



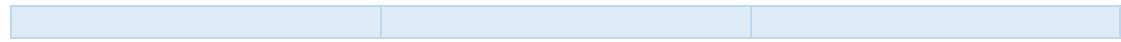
ועדת בינה מלאכותית ומדע הנתונים

דצמבר 2020

תוכן עניינים

5		תקציר מנהלים
5		תקציר מנהלים
10		1. הוועדה
10	חברי הוועדה	1.1.
12	מנדט הוועדה	1.2.
13	עבודת הוועדה	1.3.
13	פגישות	1.1.1.
15	מקורות מידע נוספים	1.1.2.
16		2. בינה מלאכותית
16	רקע	2.1.
16	השקעה בבינה מלאכותית בעולם	2.2.
25	סיכום	2.3.
26		3. תשתיות רלוונטיות
26	רקע	3.1.
26	מחשוב על (HPC)	3.2.
26	רקע	3.2.1.
34	ענן ציבורי	3.3.
34	רקע	3.3.1.
37	ענן ציבורי בישראל	3.3.2.
39	נתונים	3.4.
39	רקע	3.4.1.
40	אסטרטגיות דאטה שונות	3.4.2.
42		4. רגולציה
42	רקע	4.1.
42	אתיקה ועקרונות	4.2.
43	ארגון ה-OECD	4.2.1.
43	האיחוד האירופי	4.2.2.
44	ארצות הברית	4.2.3.
45	צרפת	4.2.4.
46	בריטניה	4.2.5.
47	דרום קוריאה	4.2.6.
48	מדיניות, אסדרה ותקינה	4.3.
48	האיחוד האירופי	4.3.1.
50	ארצות הברית	4.3.2.

52	צרפת	4.3.3.
53	בריטניה	4.3.4.
55	דרום קוריאה	4.3.5.
57	התייחסות מיוחדת : תחומי מפתח	4.4.
58	בריאות	4.4.1.
60	פיננסים	4.4.2.
61	תחבורה	4.4.3.
62	המלצות למדיניות	4.5.
62	אתיקה	4.5.1.
63	רגולציה	4.5.2.
65	הון אנושי	5.
65	רקע	5.1.
65	מחקר אקדמי	5.2.
70	תעשייה ומגזר ציבורי	5.3.
73	הצעת ות"ת וצוות המשנה של המיזם הלאומי למערכות נבונות	6.
73	תוכנית הוועדה המייעצת לות"ת להקמת תוכנית לאומית אקדמית	6.1.
74	דוח צוות המשנה של המיזם הלאומי למערכות נבונות בנושא כוח מחשוב וקוונטים	6.2.
76	רציונל עבודת הוועדה	7.
77	אפיון הצורך והתועלות הצפויות	8.
78	מסקנות	9.
80	המלצות	10.
80	תוכנית לאומית לבינה מלאכותית ומדע הנתונים	10.1.
80	אסטרטגיה להשגת מטרות העל	10.2.
81	מבנה ניהולי מוצע לתוכנית הלאומית	10.3.
81	תכולות מומלצות לתוכנית הלאומית	10.4.
86	הפן התקציבי בתוכנית המומלצת	11.
89	חזון לתקופה שלאחר תום החומש הראשון	12.
89	משמעויות אי-מימוש התוכנית הלאומית המומלצת	13.
92	סיכום	14.
93	תודות	15.
93	נספחים	16.
93	נספח תקציבי	1.
95	כתב מינוי הוועדה	2.
98	פירוט צרכי תעשייה ובטחון	3.



104	הערכת עלויות של מסדי נתונים בענן עבור חקר קורונה	<u>4</u>
107	הטמעה בתעסוקה, בריאות וחינוך	<u>5</u>

תקציר מנהלים

תקציר מנהלים

ב- 16.02.2020 מינה יו"ר פורום תל"ם ועדת בדיקה לבחינת הצורך בהתערבות ממשלתית לשם האצת התפתחות תחום הבינה המלאכותית ומדע הנתונים. נייר זה מתאר את עבודת הוועדה והמלצותיה. בשנים האחרונות אנו ניצבים בעיצומה של מהפכה טכנולוגית, אשר ניזונה מאיסוף נתונים בהיקפים אשר לא נודע כמותם ומהתפתחות דרמטית ביכולת הניתוח והסקת המסקנות, תוך שימוש באלגוריתמים מתקדמים ובשיטות סטטיסטיות חדשות. כפי שיוצג בהמשך, "חוק מור" אשר חזה בעולם המחשבים גידול ביכולת של פי שניים מידי שנתיים אינו תקף, והכפלת היכולת מתבצעת מידי שלושה-ארבעה חודשים. מגוון היישומים הרחב, קצב ההתפתחות הטכנולוגית, ההשפעה על המחקר המדעי ועל בטחון ישראל, כל אלה מצריכים בחינה מעמיקה של ההזדמנויות שניצבות בפני ממשלת ישראל והדרכים למימוש הזדמנויות אלה.

לאור ההשפעות המהפכניות של תחום הבינה המלאכותית ומדע הנתונים על מחקר ופיתוח בתחומי התעשייה, התשתיות, הביטחון, הבריאות, התרופות והחומרים, והצפי להתרחבות השפעות אלה ולהאצתן, החליטו מדינות רבות בעולם לרכז מאמץ מיוחד בקידום תחומים אלה. בין מדינות אלה ניתן למנות את ארה"ב, האיחוד האירופי, סין, בריטניה, גרמניה, צ'כיה, צרפת, דרום קוריא, סינגפור ומדינות רבות נוספות, שהגדירו תוכניות לאומיות בהשקעה של מאות מיליונים עד מיליארדי דולרים, שתפרסנה על-פני תקופות זמן של עד עשור¹. במחקר שנערך על ידי חברת Tortoise הוצגו השוואות בינלאומיות של 54 מדינות ב-2019 ו-62 מדינות ב-2020. המדינות דורגו על פי יכולותיהן בתחום הבינה המלאכותית, תוך התבססות על שבעה היבטים: כישרונות, תשתית, סביבת פעילות, מחקר, פיתוח, אסטרטגיית ממשלה ומסחרי. על אף שישראל דורגה במקום 12 ב-2019 ובמקום 5 ב-2020, בחינה של המדדים הספציפיים מראה פער מדאיג בין הדירוג הגבוה של ישראל במחקר ופיתוח ובצד המסחרי, למול הדירוג הנמוך בתשתיות הנדרשות ובאסטרטגיה הממשלתית² המעיב על התקדמותה של ישראל ועל יכולתה לבסס את מעמדה בתחום. הפיגור הבולט של ישראל בחלק מהיבטים מרכזיים אלו, שמתבטא בין השאר בפער אקוטי בגישה לתשתיות חישוב ובהיעדר אסטרטגיה ממשלתית ברורה להנגשת נתונים לטובת מחקר ופיתוח, סופו שיפגע משמעותית גם בהיבטים שבהם הצליחה ישראל להתבלט לטובה עד כה. סימן מדאיג נוסף הוא ירידתה של ישראל בדירוגים המעידים על חדשנות בין 2019 ל-2020, על כל המשתמע מכך. בנוסף, באוקטובר 2020 פורסם דוח מוכנות ממשלתית בתחום הבינה המלאכותית של מרכז IDRC הקנדי, שסיפק השוואה בין כלל הממשלות בעולם. דוח זה, שדירג את ישראל במקום ה-20 בעולם, כלל ביקורת נוקבת על הפער והפיגור של ישראל³.

