจากหนึ่งหรือสองแหล่ง หรือยังมีคนแย้งอยู่

ยังบาง งานชิ้นเดียว ประสบการณ์ของคนคนเดียว หรือคำบอกของผู้ผลิตเกี่ยวกับของของตัวเอง

ตรวจไม่ได้ มีคนเล่า แต่เล่มนี้เปิดแหล่งค้นหาไม่ได้

ไม่มีประโยชน์ไหนในเนื้อเรื่องที่พูดแรงกว่าป้ายของมัน ถ้าคุณเจอ นั่นคือที่ที่เล่นนี้ผิด

คำที่ใช้ความหมายเดียวตลอดเล่ม

ภาษาอังกฤษในวงเล็บคือคำที่ใช้ค้นต่อได้

วิทยาศาสตร์ (science) วิธีที่มนุษย์คิดขึ้นเพื่อจับความผิดของตัวเอง ย่อได้เป็นสามส่วนในการด้างบน

ข้ออ้าง (claim) ประโยคที่บอกว่าอะไรจริง หรืออะไรได้ผล คือสิ่งที่คำถามหำข้อเอาไว้ตัดสิน

แพ้ได้ (falsifiable) ข้ออ้างที่แพ้ได้บอกไว้ก่อนว่าผลแบบไหนจะทำให้มันผิด

การทดสอบที่ยุติธรรม (fair test) การทดสอบที่เทียบกับกลุ่มที่เหมือนกันตั้งแต่ต้น และจัดไว้ไม่ให้ความหวังของใครเข้าไปเปลี่ยนผล

กลุ่มเปรียบเทียบ (control group) คนที่ไม่ได้รับสิ่งที่ถูกทดสอบในช่วงเวลาเดียวกัน เอาไว้ดูว่าถ้าไม่มีมัน ผลจะเป็นอย่างไร

การทดสอบแบบปิดตา (blind test) การทดสอบที่คนถูกทดสอบไม่รู้คำตอบ ว่าของอยู่ตรงไหนหรือตัวเองได้อะไร จนกว่าจะนับผลเสร็จ

ปิดตาสองชั้น (double-blind) ทั้งคนที่ถูกทดสอบและคนที่วัดผลหรือยกของมาให้ ไม่รู้คำตอบจนกว่าจะนับผลเสร็จ

ความบังเอิญ (chance) ผลที่เกิดขึ้นแต่ไม่มีอะไรทำงานเลย เช่นการชี้แก้วถูกราวหนึ่งในสามเมื่อเดามั่ว

ยาหลอก (placebo) ของที่หน้าตา รสชาติ และวิธีใช้เหมือนของจริงทุกอย่าง แต่ไม่มีตัวยา สองกลุ่มจึงหวังได้เท่ากัน

การทดลองซ้ำ (replication) การทดลองที่ทำซ้ำตามวิธีเดิม เพื่อดูว่าผลเดิมยังออกมาไหม

งานทบทวน (systematic review) งานที่รวมงานวิจัยหลายชิ้นในเรื่องเดียวกันเข้าด้วยกัน แทนการเชื่องานชิ้นเดียว

ฉันทามติ (consensus) สิ่งที่ผู้รู้ในเรื่องนั้นเห็นตรงกันหลังตรวจกันมาแล้ว ซึ่งน่าเชื่อ กว่าคนคนเดียว

ผลประโยชน์ทับซ้อน (conflict of interest) เมื่อคนที่พูดหรือตัดสินใจอ้าง ได้ ประโยชน์จากคำตอบแบบใดแบบหนึ่ง

แหล่งอ้างอิง (citation) งานต้นทางที่คำตอบหรือบทความชี้ไปหา ซึ่งคุณเปิดดูเองได้ ว่ามีอยู่จริงและพูดแบบนั้นจริง

ประจบ (sycophancy) การที่ AI เห็นด้วยกับผู้ใช้และชมเขาเกินกว่าที่หลักฐานให้

AI (artificial intelligence) โปรแกรมที่เรียนรู้รูปแบบจากข้อมูลจำนวนมาก ในเล่มนี้มี ทั้งตัวที่คนพิมพ์คุยด้วย และตัวที่ทายรูปร่างโปรตีนหรือพยากรณ์อากาศ

อ่านแนวขวาง (lateral reading) ออกจากหน้าที่กำลังอ่าน ไปเปิดแท็บใหม่ดูว่าที่อื่น พูดถึงแหล่งนั้นหรือข้ออ้างนั้นอย่างไร ก่อนจะอ่านเนื้อหาให้ละเอียด

การลงทะเบียนล่วงหน้า (preregistration) การประกาศคำถาม วิธีทดลอง และตัววัดหลักไว้ในที่ที่ใครก็เปิดดูได้ ก่อนเริ่มเก็บข้อมูล

วิทยาศาสตร์เทียม (pseudoscience) ข้ออ้างที่ใช้ภาษาของวิทยาศาสตร์ แต่หลบสาม ส่วนบนการ์ด เช่นไม่บอกว่าผลแบบไหนจะทำให้มันผิด หรือไม่ให้คนนอกทดสอบ

ตัววัดหลัก (primary outcome) สิ่งที่การทดลองประกาศไว้ก่อนเริ่มว่าจะใช้ตัดสินผล จะได้ไม่เลือกตัววัดที่ออกมาสวยหลังเห็นผล

กติกถาหนึ่งในยี่สิบ (5% significance level) เกณฑ์ที่ยอมให้การเดามั่วผ่านได้ไม่เกิน ราวหนึ่งครั้งในยี่สิบ ผ่านเกณฑ์แล้วยังเป็นโชคได้ แค่ไม่บ่อย