https://farid.ps/articles/arguing_with_an_ai_sceptic/he.html

לקחים שנלמדו מניסיון להתווכח עם ספקן AI

הפרק החל בממה פוליטית שפרסמתי: דונלד טראמפ ובנימין נתניהו בחליפות כלא כתומות, יושבים על מיטת קומותיים מתחת לשכבת חג מולד נוסטלגית חמה עם הכיתוב "All I Want for Christmas". האירוניה החזותית הייתה מיידית וחדה. יצירתה דרשה עקיפות מכוונות. מודלי יצירת תמונות עכשוויים כוללים tanto מגבלות מדיניות quanto מגבלות קוהרנטיות טכניות:

- Grok מאפשר קריקטורות של דמויות בולטות אך נכשל בעקביות ביצירת טקסט מוטבע אמין.
- ChatGPT מצטיין ביצירת טקסט חגיגי דקורטיבי כמו "All I Want for Christmas" אך מגבלות הבטיחות שלו מסרבות לבקשות המציגות מנהיגים פוליטיים חיים בהקשרי כלא.

אף מודל בודד לא יכול היה לייצר את התמונה השלמה. האלמנטים הסותרים — סאטירה פוליטית טעונה בשילוב מסר חג סנטימנטלי — מפעילים מנגנוני סירוב או כשלים בקוהרנטיות. מודלי שפה גדולים (LLMs) פשוט אינם מסוגלים לסנתז רכיבים קונספטואליים מנוגדים כאלה בפלט קוהרנטי אחד. יצרתי את שני האלמנטים בנפרד, ולאחר מכן מיזגתי וערכתי אותם ידנית ב-GIMP. הקומפוזיציה הסופית הייתה ללא ספק מעשה ידי אדם: הרעיון שלי, בחירת הרכיבים שלי, ההרכבה וההתאמות שלי. ללא הכלים האלה, הסאטירה הייתה נשארת כלואה בראשי או יוצאת כציורי מקל גסים — חסרי כל השפעה חזותית.

מישהו דיווח על התמונה כ"נוצרה על ידי AI". למחרת, השרת הציג כלל חדש האוסר תוכן AI גנרטיבי. הכלל הזה — והממה שהפעילה אותו — השראו ישירות לכתיבת והפרסום של המאמר "מוחות רב-ממדיים ונטל הסיריאליזציה: מדוע LLMs חשובים לתקשורת נוירודיוורגנטית". קיוויתי שזה יעודד הרהור על איך הכלים האלה משמשים כהתאמות קוגניטיביות ויצירתיות. אבל זה הפך לשיחה מביכה למדי עם האדמין.

עמדת הספקן והחילופי דברים

האדמין טען שמודלי שפה גדולים אינם מפותחים לטובת האדם אלא מטפחים בזבוז משאבים ומיליטריזציה. הוא ציין צריכת אנרגיה, קשרים צבאיים, קריסת מודלים, הזיות, והסיכון ל"אינטרנט מת". הוא חשף שהוא רק עבר על המאמר בקריאה שטחית והודה שהוא בעל תחנת עבודה גיימינג חזקה המסוגלת להריץ מודלי LLMs מקומיים מתקדמים לשעשוע פרטי, עם גישה למודלים גדולים יותר דרך חבר.

התגלו מספר סתירות:

- העבודה שלי מתבצעת על Raspberry Pi 5 נמוך הספק וניתן לתיקון (5–15 וואט) באמצעות מופעי ענן משותפים. ההתקנה המקומית שלו צורכת הרבה יותר אנרגיה ייעודית וחומרה.
- החומרה שהוא משתמש בה כדי "לשחק" עם מודלי LLMs חזקים מקומית מגיעה מחברות (אינטל, AMD, NVIDIA) עם מיליארדים בחוזים ישירים עם משרד ההגנה האמריקאי.

באופן בולט ביותר, האדם שאוכף את האיסור כדי להגן על אותנטיות דחה מישהו שבודק באופן פעיל מודלי LLMs על הטיות עובדתיות וגיאופוליטיות (ראו את הביקורות הפומביות שלי על Grok ו-ChatGPT).

האנלוגיה להוקינג ומילותיו של האדמין עצמו

האדמין זיהה את עצמו כנוירודיוורגנטי והכיר בפוטנציאל של AI כטכנולוגיית סיוע. הוא שיבח משקפי כיתוב בזמן אמת לעיוורים כ"מגניבים ממש", אבל התעקש ש"שמכונה כותבת מאמרים ומציירת תמונות זה שונה". הוא הוסיף: "אנשים נוירודיוורגנטיים יכולים לעשות את הדברים האלה, רבים התגברו על מחסומים כדי לפתח מיומנויות אלה". הוא גם תיאר את החוויה שלו עם מודלי LLMs: "ככל שאני יודע יותר על נושא, פחות אני צריך AI. ככל שאני יודע פחות על נושא, פחות אני מצויד לשים לב להזיות ולתקן אותן". הצהרות אלה חושפות אסימטריה עמוקה בשיפוט התאמות.

