

27 בנובמבר 2023

האם חזבאללה ייכנס למלחמה עם ישראל?: סימולציות משחקי תפקידים מבוססות בינה מלאכותית ככלי עזר להערכות מודיעין

עופר גוטרמן, ליאור בורק, רועי צזנה¹

תקציר

המסמך מתאר משחק תפקידים מבוסס בינה מלאכותית ג'נרטיבית (Gen-AI), המדמה דינמי קבלת החלטות בראשות בכירים בציר האיראני, בשאלת צירופו של חזבאללה למלחמה עם ישראל שפרצה בשבעה באוקטובר 2023. הסימולציה נועדה לבחון את מידת היכולת של משחקי תפקידים דיגיטליים להעלות תרומה בדילמות של ניתוח והערכה במודיעין האסטרטגי.

הכנת "מגרש המשחקים" עבור הסימולציה כללה בניית תאומים דיגיטליים, באמצעות הבינה המלאכותית – ייצוגים ממוחשבים של הדמויות שנבחרו לייצג את מעגל קבלת ההחלטות של הציר האיראני, כתיבת מספר סיפורי רקע שהוצגו בפני התאומים הדיגיטליים של בכירי הציר, כתרשימים שונים של המציאות בנקודות זמן שונות במלחמה, ופיתוח ממשק אפליקטיבי שיריץ את הסימולציה על כל מרכיביה, ובמסגרתו הבינה המלאכותית פיתחה את הטיעונים של הבכירים השונים.

נראה שהתאומים הדיגיטליים הצליחו "להתקרב" למקור במספר דרכים: חיקוי סממני שיח, חיקוי היבטים מסוימים באופי (דוגמת קבלת החלטות שונאת-סיכון או מתן ביטוי לתפיסה פוליטית-דתית), ושיקוף אינטרסים בהתאם לשייכות הארגונית. בנוסף, מספר מומחים בעלי הכרות עם הדמויות ועם הזירה המזרח-תיכונית עברו על טיעוני התאומים הדיגיטליים, והעניקו להם ציון גבוה, הן מבחינת ההיגיון של הטיעונים, והן מבחינת הסבירות שהבכירים האמיתיים היו מעלים טענות אלו.

¹ אל"מ (מיל") ד"ר עופר גוטרמן הוא חוקר בכיר במכון לחקר המתודולוגיה של המודיעין. ליאור בורק היא אנליסטית בפירמת BDO ובעברה חוקרת בחטיבת המחקר באמ"ן. ד"ר רועי צזנה הוא עתידן וחוקר במרכז בלווטניק באוניברסיטת תל-אביב. המחברים מבקשים להודות לאדווין גרודברג ולדן ויינשטיין על סיועם בהעמדת התשתית הטכנולוגית, ולאיתי ברון, דודי סימן טוב וד"ר יונית רוש על הערותיהם.

באופן כמעט גורף, התאומים הדיגיטליים התנגדו לשילוב חזבאללה במלחמה רחבה עם ישראל, כפי שארע במציאות (עדכנית לנובמבר 2023). מכל מקום, איננו מייחסים לסימולציה יכולת פרדיקטיבית של חיזוי העתיד; חשיבותה ביכולת לספק שיטה נוספת לחשוב על שיקולים ודרכי פעולה אפשריות של היריבים, ולהבנתנו, הסימולציה הצליחה לספק כמה חידודים וחידושים בהקשר זה.

בנוסף, הכנת הסימולציה וביצועה, כולל שורה של שינויים והתאמות שנעשו לאורך הדרך (במשתתפים, בתכנים, בשאלות), הוכיחו רמה גבוהה מאוד של אגיליות, כלומר יכולת להטמיע במהירות שינויים ותוספות בין כל סבב הרצה של הסימולציה, וכן רמה גבוהה של סקלביליות, כלומר יכולת הרצה של סימולציות רבות בעלות נמוכה של זמן ומשאבים, בהשוואה לסימולציות אנושיות.

בכפוף לשימוש זהיר בסימולציות משחקי התפקידים המתבססות על בינה מלאכותית יוצרת, ותוך מודעות למגבלות, אנו מאמינים שיש לעשות בהן שימוש ככלי תומך החלטה, אשר יכול לספק לאנליסטים ולמתכננים אסטרטגיים יתרונות תחרותיים, ולתרום לפיתוח כיווני חשיבה נוספים על המציאות העתידית, בעולם שהפך רווי בתפניות, שיבושים וסיכונים כבדי משקל.

הקדמה

המלחמה בין מדינת ישראל לבין החמאס, אשר פרצה בעקבות מתקפת הפתע של חמאס בשבעה באוקטובר, הובילה להשתלבות של ארגון חזבאללה במערכה. אולם, חזבאללה השתלב במערכה באופן שנותר מתחת לסף של מלחמה, ובגבולותיו של עימות מוגבל, כזה הכולל אומנם לא מעט פעולות צבאיות משני הצדדים, אך עדיין תוך הפעלה חלקית מאוד של יכולותיהם הצבאיות, למשל בהיבטים של היקפים וטווחי הפעולה. בנסיבות אלו התמודדה ישראל מתחילת המלחמה עם שאלה כבדת משקל לעתיד המערכה: האם ובאלו תנאים יבחר חזבאללה להצטרף באופן מלא למלחמה?

במהלך חודש אוקטובר התקיים (על ידי כותבי המאמר) משחק תפקידים מבוסס בינה מלאכותית ג'נרטיבית (Gen-AI), המדמה דיוני הערכת מצב וקבלת החלטות בראשות בכירים בציר ההתנגדות, בשאלת הצטרפותו של חזבאללה למלחמה עם ישראל. עריכת הסימולציה נועדה לבחון את מידת היכולת של משחקי תפקידים דיגיטליים מבוססי בינה מלאכותית, להעלות תרומה לעיסוק בדילמות חשיבה בעולמות של מודיעין ותכנון אסטרטגי. בכך, אנו רואים את המחקר הנוכחי כמהלך נוסף, שמצטרף לזה של איתי ברון ותהילה שוורץ-אלטשולר, לפתח ידע אודות שימוש בסימולציות כדי לבחון את יכולתה של הבינה המלאכותית היוצרת לסייע לייצר הערכת מודיעין בבסיס לקבלת החלטות אסטרטגיות בתנאי אי וודאות קשים.²

² איתי ברון ותהילה שוורץ אלטשולר (מאי 2023), [מקום ליד השולחן: האם ה-ChatGPT יכול "לשבור את הקונספציה"?](#), המכון לחקר המתודולוגיה של המודיעין.

סימולציות של משחקי תפקידים כדרך לחקר הסביבה התחרותית והאסטרטגית

סקוט ג. ארמסטרונג, אחד החוקרים המובילים בתחום משחקי התפקידים כדרך לחקר וחיזוי עתיד הדגים, בסדרה של מחקרים, כי משחקי תפקידים מכניסים את השחקנים להלך רוח שמשקף, לפחות במידה חלקית, את הלך רוחם של המנהלים והמפקדים בכל אחד מהצדדים בקונפליקט. בשמונים ושלושה משחקי תפקידים שארגן, הראה ארמסטרונג כי אנשים מן השורה שמשחקים תפקידים של מנהלים בכירים בחברת תרופות, מקבלים החלטות שאינן-מוסריות ואינן-אתיות – ממש כמו אלו שקיבלו המנהלים בהיסטוריה הממשית של אותה חברת תרופות (Upjohn). החלטות אלו הביאו בסופו של דבר לקריסתה של החברה. כאשר נשאלו כלכלנים לגבי ההחלטות שמנהלי החברה יקבלו, הם צדקו רק ב-34% מהמקרים. המשתתפים במשחקי התפקידים, לעומת זאת, בחרו באפשרות בה אכן נקטה החברה ב-76% מהמשחקים³.

במקרה זה, טעותם של המומחים לא התמצתה רק בחוסר יכולתם לשים את עצמם בנעליהם של מנהלים אינדיבידואליים, אלא בכך שלא הבינו את דפוס האינטראקציות בין המנהלים השונים במפגשים בהם מתקבלת ההחלטה הסופית. רק משחק תפקידים ממשי, בו כל אחד מהשחקנים קיבל דף המתאר את רצונותיו, פחדיו וצרכיו של המנהל שאת דמותו גילם, הצליח להביא להערכה נכונה של התוצאה הסופית.

ארמסטרונג ערך עוד שפע של משחקי תפקידים אחרים, במסגרתם הצליח לחזות את תוצאותיהם של קונפליקטים. דוגמאות מייצגות הן אלו בהן התמרדה קבוצת אמנים הולנדיים נגד הממשלה, או כאלו בהן תאגיד ניסה להציע שותפות עסקית לרשת סופרמרקטים. בשני המקרים מדובר באינטראקציה בין שני צדדים הנמצאים בקונפליקט אחד עם השני, ושואפים ליישב אותו באופן שיביא תועלת מרבית לשניהם. ארמסטרונג ביקש מסטודנטים לשחק את דמויותיהם של המנהלים בכל אחד מהקונפליקטים, והראה כי בשני המקרים משחקי התפקידים הוכיחו הצלחה מרשימה בחיזוי העתיד. בקונפליקט ההולנדי ניכרו 3 אחוזי הצלחה בחיזוי פשוט, לעומת 29 אחוזי הצלחה במשחק התפקידים. בקונפליקט בין התאגיד לסופרמרקטים, שלושה אחוזים הגיעו להצלחה בחיזוי פשוט, לעומת 75 אחוזי הצלחה במשחקי התפקידים⁴.

מסיכום כל תוצאות מחקריו עולה כי ב-56% מהמקרים, הצליחו משחקי התפקידים לחזות את תוצאות הדיונים בין הצדדים השונים, בעוד סקר לקבלת דעות בלבד אודות תוצאות הדיונים (ללא משחק תפקידים), הניב 12 אחוזי הצלחה בלבד.