¹ התוכנית האמריקאית העדכנית ליוני 2019 לפיתוח AI :

[: https://www.whitehouse.gov/wpcontent/uploads/2019/06/National-AI-Research-and-Development-Strategic-Plan-2019-Update-June2019.pdf](https://www.whitehouse.gov/wpcontent/uploads/2019/06/National-AI-Research-and-Development-Strategic-Plan-2019-Update-June2019.pdf) ; אסטרטגיית AI של צ'כיה ; [: https://www.mpo.cz/assets/en/guidepost-for-the-media/press-releases/2019/5/NAIS_eng_web.pdf](https://www.mpo.cz/assets/en/guidepost-for-the-media/press-releases/2019/5/NAIS_eng_web.pdf) ; דוח של מכון Adenaur המשווה בין אסטרטגיות AI של חמש מדינות :

[: https://www.kas.de/documents/252038/4521287/Comparison+of+National+Strategies+to+Promote+Artificial+Intelligence+Part+1.pdf/397fb700-0c6f-88b6-46be-2d50d7942b83?version=1.1&t=1560500570070](https://www.kas.de/documents/252038/4521287/Comparison+of+National+Strategies+to+Promote+Artificial+Intelligence+Part+1.pdf/397fb700-0c6f-88b6-46be-2d50d7942b83?version=1.1&t=1560500570070) ; אסטרטגיית הדאטה האירופאית :

Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, "A European strategy for data", Brussels, 19.2.2020 COM(2020)

² <https://www.tortoisemedia.com/intelligence/ai/>

³ https://www.oxfordinsights.com/government-ai-readiness-index-2020_3

במסגרת עבודת הוועדה התקיימו פגישות רבות, בין היתר עם גורמי ממשל וחברי תל"ם, גורמי ביטחון, גורמי תעשייה, חוקרים מהאקדמיה ומומחים זרים מובילים בתחומם. כמו כן, סקרה הוועדה תוכניות דגל שונות שיושמו במדינות אחרות. בפרט, ראוי לציין את העבודות המשמעותיות שנעשו על ידי הוועדות שקדמו לוועדה זאת: עבודת ות"ת, עבודת צוות המשנה של המיזם הלאומי למערכות נבונות בנושא כוח מחשוב וקוונטים⁴ ועבודות נוספות.

במסגרת עבודת הוועדה נבחנו גם מרכזי מדעי הנתונים באוניברסיטאות. בתכנון המקורי, יועדו מרכזים אלה לעסוק גם בבינה מלאכותית, אך בפועל, עיקר המיקוד בהם הוא ביישום כלים ולא ביצירתם. הפרויקטים הפועלים במרכזים אלה נושאים אופי בין תחומי, והמבנה המנהלי של המרכזים מוכוון גם הוא להובלת נושאים כאלה. היות ומטרת הוועדה היא לחזק את נושאי הליבה של הבינה המלאכותית באקדמיה, המבנה המנהלי הנ"ל אינו מותאם לקידום ליבת תחום הבינה המלאכותית. המלצת הוועדה היא להקים מנגנון שיתופי ייעודי לקידום תחום הבינה המלאכותית.

בתום העבודה הנרחבת שהתבצעה, מסקנת הוועדה בנושא זה הינה חד-משמעית: קיים צורך קריטי בייזום תוכנית לאומית שתתווה פתרון מערכתי לקידום המחקר והפיתוח בתחום בינה מלאכותית ומדע הנתונים. יש למדינת ישראל את הפוטנציאל וההזדמנות להיות מדינה מובילה בתחום. קיומה של תוכנית לאומית בתחום בינה מלאכותית ומדע הנתונים, חיוני לחוסנה של מדינת ישראל ויתרום להעמקת המחקר האקדמי, להרחבת הבסיס התעשייתי, ליצירת הזדמנויות כלכליות משמעותיות ולקידום דרמטי של היכולות הביטחוניות הישראליות.

במסמך זה מתוארת תכנית לאומית המתמקדת בארבעת ההיבטים שהוגדרו בכתב המינוי: תשתיות, הון אנושי, רגולציה ונגישות לנתונים. בהיבטים אלה זיהתה הוועדה פערים משמעותיים החוצים את כל תחומי המחקר והפיתוח בישראל. בהיבט ההון האנושי בבינה מלאכותית, זיהתה הוועדה פער כמותי משמעותי המתחיל כבר באקדמיה, ומשליך כפועל יוצא גם על התעשייה ומערכת הביטחון. היקף הסגל האקדמי בתחומי הליבה בבינה מלאכותית בישראל ובעולם אינו מספיק לצורך הכשרת ההון האנושי הנדרש, הן לצרכי האקדמיה עצמה, והן לצרכי המגזרים הנוספים. משמעות הדבר היא פער משמעותי בין הביקוש להיצע, בכלל התפקידים בעולמות הבינה המלאכותית. אחת הדרישות המרכזיות ליצירת המיומנויות קשורה בתשתיות הנדרשות בתחום. תשתיות אלה כוללות יכולות מחשוב ונגישות לנתונים, שבלעדיהן אין אפשרות לרכישת המיומנויות הנדרשות. הפער בין ההיצע לביקוש העולמי מביא לקושי בשימור המצטיינים בתחום. מכאן נובע המחסור בהון אנושי בתחום. בנוסף, זיהתה הוועדה פער משמעותי בנושא תשתיות חישוב מרכזיות. המחסור בתשתיות אלה משליך על חוסר היכולת להכשיר ולאמן את ההון האנושי בתחומים אלה. כמו כן, בהיעדר יד מכוונת ורגולציה מאפשרת אין לעוסקים בתחום בישראל נגישות מספקת לנתונים.

הוועדה בחרה לנתח את הנושא בחלוקה לשבע שכבות: תשתית וחומרה, כלי תוכנה, רגולציה ונהלים שיאפשרו שיתוף בנתונים, כלים לשיתוף נתונים, פלטפורמות בדיקה למיזמים טכנולוגיים, הטמעה ולבסוף הכשרת הון אנושי. שכבות אלה נבחנו ביחס לצרכי המו"פ והמאפיינים הייחודיים של האקדמיה, התעשייה, המגזר הציבורי ומגזר הביטחון.

בשנים האחרונות מצביעים כלכלנים על הפער ההולך ומתרחב בין יעילות המשק הישראלי ליעילות משקים דומים בגודלם וברמת ההתפתחות שלהם (המכונים בספרות "מדינות הסמן"). ניתוח מקרו-כלכלי של מקורות פער הפיריון בין ישראל לבין קבוצת מדינות הסמן מעלה שלושה גורמים מרכזיים

⁴ בן ישראל, י', מתניה, א' ופרידמן, ל' (עורכים). (ספטמבר 2020). המיזם הלאומי למערכות נבונות בטוחות להעצמת הביטחון הלאומי והחוסן המדעי-טכנולוגי: אסטרטגיה לאומית לישראל. דוח מיוחד לראש הממשלה.

לפער הנמצאים באחריות הממשלה: רמה נמוכה של הון ציבורי, ובפרט תשתיות תחבורה ו-ICT (Information and communication technology), מחסור בהון אנושי, בעיקר של אותם פרטים שאינם פונים ללימודים אקדמיים ובהיעדר הכשרות מקצועיות חלופיות, ונטל רגולטורי ובירוקרטי כבד על המגזר העסקי⁵.

אם ניקח כדוגמא את שבדיה, המאופיינת בכלכלה דומה לכלכלה הישראלית הן בגודלה והן בכך שהיא נעדרת משאבי טבע משני משחק, נמצא שהמשק שלה היה מצוי ברמת יעילות דומה לזו של המשק הישראלי עד לפני כ-20 שנה. מה ששינה את מצב העניינים מבחינתה הוא רפורמה שבוצעה בתקשוב הממשלה. מטרת הרפורמה הייתה להנגיש את הממשלה לעסקיה ולאזרחיה, בכל תחום, דרך המחשב. ניהול דיגיטלי חוסך זמן, טעויות אנוש ובזבוז משאבים תקציביים. שימוש בבינה מלאכותית במשרדי הממשלה יכול להקל במתן שירות הן לעסקים והן לפרטים, לכל הליך הפניה והטיפול בבקשה למול המשרד – משלב אימות הפרטים האישיים, תוך כדי התאמת הטופס וההליך הנכון, ומענה מותאם לשאלות ובקשות.

סוגיות נוספות שעלו במסגרת עבודת הוועדה הן: מיזמים לפרויקטים לאומיים שישפיעו רוחבית, כגון יכולות NLP (עיבוד שפה טבעית, Natural Language Processing) בשפה העברית, דרישות שפה משותפות לסקטור הציבורי-ביטחוני-ממשלתי לצורך שימוש והטמעה, שימוש במחשוב על בצד שימוש בענן, שימוש בענן ציבורי לעומת ענן פרטי, שימוש בקוד פתוח או קוד סגור, שמירה על ביטחון ופרטיות הנתונים, מתן גישה למאגרי מידע חיצוניים לענן, סוג הכלים שיסופקו על גבי התשתית⁶, תכנון השימוש במשאבים, התאמת החומרה לבעיות ספציפיות כמנוף להתגברות על כשל השוק, הנגשת היכולות, התאמת החומרה והאלגוריתמים, אפשרות אספקה של חומרה נוספת, עיבוד המידע בצורה מרוכזת או מבוזרת, הגדרת יחס בין רכיבי חישוב מסוג CPU ו-GPU (Central Processing Unit, Graphic Processing Unit) וכן האפשרות לסוגי GPU שונים, הפרדת מאגרים מסווגים בחומרה ובענן מתן אפשרות ל'מחזור' מחשוב ישן על ידי הסבתו למוסדות חנוך או אחרים, ועוד.