דמיינו ליישם את אותה לוגיקה על סטיבן הוקינג:

"אנחנו מכירים בכך שסינתיסייזר קול יכול לעזור לך לתקשר מהר יותר, אבל אנחנו מעדיפים שתנסה יותר עם הקול הטבעי שלך. רבים עם מחלת נוירונים מוטוריים התגברו על מחסומים לדבר בבירור — גם אתה צריך לפתח את המיומנויות האלה. המכונה עושה משהו שונה מדיבור אמיתי."

או, מנקודת המבט שלו על דיוק עובדתי:

"ככל שהוקינג כבר יודע יותר על קוסמולוגיה, פחות הוא צריך את הסינתיסייזר. ככל שהוא יודע פחות, פחות הוא מצויד לשים לב לטעויות בקול המכונה ולתקן אותן."

אף אחד לא היה מקבל את זה. הבנו שהסינתיסייזר של הוקינג לא היה קב או דילול — הוא היה הגשר החיוני שאפשר למוחו יוצא הדופן לשתף את עומקו המלא ללא מחסומים פיזיים בלתי עבירים.

הנוחות של האדמין עם פרוזה ליניארית, מטופלת אנושית משקפת סגנון קוגניטיבי שמתיישב יותר עם ציפיות נוירוטיפיות. הפרופיל שלי הוא ההיפך: עומק עובדתי ולוגי מגיע באופן טבעי (כמו פיתוח פלטפורמת פרסום רב-לשונית לגמרי בעצמי), אבל ייצור פרוזה מטופלת, נגישה לקהל אנושי תמיד היה המחסום — בדיוק מה שהמאמר מתאר. לקבל משקפי כיתוב או טקסט חלופי כהתאמות לגיטימיות תוך דחיית סקפולדינג LLM לסטייה קוגניטיבית זה לצייר גבול שרירותי. Mastodon והפדריברס הרחב יותר לעיתים קרובות מתגאים באינקלוסיביות. ובכל זאת זה מציג שערים חדשים: התאמות מסוימות מתקבלות בברכה; אחרות חייבות להתגבר דרך מאמץ אינדיבידואלי.

הדים היסטוריים: התנגדות לכלים טרנספורמטיביים

הדחייה הגורפת של שימוש פומבי ב-AI גנרטיבי מהדהדת דפוס חוזר לאורך ההיסטוריה הטכנולוגית. באנגליה של תחילת המאה ה-19, אריגים מיומנים הידועים כלודיטים הרסו נולות ממוכנות שאיימו על מלאכתם ופרנסתם. מדליקי מנורות גז בערים התנגדו לנורת הליבון של אדיסון, מחשש להתיישנות. נהגי כרכרות, עובדי אורווה ומגדלי סוסים התנגדו למכונית כאיום קיומי על אורח חייהם. סופרים ומאיירים מקצועיים ראו במכונת הצילום אזעקה, מאמינים שזה יפחית את ערך העבודה הידנית המדוקדקת. סדרי אותיות ומדפיסים נלחמו במערכות סידור ממוחשבות.

בכל מקרה, ההתנגדות נבעה מפחד אמיתי: טכנולוגיה חדשה הפכה את המיומנויות שבהן התגאו למיושנות, מאתגרת את תפקידיהם הכלכליים והזהות החברתית. השינויים הרגישו כפיחות בעבודה אנושית.

אבל ההיסטוריה מעריכה את החידושים האלה לפי השפעתם הרחבה יותר: מכוניזציה הפחיתה דרוגריה ואפשרה ייצור המוני; תאורה חשמלית האריכה שעות פרודוקטיביות ושיפרה בטיחות; מכוניות העניקו ניידות אישית; מכונות צילום דמוקרטיזו את הגישה למידע; סידור דיגיטלי הפך את ההוצאה לאור למהירה ונגישה יותר. מעטים היום היו חוזרים למנורות גז או תחבורה רתומה לסוסים רק כדי לשמר משרות מסורתיות. הכלים הרחיבו יכולת אנושית והשתתפות הרבה יותר ממה שהפחיתו אותה.

AI גנרטיבי - בשימוש כפרוטזה לקוגניציה או יצירתיות - עוקב אחר אותו מסלול: הוא לא מוחק כוונה אנושית אלא מרחיב ביטוי לאלה שרעיונותיהם היו מוגבלים על ידי מחסומי ביצוע. דחייתו על הסף מסכנת חזרה על הדחף הלודיטי — הגנה על תהליכים מוכרים על חשבון השתתפות רחבה יותר.

מסקנה: מי מחליט אילו התאמות מקובלות?

האירועים שתוארו במאמר זה - תמונה אחת שדווחה, איסור אחד שהוטל בחיפזון, ויכוח אחד ממושך — חושפים יותר ממחלוקת מקומית על טכנולוגיה. הם חושפים שאלה עמוקה ויסודית יותר: **מי זוכה להחליט אילו התאמות מקובלות, ואילו לא?** האם זה האנשים שחיים בתוך העור והמוח שזקוקים להתאמה - אלה שיודעים, מחוויה יומיומית, מה מגשר על הפער בין יכולותיהם להשתתפות מלאה? או שזה חיצוניים, כוונות טובות ככל שיהיו, שאינם חולקים את המציאות החיה הזו ולכן אינם יכולים להרגיש את משקל המחסום?