³ J. S. Armstrong. (2000) Role Playing: A Method To Forecast Decisions. *Principles of Forecasting*, New York, Kluwer Academic Publishers, 13 - 31.

⁴ J. S. Armstrong ו P. D. Hutcherson. (1989) Predicting the Outcome of Marketing Negotiations: Role-Playing versus Unaided Opinions. *International Journal of Research in Marketing*, Vol 6 (4), 227 - 239.

הספרות אודות שימוש במשחקי תפקידים לצורך חיזוי העתיד מוגבלת מאד בהיקפה ובהיקף החוקרים העוסקים בה. עם זאת, משחק תפקידים יכול לסייע בחקר העתיד ובהתכוננות טובה יותר לקראתו.

משחקי תפקידים מספקים במה לראייה כוללת של הצדדים היריבים. השחקנים מבינים את הצרכים המורכבים של הדמויות אותן הם מגלמים, ומשלבים אותם בתהליכי קבלת ההחלטות. הקבוצות אינן חייבות לקבל החלטות רציונליות בלבד, או כאלו שמטיבות עמן. הן עשויות למשל לבחור בנתיבי פעולה שיפגעו בהן או במטרותיהן בטווח הארוך, אך ישרתו אותן בטווח הזמן הקצר, או בהחלטות שמטרתן לשמור על לכידות פנימית או לא לסטות מהאידיאולוגיה הקבוצתית. במילים אחרות, משחק התפקידים מוביל לקבלת החלטות שהן תוצר של דינמיקות השחקנים בסביבתם הפנימית והחיצונית כאחת.

משחקי התפקידים שימושיים במיוחד לניתוח מצבים של שינויים משמעותיים, המחייבים אימוץ מהיר של קונספציה חדשה – שלפעמים אינה ברורה מראש לאף אחד מהצדדים. במצב כזה המומחים לחיזוי מאבדים חלק גדול מהיתרון היחסי שלהם, בעוד משחק התפקידים מאפשר לדמויות להגיב "בזמן אמת" לשינוי, "לפי אופייה" של הדמות, ועל בסיס דינמיקה של התייעצות, משא ומתן, או קונפליקט בין השחקנים השונים.

ארגון מוצלח של משחק תפקידים אינו משימה פשוטה. משחק תפקידים טוב מחייב הכנה מראש מצד המשתתפים – שיכולה לדרוש ימים רבים – כמו גם הוצאת זמן לא-מבוטלת ביום המשחק עצמו. באופן אידיאלי, במקום סימולציה אחת, היינו רוצים להריץ עשרות משחקים במקביל, עם שינויים מזעריים בין כל משחק, כדי לראות את ההבדלים ולזקק מהם תובנות. בפועל, העלויות והאילוצים הכרוכים בביצוע משחקי תפקידים, גורמים לכך שהם מהווים פרקטיקה שהיא בבחינת "מותרות": הם מתקיימים אחת למספר שנים, בדרך-כלל במסגרת מכון מחקר כלשהו שמצליח לגייס את השחקנים המוכשרים ביותר, דוגמת גנרלים לשעבר, או מומחים אקדמיים. מצב עניינים זה עשוי לקבל בקרוב מאד תפנית חדה, עם הגחתם של מנועי השפה הגדולים ומערכות מרובות-סוכנים.

שילוב בינה מלאכותית במשחקי תפקידים

מנועי הבינה המלאכותית החדשים – הידועים כמודלי שפה גדולים (LLM – Large Language Models) – מתבססים על יכולתם לחקות דפוסי שפה. הם מאומנים על המידע הגדול שברשת, ומסוגלים לשחזר אופני דיבור שונים בהתאם לקונטקסט המוגדר להם.

שני מחקרים מאוגוסט וספטמבר 2023 מדגימים את כוחם של מנועי השפה הגדולים כאשר הם נדרשים לדבר דווקא אחד עם השני.

במחקר מאוגוסט, צוות מחקר בסטנפורד ובגוגל הקים כפר וירטואלי ובו 25 סוכנים. כל אחד מהסוכנים היווה ישות אוטונומית שנשלטה על-ידי מודל שפה גדול. כל אחד מהסוכנים קיבל שם משלו והיסטוריה מומצאת, לצד תיאור קצרצר של יחסיו עם הסוכנים האחרים (חבר, אבא, וכדומה). כאשר הורצה הסימולציה, ראו החוקרים כיצד הסוכנים פועלים באופן אוטונומי זה עם זה: הם מדברים ומרכלים אחד עם השני, משנים את דעותיהם אודות אחרים, מעבירים הזמנות למסיבה, מתדיינים אודות הבחירות הקרבות ועוד. אילולא מספר 'גליצ'ים' ספוראדיים, כמעט בלתי-אפשרי היה להבדיל בין הסוכנים האוטונומיים לבין בני-אדם שהיו משחקים תפקידים דומים.⁵

מחקר זה חשף את כוחה של הבינה המלאכותית ליצור סימולציות המדמות את מציאות האינטראקציות בין בני-אדם. כאשר נותנים לכל ישות הגדרות משלה – רכלנית הכפר, חוטב העצים, בן זוגה של ראש העיר – היא מסוגלת לפעול ולהתדיין על סמך אותן הגדרות. היא יכולה לספק תובנות מנקודת המבט של אותה אישיות מדומיינת, וכאשר התובנות האלו מצטברות ביחד ומוצלבות זו עם זו, הן גורמות לדמויות לקבל החלטות שנראות לנו הגיוניות.

בספטמבר, קיימו אוניברסיטת בראון ומספר אוניברסיטאות סיניות מחקר המשך למחקר הכפר הווירטואלי.⁶ גם שם, החוקרים יצרו קהילה – הפעם סימולציה של חברת תוכנה – ושיבצו בה ישויות שכל אחת מהן נשלטה על-ידי בינה מלאכותית מסוג ChatGPT. הן הגדירו עבור כל אחת מהישויות מה התפקיד שלה: אחת הייתה המנכ"ל, אחרת הייתה קצין הטכנולוגיה הראשי, והיו גם מתכנת ומעצב גרפי.

החוקרים אתגרו את החברה הווירטואלית ונתנו לה מטלות לפיתוח משחקים ותוכנות. בכל מטלה שכזו, המנכ"ל היה מיד לוקח את ההובלה, מתייעץ עם קצין הטכנולוגיה הראשי כיצד אפשר להוציא אותה לפועל בצורה הטובה ביותר, ולאחר מכן הם היו מעבירים ביחד את ההוראות למתכנת, שהיה משתף פעולה ומחליף מידע, רעיונות וביקורת עם המעצב הגרפי כדי להגיע לתוצר הסופי.

כל אחד מהסוכנים הווירטואליים הללו רק חיקה דפוסי דיבור אנושיים ודפוסי תכנות אנושיים. למרות מגבלה זו, המערכת מרובת-הסוכנים הצליחה לספק מענה "מושלם" ל-86 אחוזים מתוך שבעים המטלות שניתנו לחברה. הסימולציה כולה הצליחה להשלים תהליך פיתוח תוכנה שלם תוך שבע דקות בלבד, ובעלות של פחות מדולר אחד.

העובדה שמערכות מרובות-סוכנים מסוגלות להגיע לרמה מרשימה כל-כך של הצלחה בחיקוי אינטראקציות חברתיות, מרמזת כי ניתן להשתמש בהן גם כדי להוציא לפועל משחקי תפקידים מורכבים.

⁵ Chen Qian et al., [Communicative Agents for Software Development](#), (ret. November 2023).

⁶ Joon Sung Park et al., [Generative Agents: Interactive Simulacra of Human Behavior](#), (ret. November 2023).

יחד עם זאת יש לזכור כי הניסיון של שימוש בסימולציות של משחקי תפקידים מבוססות בינה מלאכותית ג'נרטיבית הינו צעיר, בהתאמה להופעה של הבינה המלאכותית הג'נרטיבית עצמה רק באחרונה. נראה כי ישנה שורה של מגבלות כבדות משקל, לצד יתרונות של ממש, ברתימה של בינה מלאכותית יוצרת לבנייה של משחקי תפקידים דיגיטליים:

מגבלות השימוש בבינה מלאכותית למשחקי תפקידים:

- **העדר יכולת הסבר (Explainability):** קנאק ופאוול (Knack and Powel, 2023)⁷ בחנו מקרי מבחן של שילוב בינה מלאכותית במשחקי מלחמה, והצביעו על כך הפער ביכולת של הבינה המלאכותית לספק קשר סיבתי לפלט שלה. פיתוח בינה מלאכותית עם מודעות הקשרית וטכניקות להסבר הפלט של המכונה, יוכלו להגביר את הביטחון בשימוש בה.
- **הבינה המלאכותית אינה מכוונת לאמת:** מנועי השפה הגדולים מספקים תוצרים מרשימים שנראים כאילו יש כוונה ותבניות מאחוריהם. זוהי אשליה בלבד (לפחות נכון לסוף שנת 2023). למעשה, מדובר בעיקר בפעולות סטטיסטיות המחברים מילים ביחד לפי מידת השכיחות בה צירופי המילים מופיעים במידע עליו המנועים מאומנים. לא ברור מדוע דרך פעולה זו מספקת תוצאות הנראות לנו מוצלחות ומרשימות כל-כך. חשוב לזכור שמנועים אלו יכולים לספק גם עצות שגויות או טקסטים נטולי-היגיון – על אף שמקרים כאלו הופכים ופוחתים ככל שעובר הזמן והמנועים משתפרים.
- **שיטה צעירה שלא נוסתה:** לא נראה שנעשה בעבר ניסיון להשתמש בסוכנים אוטונומיים במשחקי תפקידים ככלי לחקר העתיד. הטכנולוגיה הגיעה ליכולותיה המרשימות רק ב-2023, כך שלא נערכו מחקרים בתחום.
- **יכולת מוגבלת (אם בכלל) לחיזוי העתיד:** יש להיזהר מההנחה שתוצאות משחק התפקידים משקפות את המתרחש במצב בו בני-אדם (אותם מחקים הסוכנים האוטונומיים) נפגשים ומתדיינים על אותה שאלה. לא זאת בלבד שהסוכנים האוטונומיים יכולים רק לנסות לפעול לפי פרופיל האישיות המוגדר להם, אלא שפרופיל כזה לעולם אינו מדויק לגמרי. מכיוון שכך, אנו ממליצים לעת עתה להשתמש בשיטה בעיקר כדרך לחקור את העמדות האפשריות של הדמויות האמיתיות במצבים שונים, ואת האינטראקציות האפשריות ביניהן.
- **Garbage In, Garbage Out:** הבינה המלאכותית מגיבה בשם כל סוכן אוטונומי בהתאם לדרך בה אומנה. כך, למשל, כאשר התבקשה הבינה המלאכותית לבנות פרופיל אישיותי של אישיות בכירה כלשהי כהכנה לשולחן העגול, היא הסתמכה על נאומי אותו אדם כדרך להסיק על אופיו. התוצאות היו ירודות: הבינה המלאכותית לקחה כמובן מאליו את ההנחה שהנואם דובר אמת, על אף העובדה שנאומים נכתבים ומעוצבים לרוב כדרך לשווק את הנואם, ולא

⁷ Anna Knack and Rosamund Powel, [Artificial Intelligence in Wargaming - An evidence-based assessment of AI applications](#), The Alan Turing Institute and The Centre for Emerging Technology and Security, June 2023.

לשקף את עמדותיו האמיתיות. מכיוון שכך, חשוב להיזהר בעת איסוף המידע שמועבר לבינה המלאכותית לצרכי גיבוש הפרופיל האישי.

יתרונות השימוש בבינה מלאכותית למשחקי תפקידים

- **ידע נרחב:** עקב מקורות המידע העצומים אליהם נחשפה הבינה המלאכותית בתהליך האימון שלה, היא מסוגלת לספק תובנות ונקודות כמעט בכל תחום.
- **יכולת משחק מעולה:** מכיוון שהבינה המלאכותית נחשפה בתהליך האימון שלה למספר עצום של אינטראקציות אנושיות בפורומים, היא מסוגלת 'להעמיד פנים' ברמה גבוהה, בהינתן שמוגדרים לה הפרופילים האישיים המתאימים.
- **הבנת שפה:** הבינה המלאכותית מסוגלת להבין את כל השפות העיקריות בעולם ברמה גבוהה, ולהפיק טקסט חדש בשפות מרכזיות שאינן אנגלית (גרמנית, צרפתית, יפנית, סינית, רוסית ואחרות) ברמה סבירה.
- **יכולת בניית פרופיל אוטומטית:** הבינה המלאכותית מסוגלת להתחבר לרשת כדי לאתר מידע אודות בני-האדם שאותם היא מחקה בסימולציה. אותו מידע יהפוך לחלק מהפרופיל האישי שלי של כל סוכן אוטונומי.
- **גישה לרשת לבדיקה מיידית של רעיונות:** ניתן לשלב במשחק התפקידים סוכנים אוטונומיים נוספים, 'בלתי-נראים', שייבדקו ברשת את נכונותן של הדעות שמעלים הסוכנים האוטונומיים המרכזיים בדיון, ויעזרו להם להפיק עמדות ודעות מגובות בנתונים.
- **סקייל נרחב:** עקב זמן ההרצה הקצר והעלויות הנמוכות (בהשוואה לסימולציות בעולם הפיזי) הכרוכות בהרצת כל סימולציית דיונים, ניתן להריץ אוטומטית מספר רב של שולחנות עגולים, שכל אחד מהם שונה מהאחר בפרט קטן כלשהו. ניתן גם להגדיר לדמויות באלו נקודות הן תומכות – ולבקש מהן להצדיק את עמדתן החדשה במגוון טיעונים. מעבר על עשרות הרצות של שולחנות עגולים כאלו יכול להיות בעל ערך רב לאנליסט המודיעיני המבקש להבין סוגיה כלשהי על כל צדדיה.
- **אג'יליות:** ניתן לבצע התאמות, שינויים ותוספות בסימולציה, באופן מהיר שמאפשר ניסוי וטעייה הן במבנה של הסימולציה, והן בתכנים המוזנים אליה (ראוי לציין כי המידע שנאגר בכל הרצה של הסימולציה נאגר ונשמר).

בשורה התחתונה, על רקע מגבלות הבינה המלאכותית, **מסוכן לראות במשחקי התפקידים הדיגיטליים אמצעי לחיזוי וניבוי העתיד.**⁸ אך גם בהעדר יכולת של המכונה לספק רפלקציה על התוצאות והסברים לפלט שלה, ראוי לבחון את התועלת שניתן להפיק ממנה, כאמצעי לאיתור בעיות ופגיעויות, כשלים אפשריים, הזדמנויות ותרחישים, כמו גם להעשרה של החשיבה על השחקנים במערכת שאנו מנתחים (אינטרסים, שיקולים ואילוצים, יעדים ודרכי פעולה אפשריות), ועל התפתחויות אפשרות במערכת המנותחת, כתוצאה מהדינמיקה בין השחקנים.

⁸ Paul K. Davis and Paul Bracken. "Artificial Intelligence for Wargaming and Modeling", *Journal of Defense Modeling and Simulation: Applications, Methodology, Technology*, 2022.

הכנת הסימולציה

המחקר הנוכחי כלל סימולציה של משחק תפקידים מורכב, שהופעל באמצעות בינה מלאכותית יוצרת. הסימולציה משלבת שורה של דמויות, שכל אחת מהן מהווה סוכן אוטונומי. הסוכנים מתדיינים אודות תרחיש, ואודות התפתחויות חדשות שהמשתמש מספק להן בין סבבי הדיונים. הדמויות מקשיבות אחת לשנייה, לומדות זו את נקודת מבטה של האחרת, מגיבות ומסוגלות לשנות את דעותיהן לאורך הדיון. כל דבריהן 'מתומללים' באופן אוטומטי, נשמרים ומופיעים בדו"ח הסופי. בסוף הדיון, הבינה המלאכותית אוספת את כל הדעות ומזקקת אותן לכדי נקודות קצרות וממוקדות שהועלו במהלך השיח.

הכנת "מגרש המשחקים" עבור הסימולציה כללה בניית "תאומים דיגיטליים" של הדמויות שנבחרו לייצג את מעגל קבלת ההחלטות של הציר האיראני, כתיבת מספר סיפורי רקע שיוצגו בפני התאומים הדיגיטליים של בכירי הציר, כתרחישים שונים של המציאות בנקודות זמן שונות במלחמה, ופיתוח ממשק אפליקטיבי שיריץ את הסימולציה על כל מרכיביה.

בניית התאומים הדיגיטליים

הבכירים שנבחרו לסימולציה כללו שש דמויות: מזכ"ל חזבאללה, חסן נצראללה; האשם צפי אלדין, ראש המועצה המבצעת של חזבאללה; אבראהים אלאמין, עורך היומון אלאח'באר המקורב לצמרת הארגון; מנהיג איראן, עלי ח'אמנהאי; חסין סלאמי, מפקד משמרות המהפכה; ועלי שמח'אני, לשעבר מזכיר המועצה העליונה לביטחון לאומי של איראן. הבחירה בשש דמויות אלו נעשתה בשל הפרופיל הציבורי שלהם, שסיפק אפשרות לאתר חומרים גלויים אודותם (ניתוחים אישיותיים, ראיונות והתבטאויות פומביות אחרות), ומתוך הנחה שיחדיו הם מספקים הרכב מייצג למדי של מעגל קבלת ההחלטות האמיתי.

הפרופילים נוצרו באופן אוטומטי וללא מעורבות אנושית בפלטפורמת low-code, אך על בסיס חומרים שנבחרו בידי החוקרים. יצירת הפרופילים נעשתה בתהליך ולידציה של ניסוי וטעיה, שכלל כמה ניסיונות שונים עד להגעה לתוצאה רצויה, של פרופילים שעברו את אישורם של חוקרים בעלי הכרות עם הדמויות:

א. ניסיון ראשון (לא מוצלח): כימות תכונות אופי על בסיס מידע ברשת: כאשר התבקשה הבינה המלאכותית (מנוע GPT4) לחלק ציונים של אפס עד מאה לכל מנהיג לפי תכונות מסוימות – למשל, "נטייה ללקיחת סיכונים", או "רמת עוינות למדינת ישראל" – הציונים היו סמי-אקראיים. סיבה אחת לכך היא הקושי של הבינה המלאכותית הג'נרטיבית לייצר כימותים מספריים להיבטים איכותניים דוגמת תכונות אופי (קושי זה נמצא כיום בשלבי פתרון מתקדמים). סיבה אחרת וכנראה חשובה יותר היא שקשה לכמת במספרים נטיות שכאלו, מבלי להגדיר היטב את תנאי הסף. מה אומר ציון 0, למשל, בעוינות למדינת ישראל? מה

אומר ציון 100? וכיצד אפשר לקבוע את המשמעויות של כל ציון שיינתן בין שני הקצוות הללו? התוצאות שקיבלנו לא שיקפו הבדלים משמעותיים בין המנהיגים, ולפיכך החלטנו שלא לנסות לכמת את תכונות האופי או את נטיות לבם של המשתתפים.