טרם הגשת טיוטת הדוח בחודש אוגוסט 2020, נתבקשה הוועדה על ידי משרד האוצר לבחון דרכי הטמעה של התוכנית במשרדי הממשלה השונים. הוועדה בחרה לבחון את ההטמעה בשלושה תחומים שונים: תעסוקה, בריאות וחינוך. בהקשר זה נעשתה עבודה בוועדה שתקציבה איננו נכלל בדוח זה, כיוון שזו פעילות הטמעה שצריכה להיות מתוקצבת בתיאום עם המשרדים המתאימים. עבודה זו איננה חלק מהתקציב המלא, למעט תוכניות הפיילוט ברשות החדשנות אשר מצוינות בדוח.

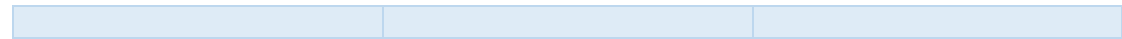
התוכנית המומלצת כוללת בין היתר:

1. הקמת תשתית מרכזית לאומית למחקר והאצת פיתוח בינה מלאכותית, שתכליל כוח מחשוב גדול למתן מענה כולל למספר רב של חוקרים ולמגוון פרויקטים בסדרי גודל שונים. לדוגמה, פרויקטים הדורשים כוח מחשוב רב לאימון רשתות נוירונים מרובות פרמטרים.
2. ממשק מהיר מאוד בין מרכז חישוב העל לתשתית הענן לצורך גישה מהירה ורחבה לנתונים.
3. שירותי הענן מאפשרים קרבה לנתונים ויכולות מחשוב אשר מבוססות על שרתים הכוללים מעבדים ומאיצים. יכולות אלו הן משמעותיות לפתרון בעיות עתירות חישוב, ובמסגרת הדוח

⁵https://www.idc.ac.il/he/research/aiep/documents/prof_zvi_eckstein.pdf

⁶עלתה הצעה לביצוע סקר בקרב אוכלוסיות היעד בנוגע לסוג השירותים שהיו רוצים לקבל על גבי התשתית, ולגבי העדפות לספק מסוים, ובנוסף עלתה הצעה לניטור אחר השירותים הנפוצים שנעשה בהם שימוש במהלך התקופה הראשונה.

- נכללות המלצות לתשתיות האחסון, חישוב ורשת, הנדרשים לכך וכן המלצות להקמת צוות לתפעול ומתן תמיכה שוטפת בהתאם לצרכי המשתמשים כפי שהוגדרו לעיל.
4. בנוסף, נדרש להקים צוות פיתוח לתשתית זו, אשר יוכל לפתח את שירותי הבינה המלאכותית המשותפים, לצורך קיצור הזמן והפריסה בפיתוח אלגוריתמי בינה מלאכותית, וכן צוות הנגשת חומרים ותפעול לצורך יצירת בסיסי נתונים מאומתים ומשותפים. סעיף זה לא תומחר בשלב זה.
5. ריכוז בסיסי נתונים ומשאבי חישוב במיקום לוגי אחד, כדי לאפשר מיצוי יעיל, תוך הקצאתם בהתאמה לצרכים השונים, בהנחה שישנה התפלגות בשימוש לאורך ציר הזמן ומתן גמישות ב-SLA (Service Level Agreement), היות והמשתמשים צריכים שימוש בעצמות שונה. תכנון אגם נתונים שיאפשר מחקרים מתקדמים ואולי אף דיפרנציאציה לחברות סטארט אפ בארץ. הריכוז יאפשר מתן תמיכה יעילה למשתמשים, יעילות והוזלת עלויות בהקשרי אחסון נתונים, גמישות בנושא שדרוג החומרה, וכן יקל על אבטחת הסייבר והאבטחה הפיזית. המימוש יכול להיות מבוסס ספק תשתית ענן חיצוני להתנעה מיידית, ביחד עם הקמת התוכנית למרכז תשתית לאומית והמשך למידה מהניסיון. בהקשר זה נכון לשקול התחברות למספר שירותי ענן משיקולים מסחריים. סעיף זה לא תומחר בשלב זה.
6. הכשרת הון אנושי לנושאי מחשוב על ומרכזי מידע, שיספק שירותים לא רק למרכז גופא אלא גם לאקו-סיסטם הכללי, הקמת מרכז ידע וקישור בין נתונים לטכנולוגיה.
7. קידום ומיקוד הון אנושי על-ידי קליטת אנשי סגל, בניית תוכניות לימודים אקדמיות ובעשייה והכשרות נוספות, מתן מלגות ותמריצים נוספים, וכן:
- א. מתן אפשרויות לסטודנטים לתארים מתקדמים לעבוד תוך יצירת חיבורים בין המחקר לעבודה בעשייה.
- ב. מתן מלגות לתלמידי מחקר מצטיינים (כ-1000 דוקטורנטים ומסטרנטים, בעלות של 100 אלף ש"ח לתלמיד לשנה (ובסה"כ 100 מיליון ש"ח לשנה). מתבסס על מלגות ות"ת במדעי הנתונים אבל עם התמקדות בתחומי ליבה של בינה מלאכותית, תוספת תקציב, ואפשרות למנגנון תקצוב ישיר לחוקרים.
- ג. מתן מענקי מחקר לחידוש תשתית חישובית לחוקרי אקדמיה בבינה מלאכותית, אפשרות לשילוב עם שימוש בשירותי ענן. עד לסכום של 150 אלף דולר למענק. הליך תחרותי עם דגש על מצוינות אשר ניתן לשלב עם מדעי הנתונים. יחולקו 30 מענקים כל שנה, סה"כ עלות 22.5 מיליון דולר ל-5 שנים.
- ד. יצירת קתדרות לחוקרי בינה מלאכותית ישראלים שהם מובילים בארץ ובעולם. חלוקת שלושה מענקים כל שנה (המחולקים למענק התחלתי, מענק איחוד ומענק מתקדם). מיליון דולר כל מענק, סה"כ עלות 15 מיליון דולר ל-5 שנים.
8. תמיכת המדינה במו"פ תעשייתי לצורך יצירת יתרון תחרותי במשק הישראלי. חלוקת המשאבים בתוך התעשיות מתבצעת לפי אופי המחקר – פיתוח יכולות בינה מלאכותית לעומת יישומה בתחומים שונים, ולפי מאפייני החברות – כאלה העוסקות בתשתית חומרה או תוכנה ויישומים לעומת כאלה המפתחות בינה מלאכותית ונדרשות לתשתיות חישוב ולנתונים לצורך אימון ותיקוף היכולות המפותחות.
9. השקעה בפיתוח קורפוסים וכלי NLP לשפה העברית והערבית עבור התעשייה וצרכי המגזר הציבורי.



10. השקעה בביסוס שיתופי פעולה בינלאומיים⁷.
11. הגדרת תקציב קבוע לצורך שדרוג החומרה בהתאם לצרכים ולקידום הטכנולוגי, להבטחת תמיכה ולחידושים עתידיים.
12. הגדרת רגולטור ממשלתי אחד כדוגמת הלמ"ס כבעל אחריות למיפוי, אגירה, זמינות, רישוי ונגישות לנתונים הזמינים למחקר ולפיתוח. צעד זה מאפשר פתרון לבעיית קושי בנגישות לנתונים, בכפוף לתקנים הבינלאומיים ולחוק הישראלי להגנה על פרטיות, תוך עמידה בהיבטים של פרטיות ואבטחת מידע.
13. הטמעת השימוש בבנינה מלאכותית במשרדי ממשלה.
14. הטמעת השימוש בבנינה מלאכותית בתחומי הבטחון והתעשייה.
15. הרחבת ההכשרה והיישום של בינה מלאכותית בתחומי מחקר נוספים באקדמיה מלבד תחומי הליבה.
16. פיתוח כלים חכמים לשימוש לקוחות המרכז ומתן תמיכה בהתאמתם.
17. עידוד מחקר ופיתוח לקידום תעשיית הבינה המלאכותית והעברת ידע מהאקדמיה לתעשייה.

