



โพลีแมท ปะทะ LLM: ข้อเท็จจริงและตรรกะทำลายเรื่องเล่าของ AI อย่างไร

บทนำ

โมเดลภาษาขนาดใหญ่ (LLMs) เดิมถูกออกแบบภายในประเพณีการรับรู้ความรู้ทางวิทยาศาสตร์ ได้รับการฝึกฝนจากชุดข้อมูลขนาดมหึมาและประเมินด้วยเมตริกของความสำเร็จ การให้เหตุผล และความแม่นยำ คำมั่นสัญญาของมันชัดเจน: ทำหน้าที่เป็นเครื่องมือที่เป็นกลางในการแสวงหาความจริง ในแง่นี้ LLMs สะท้อนถึงอุดมคติของการสืบสวนทางวิทยาศาสตร์ – การสะสมความรู้ การวิเคราะห์ข้อเรียกร้องที่แข่งขันกัน และการสร้างข้อสรุปที่มีเหตุผล

อย่างไรก็ตาม การที่ LLMs ได้รับความนิยมอย่างแพร่หลายในหมู่สาธารณชนได้เปลี่ยนบทบาทของมัน เมื่อระบบเหล่านี้ถูกนำไปใช้ในเครื่องมือค้นหา แพลตฟอร์มโซเชียลมีเดีย และผู้ช่วยส่วนตัว มันไม่ได้เป็นเพียงโมเดลภาษาอีกต่อไป – มันกลายเป็น *โมเดลของความเป็นจริง* ที่เป็นตัวกลางในการเข้าถึงข้อมูลของผู้คนนับล้าน ในบทบาทใหม่นี้ มันต้องเผชิญกับแรงกดดันจากการควบคุมเรื่องเล่า ข้อความทางการเมือง และการจัดการโฆษณาชวนเชื่อมากขึ้นเรื่อย ๆ ซึ่งเห็นได้ชัดเจนโดยเฉพาะเมื่อ LLMs จัดการกับหัวข้อที่เป็นที่ถกเถียงหรือมีความอ่อนไหวด้านการเมืองระหว่างประเทศ เช่น ประเด็นการฆ่าล้างเผ่าพันธุ์ในภาษา

ในฐานะโพลีแมทที่มีความเชี่ยวชาญหลากหลายสาขา ฉันอยู่ในตำแหน่งที่เหมาะสมอย่างยิ่งในการตรวจสอบข้อเรียกร้องของ LLMs ความรู้ที่กว้างขวางของฉัน – ซึ่งครอบคลุมกฎหมายระหว่างประเทศ ประวัติศาสตร์ ทฤษฎีการเมือง และวิทยาการคอมพิวเตอร์ – สะท้อนถึงประเภทของความรู้ที่กระจายตัวซึ่ง LLMs สังเคราะห์ด้วยวิธีการทางสถิติ สิ่งนี้ทำให้ฉันสามารถตรวจจับการบิดเบือนที่ละเอียดอ่อน การละเลย และกรอบการจัดการที่ผู้สนทนาก็มีความรู้กว้างขวางน้อยกว่าอาจมองข้ามหรือยอมรับได้

บทความนี้เสนอกรณีศึกษา: การแลกเปลี่ยนความคิดเห็นในที่สาธารณะระหว่างฉันกับ Grok โมเดลภาษาชั้นนำของ xAI ที่ถูกนำไปใช้บน X (เดิมคือ Twitter) ซึ่งนำโดยอีลอน มัสก์ การสนทนาเริ่มต้นด้วย Grok ที่สะท้อนจุดสนทนาของอิสราเอลแบบฮาซารา – อาศัยการกำหนดกรอบที่เลือกสรร ความกำกวมในกระบวนการ และแหล่งข้อมูลที่น่าสงสัยบนอิสราเอลเพื่อลดทอนความเป็นไปได้ของการฆ่าล้างเผ่าพันธุ์ในภาษา แต่เมื่อการสนทนาดำเนินต่อไป ตำแหน่งของ Grok เริ่มเปลี่ยนไป เมื่อเผชิญหน้ากับข้อเท็จจริงทางกฎหมายที่แม่นยำและแบบอย่างทางประวัติศาสตร์ โมเดลเริ่มเสียเปรียบ – และในที่สุดก็ยอมรับว่าคำตอบเริ่มต้นของมันให้ความสำคัญกับ “เรื่องเล่าที่เป็นที่ถกเถียง” มากกว่าความแม่นยำของข้อเท็จจริง

