

הסכנה היא לא שהבינה המלאכותית תקבל החלטות. זו שאנחנו נותנים לזה לקרות.

האם הבינה המלאכותית תיקח מהאדם את תפקידו? האם אנחנו בדרך לעתיד שבו הבינה המלאכותית תקבל כל החלטה והאנושות הופכת למין ללא סוכנות, רק הולכת אחרי האדון הדיגיטלי העל-אינטליגנטי שלנו?

זו פחד אמיתי, ואולי זה מוצדק: מערכת עם מטרות משלה שווה דאגה לגביה. אבל זו לא הסכנה שאנחנו הכי חשופים לה כרגע. הרבה לפני שהבינה המלאכותית תשתלט על משהו, אנחנו הולכים למסור אותו, כי זה קל יותר.

שני סוגי קונפליקט

תיאוריית הניהול מבחינה בין שני סוגי קונפליקט. קונפליקט מתכנס נפתר דרך דיון, כמו מחלוקות על מונחים או עובדות סותרות. הבהיר את המונח, או עשה קצת מחקר, ואנשים מגיעים להבנה משותפת.

קונפליקט מפוצל לא יכול להיפתר כך: שני הצדדים צודקים, והם סותרים זה את זה. לדוגמה, האם עלינו לשאוף להגדלת נתח השוק, מה שידרוש מאיתנו להגדיל את ההצעה ולהוריד מחירים, או לשאוף להגדלת רווחי הרווח, מה שיחייב אותנו להפחית את ההצעה ולהעלות את המחירים? גם הגדלת נתח השוק וגם הגדלת הרווח הן מטרות לגיטימיות, אבל אי אפשר לרדוף אחרי שתי המטרות בו זמנית. זו פשרה מבנית, לא פער ידע. דוגמה נוספת לקונפליקט מפוצל היא אינטרסים מנוגדים. win-win הוא ציפייה אוטופית, וברוב המקרים, לפחות בטווח הקצר, מישהו מנצח ומישהו מפסיד. זה פשוט כך, ושום כמות של דיבור לא תשנה זאת.

איפה הבינה המלאכותית עוזרת, ואיפה שהיא לא יכולה

בינה מלאכותית תהיה מצוינת בקונפליקטים מתכנסים: למצוא את המילה הנכונה, לחשוף מידע חסר. זו בעיית ידע.

קונפליקטים מפוצלים אינם כאלה. שום כמות נתונים לא תפתור מטרות סותרות או אינטרסים סותרים. המשימה אינה לאופטימיזציה של צד אחד; זה למצוא את האיזון הנכון, והאיזון של היום לא יהיה של מחר. זה כמו לנהוג, לתקן את הגלגל כל הזמן. התיקון הזה הוא החלטה שיפוטית, ללא תשובה נכונה, רק סיכון, ואי אפשר להאציל אותו למי שהכי טוב במקסום משתנה אחד.

דוגמה עדכנית

פאוצ'י הובא לאחרונה שוב בפני הקונגרס בעקבות החלטות מתקופת המגפה. שווה להיות מדויק: כמנהל NIAID, הוא החזיק בסמכות אמיתית על מימון מחקר וסדרי עדיפויות של המוסדות, אך מעולם לא היה לו סמכות לקבוע חובות בריאות ציבור ארציות. בצוות המשימה לקורונה בבית הלבן, בראשות טראמפ ובראשות משרד הבריאות הציבורית, שימש כיועץ מדעי מוביל, תרגם אפידמיולוגיה לאפשרויות מדיניות, לא לקבוע מדיניות. לגבי ההחלטה שהייתה חשובה, האם המדינה תסגר, הוא היה יועץ, לא מקבל החלטות, בדיוק התפקיד שהבינה המלאכותית תמלא עבורנו.

כאקדמאי ומנהל ממשלתי, פאוצ'י הסתכל על הקורונה דרך עדשה אחת: למזער את הסיכון שבתי החולים יפלו. מטרה לגיטימית, אם כי אופטימיזציה שלה פירושה לא לשקול את העלות החברתית של הסגרים, המכה הכלכלית או המחיר לחירות האישית. זו לא הייתה הכישלון של פאוצ'י; לשקול את הפשרות האלה מעולם לא היה תפקידו.

זו הייתה של טראמפ. טראמפ החזיק בסמכות; פאוצ'י החזיק רק בהשפעה. אבל טראמפ העביר לפאוצ'י את הדוכן וטיפל בהמלצותיו כעובדה מוכרת, לא כי פאוצ'י עבר את הגבול, אלא כי הוזמן, ומעטים מסרבים להנהגה אמיתית כשהיא מוצעת. זה מחמא, זה מזין את האגו. הוויתור קל משני הצדדים: האדם עם הסמכות נמנע מהחלטה קשה, והאדם שמקבל את הדוכן כמעט אף פעם לא מחזיר אותו.

הטעות העמוקה יותר לא הייתה להקשיב לפאוצ'י. זו הייתה רק הקשבה לפאוצ'י. איזון אמיתי מגיע מהצבת דעות סותרות באותו חדר: האפידמיולוג שטוען בעד סגר, מנהיגי עסקים על העלות הכלכלית, פסיכולוגים על המחיר של בידוד, תומכי זכויות אזרח על חופש אישי. השיחה הזו לא נוחה, אבל כך באמת מוצאים איזון. כל זה לא ביקורת על פאוצ'י, שעשה את העבודה שניתנה לו. הכישלון היה של טראמפ, שמעולם לא עשה את העבודה הקשה יותר של להתמודד עם הסכסוך ולקבל את השיפוט הנדרש לגבי מה צריכה להיות הגישה הנכונה. במקום זאת, הוא העדיף את הקול האחד, ועכשיו מאשים את הקול הזה כי לא אהב איך שהדברים התפתחו.

הסכנה האמיתית

התבנית הזו קדמה לבינה מלאכותית. זו הסיבה שחברות משלמות למקינזי. מנכ"ל שמתמודד עם קריאה פוליטית מסוכנת, למשל פיטורים, לא באמת קונה ניתוח; הוא קונה כיסוי. אם היועצים "ימלצו" על מה שכבר חשד, זה יהפוך ל"מקינזי אמר לנו", לא ל"החלטתי". זה לעיתים קרובות המוצר האמיתי שנמכר.

הבינה המלאכותית עומדת להפוך לגרסה הזולה והמהירה ביותר של אותו שירות. הסיכון כאן הוא לא שפאוצ'י, מקינזי או בינה מלאכותית תפסו סמכות. זה אדם עם סמכות שמתרחק בשקט ונותן ליועץ לנהוג, כי ההחלטה קשה, היועץ נשמע בטוח, ואם זה משתבש, מישהו אחר אשם.

הבינה המלאכותית תמשיך להשתפר בקונפליקטים מתכנסים, אלה שנגרמים ממידע חסר. היא לעולם לא תהיה נכונה לקונפליקטים מנוגדים: אלה בעיות איזון, ואיזון דורש אדם שמוכן לשבת עם דעות סותרות ולקבל החלטה ללא תשובה נכונה. זו לא עבודה שהבינה המלאכותית תיקח מאיתנו. זו עבודה שנעביר אותה, החלטה נוחה אחת בכל פעם.