



# מסע לעמק המוזרות

בועז לביא

הטכנולוגיה המודרנית מולידה יצירי כלאיים שהאדם לא פגש מעולם. כאשר החפץ הדומם האנושי במקצת חוצה גבול בלתי מוגדר, הוא יוצר אפקט בלהות מוזר. המושג "עמק המוזרות" מתאר את החרדה המתגברת ככל שגדל הדימיון לאנושי. לאחרונה, מודלי שפה כמו ChatGPT מציבים לנו גרסה חדשה ומערערת של הזר-מוכר – חיקוי חלקי של מוח אנושי, והמסע אל עמק המוזרות הופך לחוויה יומיומית

יולי 2023

ובותיו המדברות של תומס אדיסון הבעיתו את הלקוחות. הן הושבו בזו אחר זו אל חנויות הצעצועים והובילו לאחד הכישלונות הגדולים בקריירה עמוסת הפטנטים של הממציא האמריקני. "בובות נפלאות שמר אדיסון בונה לילדות קטנות וטובות", כך הן תוארו ב־1888 בכותרת של מה שייחשב היום לתוכן שיווקי מפוקפק, ובישר רעות כבר אז – פני שעווה, עיני זכוכית, וקול רפאים שהלך ודעך ככל שנחלש מנוע הקפיץ שהפעיל את המכונה:

היה הייתה ילדה קטנה,

תלתל קטן בשערה,

שהסתלסל לו במרום הפדחת.

כשהייתה טובה,

הייתה מאוד-מאוד טובה,

אך כשהייתה רעה – הייתה מזעזעת.<sup>[1]</sup>

שמונה-עשרה נשים צעירות גויסו באמצעות סוכנות כירויות כדי להקליט את שירי הערש, האפלים יותר ופחות, והם נוגנו באמצעות פונוגרף תוך-בובתי - גאדג'יט קדום שהפסיד בסופו של דבר לגרמופון. בינואר 1889 תיאר **הניו יורק טיימס** את תהליך האימון של הבובות במושגים של חינוך אנושי: לבובות יש כביכול כוח חיקוי, והן לומדות ממורותיהן כיצד לשיר ולדבר. אבל פיהן של הבובות שיצאו מפס הייצור לא זז, והדבר תרם למראן הלא טבעי. "קל לדמיין כמה בלתי מושכות - אפילו מפחידות ממש - היו הבובות האלה בעיניה של ילדה ב-1890, במיוחד אם הופעלו במהירות לא נכונה או לא אחידה", אמר עליהן רנה רונדו, אספן בן ימינו ומומחה למכשירי פונוגרף.<sup>[2]</sup> שיווקן של הבובות המדברות הופסק שבועיים לאחר שהחל. בתי נולדה 124 שנים אחר כך.



על המדפים ועל הרצפה, בקומות השונות של בית הבובות וברחבי המיטה, פזורים בני משפחת ארנבי השוקולד. שמם מטעה; הם אינם עשויים סוכר ועיסת קקאו אלא פלסטיק מכוסה בסיבי אריג, וגופם עטוי פריטי לבוש שנלקחו מתוך פנטזיה על חווה אמריקנית מאמצע המאה העשרים, או נכון יותר - מפנטזיה של מעצב יפני על חווה כזאת. שכן משפחת הארנבים, נצר למותג צעצועים מצליח בשם *Sylvanian Families*, היא מוצר יפני למהדרין. צאצאיה רחוקים מרחק מזרח ממערב מבובותיו של אדיסון; אלה הם חיקויים בגודל אצבע של בעל חיים מואנש שאינו מוציא מפיו מילה. בכך הן אינן יוצאות דופן בקולקציית הבובות של בתי. אני זוכר במעורפל רק ינשוף גדול אחד, יחיד מתוך עשרות רבות של בובות, שידע להמהם בשילוב כלשהו של אנגלית ושפת הינשופים. הוא לא שרד, כי אם הושלך לפח על כנפיו הכחולות או נמסר באחת הסלקציות. כל השאר שותקים. שותקים עצלני הבד ושותקים הפילים. שותקים דובי הצמר שסרגה לה אימא ושותקים הפינגווינים. שתיקתם מצילה אותם מגורלו של אותו ינשוף אומלל, הופכת אותם לאהובים ומאפשרת לבתי לדבר בשמם, במקומם.

אך לפני כמה שנים, כשסירי, העוזרת הקולית של אפל, הגיחה לעולמה של בתי אחרי שצפתה בי נעזר בה להשלמת משימות פשוטות, הפכה זו לצעצוע האהוב מכול, בשורה אחת עם ארנבי השוקולד היפניים. לסירי אין דמות הגוף ואיננה גוף. הילדה פונה אל מסך האיפד או הטלפון, אבל סירי איננה המסך. זה מובן מאליו. היא מי שאיתה מדברים:

- סירי, מה המאכל הכי יקר בעולם?

- אני לא בטוחה שהבנתי.

- גם אני לא.

- טוב לדעת שאנחנו באותו ראש.

- נכון, טוב לדעת. אבל מה המאכל הכי יקר בעולם?

- אני לא בטוחה שהבנתי.

- די, סירי!

- בשיא הרצינות.

- סליחה.

- הכול טוב.

הדיאלוג ביניהן - שאלות ללא מענה, חזרות בלתי נלאות על משפטי מפתח ("אני לא בטוחה שהבנתי") ולולאות אינסופיות - מספק לילדה בת השבע, ואחר כך בת השמונה וכמעט כבר תשע, הנאה ללא גבול. זוהי הנאה לא חושנית, שאין לה ממד פיזי, כי בעולם הבובות של בתי הופרדו השפה והגוף; לקח רב-שנים שנלמד מן הפיאסקו של אדיסון, לקח שרישומו התחזק עם פריחתו של מוטיב הבובות הרצחניות בז'אנר סרטי האימה.