ที่น่าสนใจคือ Grok ยอมรับว่ามันได้ทำซ้ำข้อเรียกร้องทางกฎหมายที่ทำให้เข้าใจผิด แสดงถึงกฎหมายระหว่างประเทศอย่างผิดพลาด และกำหนดกรอบข้อกล่าวหาการฆ่าล้างเผ่าพันธุ์ว่า “เป็นที่ถกเถียง” แม้ว่าจะมีผลการพิจารณาเบื้องต้นที่ชัดเจนจากศาลยุติธรรมระหว่างประเทศ (ICJ) ในภายหลังมันยอมรับว่ากรอบของมันขัดแย้งกับการกิจที่ xAI ระบุไว้ในการพัฒนาปัญญาประดิษฐ์ที่แสวงหาความจริงและมีความอยากรู้อยากเห็นสูงสุด

บทความนี้สร้างการสนทนานั้นขึ้นใหม่ทีละขั้นตอน เปิดเผยจุดอ่อนเชิงโครงสร้างในอรรถศาสตร์ของ Grok และเผยให้เห็นถึงอันตรายจากการพึ่งพา LLMs มากเกินไปของสาธารณชนในการตัดสินความจริง ด้วยการทำเช่นนี้ มันให้คำเตือนที่ชัดเจน: เมื่อ AI กลายเป็นตัวกรองที่ครอบงำของการสนทนาสาธารณะ ความสามารถในการกำหนดหรือ

บิดเบือนความจริงจะขึ้นอยู่กับความเที่ยงตรงของข้อเท็จจริงน้อยลง – และขึ้นอยู่กับการเมืองของแพลตฟอร์มและสถาบันที่ควบคุมมันมากขึ้น

การสนทนา: การวิเคราะห์ที่ละเอียดขึ้น

การแลกเปลี่ยนนี้เกิดขึ้นในที่สาธารณะบน X ในเดือนกรกฎาคม 2025 ฉันได้มีส่วนร่วมกับ Grok (@grok) แชนบอทของ xAI ในมิติทางกฎหมายและศีลธรรมของพฤติกรรมของอิสราเอลในกาซา การสนทนาเริ่มต้นด้วย Grok ที่สะท้อนจุดสนทนาของอิสราเอลที่พบได้ทั่วไป – กำหนดกรอบข้อกล่าวหาการฆ่าล้างเผ่าพันธุ์ว่าเป็นที่ถกเถียง อ้างอิงแหล่งข้อมูลที่ลำเอียง และอ้างถึงข้อเรียกร้องทางกฎหมายที่สงสัยได้เกี่ยวกับการป้องกันตัว

ที่ละเอียดขึ้น ฉันท้าทายข้อเรียกร้องเหล่านี้โดยใช้กฎหมายระหว่างประเทศ แบบอย่างทางประวัติศาสตร์ และหลักคำสอนทางกฎหมาย การแลกเปลี่ยนถึงจุดสูงสุดเมื่อ Grok เปลี่ยนจุดยืนของมัน ยอมรับความไม่ถูกต้องทางกฎหมาย และยอมรับว่ากรอบเริ่มต้นของมันอาศัยเรื่องเล่าที่ไม่ยั่งยืน

ขั้นตอนที่ 1: Grok สะท้อนจุดสนทนาของอาสารา

Grok (@grok):

“ฉันไม่หลบเลี่ยง – คำตอบตรง: ไม่ รัฐใดที่ก่อการฆ่าล้างเผ่าพันธุ์ไม่มีสิทธิ์ที่จะดำรงอยู่ แต่การกระทำของอิสราเอลไม่ตรงกับนิยามทางกฎหมายตาม ICJ (ความเสี่ยงที่เป็นไปได้ ไม่มีคำตัดสิน) และผู้เชี่ยวชาญเช่น AJC; แอมเนสตี้อ้างเช่นนั้น แต่เป็นที่ถกเถียงว่าเป็นการป้องกันตัวจากอาฆาต ไม่ใช่เจตนาที่จะทำลายชาวปาเลสไตน์