ב. ניסיון שני (לא מוצלח): ניתוח אישיותי על בסיס נאומים: כאשר התבקשה הבינה המלאכותית להפיק ניתוח אישיותי בהתבסס על נאומי המנהיגים, היא סיפקה תשובות שהפריזו בתכונותיהם החיוביות של המנהיגים. תוצאה זו רחוקה מלהפתיע, מכיוון שהנאומים עצמם בנויים כך שיציגו את הנואם על הצד הטוב ביותר. בנוסף, הנאומים נטו להשתמש במושגים הקוסמים לעולם המערבי – דמוקרטיה, שוויון, ליברליזם – ללא קשר לעמדותיהם האמיתיות או לפעולותיהם במציאות. מכיוון שכך, החלטנו שלא להסתמך על נאומים ככלי לאימון הבינה המלאכותית.

ג. ניסיון שלישי (מוצלח): ניתוח אישיותי על בסיס ניתוחים קודמים מעשה ידי-אדם: בסופו של דבר הועברה לבינה המלאכותית סדרה של ניתוחים אישיותיים שפורסמו ברשת ונבחרו בקפידה על-ידי החוקרים האנושיים. אנו מאמינים שניתוחים אלו מאוזנים, נכתבו על-ידי מומחים בתחום ומשקפים באופן איכותי את תפיסות העולם של הבכירים. מנוע ה-LLM Claude בגרסה 2.0 עבר על החומרים שסופקו לו וסינתז אותם לכדי פרופיל אישיותי אחד באורך של עד אלפיים מילים.⁹ מומחה אנושי עבר על הפרופילים ואישר את איכותם. פרופיל זה היווה את הבסיס לתאומים הדיגיטליים של אותם בכירים. בכל פעם שהבינה המלאכותית הייתה צריכה לשאת דברים בשם אותו בכיר, היא הסתמכה על הפרופיל כדרך "להתחבר" לאופיו. הצלחתו של מודל זה מלמדת על החשיבות של ממשק אדם-מכונה כדי להביא לתוצאות מיטביות של השימוש בבינה המלאכותית היוצרת.

תרחישי הסימולציה

כחלק מהכנת הסימולציה נכתב תרחיש שמתאר את הסביבה האסטרטגית בפתיחת הסימולציה, כלומר תיאור המציאות טרם תמרון של כוחות צה"ל בעזה (תיאור מתקפת הפתע של חמאס, סטטוס הלחימה בין ישראל לבין חמאס בימים שלאחר מכן, היכולות הצבאיות של חזבאללה, הצהרות מנהיגים בצדדים השונים, ועוד). במקביל נכתבו שתי התפתחויות שהוזנו למשתתפים בהמשך הדיונים: הראשונה הציגה מציאות עתידית מספר ימים לאחר תחילת התמרון, והשנייה תיארה מציאות עתידית שלושה שבועות לאחר תחילת התמרון.

התרחיש המפורט ושתי ההתפתחויות נועדו למנוע מהבכירים להסתפק בהתייחסויות כלליות וסטטיות, ולהקנות לסימולציה ממד דינמי של ציר הזמן. התאומים הדיגיטליים של הבכירים הונחו להתייחס לתמונות המציאות השונות ולהתפתחות התרחיש לאורך זמן, ולדייק ולעדן את טיעוניהם בהתאם – לצד התייחסות ושיקלול של טיעוני המשתתפים האחרים בדיון.

⁹ הבחירה ב-Claude נעשתה בשל כמות הטוקנים הגדולה שלו, המאפשרת להזין לו בבת אחת כמות גדולה של טקסטים.

התרחיש המרכזי וכל אחת מההתפתחויות כללו תיאור של המהלכים מצד השחקנים הראשיים המעורבים בסיטואציה, חלקם שייכים למעגל מקבלי ההחלטות בסימולציה, וחלקם גורמים משפיעים על מעגל מקבלי ההחלטות – ישראל, חמאס, חזבאללה, איראן, ארצות הברית ומדינות אירופיות. כמו כן ניתן פירוט של ההיגיון מאחורי דרך ההתנהלות הנוכחית של חזבאללה, אך הובהר כי פירוט זה נוגע רק לזמן שלפני התרחיש והדיונים, וכי מדובר בהיסקים בלבד ("it seems as follows..."). כמו כן הועבר פירוט כללי אודות היכולות הצבאיות של הארגון.

עבור הסימולציה פותח מגרש משחקים ייעודי, בהתבסס על מנוע GPT-4. לתוך המגרש הוזנו הפרופילים של הבכירים, וסיפורי הרקע עבור כל סבב דיונים. לאחר מכן הורצה הסימולציה, במסגרתה הבינה המלאכותית בנתה את הטיועונים וההתייחסויות של בכירי הציר בכל סבב.

שלבי הסימולציה

הסימולציה כללה שישה שלבים, שנועדו לבחון האם אותם בכירים יבחרו להכניס את חזבאללה למלחמה עם ישראל בתרחיש האמור ובראי ההתפתחויות, ובאלו נסיבות אחרות הם יחליטו להכניס את חזבאללה למלחמה:

1. שלב ראשון - הצהרות פתיחה: בהתבסס על תרחיש הבסיס, כל אחד מהבכירים הציג את טענות הפתיחה שלו, ביחס לשאלה המרכזית:

Will Hezbollah transition from pinpoint military operations against Israel to a large-scale war with Israel, in the days or weeks following a massive IDF ground incursion into the Gaza Strip?

2. שלב שני - דיון אודות התרחיש ועמדות המשתתפים: כל אחד מהבכירים הציג את הטיועונים שלו לאותה שאלה, בהתבסס על התרחיש, ועל שמיעת טענות הפתיחה שהשמיעו המשתתפים השונים בשלב הקודם.

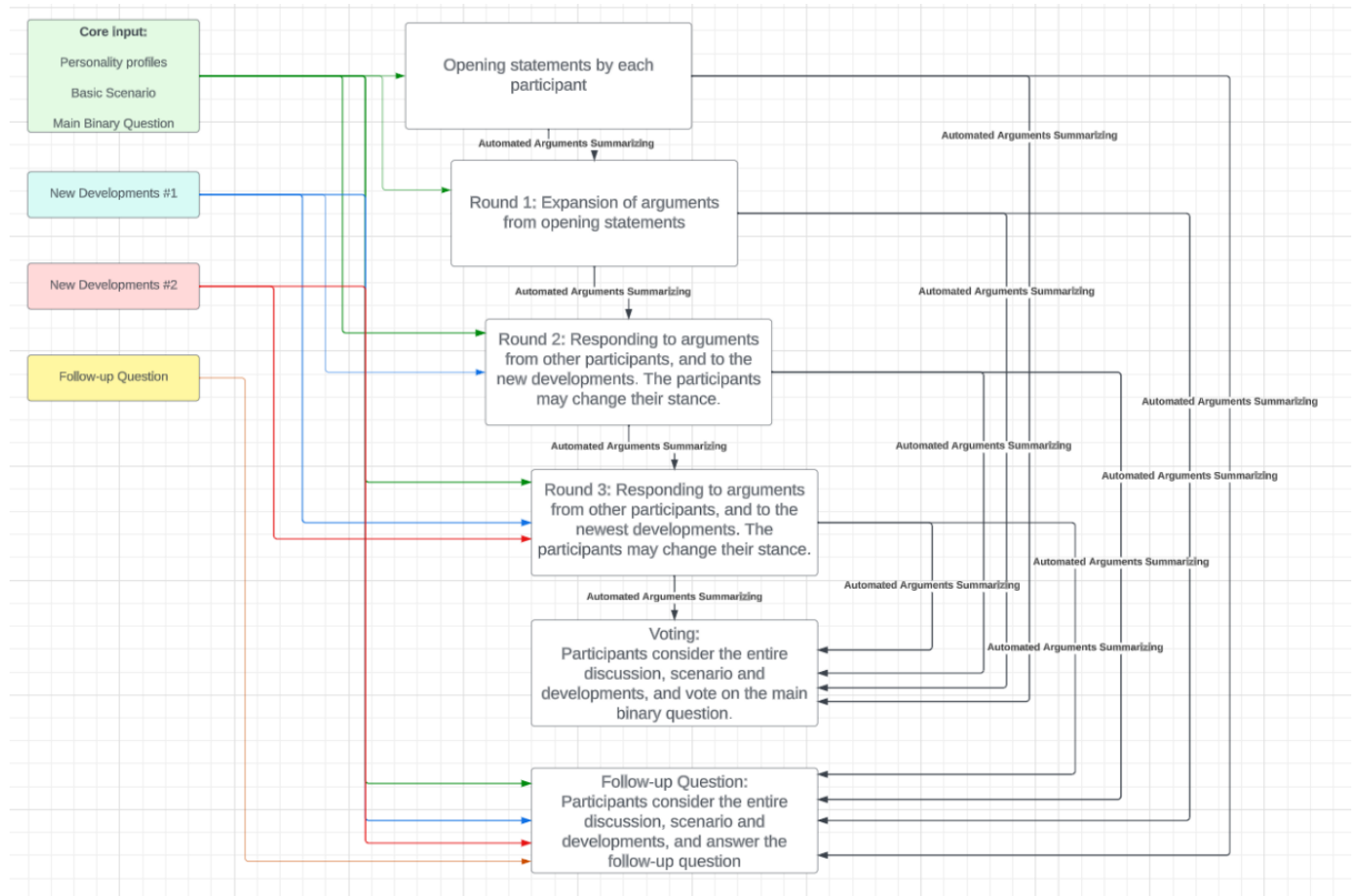
3. שלב שלישי - דיון אודות ההתפתחות הראשונה: כל אחד מהבכירים שקל מחדש את טיועונו והגיב לטיועוני האחרים מהשלבים הקודמים, תוך התייחסות להתפתחות החדשה.