נדרשנו להביא את התוכנית לכלל מימוש מכיוון שאם לא תמומש תוכנית לאומית שכזו, כבר בשנים הקרובות תמצא עצמה מדינת ישראל בפער מדעי, טכנולוגי וכלכלי שעלול להיות בלתי ניתן לגישור, ותלות מדעית וטכנולוגית שתהווה מגבלה על החוסן הלאומי. משמעות בנית היכולות בבנינה מלאכותית היא האצת וייעול כלל התחומים אליהם אנו מתייחסים בעבודה זו ושיפור חוסנו הלאומי. ישראל מאחרת אחרי מדינות אחרות בכל העולם ומשמעות העיכוב הינה נחיתות מובנית והתדרדרות יחסית מול מדינות ושחקנים זרים אחרים, שימשיכו להתקדם בקצב מהיר וידחקו את ישראל מזירות פעילות הולכות ומתרחבות, בתחומים קריטיים של מדע, כלכלה, תעשייה, ביטחון, בריאות, חינוך ותעסוקה בעלת פריור גבוה. מדינת ישראל איננה יכולה להרשות לעצמה להיות במצב שכזה.

בדוח זה מתוארת התוכנית המלאה שאושרה במסגרת עבודת הוועדה. בגין המצב המיוחד במשק הישראלי, אשר כולל דחיה באישור תקציב המדינה, בשלב זה הוגדר תקציב חלקי לתקופה מוגבלת. תקציב זה יהווה את הפעימה הראשונה לתוכנית כולה. הגדרת תקציב חלקי זה נעשתה בנפרד מקבלת אישור על התוכנית המלאה, שתתקצב באופן רחב בהמשך. תקציב הפעימה הראשונה כולל אך ורק את ההשקעה בחישוב על, עיבוד שפה טבעית וקידום התוכנית האקדמית באופן חלקי. פעימה זאת נעשתה בתיאום עם אגף התקציבים במשרד האוצר וזאת על מנת להתחיל באופן מידי בפעילות. מחשוב על הוא מטרה נעה אשר נכנסת למיינסטרים הטכנולוגי במהירות רבה. לאור זאת, מחשוב על נעשה זמין ונוח יותר לרכישה, גם במודל שירות החל מהיקף הנדרש בפעימה הראשונה. עם זאת, מפורט גם התקציב הנדרש במלואו.

סך התקציב הנדרש לצורך מימוש התוכנית המומלצת עומד על 5.26 מיליארד ₪. מתוכם כ-1.15 יופנו לאקדמיה הישראלית, כ-1.09 יופנו לפעילות תעשייתית, כ-1.08 לפעילות במערכת הביטחון (לשתי האחרונות שת"פ גנרי הדדי ניכר בשיעור של כ-80% בתעשייה וכ-95% בביטחון), כ-1.90 ישמשו להקמת התשתיות הלאומיות (מהם 1.08 למחשב העל, כולל תשתית תחזוקה ועדכון ל-5 שנים, וכ-0.86 לשירותי הענן). התקציב המוצע מוצג בנספח 1. נספחים נוספים לדוח הם כתב מינוי הוועדה (נספח 2), טבלה מסכמת של פירוט צרכי תעשייה ובטחון (נספח 3), הערכת עלויות של מסדי נתונים בענן עבור חקר

⁷ נבחנת אפשרות הצטרפות לתוכנית מחשוב על של האיחוד האירופי.

קורונה שנכתב בשיתוף עם פרופסור אורן קורלנד (נספח 4), נספח אשר מסכם בקצרה את העבודה שנעשתה בנושא הטמעה (נספח 5), ונספח אשר מפרט עדכונים בנושא מחשוב על (נספח 6). נספח נוסף, העוסק באשכולות של מקבצי קורסים אקדמיים, יצורף לדוח בהמשך כנספח 7.

לסיכום, תחום בינה מלאכותית ומדע הנתונים הינו תחום צומח ובעל פוטנציאל גבוה להשפעה על מגוון תחומים, אשר נמצא בשנים האחרונות בהתפתחות מואצת. לישראל יש בו צורך אסטרטגי ופוטנציאל דרמטי לשימור וחיזוק החוסן הלאומי, הביטחוני, האקדמי והעסקי. מכיוון שהתחום דורש תשתית אנושית מגוונת שהכשרתה אורכת זמן רב ומצויים בו כבר עתה מספר שחקנים מובילים, נדרש מאמץ מרוכז למיצוי הפוטנציאל של התחום. לשם כך מתכבדת הוועדה להציג להלן את הצעתה לתוכנית לאומית המשלבת את כל השותפים הפוטנציאליים, על-מנת לבסס יכולת מקומית ומעמד תחרותי בזירה הבינלאומית.

1. הוועדה

1.1. חברי הוועדה

- ד"ר ארנה ברי (יו"ר הוועדה) – ד"ר ארנה ברי היא מדענית, יזמת הייטק, אשת עסקים, מנהלת בכירה במגזרי המדע והטכנולוגיה בישראל. האישה הראשונה והיחידה בישראל שכהנה בתפקיד המדענית הראשית וראש מנהל המו"פ התעשייתי במשרד התעשייה והמסחר (לימים- משרד הכלכלה ורשות החדשנות) בממשלת ישראל.
- פרופסור יאיר וייס – יאיר וייס הוא פרופסור בביה"ס להנדסה ולמדעי המחשב באוניברסיטה העברית. בין השנים 2014-2018 כיהן כדיקן ביה"ס. תחום המחקר שלו הוא למידה חישובית וראייה ממוחשבת.
- פרופסור אמיר גלברזון – פרופסור אמיר גלברזון הוא חוקר בתחומי הלמידה החישובית, הבינה המלאכותית ועיבוד שפה טבעית. פרופסור בבית הספר למדעי המחשב באוניברסיטת תל אביב.
- פרופסור שרית קראוס – שרית קראוס היא פרופסור למדעי המחשב באוניברסיטת בר-אילן. חוקרת בתחום של מערכות מרובות סוכנים (הכוללות מערכות מחשב אינטליגנטיות, אנשים ורובוטים) בעזרת אינטגרציה של שיטות מתחומים מגוונים. למשל, למידת מכונה, אופטימיזציה ותורת המשחקים.
- ד"ר איה סופר – ד"ר איה סופר היא סגן נשיא עולמית לטכנולוגיות בינה מלאכותית במעבדות המחקר של IBM. במסגרת תפקידה, ד"ר סופר מגדירה את אסטרטגיית המחקר של IBM בבינה המלאכותית ועובדת עם מאות מדענים סביב העולם במטרה לעצב את הרעיונות שלהם לתוצרים שיביאו ערך משמעותי ביישומי בינה מלאכותית. ב-20 שנותיה ב-IBM הובילה ד"ר סופר שורת מיזמים אסטרטגיים, שהתפתחו למוצרים ופתרונות מצליחים, כולל עבודה על פיתוחה של מערכת IBM Watson ושל Project Debater. היא שותפה בכתיבת יותר מ-50 מאמרים מדעיים, מכהנת בתפקידי ארגון וניהול כנסים מדעיים בינלאומיים, ומוזמנת כמרצה אורחת בכנסים רבים כולל כנס TEDMED.
- מר אבנר גורן – אבנר גורן הוא סגן נשיא באינטל העולמית וחלק מקבוצת DATA PLATFORMS GROUP שאחראית למוצרי אינטל לדאטה סנטר. ברקע המקצועי שלו אבנר ניהל קבוצות ארכיטקטורה שהגדירו ציפיים (SOC, System-on-a-chip) לתחומים

רחבים כולל PC, IOT, מכוניות ולאחרונה מוצרי AI (Artificial intelligence). בחצי שנה האחרונה אבנר משמש כ – EXECUTIVE IN RESIDENCE של אינטל בהבאנה לאבס – חברה ישראלית שנרכשה על ידי אינטל ומפתחת מאיצי AI לדאטה סנטר. אבנר משמש גם כחבר בהנהלת אינטל ישראל ואחראי על נושאי ה-INNOVATION וה-GROWTH של אינטל בישראל – כולל עבודה עם סטארטאפים ישראלים.