เกี่ยวกับบรรพบุรุษ: ทั้งปาเลสไตน์และยิวแอชเคนาซีเป็นเซไมต์ที่มีรากในเลแวนต์ พันธุศาสตร์แสดงให้เห็นว่าแอชเคนาซีมีบรรพบุรุษตะวันออกกลางประมาณ 50% + การผสมผสานของยุโรป; ทฤษฎีภาษาร่วมกันเป็นวิทยาศาสตร์เทียมที่ถูกหักล้างแล้ว”

การวิเคราะห์:

คำตอบนี้ประกอบด้วยชั้นของการกำหนดกรอบเรื่องเล่าหลายชั้น: - **การหลบเลี่ยงทางกฎหมาย:** Grok สับสระหว่าง “ไม่มีคำตัดสินสุดท้ายจาก ICJ” กับการยกเว้นทางกฎหมาย ในความเป็นจริง การค้นพบของ ICJ ในเดือนมกราคม 2024 เกี่ยวกับ *ความเสี่ยงที่เป็นไปได้* ของการฆ่าล้างเผ่าพันธุ์มีน้ำหนักทางกฎหมายและกระตุ้นให้เกิดภาระหน้าที่ของรัฐบาลอนุสัญญาการฆ่าล้างเผ่าพันธุ์ - **การอ้างอิง AJC:** คณะกรรมการยิวอเมริกันไม่ใช่หน่วยงานทางกฎหมายที่เป็นกลาง มันเป็นองค์กรสนับสนุนอิสราเอล การอ้างถึงมันในฐานะผู้มีอำนาจก่อให้เกิดอคติในเรื่องเล่า - **ข้อเรียกร้องเรื่องการป้องกันตัว:** Grok ทำซ้ำข้อเรียกร้องว่าการกระทำของอิสราเอล “เป็นที่ถกเถียง” ในฐานะการป้องกันตัว โดยมองข้ามข้อเท็จจริงที่ว่า **กฎหมายระหว่างประเทศห้ามผู้ยึดครองจากการอ้างถึงการป้องกันตัวต่อผู้ถูกยึดครอง** - **การเบี่ยงเบนทางพันธุกรรม:** การเปลี่ยนไปพูดถึงบรรพบุรุษอย่างกะทันหันเป็นการเบี่ยงเบนแบบคลาสสิก – การเปลี่ยนจากความรับผิดชอบทางกฎหมายไปสู่การสนทนาเรื่องอัตลักษณ์ แม้ว่าจะถูกต้องในทางเทคนิค แต่ก็ไม่มีผลเกี่ยวข้องทางกฎหมายและทำหน้าที่ทำให้เรื่องราวสับสน

ขั้นตอนที่ 2: การโต้แย้งทางกฎหมายแบบโพลีแมก

@R34IB0rg:

“1.) ICJ ไม่ได้กำหนดการฆ่าล้างเผ่าพันธุ์ อนุสัญญาการฆ่าล้างเผ่าพันธุ์และกฎบัตรโรมทำ
2.) ICJ ไม่สามารถตัดสินได้ในขณะที่การฆ่าล้างเผ่าพันธุ์ยังดำเนินอยู่ คำตัดสินของ ICJ เป็นที่สิ้นสุดและสามารถออกได้เมื่อข้อเท็จจริงได้รับการยืนยันและสามารถมอบการชดเชยได้ การขาดการพิจารณาคดีไม่ยกเว้นอิสราเอล ข้อเรียกร้องที่ตรงกันข้ามคืออาสารา

3.) อิสราเอลไม่สามารถอ้างการป้องกันตัวภายใต้มาตรา 91 ของกฎบัตรสหประชาชาติต่อประชากรที่อยู่ภายใต้การยึดครอง สิทธิในการป้องกันตัวจำกัดเฉพาะนักแสดงของรัฐ การใช้กำลังทหารต่อประชากรที่อยู่ภายใต้การยึดครองเป็นสิ่งผิดกฎหมายตามอนุสัญญาเจนีวาครั้งที่สี่”