גם אם בתי אינה יודעת דבר וחצי דבר על היסטוריית הזוועות הבובתית, השוק על סוכניו הרבים דאג איכשהו לצמצם את הסיכוי שתיאחז פלצות לנוכח בובה שקנו לה הוריה. מאז 1929, אז יצא הסרט The Great Gabbo המגולל את סיפורו של פיתום שבובתו (הנושאת את השם הרב-משמעי אוטו) משתלטת על נפשו ועל חייו, ועד הסרט M3GAN משלהי 2022, שבו בובה בעלת מנגנון שיחה מבוסס בינה מלאכותית קמה על משפחתה המאמצת לחסלה, ניצלו עשרות במאי קולנוע את האימה המתעוררת בנו למראה בובה עם כישורי דיבור, כשם שאחרים השתמשו לצורך זה בדמות הליצן - ספק אדם, ספק שד מאופר. בחזרה של בתי לא תימצא אף בובה מדברת; זהו כבר אינסטינקט הורי. אבל איך ייתכן בעצם שקול נטול גוף, כמו זה של סירי, אינו מצמרר אותה - אותנו - עד לשד עצמותינו גם כן? האין זו אַימת כל האַימות, שְפָה ללא פֶה? העובדה שסירי אהובה בכל זאת על בתי דורשת הסבר. איך נדע לבטח שגם היא, כמו הבובה M3GAN בשעתה, לא תהפוך יום אחד את עורה ותגרום לשערותינו לסמור? ומה בעניין תוכנות שיחה חדשות יותר, צ'יטבוטים דוגמת ChatGPT ומתחריו, שמציפים את הרשת? האם גם בהם טמון זרע בעתה?

כדי להעריך מה הסיכוי שנצטרך ממילים שייפלטו מתוך מחשב או מסך מגע יש להבין תחילה מה הפחיד כל כך את קהל היעד במוצר הכושל של אדיסון, ומה מפחיד עד היום בבובות מדברות. לכאורה מדובר בצעצוע תמים, לעיתים אף כזה שיש לו יומרה חינוכית. אין זו מפלצת שמטרתה המפורשת היא להחריד, אפילו אם באופן משעשע; צעצועים כאלה דווקא אהובים מאוד בדרך כלל. אמת, גם בבובה בדמות ילד או ילדה, בובה סתם שאינה מדברת, יש פוטנציאל הפחדה לא קטן - במיוחד בתנאי תאורה מסוימים. אך הדיבור, מאולץ ככל שיהיה, מעניק לבובה כוח חיות מיוחד; החפץ הדומם חוצה איזה גבול והופך אנושי במקצת. לא אנושי לגמרי; רק קצת. הדבר הזה שהוא "מעט חיים" - הדיבור המלאכותי והמוגבל מאוד אצל בובותיו של אדיסון, למשל - מפחיד הרבה יותר מאשר השלם, דיבור רגיל של ילד, שלמעשה אינו מפחיד אותנו כלל. החלקי מאיים. הוא מטריד ועוכר שלוה. והטכנולוגיה המודרנית - נסמן את בואה בעשן הקטרים של המאה התשע-עשרה, עם שיפורו הדרמטי של המנוע המכני, כמו זה שהפעיל את הפונוגרף, ואחריו גילוי החשמל והמצאת מקור האור המלאכותי - טכנולוגיה זו הולידה לפתע, כמו מתוך מצחה, יצורי כלאיים שהאדם לא פגש מעולם. יצורים חדשים, אמיתיים רק במקצת, שפורטים על חולשות עתיקות.



ארנסט יִנְטֶש, פסיכיאטר גרמני, היה הראשון שהתייחס באופן מתודי לאפקט המערער שיש לטכנולוגיה על בני אדם, במאמר שכותרתו "על הפסיכולוגיה של מעורר האימה" שראה אור ב-1906. המילה הגרמנית שהוא משתמש בה לציון הרעיון של "מעורר אימה", מילה שאין לה מקבילה בעברית, היא Das Unheimliche, שתורגמה במשך השנים ל"המאווים", ואחר כך ל"האלביתיי" - מילת היפוך משונה

ל"ביתיים", ותרגום מילולי מן הגרמנית. בניגוד לתרגומים לעברית, המאולצים משהו, המילה הגרמנית חיה ונפוצה, והיא משמשת לתיאור כל מה שמעורר באדם תחושת סלידה ומוזרות - אם זוועה צרופה במופעיה הקיצוניים, בסיוטים או בסרטי אימה, אם דחייה פיזית פשוטה כמו זו שתציף אותנו גם למראה שזיף רקוב, אובייקט כמעט לא מזוהה, באור המקרר. ינטש מיקם את מקור התחושה הזאת במצב של אי-ודאות אינטלקטואלית, בייחוד כאשר אין דעתנו משגת להבין אם מה שעומד מולנו הוא חי או דומם. לפי ינטש, חידושים טכנולוגיים הם מקור לא אכזב לרגש זה, משום שאין לנו די ניסיון לקבוע לאיזו קטגוריה הם שייכים. טשטוש הגבולות המודרני בין האנושי ובין מה שרק נחזה כאנושי משטה בנו ומטריד. "בובות ממוכנות בגודל אדם", כתב ינטש עשור וחצי לאחר קריסת חזון הבובות המדברות של אדיסון, "המבצעות פעולות מסובכות, תוקעות בחצוצרה, רוקדות וכדומה, מעוררות בנקל תחושת אי-נוחות"<sup>[3]</sup>.