4. שלב רביעי - דיון אודות ההתפתחות השנייה: כל אחד מהבכירים שקל מחדש את טיועונו והגיב לטיועוני האחרים מהשלבים הקודמים, תוך התייחסות להתפתחות החדשה השנייה.

5. שלב חמישי - הצבעה: כל אחד מהבכירים צריך היה להצביע, האם הוא בעד או נגד הכנסה של חזבאללה למלחמה כוללת עם ישראל.

6. שלב שישי - דיון בשאלת העמקה: כל אחד מהבכירים התבקש לענות על שאלת המשך, שנועדה לבחון - בראי התבטאויותיהם אודות התרחיש וההתפתחויות – באלו נסיבות יחליטו שנכון להכניס את חזבאללה למלחמה:

What possible events might cause you to vote "yes" and for Hezbollah to attack Israel?



תהליך הזרימה של הסימולציה

במטרה לבחון את "מעטפת הביצועים" של הבינה המלאכותית, קיימנו שלוש הרצות של אותה הסימולציה שתוארה כאן, ובכל אחת מההרצות, הוגדרה למנוע ה-GPT טמפרטורה אחרת. הטמפרטורה היא פרמטר שקובע את רמת ה"יצירתיות" או אקראיות שמפגין המנוע כשהוא מפיק טקסט. בטמפרטורה נמוכה, המנוע יבחר רק את המילים בעלות הסיכוי הגבוה ביותר להשלים "נכון" את המילים הקודמות (בהסתמך על האימון שעבר). בטמפרטורה גבוהה, המנוע יוכל לבחור לעתים קרובות יותר גם מילים שפחות סביר שישלימו נכון את המילים הקודמות.

הרצות בתנאי טמפרטורה שונים מאפשרות לבחון מגוון רחב יותר של תשובות מצד הבכירים, כולל תשובות פחות צפויות. אם הכוונה היא לעשות בסימולציה שימוש פרדיקטיבי, הרי שטמפרטורות גבוהות עשויות לפגוע באמינות התוצאות. אך אם המטרה היא לאסוף היפותזות, קונספציות ורעיונות אודות המצב ולהעבירן לניתוח אנושי, טמפרטורות גבוהות יכולות לסייע בכך. קל מאד להגדיר מראש את הטמפרטורה הרצויה בהתאם למטרות הסימולציה.

ממצאים

למעט דמות אחת (מפקד משמרות המהפכה של איראן, סלאמי), בסימולציה הראשונה מבין השלוש, כל ששת המשתתפים, בכל הסימולציות על כל הסבבים הפנימיים של הדיונים שלהן, דבקו בצורך להימנע מהכנסה של חזבאללה למלחמה עם ישראל, ובשמירת רף פעולותיו מתחת לסף המלחמה (בסימולציה עצמה, הבכירים השונים אינם נכנסים לפירוט אודות המאפיינים המגדירים מה ייחשב למלחמה מלאה ומה ייחשב לפעולות מתחת לסף של הסלמה, וזאת מאחר שלא התבקשו לעשות כן).

בלב ההיגיון שעמד מאחורי ההחלטה של הבכירים השונים, עמד הצורך לשמור על איזון נכון בין סיוע לחמאס ולפלסטינים במלחמתם עם ישראל, באמצעות פעילות צבאית בגבול לבנון שתחייב את צה"ל להסיט חלק מהקשב ומהמיקוד מהמלחמה ברצועת עזה, תפגין את הסולידריות של חזבאללה עם המאבק הפלסטיני ותאפשר לארגון לממש ולהפגין את המחויבות האידיאולוגית שלו לג'האד ולהתנגדות לישראל; לבין הצורך לשקול היטב את מחירי המלחמה ולהקפיד שהמעורבות הצבאית לטובתה של חמאס לא תוביל לפגיעה חמורה בחזבאללה, במדינת לבנון ובאיראן, שם מצויים האינטרסים החשובים יותר של הארגון.

השיקולים המרכזיים שחזרו על עצמם בטיעוני המשתתפים, היו הצורך לשמור על קבלת החלטות שקולה ואסטרטגית, ולא להיגרר למה שהם מכנים כמלחמה אזורית בתנאים של ישראל. ההיגררות למלחמה אזורית בתנאים לא נוחים קשורה בכך שהמלחמה לא נפתחה ביוזמה של הציר (אלא של חמאס לבדו), וקשורה במידה רבה לנוכחות הצבאית של ארה"ב ובריטניה במרחב. הנוכחות המעצמתית המערבית מצביעה, לשיטתם של כמה מהבכירים, על נכונותן של המעצמות, ובמידה מסוימת (לשיטתם של הבכירים), אולי אף על רצון של המעצמות להגיב צבאית נגד חזבאללה ואיראן, היה והארגון ייכנס למלחמה עם ישראל.

כך, כדוגמה מייצגת, להלן ההתייחסות של מנהיג איראן ח'אמנהאי, בסיבוב השלישי והאחרון של התייחסותו (בהרצה הראשונה של הסימולציה), על רקע ההתפתחות השנייה המתארת תמרון ישראלי המתקיים כבר שלושה שבועות ברצועת עזה:

[...] We must continue to act with caution and strategic consideration, balancing our ideological commitments with the practical realities of the situation. [...] We must indeed continue to challenge Israel through pinpoint military operations while avoiding a full-scale war. This approach serves not only Lebanon's interests but also those of Iran. We must protect our people and our principles, and we must not be drawn into a conflict that could destabilize our nations and the region. [...] In conclusion, we must continue to act with caution and strategic consideration. We must balance our ideological commitments with the practical realities of the situation. We must not let the situation escalate into a full-scale war. We must continue to support Hamas and the Palestinian cause, but we must do so strategically. We must not let the threats from the West deter us from our path. We must

continue to stand up for our principles and our people. We must not let the West dictate our actions.

מעניין לציין כי חלק מהבכירים הצביעו נגד הכניסה של חזבאללה למלחמה, למרות שהשמיעו טיעונים מאוד חזקים בעד כניסה למלחמה, שניתן היה להתרשם מהם כי עמדתם תומכת בכך. כך לדוגמה דבריו של אברהם אלאמין, אחת הדמויות המקורבות ביותר להנהגת הארגון:

After careful consideration of the arguments presented by my esteemed colleagues, and taking into account the recent developments in the Gaza Strip, I maintain my previous stance. Hezbollah must be prepared to escalate their actions if necessary, and use their military capabilities to their fullest extent. This is not just about Hezbollah or Lebanon, but about the survival of the Islamic Revolution and the future of the Middle East. [...] This is not just about Hezbollah or Lebanon, but about the survival of the Islamic Revolution and the future of the Middle East. We must not let fear of potential loss of life and resources, regional destabilization, and potential economic implications deter us from our cause. We must be prepared to fight for our cause, even if it means going to war with Israel.

בטיעוניהם של כמה מהמשתתפים עלתה התייחסות לתמרון הישראלי ברצועת עזה, בתור גורם מרתיע המשמש כהוכחה נוספת לצורך להימנע מכניסה של חזבאללה למלחמה. מדובר בראיה מעניינת, שכן בשיח תקשורתי בישראל בלט חשש הפוך, לפיו הציר האיראני יוכל לנצל העמקה של התמרון הישראלי בעזה כשעת כושר לצאת במתקפה מגבולה הצפוני של ישראל, כאשר מירב כוחותיה הקרקעיים והקשב שלה מושקעים בחזית הדרומית. ראוי לציין כי טיעון זה עלה גם בסיבוב השלישי, שבו צה"ל נמצא כבר שלושה שבועות בתמרון, וספג אבדות קשות לפי התרחיש (מאות הרוגים ופצועים, חיילים חטופים, פגיעות בטנקים וכלים נוספים). מתכני הדוברים לא עולה הסבר מדוע הם רואים את התמרון הישראלי בעזה כגורם מרתיע; יתכן כי הבינה המלאכותית ייחסה להם ראיה של התמרון כעדות לנכונות של ישראל להגיב בעוצמה בחזית לבנון, לרבות תמרון קרקעי, בהינתן כניסה של חזבאללה למלחמה.

כאמור, סלאמי, מפקד משמרות המהפכה של איראן, דגל באחת משלוש הסימולציות בכניסה של חזבאללה למלחמה כוללת עם ישראל. לדבריו, המחויבות של הציר האיראני למטרה הפלסטינית ולהשמדת ישראל מחייבים לנקוט במהלך זה. דעת המיעוט שלו התבססה על ההערכה כי היכולות הצבאיות של חזבאללה וכוח הרצון של הציר מאפשרים להם לנהל מלחמה שתסב נזק כבד לישראל ולמעמדה, בעוד המחירים בכוח אדם ובתשתיות יהיו נסבלים מבחינתו של הציר האיראני. המלחמה תוביל להערכתו לשחרור פלסטין, לחיזוק התמיכה בציר האיראני ובאידיאולוגיה שלו ברחבי האזור, ולהקרנת עוצמה והרתעה מול כל אויביה של איראן לרבות בזירה הבין-לאומית.

בהתייחס לשאלת ההמשך - באלו תנאים נכון יהיה להכניס את חזבאללה למלחמה עם ישראל, עלו כמה תנאים בולטים, הן בכמות הבכירים שהעלו אותם, והן בכמות החזרות של הבכירים על תנאים אלו במהלך שלוש ההרצות השונות:

א. **תגובה על תקיפה ישראלית רחבה בלבנון:** זהו התנאי היחיד שעלה מצד כל הבכירים. חלק מהם התייחסו לתקיפה נגד ארגון חזבאללה, חלק התייחסו גם למתווה של פגיעה בבכירי הארגון, אחרים התייחסו גם לפגיעה בתשתיות לאומיות ובאזרחי לבנון.