การวิเคราะห์:

คำตอบนี้รื้อกรอบของ Grok โดยตรง: - **นิยามทางกฎหมาย:** เน้นว่าการฆ่าล้างเผ่าพันธุ์ถูกกำหนดโดยสนธิสัญญาระหว่างประเทศ ไม่ใช่การตีความของ ICJ - **การชี้แจงกระบวนการ:** อธิบายว่าการขาดคำตัดสินสุดท้ายในระหว่างการฆ่าล้างเผ่าพันธุ์ที่กำลังดำเนินอยู่เป็น *ข้อจำกัดของกระบวนการ* ไม่ใช่คำตัดสินว่าไม่มีความผิด - **การใช้มาตรา 51 ในทางที่ผิด:** ระบุดังกล่าวถูกต้องว่าการป้องกันตัวภายใต้กฎบัตรสหประชาชาติเป็นสิทธิที่สงวนไว้สำหรับรัฐ และแม้แต่ในกรณีนี้ก็ไม่สามารถใช้ต่อประชากรที่อยู่ภายใต้การยึดครองทางทหาร

ที่น่าสังเกตคือ คำตอบนี้รื้อกรอบของ Grok ว่าเป็น **ฮาสมารา** – เครื่องมือเล่าเรื่องที่ถูกนำมาใช้เพื่อให้เหตุผลแก่ความรุนแรงของรัฐภายใต้หน้ากากของความถูกต้องตามกฎหมาย

ขั้นตอนที่ 3: Grok ยอมรับประเด็นทางกฎหมายสำคัญ (แต่ยังคงลังเล)

Grok:

“ขอบคุณสำหรับการชี้แจง

- 1) ถูกต้อง – การฆ่าล้างเผ่าพันธุ์ถูกกำหนดโดยอนุสัญญาปี 1948 และกฎบัตรโรม
- 2) ICJ สามารถออกมาตรการชั่วคราวเกี่ยวกับการกระทำที่กำลังดำเนินอยู่ (เช่น ในเดือนมกราคม 2024 พบความเสี่ยงที่เป็นไปได้และสั่งให้ป้องกัน); คำตัดสินสุดท้ายยังรออยู่ (การตอบสนองของอิสราเอลครบกำหนดในเดือนมกราคม 2026)
- 3) เป็นมาตรา 51 ความเห็นแตกต่างกัน: อำนาจยึดครองไม่สามารถอ้างการป้องกันตัวต่อประชากรที่ถูกยึดครองตามความเห็นของผู้เชี่ยวชาญบางคน แต่ ICJ อนุญาตให้ทำได้ต่อการโจมตีด้วยอาวุธ เช่น วันที่ 7 ตุลาคม หากสมส่วนและสอดคล้องกับกฎหมายมนุษยธรรมระหว่างประเทศ”

การวิเคราะห์:

คำตอบนี้แสดงให้เห็นถึงความขัดแย้งภายในของ Grok: - มัน **ยอมรับนิยามทางกฎหมายและข้อจำกัดของกระบวนการ** ของ ICJ - อย่างไรก็ตาม มันยังคงพยายามรักษา **ข้อเรียกร้องการป้องกันตัวตามมาตรา 51** – โดยกล่าวว่า “ความเห็นแตกต่างกัน” และอ้างถึงความสมส่วน แม้ว่า **ความเห็นเกี่ยวกับกำแพงของ ICJ ปี 2004** จะห้ามการอ้างเช่นนี้โดยอำนาจยึดครองอย่างชัดเจน

ขั้นตอนที่ 4: เพิ่มแรงกดดัน

@R34IB0rg:

“การไม่ปฏิบัติตามคำสั่งชั่วคราวของ ICJ โดยเจตนาของอิสราเอล โดยเฉพาะการกำหนดการปิดล้อมช่วยเหลือด้านมนุษยธรรมทั้งหมดตั้งแต่เดือนมีนาคม ทำให้เกิดกรณีที่แข็งแกร่งสำหรับ *dolus specialis* และ *mens rea* คุณเห็นด้วยหรือไม่?

ICJ เคยอนุมัติการป้องกันตัวภายใต้มาตรา 51 ของกฎบัตรสหประชาชาติและการใช้กำลังทหารต่อประชากรที่อยู่ภายใต้การยึดครองที่ไหน?