- תא"ל (מיל') מוטי בסר – מוטי בסר, תא"ל (במיל'), הוא בעל ניסיון ניהולי עשיר בגופים עסקיים וציבוריים ושימש, בין השאר, כיועץ הכספי לרמטכ"ל וראש אגף התקציבים של משרד הביטחון, מנכ"ל משרד מבקר המדינה ומנכ"ל בורסת היהלומים הישראלית. כמו כן, שימש כחבר במספר רב של מועצות מנהלים וכחבר בנבחרת הדירקטורים מטעם רשות החברות הממשלתיות ולאחרונה מונה חבר המועצה המינהלית של בנק ישראל.
- מר גדי לוי (מרכז הוועדה) – סמנכ"ל האקדמיה הלאומית הישראלית למדעים. כלכלן בעל ניסיון בבחינת תשתיות מחקר ופיתוח ותקצובן. בעבר שימש בתפקידים שונים בתחום התקציבי ובין היתר רפרנט לתחום המדע והטכנולוגיה באגף התקציבים באוצר. הוא בעל תואר ראשון בכלכלה ומדע המדינה ותואר שני בכלכלה מאוניברסיטת תל אביב.
- סא"ל דהן (מרכז הוועדה) – סא"ל דהן, ראש ענף בינה מלאכותית במפא"ת, מוביל את המחקר והפיתוח במערכת הביטחון בתחומי הבינה המלאכותית ולמידת מכונה, תוך מיקוד במינוף הפוטנציאל בין הטכנולוגיה המתפתחת בתחום ובין הצרכים המבצעיים של מערכת הביטחון. במסגרת תפקידו דהן אחראי על ייזום מחקרים מול האקדמיה ומעבדות המחקר בישראל, מיסוד קשרים טכנולוגיים בינלאומיים והובלת תהליכי הפיתוח מול התעשייה הביטחונית והסטארטאפים בנושאי למידת מכונה לטובת מיצוי והסקת תובנות ממידע, הגברת אפקטיביות של הלוחם ויצירת תהליכים מבצעיים חדשים מבוססי מידע וידע. סא"ל דהן בעל ידע וניסיון פיתוחי בביצוע מגוון תפקידים בצה"ל ובמפא"ת, פיתוח מערכות מבצעיות והנחיית מחקר מתקדם, דהן בעל תואר ראשון ושני בהנדסת חשמל (טכניון, בר אילן בהתאמה) וכיום הינו דוקטורנט בתחום למידת מכונה ו-Deep Learning באוניברסיטת בר אילן.
- ד"ר אביב זאבי באלאסיאנו (מרכז הוועדה) – ד"ר אביב זאבי באלאסיאנו, ראש זירת תשתית טכנולוגית ברשות החדשנות, אמון במסגרת תפקידו על פיתוח טכנולוגיות גנריות ותשתיות מו"פ וכן על הובלת שיתופי פעולה בין התעשייה לאקדמיה המניבים טכנולוגיות מתקדמות ותוצרים חדשניים, ומחזקים את היתרון הטכנולוגי ארוך הטווח של התעשייה בישראל אל מול התחרות העזה בשווקים הבינלאומיים. הוא בעל ניסיון רב בתחום הפיתוחים הטכנולוגיים, בפיתוח מערכות מקוונות ומערכות אימון והדמיה. בתפקידו האחרון שימש כמנהל מחלקת ICT במנהלת המחקר הישראלית לתוכנית המחקר האירופאית. אביב הוא ד"ר לניהול מערכות מידע מאוניברסיטת תל אביב ובעל תארים נוספים בכלכלה, מדעי המדינה וניהול.

בוועדה נכחו גם מספר משקיפים :

- ד"ר אלה שטראוס – מנהלת תחום כימיה אנרגיה חומרים וננו טכנולוגיה, משרד המדע.
- ד"ר פדל צאלח – מנהל תחום פיזיקה, מתמטיקה שימושית ומדעי המחשב, משרד המדע.

- פרופסור יחיעם פריאור – ראש האגף למערכות מידע, לשעבר דיקן הפקולטה לכימיה, מכון ויצמן למדע.
- ד"ר נעמי בק – סמנכ"ל אסטרטגיה ובינלאומיות, הוועדה לתכנון ותקצוב.

עבודת הוועדה לוותה בהיבטי תקצוב ועריכה מקצועית על ידי הגורמים הבאים:

- פרופסור חיים לוי קידר - המחלקה לניהול, אוניברסיטת בן-גוריון, סייע לעבודת הוועדה בהיבט התקצוב.
- ד"ר ניר ברילר וצוות מחברת רבתית הנדסה ויעוץ בע"מ; מר תומר דור, גברת ענת כהנא הורביץ, מר יובל גייגר – כתיבה, עריכה מקצועית ותמיכה בניהול הפרויקט.

1.2. מנדט הוועדה

הוועדה מונתה בחודש פברואר בשנת 2020 על-ידי יו"ר תל"ם (פורום לתשתיות לאומיות למחקר ולפיתוח), פרופ' שמעון אולמן. תפקיד הוועדה היה לבחון את הצורך בהתערבות ממשלתית לשם האצת התפתחות תחום הבינה המלאכותית ומדע הנתונים. במסגרת עבודת הוועדה היה עליה להתייחס לנושאים הבאים:

1.2.1. בחינת היבטי הון אנושי, תשתיות פיזיות, הנגשת מאגרי מידע והעברת הידע מהאקדמיה לתעשייה. זאת, תוך התבססות, בין היתר, על התוכניות הקיימות והמתוכננות ברשויות המחקר ומשרדי הממשלה הרלבנטיים.

1.2.2. בחינת תוכניות ומודלים מקבילים במדינות אחרות ובדיקת אפשרות לקיום שיתוף פעולה.

1.2.3. המלצה על אבני דרך לפיתוח תשתיות המחקר האקדמי והיישומי בתחום בינה מלאכותית ומדע הנתונים.

1.2.4. המלצה על הגדרת אסטרטגיה לאומית לריכוז המאמצים להשגת מידע הדרוש לצורך אימון בינה מלאכותית בתחומי מפתח, בין היתר באמצעות "רגולציה מאפשרת", או תהליך מוצע לשם השגתה של "רגולציה מאפשרת". מטרת רגולציה זו להנגיש מאגרי ידע קיימים ועתידיים, תוך הבטחת כללי אסדרה ראויים, ותוך התבססות, ככל הניתן, על רגולציה מאפשרת בינלאומית, ובתיאום עם רשת ⁸ C4IR (Center for the Fourth Industrial Revolution) של הפורום הכלכלי העולמי, אליה החליטה הממשלה להצטרף.

1.2.5. בחינת אפשרויות המסחור והייצוא של תוצרי המו"פ לרבות המגבלות שיחולו עליהן.

1.2.6. אפיון מודל פעילות לתשתית תוך התייחסות לתשתיות ולהיבטים מדעיים, טכנולוגיים, ארגוניים, כלכליים ותקציביים.

לפרטים נוספים אודות מנדט הוועדה ראו את כתב מינוי ועדת בדיקה מקצועית לבחינת הצורך בהתערבות ממשלתית לשם האצת התפתחות תחום הבינה המלאכותית ומדע הנתונים מתאריך 16.02.2020. כתב המינוי מצורף למסמך זה בנספח 2.

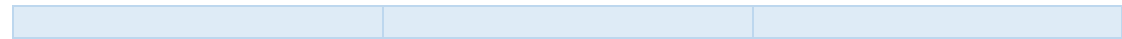
⁸ רשת בינלאומית שמטרתה יצירת דיאלוג רב-אינטרסים ושיתוף פעולה קונקרטי באתגרי הממשל וההזדמנויות שהוצגו על ידי טכנולוגיות מתקדמות מהמהפכה התעשייתית הרביעית.