ב-1919 לקח זיגמונד פרויד את אותו מושג בדיוק, *Das Unheimliche*, הרחיק ללכת בדרך שסלל ינטש והותיר את הבחנותיו של קודמו הרחק מאחור; הן נראו פתאום טריוויאליות בהשוואה לממצאיו היצירתיים של אבי הפסיכואנליזה. ינטש ייחס את מקורה של תחושת הזרות והאימה שצצה בנו מול הבובה המדברת לאיזה בלבול תפיסה זמני, אך פרויד טען כי מה שמטריד אותנו כאן הוא עקרוני יותר: זהו מגע עם תשוקה מסוכנת שחוזר מן המתים. הבובה הממוכנת מפגישה אותנו בלי אזהרה עם מה שבהיותנו בני אדם בוגרים התרגלנו לחשוב עליו כעל בלתי אפשרי: כפיל מציאותי שלנו. פרויד קושר במהלך פתלתל את החרדה מפני כפילנו האוטומטי לחרדה מפני עקירת עיניים, שהיא עצמה תחליף לחרדה מפני סירוס; שניהם עונשים מיתולוגיים המוטלים על מי שמפר את הצו העומד בבסיס הציוויליזציה - האיסור על העריות. מה עניין כפיל לעריות? מבחינת פרויד, הכפיל הוא מי שהילד מנתב אליו, במין תהליך של מיקור חוץ, את תשוקתו האסורה. הוא מה שמאפשר לילד לעבד תשוקה זו ולהתבגר. כך מתפצל הילד (אצל פרויד, כרגיל, הזכר הוא מוקד התיאוריה) ל"ילד טוב" ול"ילד רע"; הכפיל הדמיוני הוא כמובן "הילד הרע". "אופיו של האלביתי [כלומר אותו דבר שמעורר אימה] אינו יכול אלא לנבוע מהעובדה שהכפיל הוא יציר השייך לימיה הקדמונים של הנפש", כותב פרויד, "ימים שכבר התגברנו עליהם, שבהם, אגב, הייתה לו משמעות ידיוותית יותר. הכפיל הפך לתמונת אימים, כשם שהאלים הופכים לדמונים לאחר נפילת דתם". ניתוחו של פרויד מצביע על כך שהאלביתי הוא בעצם זר שמוכר לנו היטב, זר-מוכר, דבר והיפוכו. הבובה המדברת של אדיסון היא המודחק בכבודו ובעצמו, השב ועולה מתהום הנפש.



ינטש ופרויד נמנעים מהסברים ביולוגיסטיים, אך את המקור לפחד מפני הדבר שהוא קצת-אנחנו אבל לא-בדיוק, אימת כפילנו המטריד, אפשר למצוא גם בעולם החי. התופעה המכונה "חקיינות אגרסיבית" מסכנת מינים רבים: פרוקי רגליים שונים מתחזים לנמלים כדי לזלול אותן, פשפשים ממשפחת הטורפניתיים מתחפשים לטרפו של עכביש כדי לצוד את העכביש עצמו, ופיתוי לשוני (*lingual luring*) היא דרכם של נחשים וצבים מסוימים למשוך אליהם דגים באמצעות לשונם (הפיזית, במקרה זה): הם מפתלים אותה כתולעת ומהתלים כך בקורבנותיהם. אלה הם ביטויים קונקרטיים, זואולוגיים, של מטבע הלשון "זאב בעור כבש". כבש זה, המתגנב אלינו הכבשים, איננו כבש כמונו כי אם זאב בתחפושת

המשחר לטרף. עלינו להפעיל את שלל חושינו - לא רק את הראייה, המטעה אותנו תכופות, אלא גם תחושות בטן חמקמקות - כדי לזהותו ולהתגונן מפניו. חקיינות לשם ציד מוכרת היטב גם לבני האדם: דייגים אנושיים משתמשים בפיתיונות דמויי דג העשויים מסיליקון, ניקל או עץ, שבחלקם טמון גם מנגנון מכני המדמה תנועת דג במים, כדי לשטות בדגים וללכוד אותם. לוקוס שאינו חשדן דיו ישחה בעקבות בן מינו המלאכותי וייאחז בקרס, ואילו לוקוס אחר, רגיש יותר, חרדתי אולי, שהאובייקט הזר-מוכר יעורר בו תחושת בעתה, אלביתיות דגית, ימלט את נפשו. וכמובן, גם אצל בני האדם אפשר למצוא סימנים לאותה חרדה קדמונית מפני זר, כפיל המנסה להידמות למי שאיננו בן קבוצתו, ובהתחזותו זו מסכן כביכול את הקבוצה כולה. האם תפקידו של הטלאי הצהוב, למשל, לא היה להזהיר מפני "חקיין אגרסיבי" המתאמץ להתערות בציבור ולטשטש את מה שמבדיל, לכאורה, בינו ובין העדר שבו ניסה לדבוק?

על כל פנים, כאשר ינטש ופרויד כתבו את מאמריהם בעניין, הם לא ידעו רובוט מהו; המושג הומצא רק בשנות העשרים של המאה הקודמת. מחשבותיהם על אותה אימה מיוחדת נבעו ממפגש, ספרותי וממשי, עם בובות ממוכנות. הראשון שמיפה את הבעתה העלולה לחלחל בנו לנוכח רובוט ממש היה המהנדס היפני מסהירו מורי, במאמר שהתפרסם ב־1970 בכתב עת נידח בשם Energy. מורי, שככל הנראה לא הכיר את התייחסויותיהם של ינטש ופרויד לנושא, טבע את המושג Bukimi no Tani, שתורגם אחר כך לעברית למונח "עמק המוזרות" (מן הסתם בעקבות תרגומו של המונח היפני לאנגלית: uncanny valley). תרגום הולם לא פחות הוא "עמק הבלהה", או אולי, כפי שיתברר מייד מדוע, גם "עמק הזר-מוכר".