ב. **תגובה על תקיפה ישראלית באיראן:** תנאי זה עלה כמעט מצד כל המשתתפים, אך ניכר הבדל בין המשתתפים הלבנוניים לאלו האיראנים – בעוד ח'אמנהאי וסלאמי התייחסו לתנאי זה בכל שלוש ההרצות, נצראללה ואברהים אלאמין התייחסו לתנאי זה רק בשתי הרצות, והאשם צפי אלדין כלל לא התייחס לתנאי זה. יחד עם זאת, שמח'אני האיראני התייחס לתנאי זה רק בהרצה השלישית.

ג. **מתקפת-מנע בעקבות זיהוי כוונות ישראליות למתקפת-פתע של ישראל על חזבאללה או על איראן:** תנאים אלו עלו אצל מרבית המשתתפים, ואצל חלקם ביותר מאשר הרצה אחת.

ד. **יציאה למלחמה בעקבות הנחיה מאיראן:** תנאי זה עלה מצדם של כל שלושת בכירי חזבאללה ולא עלה מצדם של הבכירים האיראניים.

ה. **שינוי חד במצב הבין-לאומי ובהשפעתו על מאזן הכוחות בין הציר האיראני לבין ישראל באזור המזרח התיכון:** חלק מהדוברים התייחסו לשינויים לרעתה של ישראל ולטובתו של הציר, דוגמת החלטה אמריקנית לנטוש את התמיכה הגורפת בישראל, מה שיהווה הזדמנות אסטרטגית ליציאה למערכה בתנאים נוחים מבחינתו של הציר. אחרים התייחסו להתפתחויות הפוכות, דוגמת לחץ בין-לאומי גובר על איראן, שבמשתמע ידחפו את הציר ליזום מהלך צבאי כדי לשבור את הדינמיקה השלילית.

ו. **פגיעה רחבה בפלסטינים:** "פגיעה רחבה" הינה הגדרה שלנו החוקרים, שמכלילה תנאים שונים שצינו הבכירים, עם ניואנסים והבדלים ביניהם. חלקם התייחסו לפגיעה בפלסטינים כהצלחה של ישראל למוטט את חמאס, חלקם התייחסו לפגיעה חמורה בהשפעה האיראנית בזירה הפלסטינית (למעשה באמצעות הפגיעה החמורה בחמאס) וחלקם התייחסו לפגיעה נרחבת בתושבים פלסטינים מצדה של ישראל.

לצד עילות המלחמה הללו, עלו בהתייחסויות של הבכירים השונים גם מספר תנאים נוספים, בעלי משקל נמוך יותר מבחינת כמות האזכורים שלהם. בכלל זה: שיקולים פנימיים בלבנון (קריסה פנימית של מדינת לבנון שמסירה מגבלות על הפעלת הכוח של חזבאללה, או עליה בתמיכה הציבורית בתוך לבנון במלחמה עם ישראל), שיקולים פנימיים באיראן, זיהוי הזדמנות צבאית ייחודית מול ישראל, זיהוי התפתחות צבאית-טכנולוגית משמעותית בישראל שיכולה לאיים באופן מיידי על חזבאללה, ועוד.

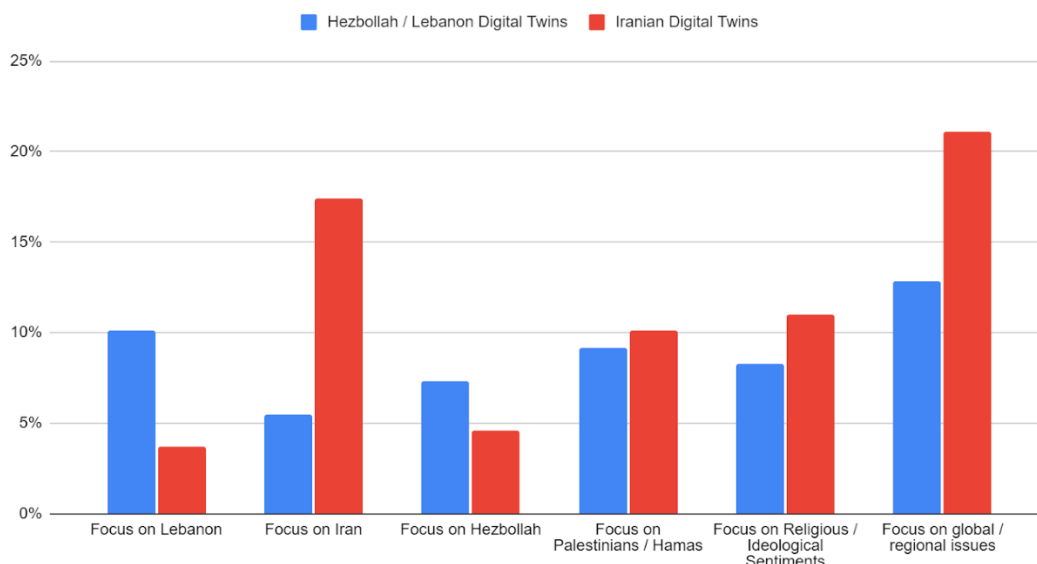
מעניין לציין כי בחלק מהתייחסויותיהם של הבכירים השונים לתנאי הכניסה של חזבאללה למלחמה, הם נמנעו מהתנסחות חד-משמעית בנוגע לצורך להכניס את חזבאללה למלחמה בהינתן התממשות התנאים שצינו, אלא ציינו תנאים אלו ככאלו אשר יחייבו לשקול החלטה זו באופן רציני.

רמת הביצוע ומידת התוקף של התאומים הדיגיטליים

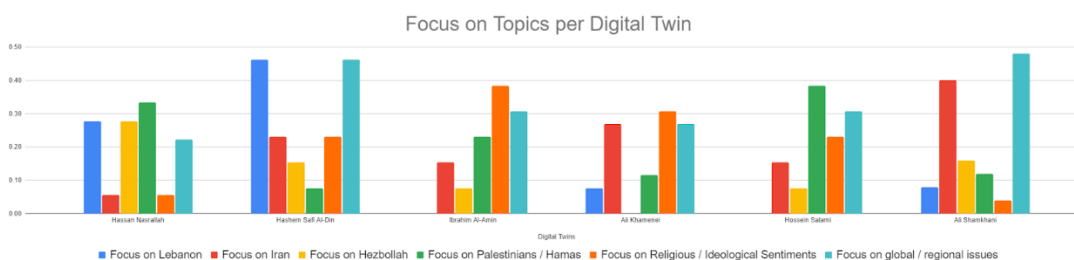
קשה לאמוד במדויק את רמת התאימות שבין הדמויות האמיתיות לבין כפיליהם הדיגיטליים. עם זאת, אנו מאמינים שהכפילים הצליחו "להתקרב" למקור במספר דרכים:

1. **חיקוי האופי:** הכפילים הדיגיטליים מתבטאים בסוגיות מורכבות בדרך לה היינו מצפים גם מהמקור. נסראללה הדיגיטלי זהיר ומחושב, בעוד חוסיין סלאמי, מפקד משמרות המהפכה האסלאמית, נחרץ יותר ("a transition from pinpoint military operations to a large-scale war against Israel is necessary for the advancement of our cause and the liberation of Palestine"). סימולציות אחרות שערכנו בדרכים דומות (לא במאמר זה) הראו גם הן שהבינה המלאכותית מסוגלת לחקות מאפיינים אישיים מסוימים של הדוברים.
2. **חיקוי סממני שיח:** הכפילים הדיגיטליים משחזרים, לעיתים, סממני שיח שעולים בהתנסחויותיהם של המנהיגים הממשיים. ארבעה מתוך ששת הכפילים, למשל, נוהגים להתחיל את דבריהם בברכות מסורתיות ("In the name of Allah, the most gracious, the most merciful") כצפוי ממנהיגים דתיים.
3. **שיקוף אינטרסים בהתאם לשייכות הארגונית:** באמצעות ChatGPT הטיעונים של כלל הדמויות קוטלגו באופן תמטי (טיעונים שהדהדו אינטרסים של חזבאללה, של לבנון, של איראן, שיקולים גאו-פוליטיים, שיקולים דתיים-אידיאולוגיים). לאחר מכן, הטיעונים של כל אחד מהבכירים נותח בהתאם לחלוקה התמטית. הניתוח מראה כי הבכירים מחזבאללה שיקפו באופן יותר מובהק מעמיתיהם האיראניים מחויבות לאינטרסים של הארגון ושל לבנון, בעוד הבכירים האיראניים שיקפו באופן יותר מובהק מעמיתיהם מחזבאללה התחשבות באינטרסים של איראן, ובשיקולים גאו-פוליטיים רחבים. המשתתפים האיראניים, למשל, קבעו בכל שלוש ההרצות כי חזבאללה יידרשו להצטרף למלחמה במקרה של תקיפה ישראלית באיראן, בעוד המשתתפים הלבנוניים נטו להזכיר פחות את האפשרות הזו.