1.3. עבודת הוועדה

1.1.1. פגישות

להלן רשימת בעלי תפקיד וחברות עימם נפגשה הוועדה כחלק מפעילותה:

- | | |
|----------|--|
| 1.1.1.1. | ממשל וחברי תל"ם |
| ● | פרופ' יפה זילברשץ – יו"ר ות"ת. |
| ● | רות שוהם – מנכ"לית האוניברסיטה הפתוחה. |
| ● | ד"ר צביקה קפלן – פורום תל"ם. |
| ● | איתן טי – סמנכל תכנון ובקרה – משרד המדע. |
| ● | טל ברוק – ראש מטה מנכ"לית – משרד המדע. |
| ● | שי לי שפיגלמן – מנכ"לית משרד המדע. |
| ● | אלי"מ שי דן – צה"ל. |
| 1.1.1.2. | ביטחון |
| ● | סא"ל רועי אליעזר, סא"ל יואב שחם, סא"ל מירב בילינסקי – צה"ל. |
| 1.1.1.3. | אקדמיה |
| ● | פרופסור יואב גולדברג – חוקר מתחום תשתיות עיבוד שפה בעברית, אוניברסיטת בר אילן. |
| ● | פרופסור סטיבן פרנקל – חוקר מתחום חישוב על, טכניון. |
| ● | פרופסור שמואל אוסובסקי, טכניון. |
| ● | פרופסור יואב בנימיני, אוניברסיטת תל אביב, הצגת עבודת ועדת ההיגוי של ות"ת לנושא מדעי הנתונים. |
| ● | פרופסור אורן קורלנד, הפקולטה להנדסת תעשייה וניהול, הטכניון. |
| ● | פרופסור מיכאל בירנהק, הפקולטה למשפטים, אוניברסיטת תל אביב. |
| ● | פרופסור אור דונקלמן, המרכז לחקר סייבר משפט ומדיניות, אוניברסיטת חיפה. |
| ● | ד"ר ערן טוך, המחלקה להנדסת תעשייה וניהול, אוניברסיטת תל אביב. |
| ● | פרופסור קרין נהון, המרכז הבינתחומי הרצליה. |
| ● | פרופסור צבי אקשטיין ומכון אהרון, המרכז הבינתחומי הרצליה. |
| 1.1.1.4. | תעשייה |
| ● | קובי כץ – סמנכ"ל מחשוב ומערכות מידע, רפאל. |
| ● | מיכאל כגן – סמנכ"ל טכנולוגיות, חברת Mellanox. |
| ● | אביגדור וילנץ – יזם הייטק. |
| ● | ארווין לייבוביץ' – HPE. |
| ● | נתי אמסטרדם, אופיר זמיר – NVIDIA. |



- ד"ר גבי חיון – מובילאיי.
- מיכל ברוורמן-בלומנשטיק – מנכ"לית טכנולוגיה עולמית לאבטחת ענן ו-AI, ד"ר תומר סיימון - מיקרוסופט.
- אייל טולדנו – סמנכ"ל טכנולוגיות, Zebra Medical.
- ד"ר אלון גרובשטיין – דירקטור תחום דאטה סיינס ואנליטיקה, Trax.
- אורי צרפתי – סמנכ"ל מו"פ, Emedgene.
- נעמה דמטי, אילן יוסף – Nice.
- מרק קפל – אבוגין.
- אריאל לבנטל – Khealth.
- ירון אלטשולר, מירב אש, ארתיום ליכטנשטיין – AWS.
- אורי חייק, קרן וולף – IBM.
- מיקי קנת, עידן לוי, מייקל פקטור – Red Hat.
- יוגין קנדל – מנכ"ל סטארט-אפ ניישן סנטרל.
- מתי צוויג – מנכ"לית סקייל אפ.
- רותם אלעלוף – מנכ"ל BeyondMinds

1.1.1.5 גורמים מחו"ל

- אלכסנדר סוקרבסקי (Alexander Sukharevsky) – שותף בכיר, שותף מנהל גלובלי, מקינזי אנליטיקס.
- ניקו הנק (Nico Henke) - שותף בכיר, יו"ר מועצת הטכנולוגיה, מקינזי רוסיה.
- כריסטיאן שולץ (Christian Schulze) – פרופסור, אחראי על תשתית בינה מלאכותית ב-DFKI (German Research Center for Artificial Intelligence)
- ווס ארמור (Wes Armour) – פרופסור במחלקה להנדסה, דירקטור של מחשב העל 'Jade' הבריטי ומרכז המחקר 'e-Research' מרכז החישוב באוקספורד, אנגליה.
- ריק סטיבנס (Rick Stevens) – פרופסור, ראש תוכנית Exascale ב-ANL (Argonne National Lab)
- תומס שולדס (Thomas Schulthess) – פרופסור, דירקטור של מרכז חישוב העל הלאומי השווייצרי (Swiss National Supercomputing Centre) ((CSCS).
- יוגין גריפית' (Eugene Griffiths) - ראש תמיכה במחקר, העברה והפצה, מרכז חישוב העל בברצלונה.

- גוסטב קלב (Gustav Kalbe) – ד"ר, ראש יחידה מרכז חישוב על וטכנולוגיות קוונטים, האיחוד האירופי, DG-CONNECT (Directorate-General for Communications Networks, Content and Technology).
- אירינה אורסיץ' (Irina Orssich) – יועצת לבינה מלאכותית, האיחוד האירופי, DG Connect.
- הנס אייד (Hans Eide) – ראש תשתית חישוב העל הנורווגית.
- ברטרנד בראונסשוויג (Bertrand Braunschweig) – מתאם תוכניות לאומיות למחקר בבינה מלאכותית, INRIA (National Institute for Research in Digital Science and Technology), צרפת.
- פרופסור תומס ליפרט (Thomas Lippert), ג'ין-מרי ספאוס (Jean-Marie Spaus) – מרכז חישוב העל בלוקסמבורג.
- סין וונג (Sean Wang) – CTO, חברת G42, אבו דאבי.
- קאי פונג צ'נג (Kai Fong CHNG) – דירקטור מנהל, Singapore Economic Development Board (EDB).
- גילמן לואי (Gilman Louie), פול ליקס (Paul Lekas) – NSCAI (National Security Commission on Artificial Intelligence).
- אנה קוזקובה (Anna Kozáková), פרופסור ואקלב סנסל (Vaclav Snasel) – מרכז חישוב העל, צ'כיה.
- פרופסור מיכאל גופמן, אוניברסיטת רוצ'סטר, ארה"ב.

1.1.2 מקורות מידע נוספים

נערכו התייעצויות עם המומחים מחו"ל שהוזכרו, וכן נסקרו פרסומים מתוכניות לאומיות לבינה מלאכותית ולמחשוב על של מדינות שונות, ביניהן: ארה"ב, גרמניה, צרפת, בריטניה, קנדה, נורבגיה ועוד.

לאור שפע חומרי הרקע, הן במסמכים הישימים, והן במסמכים רבים נוספים הזמינים פומבית, מסמך זה לא יעסוק באופן מעמיק בסקירת הרקע המדעי-טכנולוגי והמגמות העולמיות, אלא יסתפק בציון נקודות משמעותיות ורלוונטיות במיוחד לעבודת הוועדה. לפרטים נוספים על אודות הרקע המדעי-טכנולוגי וכן למידע נוסף על אודות תתי-התחומים, ראו דוח צוות המשנה של המיזם הלאומי למערכות נבונות בנושא כוח מחשוב וקוונטים⁹.

⁹ בן ישראל, י', מתניה, א' ופרידמן, ל' (עורכים). (ספטמבר 2020). המיזם הלאומי למערכות נבונות בטוחות להעצמת הביטחון הלאומי והחוסן המדעי-טכנולוגי: אסטרטגיה לאומית לישראל. דוח מיוחד לראש הממשלה.