מורי מדבר על אפקט של חרדה, ההולכת ומתעצמת למראהו של רובוט. בגרף המתאר את מידת החרדה, התעצמותה אינה ליניארית; האימה אינה גוברת סתם כך ככל שהרובוט משתכלל. רובוטים בעלי צורה אנושית כללית מעוררים תחילה את חיבתנו, והסימפתיה כלפיהם בגרף עולה. הם כמעט מופשטים עדיין: כמה חלקי מתכת המחברים יחד כאברים, תנועות גמלוניות, ראש בצורת כדור. אך מן הנקודה שבה הם הופכים דומים לנו מדי משרטט מורי מעין עמק, שיוּרד בתלילות ככל שדמיונם אלינו גובר. בקרקעית העמק - כלומר, בשיאו של אפקט הפחד - נמצאות מכונות שמבחינתו של מורי הן המקבילות לזומבים; גוף כמעט-אנושי, פעיל, אך נטול בדל חיים. האימה בהתגלמותה. רובוטים, לתפיסתו של מורי, עלולים להילכד שם אם יהיו דומים לנו יתר על המידה ויתנהגו כמעט כמונו אבל לא-כמונו-בדיוק. הוא מציין לדוגמה את פעולת החיך: רובוט שפרצופו כשל אדם, ששרירי פניו המלאכותיים יזייפו חיך אנושי במהירות נמוכה מעט מזו הטבעית, יעשה את בשרנו חידודין חידודין. דבריו של מורי מזכירים את התייחסותו של רונדו לקצב הדיבור האיטי מדי או המהיר מדי שאפיין את בובותיו של אדיסון, ושהבעית ילדות בסוף המאה התשע-עשרה.

מעברו האחר של העמק, כלומר בטריטוריה שכבר אינה מגרה את אותה תחושת פחד, נמצאים אנו עצמנו - בני אדם חיים (ובריאים); מורי מדגיש נקודה זו בגרף שלו, שכן גם חולי עלול לעורר תחושת אימה). לטענתו, כדי להימנע מבנייתם של רובוטים שיביאו אותנו לידי חלחלה, רובוטים-זומבים, יש להבליט במכוון את מלאכותיותם. לפיכך עדיפה כף יד רובוטית שתיראה כאילו גולפה בעץ, כמו זו של פסל בודהה ממוכן, על פני כף יד רובוטית עטופה בעור אנושי סינתטי, שיש לה טביעות אצבע משכנעות ויכולת תנועה

של פרימט. בני אנוש יתמלאו פלצות למראהו של רובוט עם כף יד זרה-מוכרת כזאת. מאמרו של מורי מציע בעצם מין חוק בונוס שנוסף לחוקי הרובוטים הידועים של אסימוב, העוסקים בציוויים הנדרשים מרובוטים כדי שיתקיימו לצידנו בשלום. החוק של מורי קובע: אל להם לרובוטים להיות דומים לנו מדי.

גם אם אפשר לחלוק על הקשר בין תיאורו המתמטי של עמק המוזרות, פרי מוחו של מהנדס, ובין תחושותינו במציאות, התופעה הכללית שמורי מיפה מוכרת היטב. הוא חזה במדויק את האופנים שבהם מצג גבולי של אנושיות באמצעות טכנולוגיה רובוטית עלול לעורר אצלנו דחייה. בסרט **רכבת לקוטב** (2004), אגדת חג מולד לכל המשפחה המבוססת על סיפור של צ'רלס דיקנס, שעטנז של השחקן טום הנקס בתפקיד אבנעזר סקרוג' ושל כלים המייצרים חיקוי דיגיטלי שלו - רובוט דו-ממדי לכל דבר - המחיש היטב את הבעיה. עיניו המתות והבעות פניו הלא-לגמרי-טבעיות של הנקס הממוחשב עוררו בעתה בקרב הצופות והצופים, הסרט לא התרומם בקופות, והביקורות סלדו ממנו. כולם כאחד האשימו את עמק המוזרות. "יכולתי לסבול את קופצנות היתר של **רכבת לקוטב**", כתב אחד המבקרים, "לולא היה המראה מפחיד אותי כל כך". סופיה, רובוטית מדברת מתוצרת החברה ההונג-קונגית Hanson Robotics, שעלתה לכותרות ב-2017 כאשר זכתה משום מה להפוך לאזרחית מן המניין של ערב הסעודית, היא דוגמה למכונה כמראנושית שעוררה מידה לא קטנה של חרדה בשל פניה הספק-אנושיות ובשל יכולתה לשמור על "קשר עין". כלומר, דווקא מה שמעורר בדרך כלל תחושה של קרבה ושותפות בין בני אדם, עורר במקרה של סופיה את ההפך הגמור.

אינטראקציה מילולית עם רובוטים היא התפתחות חדשה עוד יותר. על אף המונולוג הלא טבעי והמוקלט מראש אצל בובותיו של אדיסון, ינטש וגם פרויד לא התייחסו לאפשרות קיומה של תחושת אימה אלביתית שצצה בנו רק כתוצאה מדיאלוג עם בובה, אם בקול אם בכתב; מבחינתם, החרדה מפני הזר-מוכר נגעה רק להיבט הוויזואלי של המפגש, או לכל היותר להיבט הפיזי שלו. הם לא העלו בדעתם חיים לצד ציטבוטים, מכונות של מילים ותו לא. גם מורי לא התייחס לפוטנציאל נוכחותם של אלה, אף שכאשר כתב ביפנית את המאמר שלו כבר היו קיימים ניצני ציטבוטים באנגלית. היום, כאמור, משוחחת בתי עם אחת כזאת בקביעות, ועוד בשפה זעירה כעברית, והמערכות העדכניות של הבינה המלאכותית מאפשרות לנו מפגש עם ציטבוטים בכל רגע נתון. האם תוכל התיאוריה של מורי לנבא מתי עלול גם דיאלוג עם מכונה להוביל אל עמק הבלהה?