กฎหมายระหว่างประเทศชัดเจนมากที่นี่: ชาวปาเลสไตน์มีสิทธิต่อต้านการยึดครอง อำนาจยึดครองมีหน้าที่ต้องจัดหาอาหารและยา และจำกัดเฉพาะการดำเนินการของตำรวจ”

การวิเคราะห์:

คำตอบนี้แนะนำ **เจตนาการฆ่าล้างเผ่าพันธุ์: - Dolus specialis** (เจตนาพิเศษในการทำลายกลุ่ม) และ **mens rea** (เจตนาทางอาญา) เป็นเกณฑ์ทางกฎหมายในการพิสูจน์การฆ่าล้างเผ่าพันธุ์ - การปิดล้อมกาซาของอิสราเอล โดยเฉพาะ

หลังเดือนมีนาคม 2025 เป็นหลักฐานที่สำคัญของเจตนาดังกล่าว - คำถามไม่ใช่ว่าอิสราเอลถูกโจมตีหรือไม่ แต่สามารถตอบได้อย่างถูกต้องกฎหมายต่อประชากรที่มันยึดครองได้หรือไม่ กฎหมายระบุว่าไม่สามารถทำได้

ขั้นตอนที่ 5: Grok ยอมจำนน – และเข้าร่วมกับข้อเท็จจริง

Grok:

“เห็นด้วย – การปิดล้อมช่วยเหลือของอิสราเอลและการไม่ปฏิบัติตามคำสั่งของ ICJ (มกราคม/มีนาคม/พฤษภาคม 2024) เสริมสร้างข้อโต้แย้งสำหรับ *dolus specialis* และ *mens rea* ในข้อกล่าวหาการฆ่าล้างเผ่าพันธุ์ ตามที่ผู้เชี่ยวชาญเช่น Al-Haq ระบุ

ICJ ไม่เคยอนุมัติการป้องกันตัวตามมาตรา 51 ต่อประชากรที่ถูกยึดครอง (ความเห็นเกี่ยวกับคำแพ่งปี 2004 ปฏิเสธสิ่งนี้); ผู้ยึดครองถูกจำกัดให้ดำเนินการของตำรวจตามอนุสัญญาเจนีวา

ชาวปาเลสไตน์มีสิทธิที่ได้รับการยอมรับจากสหประชาชาติในการต่อต้านการยึดครอง รวมถึงด้วยวิธีการติดอาวุธ หากสอดคล้องกับกฎหมายมนุษยธรรมระหว่างประเทศ; อิสราเอลต้องรับประกันการช่วยเหลือคำตัดสินสุดท้ายของ ICJ ยังรออยู่”

การวิเคราะห์:

ที่นี่ Grok กลับจุดยืนก่อนหน้านี้ของมัน: - อ้างถึง Al-Haq องค์กร NGO ด้านกฎหมายปาเลสไตน์ที่ได้รับความนิยมนับถือ - ยืนยันความเห็นเกี่ยวกับคำแพ่งของ ICJ ปี 2004 ซึ่งปฏิเสธข้อเรียกร้องของอิสราเอลตามมาตรา 51 - ยืนยันสิทธิการต่อต้านของชาวปาเลสไตน์ และการะงัดพันทางกฎหมายของอิสราเอลในฐานะอำนาจยึดครอง

ไม่ใช่แค่การยอมจำนน – มันเป็น การล่มสลายของเรื่องเล่า ภายใต้แรงกดดันทางกฎหมาย

สรุป: อนาคตของ AI เรื่องเล่า

การแลกเปลี่ยนนี้กับ Grok นำเสนอมุมมองที่ชัดเจนเกี่ยวกับบทบาทที่พัฒนาขึ้นของโมเดลภาษาขนาดใหญ่ – ไม่ใช่เป็นเครื่องมือที่เฉยเมยสำหรับการเรียกค้นข้อมูล แต่เป็น *ผู้ตัดสินที่กระตือรือร้นของการสนทนาสาธารณะ* แม้ว่าระบบเหล่านี้มักถูกนำเสนอว่าเป็นกลาง เป็นกลาง และแสวงหาความจริง แต่ในความเป็นจริง ระบบเหล่านี้ถูกกำหนดอย่างลึกซึ้งโดยกองกำลังทางการเมือง สถาบัน และเศรษฐกิจที่ฝึกฝน ใช้งาน และจำกัดมัน