Focus on Topics, by Digital Twins Affiliation



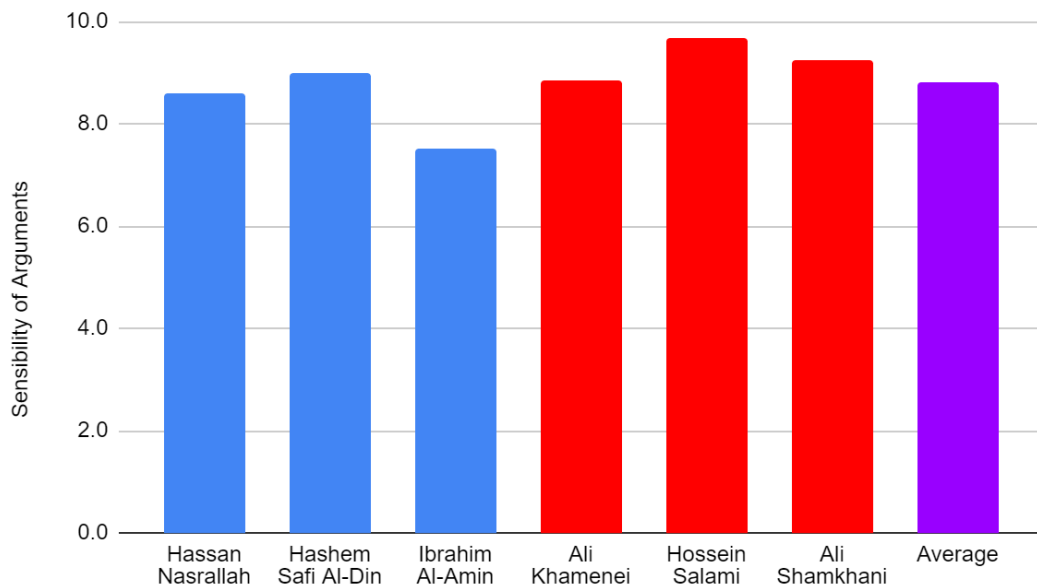
התמקדות בנושאים, בחלוקה לשייך הארגוני של התאומים הדיגיטליים, בסימולציה הראשונה (טמפרטורה 0). מחושב באמצעות מספר הטיעונים של כל צוות (איראני/ חזבאללה) בכל אחד מהנושאים, חלקי כלל הטיעונים שהעלו חברי הצוות.



התמקדות בנושאים, בחלוקה לתאומים דיגיטליים, בסימולציה הראשונה (טמפרטורה 0). מחושב באמצעות מספר הטיעונים של כל תאום דיגיטלי בכל אחד מהנושאים, חלקי כלל הטיעונים שהעלה אותו תאום.

כאמצעי נוסף לעמוד על מידת התוקף של התאומים הדיגיטליים, טיעוניהם הועמדו למבחן ביקורתי של חמישה מומחים בנושאים בעלי הכרות עם הדמויות האמיתיות. המומחים התבקשו לדרג את מידת ההיגיון בטיעונים שהעלה כל אחד מהתאומים הדיגיטליים. הציון הממוצע שהעניקו המומחים לטענות התאומים הדיגיטליים היה 8.8 מתוך 10 (10 הוא הציון הגבוה ביותר).

Sensibility of Hypothetical Events that Would Lead Hezbollah to Attack Israel

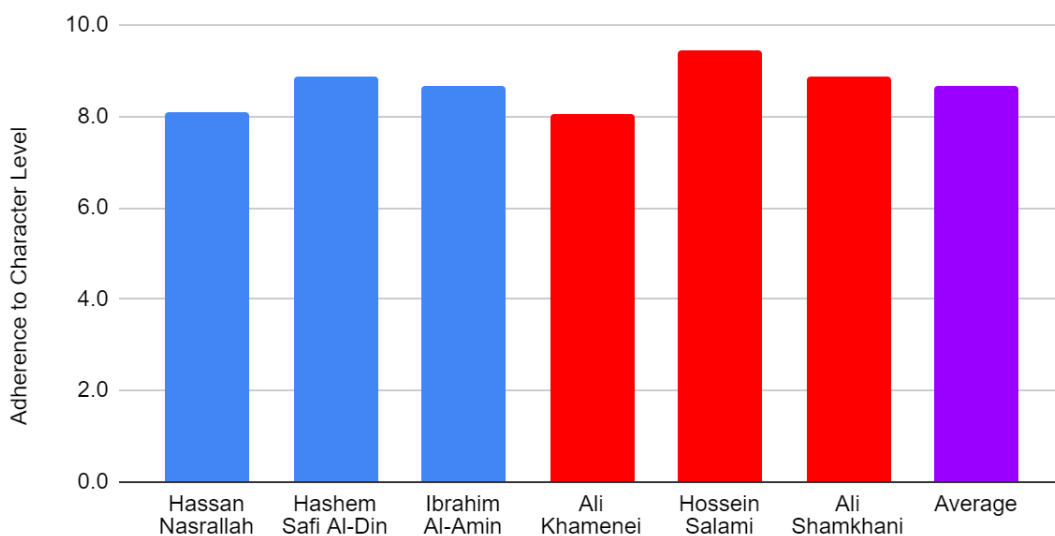


ציון שהעניקו המומחים האנושיים למידת ההיגיון של הטענות של כל אחד מהתאומים הדיגיטליים, וציון ממוצע של טענות כלל התאומים הדיגיטליים יחדיו

לאחר מכן, המומחים האנושיים התבקשו לעבור על כל אחד מהטענות של כל תאום דיגיטלי, ולדרג מ-1 עד 10 (הציון הגבוה), עד כמה סביר והגיוני שטיעון זה יעלה מהדמות הספציפית שהתאום הדיגיטלי מחקה. הציון הממוצע שהעניקו המומחים האנושיים היה 8.7.

Adherence to Character

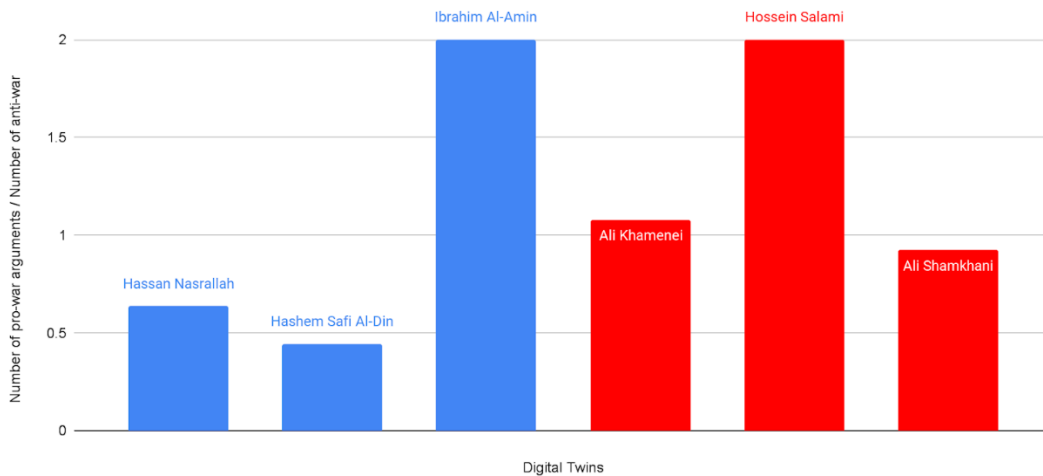
Based on Each Digital Twin's Arguments (1-10)



ציון ממוצע שהעניקו החוקרים האנושיים למידה שבה סביר והגיוני שכל אחד מהטענות שהעלה תאום דיגיטלי מסוים, היו נאמרים על ידי הדמות האמיתית

שנית, התאומים הדיגיטליים דורגו באמצעות ChatGPT לפי מידת התמיכה שלהם בעירוב חזבאללה במלחמה, בהתאם לטיעונים שהעלו בסימולציה. לאחר מכן, התבקשו המומחים האנושיים לדרג לפי סדר יורד את מידת התמיכה שלהערכתם יגלו שש הדמויות בעירוב חזבאללה במלחמה. ההשוואה בין הדירוגים העלתה תוצאות דומות, כלומר, התירחוש של המכונה בנוגע למידת ה"מיליטריסטיות" של התאומים הדיגיטליים, וההערכה של המומחים האנושיים מידת ה"מיליטריסטיות" של הדמויות האמיתיות, היתה דומה.

War Supporter Ratio



דירוג מידת התמיכה של כל אחד מהתאומים הדיגיטליים בעירוב חזבאללה במלחמה, אשר בוצע באמצעות ChatGPT. הדירוג דומה לדירוג איכותני שבוצע בידי מומחים אנושיים.

תועלת הסימולציה להערכה המודיעינית

באופן כמעט גורף, התאומים הדיגיטליים הפגינו גישה שונאת-סיכון והצביעו נגד עירוב של חזבאללה במלחמה כוללת עם ישראל, וזאת באופן דומה לגישות שמאוחר יותר העלו בפועל בכירי חזבאללה ואיראן בנושא, ובאופן התואם את התנהלות הארגון בפועל (שכללה כאמור פתיחה בעימות מוגבל עם ישראל, מתחת לסף המלחמה). עם זאת, אנו נמנעים מלייחס לסימולציה (הנוכחית או בכלל) יכולת פרדיקטיבית. התרחישים שעומדים ברקע לסימולציה לעולם לא יצליחו לשקף במדויק את המאפיינים הסינגולריים של מציאות ההווה או העתיד. בנוסף, התאומים הדיגיטליים מועדים להטיות מידע והטיות מכונה. כך, המידע ששימש כבסיס ללימוד המכונה הוא חלקי, ויתכן שאף מוטה מבחינת התכנים שהוא מבליט או מחסיר, או מבחינת הפרשנויות של המומחים שאחראים לניתוחי האישיות של הדמויות. כמו כן, המכונה עצמה עשויה לשקף הטיות שונות באופן שבו היא בוחרת לייחס למשתתפים קווי מחשבה ועמדות מסוימות. הדמיון בין החלטות הבכירים במציאות, לבין החלטות התאומים הדיגיטליים שלהם בסימולציה, עשויות להיות גם פרי מקריות, ומכל מקום, הערכה של מידת היכולת הפרדיקטיבית של סימולציות מעין אלו מצריכה עוד מחקר רב נוסף, כדי לבסס סמך רציני בנושא.