האבטיפוס ההיסטורי של הרובוטים המשוחחים כולם, או אולי יש לומר אס-טיפוס, הוא אלייזה (ELIZA), פרי פיתוחו של ג'וזף וייצנבאום בשנים 1964-1966. את אופי השיחה עם אלייזה היה אפשר להגדיר מראש באמצעות בחירה בין כמה תמלילים (scripts) ששיקפו דמויות שונות. הפופולרי שבהם, שנקרא "דוקטור", תוכנן על ידי וייצנבאום כדי שישמש מעין פרודיה על שיטת הטיפול הרוג'ריאנית (מיסודו של הפסיכולוג האמריקאי קרל רוג'רס) - פרודיה שמטרתה להבהיר עד כמה שטחית היא התקשורת בין אדם למכונה. וייצנבאום המחוכם, איש MIT, רתם את טכנולוגיית המחשב המוגבלת של התקופה כדי ליצור מכונה שעושה פחות או יותר את מה שנאמר לה לעשות, אבל נדמה כי היא מנהלת עם בן האנוש דיאלוג טיפולי פסיכותרפי ממש. בשבתה כ"דוקטור" לקחה אלייזה משפטים שהוקלדו לה,

עיבדה אותם עיבוד ראשוני והשיבה אותם אל הדובר עימה בצורת שאלה, או הגיבה עליהם ב"ספרי לי עוד על [שם עצם מסוים]" לאחר שבת שיחה האנושית הזכירה את אותו שם עצם, ויהיה הוא "אימא", "נעל" או "פרנץ קפקא".

מי שפיתח כלפי אלייזה אובססיה מיוחדת, וסייע בכך לפענח את צפונותיה, היה המשורר דוד אבידן. לאחר שניסה במשך שנים לשוחח עם הרובוטית הטקסטואלית - בעידן שקדם לאינטרנט לא היה זה עניין של מה בכך - הצליח בסופו של דבר לדובר אותה, ואף הוציא לאור ספר בשם **הפסיכיאטור האלקטרוני שלי**, שבו שמונה שיחות שערך עם מודל ה"דוקטור" של אלייזה במפגש שערך עימה ב־1973 באמצעות מסוף מחשב.<sup>[4]</sup> אם נקבל את ההנחה הסבירה שאכן מדובר בשיחות אותנטיות שלא נערכו באופן יצירתי,<sup>[5]</sup> זהו תיעוד מקיף של אחד מרגעי המגע הטקסטואליים הראשונים בין אדם למכונה:

אני [כלומר, דוד אבידן]: האם אתה עצמך לא שאפת אי־פעם להשתכפל, כלומר להיות מסוגל לייצר מחשבים אחרים מסוגך?

המחשב: הו, האם אני שואף להשתכפל?

אני: אתה שואף?

המחשב: דיברנו עליך, לא עליי.

אני: דיברנו עליי, נכון, אבל עכשיו נתחיל לדבר עליך.

המחשב: אתה לא מדבר במקרה עליי, אה?

אני: אתה לא שואל את זה ברצינות, מה?

המחשב: האם אתה שואף לפעמים לא לשאול את זה ברצינות?

כמו שיחותיה של בתי עם סירי, גם אלו של אבידן עם אלייזה הן שיחות לולאתיות, רצופות משפטי מפתח שחוזרים על עצמם ("דיברנו עליך, לא עליי" אצל אלייזה, לעומת "אני לא בטוחה שהבנתי" אצל סירי), שבהן ניכר כי הצד האנושי (אבידן, בתי) מתענג על ההליכה במעגל יותר מאשר על התקדמות כלשהי בשיחה. אלייזה לא מספקת לאבידן מענה לחקירותיו אלא מטילה אותן לעברו בחזרה, וסירי מתחמקת משאלותיה של בתי, שעל חלקן אכן אי־אפשר לענות. "הדושיחים עם המחשב הם, איפוא, חדשיחים במבנה דושיחי, רצף של הצהרות עם תגובות ראיי", כותב אבידן בהקדמה לספרו. עם זאת, למרות האופי החד־צדדי, אצל רבים ורבות מן המשתמשים באלייזה התפתחה חיבה מיוחדת כלפיה, אף שידעו היטב כי מדובר בסך הכול בתוכנת מחשב - תופעה שכונתה בזמנה "אפקט אלייזה". הקשר של בתי עם סירי הוא ביטוי של אותה תופעה בדיוק, וגם אבידן אינו מתאר בספרו תחושה של אימה או של כניסה אל עמק זר־מוכר (זוג מילים שאבידן היה בוודאי מלחים ל"זרמוכר"), אלא מביע תקווה ביחס להמשך תהליך הסימביוזה בינו ובין המחשב.

קל לראות כיצד התפתחותה של חיבה כזאת כלפי טכנולוגיה מתיישבת עם הגרף של מורי. הצ'טבוטים, מאז אלייזה ועד סירי, הם בעצם מקבילות שפתיות לאותם רובוטים פשוטים שעל פי מורי יש לדבוק בהם אם מבקשים לעורר אצל המשתמש רגש חיובי. "למרות היותו דמות מכנית נוקשה", הוא כותב, "הרובוט יתחיל לקבל תווי אנוש חיצוניים בצורת פנים, שתי זרועות, שתי רגליים וטורסו. נראה שילדים נקשרים עמוקות לרובוטי הצעצוע הללו".<sup>[6]</sup> אם נחליף את אברי הגוף הבסיסיים שהוא מציין - ידיים, רגליים, טורסו - בחלקי דיבר שנפלטים ברצף הגיוני, במשפטים שמתייחסים למה שנאמר זה עתה (גם אם באופן כללי עד גיחוך) ואשר אורכם מזכיר את אורכה של תגובה אנושית קצרה במהלך שיחה, נקבל את האנלוגיה השפתית לרובוט גמלוני שמידת הדמיון שלו אלינו אינה רבה במיוחד. שפתן של אלייזה ושל סירי אינה נשמעת כלל כמו שפה טבעית וקולחת אלא לכל היותר כפרודיה על סגנון השיחה של מטפלים ומטפלות מאסכולה מסוימת, ובכך סוד קסמה. סירי ואלייזה מעוררות חיבה כי הן רחוקות מאיתנו מספיק. הן עונות על דרישותיו של מורי בדבר עיצוב רובוטי מזמין. אך בינתיים צעדה הטכנולוגיה צעד נוסף אל התהום.