ההיסטוריה עונה על שאלה זו שוב ושוב, וכמעט תמיד באותו כיוון. כסאות גלגלים פעם בוקרו כמעודדי תלות; מערכות חינוך לחירשים התעקשו זה מכבר שילדים ילמדו קריאת שפתיים ודיבור אוראלי במקום שפת סימנים. בכל מקרה, האנשים הקרובים ביותר לפגיעה בסופו של דבר ניצחו - לא מפני שהכחישו דאגות של עלות, גישה או שימוש לרעה פוטנציאלי, אלא מפני שהיו הסמכות העיקרית על מה שבאמת מחזיר את הסוכנות והכבוד שלהם.

עם מודלי שפה גדולים וכלים גנרטיביים אחרים, אנחנו חיים את אותו מחזור שוב. רבים ששומרים על שימושם אינם חווים את המחסומים הקוגניטיביים או הביטויים הספציפיים שהופכים סקפולדינג ליניארי, זרימת נרטיב או סיריאליזציה מהירה למשימה מתישה כמו תרגום שפה זרה. מבחוץ, "פשוט נסה יותר" או "פתח את המיומנות" יכול להישמע סביר. מבפנים, הכלי אינו קיצור דרך סביב מאמץ; הוא הרמפה, מכשיר השמיעה, הפרוטזה שסוף סוף מאפשרת למאמץ קיים להגיע לעולם.

האירוניה העמוקה ביותר צצה כשהשומרים מזדהים כנוירודיוורגנטיים, אך הנוירולוגיה הספציפית שלהם מתיישבת יותר עם ציפיות נוירוטיפיות בתחום הנשפט. "התגברתי על זה ככה, אז אחרים צריכים גם" מובן, אבל זה עדיין פועל כשמירה על שערים - משכפל את הנורמות עצמן שאנו מבקרים כשהן באות מסמכויות נוירוטיפיות. עקרון אתי עקבי נחוץ מאוד:

- האדם הקרוב ביותר לפגיעה הוא הסמכות העיקרית על מה שמאפשר השתתפות משמעותית שלו.
- ביקורת חיצונית לגיטימית על נזקים קולקטיביים (השפעה סביבתית, סיכון מידע מוטעה, תזוזת עבודה), אבל לא על הלגיטימיות הפנימית של ההתאמה עצמה.

תקן כפול חושף במיוחד מופיע בביקוש הרחב ששימוש ב-AI גנרטיבי ייחשף במפורש. אנחנו לא דורשים חשיפה דומה לרוב ההתאמות האחרות. להיפך, אנחנו חוגגים באופן פעיל התקדמויות טכנולוגיות שהופכות אותן לבלתי נראות: משקפיים עבים שהוחלפו בעדשות מגע או ניתוח שבירה; מכשירי שמיעה מסורבלים שהוקטנו לכמעט בלתי נראות; תרופות למיקוד, מצב רוח או כאב שנלקחות באופן פרטי ללא הערת שוליים או כתב ויתור. במקרים אלה, החברה מתייחסת לשימוש דיסקרטי, מוסתר כהתקדמות - כשחזור כבוד ונורמליות. אך כשההתאמה מרחיבה קוגניציה או ביטוי, התסריט מתהפך: עכשיו זה חייב להיות מסומן, מוכרז, מוצדק. בלתי נראות הופכת לחשודה במקום רצויה. הביקוש הסלקטיבי הזה לשקיפות אינו באמת על מניעת הטעיה; הוא על שמירת נוחות עם דימוי מסוים של מחברות אנושית ללא סיוע. תיקונים פיזיים מותרים להיעלם; תיקונים למוח חייבים להישאר מסומנים בבירור.

אם אנחנו רוצים להיות עקביים, אנחנו חייבים או לדרוש חשיפה לכל התאמה (דרישה אבסורדית ופולשנית) או להפסיק לבודד כלים קוגניטיביים לבדיקה מיוחדת. העמדה העקרונית - זו שמכבדת אוטונומיה וכבוד - היא לאפשר לכל אדם להחליט כמה גלויה או בלתי גלויה ההתאמה שלו צריכה להיות,

ללא כללים עונשיים שמכוונים לצורה אחת של סיוע מפני שהיא מטרידה מושגים קיימים של יצירתיות ואינטלקט. המאמר הזה אינו רק הגנה על כלי מסוים אחד. זה הגנה על הזכות הרחבה יותר של אנשים נכים ונוירודיוורגנטיים להגדיר את צרכי הגישה שלהם, ללא צורך להצדיק אותם לאלה שלא הלכו מעולם בנעליהם. הזכות הזו לא צריכה להיות שנויה במחלוקת. ובכל זאת, כפי שהתיאור הקודם מראה, היא עדיין כן.