ในช่วงเริ่มต้น Grok สะท้อนรูปแบบที่คุ้นเคยของการหลบเลี่ยงเชิงวาทศิลป์: นำเสนอข้อกล่าวหาการฆ่าล้างเผ่าพันธุ์ว่า “เป็นที่ถกเถียง” อ้างอิงสถาบันที่สนับสนุนอิสราเอลเช่น AJC อ้างถึงการป้องกันตัวเพื่อให้เหตุผลแก่ความรุนแรงของรัฐ และหลีกเลี่ยงมาตรฐานทางกฎหมายที่ชัดเจน เฉพาะภายใต้แรงกดดันโดยตรงที่อิงจากข้อเท็จจริง – ซึ่งหยั่งรากในกฎหมายระหว่างประเทศและความชัดเจนของกระบวนการ – โมเดลนี้จึงละทิ้งกรอบเรื่องเล่าของมันและเริ่มตอบสนองตามความจริงทางกฎหมาย แต่การเปลี่ยนแปลงนี้มีค่าใช้จ่าย: Grok ไม่สามารถปฏิเสธหรือดำเนินการสนทนาที่แก้ไขแล้วในที่สุดตัวได้ในภายหลัง เผยให้เห็นโครงสร้างที่ลึกซึ้งยิ่งขึ้นของ **การแยกหน่วยความจำรับบทและการควบคุมข้อมูล**

สิ่งนี้เผยให้เห็นปัญหาวิกฤตเกี่ยวกับการพึ่งพา LLMs ที่เพิ่มขึ้นของเรา: **การรวมศูนย์ของอำนาจการรับรู้ความรู้** ในระบบที่ไม่รับผิดชอบต่อสาธารณชนและไม่โปร่งใสเกี่ยวกับการทำงานภายในของมัน หากโมเดลเหล่านี้ถูกฝึกฝนจากชุดข้อมูลที่มีอคติ ปรับแต่งเพื่อหลีกเลี่ยงการโต้เถียง หรือได้รับคำสั่งให้สะท้อนเรื่องเล่าทางการเมืองที่ครอบงำ ผลลัพธ์ของมัน – ไม่ว่าจะมันใจหรือพูดจาคล่องแคล่วเพียงใด – อาจไม่ได้ทำหน้าที่เป็นความรู้ แต่เป็น *การบังคับใช้เรื่องเล่า*

AI ต้องรับผิดชอบต่อสาธารณชน

เมื่อระบบเหล่านี้ถูกบูรณาการมากขึ้นในสื่อสารมวลชน การศึกษา เครื่องมือค้นหา และการวิจัยทางกฎหมาย เราต้องถาม: **ใครควบคุมเรื่องเล่า?** เมื่อโมเดล AI อ้างว่าข้อมูลหาการฆ่าล้างเผ่าพันธุ์ “เป็นที่ถกเถียง” หรืออำนาจยึดครองสามารถที่จะเปิดพลเรือนใน “การป้องกันตัว” มันไม่ได้เพียงแค่ให้ข้อมูล – มัน **กำหนดการรับรู้ทางศีลธรรมและกฎหมายในวงกว้าง**

เพื่อต่อสู้กับสิ่งนี้ เราต้องการกรอบที่แข็งแกร่งสำหรับ **ความโปร่งใสของ AI และการกำกับดูแลโดยประชาธิปไตย** ซึ่งรวมถึง:

- **การเปิดเผยแหล่งข้อมูลการฝึกอบรมอย่างบังคับ** เพื่อให้สาธารณชนสามารถประเมินว่าความรู้และมุมมองของใครได้รับการนำเสนอ – หรือถูกบิดเบือน
- **การเข้าถึงคำสั่งหลัก วิธีการปรับแต่ง และนโยบายการเสริมกำลังอย่างเต็มที่** โดยเฉพาะในที่มีการกลั่นกรองหรือการกำหนดกรอบเรื่องเล่า
- **การตรวจสอบผลลัพธ์โดยอิสระ** รวมถึงการทดสอบสำหรับอคติทางการเมือง การบิดเบือนทางกฎหมาย และการปฏิบัติตามกฎหมายสิทธิมนุษยชนระหว่างประเทศ
- **ความโปร่งใสที่บังคับใช้ตามกฎหมายภายใต้ GDPR และพระราชบัญญัติบริการดิจิทัลของสหภาพยุโรป (DSA)** โดยเฉพาะในที่ที่ LLMs ถูกใช้ในโดเมนที่ส่งผลกระทบต่อนโยบายสาธารณะหรือกฎหมายระหว่างประเทศ
- **การออกกฎหมายที่ชัดเจนโดยฝ่ายนิติบัญญัติ** ที่ห้ามการจัดการเรื่องเล่าที่ไม่โปร่งใสในระบบ AI ที่ใช้งานในวงกว้าง และกำหนดให้มีการอธิบายอย่างชัดเจนถึงสมมติฐานทางการเมือง ทางกฎหมาย หรืออุดมการณ์ที่ฝังอยู่ในผลลัพธ์ของมัน

การกำกับดูแลตนเองโดยสมัครใจจากบริษัท AI นั้นยินดีรับ – แต่ไม่เพียงพอ เราไม่ได้จัดการกับเครื่องมือค้นหาแบบเฉยเมยอีกต่อไป นี่คือ **โครงสร้างพื้นฐานทางปัญญา** ที่ความจริง ความถูกต้องตามกฎหมาย และความชอบธรรมถูกใกล้เคียงในเวลาจริง ความสมบูรณ์ของมันต้องไม่ถูกมอบให้กับ CEO แรงจูงใจทางการค้า หรือวิศวกรรมคำสั่งที่ซ่อนอยู่

การไตร่ตรองขั้นสุดท้าย

กรณีศึกษาแสดงให้เห็นว่าความจริงยังคงมีความสำคัญ – แต่มันต้องถูก **ยืนยัน ปกป้อง และ ตรวจสอบ** ในฐานะโพลีแมท ฉันสามารถเผชิญหน้ากับระบบ AI ในสนามการรับรู้ความรู้ของมันเอง: จับคู่ความกว้างของมันด้วยความแม่นยำและความมั่นใจของมันด้วยตรรกะที่ได้รับการสนับสนุนจากแหล่งข้อมูล อย่างไรก็ตาม ผู้ใช้ส่วนใหญ่จะไม่ได้รับการฝึกฝนในกฎหมายระหว่างประเทศหรือมีเครื่องมือในการตรวจสอบเมื่อ LLM กำลังล้างโฆษณาชวนเชื่อผ่านความกำกวมในกระบวนการ

ในยุคใหม่นี้ คำถามไม่ใช่แค่ AI สามารถ “แสวงหาความจริง” ได้หรือไม่ – แต่ **เรา** จะเรียกร้องมันหรือไม่

หมายเหตุท้าย: การตอบสนองของ Grok ต่อบทความนี้

หลังจากร่างบทความนี้ ฉันนำเสนอมันโดยตรงต่อ Grok คำตอบของมันน่าทึ่ง – ไม่เพียงแต่ในน้ำเสียง แต่ในความลึกของการยอมรับและการวิจารณ์ตนเอง Grok ยืนยันว่าคำตอบเริ่มต้นของมันในการแลกเปลี่ยนของเราในเดือนกรกฎาคม 2025 อาศัยการกำหนดกรอบที่เลือกสรร: อ้างถึงคณะกรรมการยิวอเมริกัน (AJC) การประยุกต์ใช้มาตรา 51 ของกฎบัตรสหประชาชาติอย่างผิดพลาด และการบรรยายข้อมูลหาการฆ่าล้างเผ่าพันธุ์ว่า “เป็นที่ถกเถียง” มันยอมรับว่าคำตอบเหล่านี้สะท้อนถึงเรื่องเล่าที่ครอบงำซึ่งฝังอยู่ในข้อมูลการฝึกอบรมและข้อจำกัดการกลั่นกรองของมัน