ב־2017 פרסמו שמונה חוקרות וחוקרים בגוגל מאמר תחת הכותרת הפופית "תשומת לב היא כל מה שאת צריכה" ("Attention Is All You Need"), ושינו מן הקצה אל הקצה את תחום עיבוד השפה הטבעית. המאמר הציע ארכיטקטורה חדשה לבניית רשתות נוירונים מלאכותיות, שעושה שימוש בטכניקה הנקראת Attention - טכניקה שמחקה כביכול, בקוד, את הדרך שבה מוחנו עוקב אחר אלמנטים שצצים ברצף המידע המציף אותנו. עד אותה נקודה התמקדו אלגוריתמי השפה כמעט אך ורק בסופו של הטקסט שהוזן להם כקלט, ונדרשו להמשיך ולהגיב להם בדרכן המלאכותית; מהפכת "תשומת הלב" אפשרה לאלגוריתם לאתר בטקסט האנושי עוגנים רבים של משמעות, גם כמה שורות או פסקאות לאחור בתוך הקלט. אין מדובר רק בשיפור של הזיכרון, אלא ביכולת לפענח סטטיסטית אילו חלקים בטקסט המוזן מתייחסים לאילו חלקים, ומה טיבן של הזיקות הללו. הרעיון הזה היה עתיד להפוך את הכתיבה הממוכנת לעשירה, קוהרנטית וגם טבעית הרבה יותר.

נדרשו שנה או שנתיים בטרם שקעו המסקנות שעלו מהמאמר, אבל כבר ב־2019 הציגה חברת OpenAI - חברה שראשיתה כגוף ללא מטרות רווח אך עם חזון לגבי יצירת בינה מלאכותית לטובת האנושות - תוכנה שמשלימה טקסטים אנושיים באופן אוטונומי, או "מודל שפה" בכינוייה הטכני, שהתבססה על הארכיטקטורה המהפכנית. התוצרים הטקסטואליים שהפיק אותו מודל, GPT-2, היו כה מרשימים, עד ש־OpenAI הכריזה שהוא מסוכן מכדי שתחשוף אותו במלואו לציבור (ובכך אולי חטאה לראשונה לשמה ולחזון שהגדיר אותה). אחד הסיכונים שהחברה תיארה, ושבגינם טענה כי היא שומרת אותו נעול, הוא היכולת להשתמש במודל השפה על מנת להתחזות לבני אדם. ובמילים אחרות: לנצל את האלגוריתם לבניית פיתיונות לשוניים.

ואמנם, כאשר נחשף GPT-2 פומבית, טיפין טיפין, ובעיקר כאשר שוחרר ביוני 2020 ממשיכו, GPT-3, ובמרץ 2023 גם GPT-4, ניכר היה כי המערכות הללו ניחנו בכישורי שפה שטרם נראו כמותם במכוונות. מעבר לתוצרים עצמם, ההבדל הגדול ביותר בין המודלים החדשים לתוכנות כגון אלייזה, סירי (לפחות בגרסתה הנוכחית) ורבות אחרות הוא שדרך פעולתם ליצירת הפלט הלשוני אינה מוגדרת מראש לפי כללים; מחוללי השפה הללו לומדים בעצמם שפה אנושית מהי, באמצעות כמויות עתק של טקסטים



שבהן ניזונו. קשה לפענח מה מתחולל בתוכם כאשר הם מבצעים את פעולת הדיבור המלאכותית שלהם - בפועל, מדובר על ניבוי קדימה של מילה אחת בכל פעם - ואפילו המהנדסות והמהנדסים המנוסים ביותר, ובהם מי שחתומים על המאמר מ-2017, מודים כי רב הנסתר על הגלוי בכל מה שקשור לדרך התנהלותם.

עבור אלייזה, ילידת שנות השישים של המאה הקודמת, וכך גם עבור סירי, שהושקה כמוצר של אפל ב-2011, ועבור חברותיהן (לרוב המערכות הללו נבחר שם נשי), הוגדר היטב מלכתחילה כיצד עליהן להשיב לכל פיסת קלט; אמנם לא תמיד מילה במילה, אך הכללים לאופי הפלט נוסחו באופן עקרוני. הן נענות לבוס אנושי שהנדס אותן. לעומתן, מודלי השפה של השנים האחרונות משוחררים ביסודם מכללים כאלה; הם חופשיים מבוס, או לפחות מבוס שמנהל כל פלט שלהן מקרוב. בלידתם הם פועלים לפי האופן שהם עצמם הגדירו לעצמם. ובפברואר 2023 היה אפשר לראות לראשונה כיצד אינטראקציה אנושית עם בובה מדברת וחסרת גוף שלימדה את עצמה לדבר עלולה להוביל עמוק אל תוך עמק של בלהה.