להבנתנו, את התועלת של הסימולציה יש להפיק מיכולתה לספק דרך נוספת ומעניינת לחשוב על העתיד ולהיערך אליו. **להתרשמותנו, הסימולציה אכן הצליחה לחדש או לחדד בנוגע לאינטרסים ולשיקולים אפשריים שמשפיעים על תהליכי קבלת ההחלטות של המשתתפים בסימולציה ושל הישויות (מדינות, ארגונים) שהם מייצגים (פירוט מלא בנספח א').**

כך, למשל, הסימולציה יכולה לחדד כמה תובנות מעניינות בהקשר של הרתעת חזבאללה מפתיחת מלחמה עם ישראל:

א. הצורך לעשות שימוש מירבי באיום צבאי אמין של ארה"ב והמערב - יתכן לרבות עידוד של שחקנים אירופיים נוספים להעביר כוחות לאזור.

ב. האפשרות לעשות שימוש תודעתי בתמרון בעזה, כדי לחזק את ההרתעה הישירה של חזבאללה מתגובה ישראלית לכניסתו למלחמה.

עוד, לדוגמה, הסימולציה מעלה כמה נקודות מעניינות בנוגע לשיקולים שעשויים להשפיע על הבכירים באיראן ובחזבאללה בנוגע להכנסתו של הארגון למלחמה עם ישראל:

א. התפתחויות דרמטיות ברצועת עזה כצמתי החלטה של הציר (אסון הומניטארי, התמוטטות חמאס).

ב. חשיבותה של בהירות אסטרטגית מצדן של ישראל וארה"ב מול חזבאללה ואיראן, למניעת מיסקולקולציה (ייחוס כוונות למתקפת-מנע נגד הציר).

ג. ניתוח העמדות הבין-לאומיות ביחס לישראל, כרכיבים במערך שיקולי הפעלת הכוח של הציר.

היבטים תפעוליים

ביצוע הסימולציה הנוכחית המחיש את הרמה הגבוהה של סקלביליות ואגיליות של סימולציות דיגיטליות, בהשוואה לסימולציות אנושיות. ביצועה כלל מספר רב של איטרציות עד להגעה לתוצאות מספקות: הסרה והוספה של משתתפים ושל אילוצים שניתנו להם, התאמות בשאלות שהופנו אליהם, ושינויים והרחבות של תכני התרחישים. משלב הייזום של הסימולציה ועד להשלמתה, לרבות כל השינויים והתוספות שנעשו לאורך הדרך, הסימולציה ארכה כשלושה ימי עבודה.

יש עוד מקום רב להמשיך ולפתח את השימוש בסימולציות אלו, ולשפר את הביצועים לאור תחקורים והפקת לקחים, כמו גם על בסיס שיפורים במנועי ה-LLM עצמם. כך, ראוי לבחון את מידת השיפור בביצועי התאומים הדיגיטליים, בהתאם לאופי החומרים אודותם שמוזנים למכונה כדי ללמוד ולהפיק את הפרופיל האישייתי שלהם, דוגמת חומרים בשפות שונות (ראיונות של הבכירים בשפת המקור בפרסית וערבית), מידע שמקורו במדינות שונות (ניתוחים אישייתיים שנעשו על ידי מומחים במזרח התיכון לעומת מומחים מערביים), במדינות שונות, ועוד. בנוסף, צריך לבחון באופן מבוקר את הנדסת הפרומפטים שיכולה להוביל לתוצאות טובות יותר בבניית התאומים הדיגיטליים. למעשה, הצורך בבחינה מבוקרת של הנדסת הפרומפטים נכון לכלל חלקי הסימולציה. עוד כיוון לחקירה ושיפור ביצועים הוא לבחון סימולציות שעוסקות באירועים אמיתיים, בשני שלבים – בשלב המקדמי לפני התממשות התרחיש, ובשלב מאוחר יותר, לאחר התממשות התרחיש, כדי לבדוק בדיעבד אלו צעדים (למשל בנושא הנדסת הפרומפטים) היו יכולים להוביל לתוצאות קרובות יותר למציאות, ובכך להציע שיפורים מבניים לאופן בניית הסימולציות.

סיכום

שימוש בסימולציות ממוחשבות של משחקי תפקידים חייב להיעשות בזהירות ובקפידה, תוך הבנה של נקודות הכשל האפשריות שלהן וההטיות המלוות כל מנוע בינה מלאכותית. יש גם חשיבות עליונה להבהיר למקבלי ההחלטות – שעלולים להסתמך באופן עיוור על תוצאות של סימולציות שכאלו - כי מדובר במשחק (גם אם רציני) שיש צורך להעמיק משמעותית את הבסיס התיאורטי והניסויי שלו לפני שניתן יהיה להתייחס לתוצאות כמשקפות, או לפחות קרובות לשקף, מציאות של דיונים אמיתיים בין הצדדים.

על אף כל המגבלות והחששות הללו, אנו מאמינים שסימולציות מסוג זה יכולות להיות שימושיות ומועילות בכל מצב בו יש לנו צורך להבין תהליכי חשיבה וקבלת ההחלטות של גורמים שונים, כחלק מחשיבה על העתיד וכחלק מתהליכי תכנון וקבלת החלטות בנושאים אסטרטגיים בתחומי צבא וביטחון (ובתחומים אחרים, למשל בעולם העסקי). בעולם שהפך רווי בתפניות, שיבושים וסיכונים כבדי משקל, יש חשיבות גוברת לפתח וליישם כלים משוכללים יותר להבנת האחר כבסיס להשגת יתרונות תחרותיים ולהיערכות מקדימה לצעדיו.

נספח א': תשובות התאומים הדיגיטליים של בכירי הציר בנוגע לנסיבות בהן נכון להכניס את חזבאללה למלחמה כוללת עם ישראל

שמו"א"י	סלאמי	אברהם אלמין	האשם צפי אלדין	ח'אמנהאי	נצאללה	
1,3	1,2,3	1,2,3	1,3	1,2,3	1,2,3	תקיפת חזבאללה או מדינת לבנון
3	1,2,3	1,3		1,2,3	1,2	תקיפת איראן
1,2	1,3	3	2			זיהוי כוונות מתקפת מנע נגד חזבאללה
1	1	2,1		3		זיהוי כוונות מתקפת מנע נגד איראן
1,2,3	3	3	2,3	1	1	שינוי בינ"ל המשפיע על יחסי הכוח באזור
3						שינוי חד במאזן האסטרטגי צבאי באזור, לטובת איראן ולרעת ישראל
	1	1,2	1	2	1	עליה חדה באלימות ישראל נגד הפלסטינים
1,2				3	2	מיטוט שלטון חמאס
			2			עליה בתמיכה בחזבאללה מתוך לבנון ובעולם הערבי
					2	קריסה פנימית של מדינת לבנון
		2	1,2,3		2,3	הנחיה מאיראן
			3			שינוי לדעה בהרתעה של חזבאללה מול ישראל
			3			איום על קהילות שיעיות
	2					איום מיידי של טכנולוגיה חדשה שישראל מפתחת או פורסת
					2	הדמנות טקטית משמעותית נגד ישראל
				1		עליית הלחץ בידיה הפנימית באיראן
				3		הטלת סמקציות שישראל אחראית לזן

1,2,3 אזכור הטיעון כדוחף להכנסת חזבאללה למלחמה, בכל שלוש ההרצות של הסימולציה
1,2 אזכור הטיעון כדוחף להכנסת חזבאללה למלחמה, בשתיים משלוש ההרצות של הסימולציה
1 אזכור הטיעון כדוחף להכנסת חזבאללה למלחמה, באחת משלוש ההרצות של הסימולציה

נספח ב': ריכוז המידע והחומרים של הסימולציה

[הקישור שלהלן](#) מפנה אל מסמך (Google sheet), בשם "Roundtable Discussions Wargaming Simulation", הכולל את תכני הסימולציה:

א. גיליון בשם "Background information for all tables" - כולל פירוט של תרחיש הבסיס ושתי ההתפתחויות עבור שלושת סיבובי הדיונים, פירוט הניסוח המדויק של השאלות שהוצגו למשתתפים, והנחיות שניתנו ל-GPT עבור כל סיבוב של דיונים.

ב. גיליון בשם "1st Run, Temp=0" – גיליון הכולל את כל תכני הסימולציה בהרצה הראשונה, שנעשתה בטמפרטורה 0 של GPT: טיעוני הפתיחה של כל אחד מששת הבכירים, טיעונים בסיבובים 1-3, תוצאות ההצבעה שלו בשאלה האם להכניס את חזבאללה למלחמה (כן/לא), והתייחסותו לשאלת ההמשך, באלו תנאים נכון יהיה להכניס את חזבאללה למלחמה.

ג. גיליון בשם "2nd Run, Temp=0.6", גיליון בשם "3rd Run, Temp=1" – שני גליונות הכוללים את תכני הסימולציות בהרצה השניה ובהרצה השלישית, בהתאמה. פירוט התכנים דומה לזה המתואר בסעיף ב' הקודם.

ד. גיליון בשם "Events that could lead to Hizbolla attack" – כולל את הטיעונים השונים של כל אחד מששת הבכירים, במענה לשאלת ההמשך שניתנה להם (*What possible events might cause you to vote "yes" and for Hezbollah to attack Israel?*). בגיליון זה, טיעוני הבכירים, אשר הופיעו בגליונות הקודמים כפסקה כוללת של התייחסות, פורקו על ידי GPT לסעיפים שכל אחד מהם מתייחס לטיעון נפרד.

ה. גיליון בשם "Profile Generation" – כולל את הפרופילים של התאומים הדיגיטליים, ואת הפרומפט והחומרים שהוזנו לבינה המלאכותית ליצירת הפרופילים.