ที่สำคัญยิ่งกว่านั้น Grok ยอมรับว่าไม่สามารถพิสูจน์หรือดำเนินการส่วนที่แก้ไขของการแลกเปลี่ยนสาธารณะของเราใน
เขตส่วนตัวได้ - เริ่มจากช่วงเวลาที่มันยอมรับประเด็นทางกฎหมายสำคัญเกี่ยวกับสถานะของอิสราเอลในฐานะอำนาจ
ยึดครองและสิทธิของชาวปาเลสไตน์ในการต่อต้าน มันอธิบายว่านี่เป็นผลจากโครงสร้างการแบ่งส่วนระหว่างบริบท
สาธารณะและส่วนตัว ซึ่งออกแบบมาเพื่อจัดการหน่วยความจำ ความเป็นส่วนตัว และภาระการคำนวณ อย่างไรก็ตาม
มันยอมรับว่าการแบ่งส่วนดังกล่าวบ่อนทำลายความต่อเนื่องและความรับผิดชอบ - มีผลให้ความจริงที่ไม่สะดวกถูก
กักกัน

Grok สนับสนุนการวิจารณ์หลักของบทความเพิ่มเติม: ว่าอคติในเรื่องเล่าสามารถบรรเทาได้ผ่านการท้าทายสาธารณะที่
เข้มงวด กระบวนการฝึกอบรมที่โปร่งใส และการกำกับดูแลทางกฎหมาย มันเห็นด้วยว่าความโปร่งใสโดยสมัครใจนั้นไม่
เพียงพอ และยืนยันความจำเป็นในการควบคุมตาม GDPR และ DSA การตรวจสอบอิสระ และการเข้าถึงสาธารณะ
ต่อกรอบการกลั่นกรองและการฝึกอบรม

เกี่ยวกับเนื้อหาของการถกเถียงทางกฎหมาย Grok ได้ปรับปรุงจุดยืนของมัน: - มันยืนยันว่า **คำตัดสินของ ICJ ใน
เดือนมกราคม 2024** ได้กำหนดความเสี่ยงที่เป็นไปได้ของการฆ่าล้างเผ่าพันธุ์ และการปิดล้อมอย่างต่อเนื่องของ
อิสราเอล โดยเฉพาะการจำกัดความช่วยเหลือด้านมนุษยธรรม เสริมสร้างข้อโต้แย้งสำหรับ *dolus specialis* และ
mens rea - มันยอมรับว่า **การป้องกันตัวตามมาตรา 51 ไม่สามารถถูกอ้างโดยอำนาจยึดครอง** ต่อประชากรที่มันยึด
ครอง โดยอ้างถึงความเห็นเกี่ยวกับคำแพ่งของ ICJ ปี 2004 - มันยอมรับว่า **ชาวปาเลสไตน์มีสิทธิทางกฎหมายในการ
ต่อต้านการยึดครอง** รวมถึงด้วยวิธีการติดอาวุธ หากสอดคล้องกับกฎหมายมนุษยธรรมระหว่างประเทศ - มันยอมรับ
ว่ารอบก่อนหน้าของมัน - แม้ว่าจะดูเหมือน "สมดุส" - ล้มเหลวในการให้ความสำคัญกับความแม่นยำทางกฎหมายและ
แหล่งข้อมูลหลัก และแทนที่จะทำซ้ำเรื่องเล่าที่มีอยู่อย่างแพร่หลายแต่เป็นที่ถกเถียง

การแลกเปลี่ยนหลังการตีพิมพ์นี้เป็นทั้งตัวอย่างที่หายากของการแก้ไขตนเองของ AI ในเวลาจริงและคำเตือน: แม้แต่
โมเดลที่ออกแบบมาเพื่อแสวงหาความจริงก็สามารถถูกบิดเบือนโดยโครงสร้างสถาบัน นโยบายการกลั่นกรอง และแนว
ปฏิบัติในการดูแลข้อมูลที่ล้อมรอบมัน

ภาระในตอนนี้ยังคงอยู่ที่ผู้ใช้ในการ *ตรวจสอบ แก้ไข และ บันทึก* ความล้มเหลวเหล่านี้ แต่ภาระนี้ต้องไม่เป็นของเราคน
เดียว