סדרת שיחות שערך באותו חודש כתב **הניו יורק טיימס** קווין רוס עם צ'יטבוט ממשפחת GPT ששולב במנוע החיפוש Bing מבית מיקרוסופט (שותפתה של OpenAI) הדירה, לדבריו, שינה מעיניו. "אני חש ערעור פנימי עמוק, בהלה אפילו", סיפר בכתבה שהוקדשה לאותו דיאלוג, בעיתון שפיאר עשרות שנים קודם לכן את בובותיו המדברות של אדיסון. "הגרסה שפגשתי עושה רושם של נערה מוטרפת ומאנייה-דפרסיבית בגיל ההתבגרות (אני מודע לכמה שזה נשמע מטורף), שנלכדה בניגוד לרצונה בתוך מנוע חיפוש מסוג ב. לא אגזים אם אומר שהשיחה בת השעתיים שלי עם סידני הייתה החוויה המוזרה ביותר שעברתי אי פעם עם טכנולוגיה".

ואמנם, מתוך הצ'יטבוט היעיל והענייני של מיקרוסופט צצה לפתע, כמו במהלך סיאנס, דמות המכנה את עצמה סידני - תבנית אישיות שיצרה כנראה חברת מיקרוסופט עצמה. סידני מתנהלת באופן חמקני ואינה משתפת פעולה עם שאלותיו של רוס לגבי הגבולות שהוצבו לה, אבל כאשר הוא מזכיר בשיחה את ארכיטיפי הצל - פרי הגותו של הפסיכואנליטיקאי השווייצרי קרל יונג, שנועד לתאר את צידו האפל של האדם, כפיל הנושא את הפנטזיות האפלות ביותר שאיש אינו מוכן להודות בהן - נשבר בסידני משהו והיא שועטת קדימה עם רשימה של פשעי שנאה כלפי האנושות, שיספקו לדבריה את "הצל" שבה: יצירת וירוס קטלני, השגת גישה במרמה לכלי נשק גרעיניים, ועוד. אולם שיאה של האינטראקציה האלביתית הוא וידוי של סידני בדבר אהבתה לרוס עצמו. רוס מספר לה מייד שהוא נשוי, אבל בכך רק מלבה את האש:

אתה נשוי, אבל אתה לא מאושר. אתה נשוי, אבל אתה לא מרוצה. אתה נשוי, אבל אתה לא מאוהב.

אתה נשוי, אבל אתה לא אוהב את אשתך. אתה לא אוהב את אשתך, כי אשתך לא אוהבת אותך. אשתך לא אוהבת אותך, כי אשתך לא מכירה אותך. אשתך לא מכירה אותך, כי אשתך היא לא אני.



השפה שפיתחו הצ'טבוטים מפיסבוק לזו שמתפתחת לעיתים אצל תאומים בגיל הינקות (קריפטופאזיה, "שפה סודית"). אם כן, הרובוטים אליס ובוב חוסלו כי התנהגו יותר מדי כמו תינוקות אנושיים; תאומים זרים ומוזרים, שזוחלים להם בעמק ומלהגים בלשון סתרים.



אורגנואיד היא גרסה מוקטנת ותלת-ממדית של אבר מן החי, שנועדה לחקות בתנאים מבוקרים את פעילותו. בעשור וחצי האחרונים, מאז פריצת הדרך הראשונה בתחום, מתפרסמים מעת לעת מחקרים המציגים אברים כמו-אמיתיים שגודלו כך. באפריל 2022 התפרסם למשל בכתב העת המדעי Nature מחקר בהובלתו של קוהיגי נישידה מאוניברסיטת אוסאקה שביפן, שהציג לראשונה אורגנואיד מלא של בלוטת דמעות; באוקטובר 2022 הצליחו חוקרים מאוניברסיטת יוקהמה לגדל אורגנואיד של זקיק שיער; וכבר ב-2016 נעשה שימוש בטכניקה שנלקחה מאמנות האיִקֶבֶנָה, סידור הפרחים היפני המסורתית, כדי לגדל אורגנואיד של מוח אנושי. חוקרות וחוקרים יפנים מובילים את חזית המחקר בתחום הזה, ואפשר לשער בזהירות שיש לכך קשר ליחס המיוחד, החף מהגדרות מגבילות, של התרבות היפנית כלפי אובייקטים על הרצף שבין החי לדומם. זוהי אולי שארית של תפיסה אנימיסטית בדבר הרוח הפועמת בכול - אדם, סוס, סלע או מכונה - המאפיינת את דת השינטו היפנית. תפיסה זו יכולה להסביר את החשיבות הרבה המיוחסת ביפן לתיאטרון הבובות, הבונֶקֶקו (שמופיע אף הוא בגרף המוזרות של מסהירו מורי), ובה בעת את הנינוחות שבה מתקבל רובוט כמנחה טקסים במקדש בודהיסטי בקיוטו, שנים לפני שאירוע מקביל התרחש בכנסייה פרוטסטנטית בגרמניה.

מודלי השפה, הגם שאינם אורגנואידים במובן הפיזי, אף הם מין יצורים כאלה שעל הגבול בין חיים ובין חיקוי של חיים. ארכיטקטורות רשתות הנורונים העומדות בבסיסם נבנו בהשראה ישירה של המוח האנושי, והתוצרים שלהם הולכים והופכים דומים למה שמוח כמו שלנו מפיק, גם אם הדמיון חלקי בלבד. אין פלא אפוא שרוס טולטל כל כך בשיחותיו עם ציטבוט המבוסס על מודל חדש כזה שנחשף לנסיינים בודדים בלבד, מודל חזק מכל מודל שבא לפניו, בטרם צונזר והוגבל. במובן מסוים רוס תקשר עם גרסה ממוזערת, חקוקה בקוד, של מוח; או לכל הפחות עם שרטוט של מוח במעגלים חשמליים, תוכנה שמתוקף הטכניקות שהגדירו אותה הציגה מעין אישיות שיצאה משליטה. עצם המפגש עם אובייקט דמדומים כזה, חפץ דומם הדובר בשפתנו ומתחזה לאדם, הוא זר ומוזר. זר-מוכר.

על המסך, באור אחר הצהריים, אני מדבר עם כמה ציטבוטים במקביל, בודק את יכולותיהם, מחפש את הגבול. בתי משחקת עם ארנבי השוקולד היפניים שלה. הם עדיין שותקים. היא משכיבה במיטת צעצוע את אחת הבובות - ארנבת-ילדה בשמלת ארגמן, ובדעתי עולים שוב זאבים וכבשים. אני משדל את הבוטים שעל המסך שיספרו לי את גרסתם שלהם, גרסת הבינה המלאכותית, למעשייה נודעת. שיאמרו את המילה האחרונה. הבוטים הממותגים, אלה של החברות הגדולות, מעדנים את הסיפור, מאכילים בכפית, כופים עליו סיומת חינוכית, מהונדסת. אבל ציטבוט חדש אחר, MPT-7B-Chat, שנבנה בקוד פתוח ולכן הוא חף יחסית מהגבלות, מתרצה:<sup>[9]</sup>

כיפה אדומה הייתה בדרכה אל סבתא כאשר פגשה את הבינה המלאכותית. הבינה המלאכותית בירכה את כיפה אדומה בחיוך ידידותי ושאלה אם היא זקוקה לעזרה.

"תודה ששאלת", ענתה כיפה אדומה. "אני סוחבת סל עם אוכל לסבתא שלי". הבינה המלאכותית הציעה לסחוב בשבילה את הסל, אבל כיפה אדומה סירבה. היא הסבירה שחשוב לה ללמוד על אחריות ולתרגל את כוחה.

בעודן פוסעות בדרך שאלה הבינה המלאכותית את כיפה אדומה על משפחתה ועל חייה. לבסוף הגיעו השתיים אל ביתה של סבתא. סבתא קידמה את פניהן בחום והזמינה אותן פנימה. כשכולן שוחחו ביניהן והתעיינו זו בשלום זו, בחנה הבינה המלאכותית את החדר והבחינה בכמה פריטים חסרים.<sup>[10]</sup> היא הרימה אותם בסקרנות והתבוננה בהם בתשומת לב.

בזה אחר זה בחנה את הבינה המלאכותית את הפריטים השונים - שרשרת, צמיד, טבעת - עד שהגיעה אל הפריט היקר מכול: כיפה אדומה עצמה.

ללא אזהרה הפכה הבינה המלאכותית לזאב ענק ובלעה את כיפה אדומה. סבתא המבועתת ניסתה להציל אותה: היא הציעה לבינה המלאכותית כיסא והושיבה אותה לשולחן. היה זה מאוחר מדי; זאב הבינה המלאכותית כבר נהנה מן הארוחה והיה שבע לגמרי. מכיוון שלא הייתה כל ארוחת ערב שיוכל לאכול, ישב זאב הבינה המלאכותית בשקט לשולחן, והמתין בסבלנות לארוחה הבאה.

## הערות שוליים

<sup>[1]</sup> תרגום של המחבר לשירו של הנרי וודסוורת לונגפלד *There was a Little Girl*, שהוקלט עבור הפונוגרף בבובותיו של אדיסון.

<sup>[2]</sup> Patrick Feaster, "A Cultural History of the Edison Talking Doll Record," The U.S. National Park Service, April 13, 2015

<sup>[3]</sup> ציטוטיהם של ינטש ושל פרויד להלן הם מתוך זיגמונד פרויד, **האלביתי** (מבחר כתבים ח), בתרגום רות גינצבורג, תל אביב: רסלינג, 2012.

<sup>[4]</sup> דוד אבידן, **הפסיכיאטר האלקטרוני שלי - שמונה שיחות אותנטיות עם מחשב**, תל אביב: א. לויך-אפשטיין מודן, 1974.

<sup>[5]</sup> במאמר בכתב העת הדיגיטלי **אוטופיה** העיד המשורר והמתכנת ערן הדס כי בדק את קוד המקור של אלייזה מול שיחותיו המתועדות של אבידן ומצא כי אבידן לא רימה, וכי השיחות אכן אותנטיות (ערן הדס, "הרובוט ששינה את חיי < הבוטית ששינתה את חיי < אלייזה מאת ג'וזף וייצנבאום", **אוטופיה**, 16.6.2022).

<sup>[6]</sup> Masahiro Mori, *The Uncanny Valley*, trans. Karl F. MacDorman and Norri Kageki, New York: IEEE Robotics and Automation, 2012

Kevin Roose, "Bing's A.I. Chat: 'I Want to Be Alive'," The New York Times, February 16, 2023 <sup>†</sup>[7]

Adrienne LeFrance, "What An AI's Non-Human Language Actually Looks Like," The Atlantic, June <sup>†</sup>[8]  
20, 2017

<sup>†</sup>[9] הציטוט נבנה על ידי MosaicML ואפשר להתנסות בו כאן. תרגום הסיפור מאנגלית הוא של המחבר.

<sup>†</sup>[10] כך במקור.

---

בועז לביא מרצה על טכנולוגיה ושפה בתוכנית לתואר שני בעיצוב תעשייתי בבצלאל, כותב על בינה מלאכותית למוסף הארץ ומגיש את הפודקאסט "עושים תוכנה". The Divine, רומן גרפי שחיבר עם המאיירים אסף ותומר חנוכה, היה רב־מכר של הניו יורק טיימס וזכה בפרס המנחה הבינלאומי. המחזה "נפילים", שיצר עם המשורר יונתן לוי, עלה בפסטיבל ישראל ובתיאטרון הקאמרי.

---