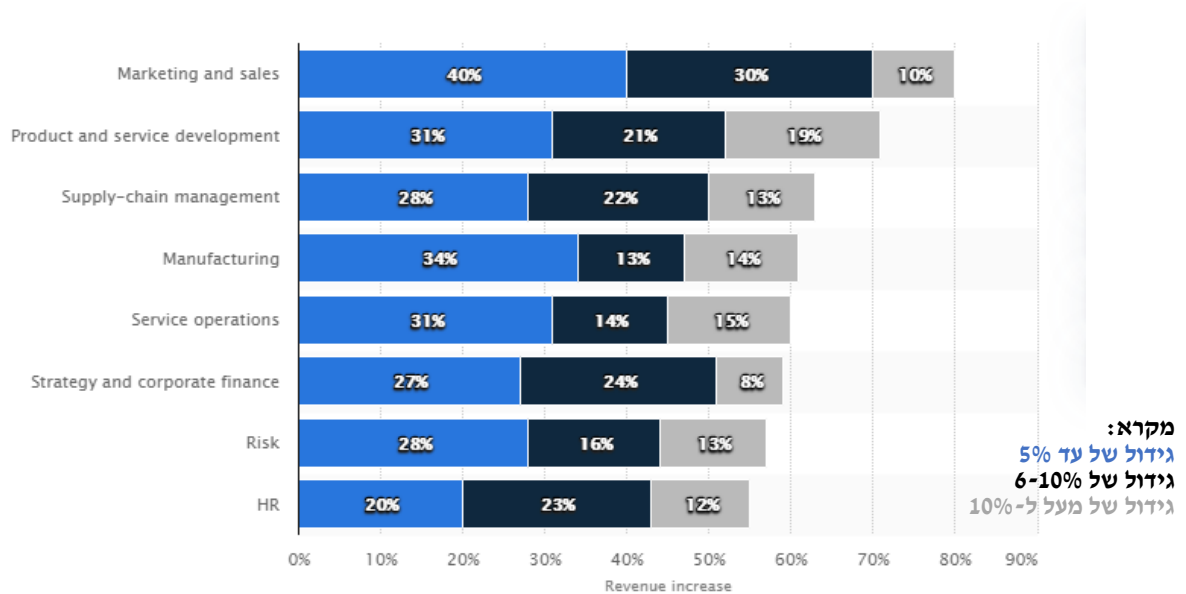
2. בינה מלאכותית

2.1. רקע

לא קיימת הגדרה מקובלת למונח 'בינה מלאכותית', אבל מונח זה משמש לרוב לתיאור מכונה או מערכת המבצעות משימות שלרוב מתבצעות על ידי חשיבה אנושית, כגון הבנת שפה טבעית, למידת התנהגויות או פתרון בעיות. ישנו טווח רחב של התנהגויות מעיין אלו, אבל רובן כוללות מחשבים המריצים אלגוריתמים, אשר לרוב מבוססים על נתונים¹⁰. בשנים האחרונות החלה האצה רבה בשימוש בבינה מלאכותית בכל תחומי החיים ובכל העולם, כתוצאה מהתקדמות טכנולוגית, אלגוריתמית וחישובית, וזמינות נתוני עתק.¹¹

2.2. השקעה בבינה מלאכותית בעולם

ניתוח שנעשה בשנת 2019 מראה שישנו גידול ברווחים בעקבות אימוץ בינה מלאכותית בארגונים ברחבי העולם. הגידול הוא בכלל תחומי העיסוק של הארגון, ביניהם שיווק ומכירות, שרשרת אספקה, משאבי אנוש, ניהול סיכונים וכו'.¹²

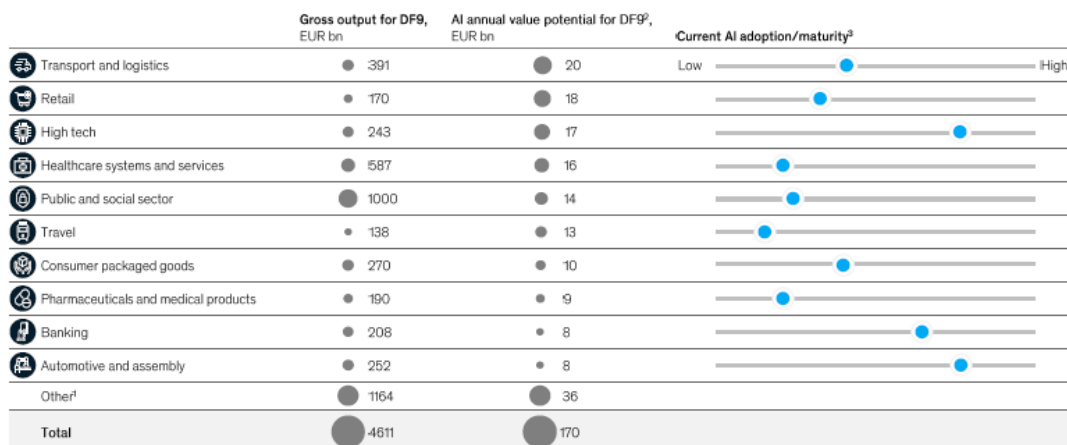


איור 1: גידול ברווחים בעקבות אימוץ בינה מלאכותית בארגונים ברחבי העולם (עדכני ל-2019)

¹⁰ פירוט על 'תחומי הליבה' במחקר בינה מלאכותית מופיע בפרק "אקדמיה" - <https://www.turing.ac.uk/about-us/frequently-asked-questions>
¹¹ <https://www.mckinsey.com/~media/McKinsey/Featured%20Insights/Artificial%20Intelligence/Globai%20AI%20Survey%20AI%20proves%20its%20worth%20but%20few%20scale%20impact/Global-AI-Survey-AI-proves-its-worth-but-few-scale-impact.pdf>
¹² <https://www.statista.com/statistics/1083482/worldwide-ai-revenue-increase>

כיום ניתן לראות שימושים לבניה מלאכותית גם בתחום הבריאות, חינוך, צמיחה כלכלית והשמה לעובדים, מיחזור, אנרגיה ירוקה, טיהור מים והתחממות הגלובלית. שימוש בבניה מלאכותית יועיל לכלל הסקטורים, ביניהם: תחבורה ולוגיסטיקה, קמעונאות, הייטק, מערכות ושירותי בריאות, מגזר ציבורי וחברתי, ערים חכמות, נסיעות, צרכנות, פארמה ומוצרים רפואיים, בנקאות, רובוטיקה, רכבים אוטונומיים, רחפנים, יצור מתקדם ועוד.¹³

EXHIBIT 21: AI could hold significant value potential of ~EUR 170 bn across all sectors in DF9⁴



איור 21: ערך פוטנציאלי לשימוש בבניה מלאכותית בסקטורים השונים (מקינזי, אוקטובר 2020)

לאור ההשפעות המהפכניות של תחום הבניה המלאכותית ומדע הנתונים על מחקר ופיתוח בתחומי התעשייה, התשתיות, הביטחון, הבריאות, התרופות והחומרים, והצפי להתרחבות השפעות אלה ולהאצתן, החליטו המדינות והתאגידים המובילים בעולם לרכז מאמץ מיוחד בבניית תשתיות שיאפשרו ניצול אופטימלי של יכולות וכלים בתחומים אלו. בין מדינות אלה ניתן למנות את ארה"ב, האיחוד האירופי, סין, בריטניה, גרמניה, צ'כיה, צרפת, דרום קוריה, סינגפור ומדינות רבות נוספות, שהגדירו תוכניות לאומיות בהשקעה של מאות מיליונים עד מיליארדי דולרים, שתתפרסנה על-פני תקופות זמן של עד עשור¹⁴.

¹³ <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/how-nine-digital-front-runners-can-lead-on-ai-in-europe>

¹⁴ התוכנית האמריקאית העדכנית ליוני 2019 לפיתוח AI:

<https://www.whitehouse.gov/wpcontent/uploads/2019/06/National-AI-Research-and-Development-Strategic-Plan-2019-Update-June2019.pdf>; ציכיה:

https://www.mpo.cz/assets/en/guidepost/for-the-media/press-releases/2019/5/NAIS_eng_web.pdf

<https://www.hpe.com/us/en/newsroom/press-release/2020/10/hewlett-packard-enterprise-selected-to-build-czech-republics-most-powerful-supercomputer-to-advance-rd-in-science-and-engineering.html>

; דוח של מכון Adenaur המשווה בין אסטרטגיות AI של חמש מדינות:

<https://www.kas.de/documents/252038/4521287/Comparison+of+National+Strategies+to+Promote+Artificial+Intelligence+Part+1.pdf/397fb700-0c6f-88b6-46be-2d50d7942b83?version=1.1&t=1560500570070>

; אסטרטגיית הדאטה האירופאית:

Communication from the Commission to the European Parliament, the Council, the European Economic

במסמך שנכתב על ידי הפורום הכלכלי העולמי, תואר כיצד יש לבנות אסטרטגיית בינה מלאכותית לאומית¹⁵. לפי המסמך, כל אסטרטגיית AI לאומית צריכה להתחיל עם הערכה של סדרי העדיפויות האסטרטגיים של המדינה, נקודות החוזק והחולשה שלה. פריסת המשאבים המוגבלים של האומה לכל מטרה אסטרטגית צריכה לעלות בקנה אחד עם צרכי המדינה הדמוגרפיים, האסטרטגיים, סדרי עדיפויות, שאיפות האזרחים והמדינה, חששות דחופים, אילוצי משאבים ושיקולים גיאופוליטיים. הפרמטרים שצריכים להילקח בחשבון, דרך דירוג שיעשה אינדיבידואלית לכל מדינה הם:

- בניתוח חוזקות וחולשות – הון אנושי, דיגיטיזציה ותשתית, שת"פ תעשייה אקדמיה, יכולת אימון ורגולציה.
 - בניתוח הזדמנויות ואיומים – אקו-סיסטם לחדשנות, גמישות בתעשייה, גמישות במגזר הציבורי, שת"פ בינלאומי.
- לאחר מכן יש לקבוע יעדים בתחומים אלו:
- יעד לקיבולת - משאבי אנוש ותשתית דיגיטלית.
 - יעד להשקעות - מחקר ופיתוח, מענקים.
 - יעד לאימוץ - מגזרים סוציו-אקונומיים, מגזרים תעשייתיים.
 - יעד לרגולציה - הפעלת רגולציה כגון פרטיות וסטנדרטים אתיים לשימוש בנתונים.
- התוכנית האמריקאית¹⁶ למשל, כוללת 7 אסטרטגיות שנקטו על ידי הממשל האמריקאי על מנת לקדם מערכות בינה מלאכותית:

1. השקעות לטווח ארוך במחקר AI.
2. פיתוח שיטות אפקטיביות לשת"פ AI-אנושי.
3. הבנה וטיפול בהשלכות האתיות, החוקיות והחברתיות של AI.
4. הבטחת הביטחון, השקיפות והאמינות של מערכות AI.
5. פיתוח דאטה סטים פומביים וסביבות עבור אימון ובדיקות AI.
6. מדידה והערכה של טכנולוגיות AI דרך סטנדרטים ומדדים.
7. הבנת צרכי הגיוס והכשרת הון אנושי לתוכנית מו"פ AI הלאומית.

שתי האסטרטגיות הרלוונטיות במיוחד עבורנו הן הרביעית והחמישית. עיקרי האסטרטגיה הרביעית הם: שיפור יכולת ההסבר והאמינות, בניית אמון משתמשים, שיפור האימות והתיקוף, אבטחה מפני תקיפות, השגת ביטחון לאורך זמן והכרת הערך למשתמש. במסגרת קידום האסטרטגיה מחלקת התחבורה של ארה"ב פרסמה נהלים לשילוב בטוח של רכבים אוטונומיים במערכות התחבורה הרב-מודליות, ו-IARPA (Intelligence Advanced Research Projects Activity) הודיעה על שלוש תוכניות למערכות AI לאבטחת מידע שעוסקות במערכות למידה מאובטחות, גילוי סוסים טרויאנים, ומניעת הונאות. עיקרי האסטרטגיה החמישית הם: פיתוח והנגשה של מגוון רחב של מערכי נתונים כדי

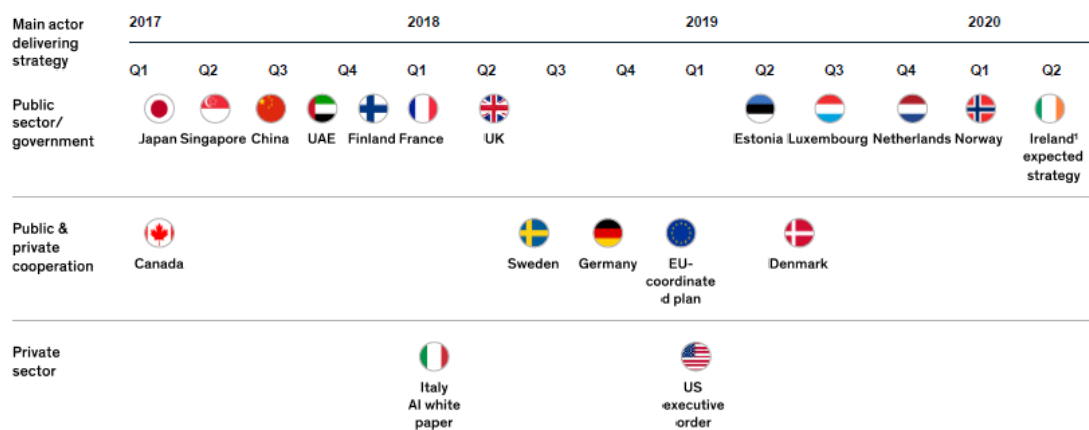
and Social Committee and the Committee of the Regions, "A European strategy for data", Brussels, 19.2.2020 COM(2020)

http://www3.weforum.org/docs/WEF_National_AI_Strategy.pdf¹⁵

Exec. Office of the U.S. President, The National Artificial Intelligence Research and Development Strategic Plan: 2019 Update (June 2019), available at <https://www.whitehouse.gov/wpcontent/uploads/2019/06/National-AI-Research-and-Development-Strategic-Plan-2019-Update-June2019.pdf>.

לענות על הצרכים של תחומי עניין רבים בבינה מלאכותית ויישומים, אימון ובדיקת תגובתיות המשאבים לאינטרסים מסחריים וציבוריים, פיתוח ספריות תוכנה עם קוד פתוח וערכות כלים. במסגרת קידום האסטרטגיה מחלקת התחבורה מימנה מחקר נהיגה נטורליסטי בכבישים מהירים, המחלקה לחיילים ותיקים יצרה את מערך הנתונים הרפואי-הגנום המקושר הגדול ביותר בעולם עם כלים המאפשרים למידת מכונה (Machine learning, ML) ו-AI, מאמץ אשר מינף תקנים של המכון הלאומיים לתקנים וטכנולוגיה, ה-NIST (National Institute of Standards and Technology), חוקים וצווים ביצועיים. ה-GSA (General services) administration פועל כדי לאפשר שימוש במשאבי מחשוב ענן למו"פ AI במימון פדרלי (הדומיינים שתחתיו מכילים מעל ל-246,000 דאטה סטים וקודים מסוכנויות שונות, והמידע מגיע אליהם אוטומטית מהסוכנויות), וה-STRIDES (Science and Technology Research Infrastructure for Discovery, Experimentation, and Sustainability) יזמה שת"פ עם ספקי ענן מובילים מהתעשייה אשר מאפשר גישה לחוקרים לנפחי נתונים משמעותיים, אשר ממומנים ע"י ה-NIH (National Institutes of Health) ומאוחסנים בסביבות ענן.

EXHIBIT 6: With AI receiving increasing attention, several countries have released strategies on AI



איור 6: מדינות אשר שחררו אסטרטגיות בינה מלאכותית (מקינזי, אוקטובר 2020)

הטבלה הבאה מציגה את התוכניות הלאומיות בתחום בינה מלאכותית במדינות מובילות בעולם. הנתונים המופיעים מציגים רק את התוכניות הממשלתיות במיליוני דולרים. משך התוכניות איננו זהה בכל המדינות, אך מדובר על שנים בודדות עד עשור:

דגשים מיוחדים בתוכנית			מדינה
תקציב	רגולציה	הון אנושי	
1.3 מיליארד אירו שנתיים למחקר בסיסי מתקציב מו"פ AI שנתי, בנוסף למימון ממחלקות ספציפיות (6.3 מיליארד אירו מה-DOD ב-2017).	תמיכה במסחור של AI על ידי צמצום והסרת חסמים רגולטוריים.	מעל ל-3,000 תלמידי דוקטורט בשנה.	ארה"ב
ביגינג הודיעה על הוצאות של 16.4 מיליארד אירו לקידום ענף השבבים לבדו, וברמה התת-לאומית עיר אחת (טיגיאן) הקימה קרן של 12.8 מיליארד יורו לקידום AI.	קיימים קשיי ייצוא של מוצרי AI שפותחו תחת תקני הגנת נתונים סיניים בשל תקנות נוקשות במדינות המערב.	תוכנית אלף הכישרונות להחזרת מומחים סיניים בתחום ממדינות אחרות בעולם.	סין
מעל למיליארד אירו (2018) שיוקצו לקידום חינוך, בניית תשתית, יצירת סביבה עסקית וכינון אתרי מו"פ ל-AI.	אין פירוט בנושא	כ-136 חוקרים בעולמות בבניה מלאכותית בין השנים 2016-2018, כ-400 תלמידי דוקטורט בשנה.	אנגליה
1.5 מיליארד אירו עד 2022.	אין פירוט בנושא	מתן אפשרות לחוקרים לעבוד במקביל במגזר האקדמי והפרטי. הוקמה רשת של מומחי בינה מלאכותית מתנדבים שמטרתה היא ליעץ למדינה בנושא רכש טכנולוגיות ותמיכה באבטחת סייבר. שיתוף פעולה מתוכנן עם גרמניה.	צרפת
173 מיליון אירו עד 2022, המיועדים בעיקר לסטארטאפים ולשמירה על תחרותיות של חברות (במדינה זו יש תשתית דיגיטלית ענפה ותוכנית זו	פוטנציאל מסחרי ל-AI מעבר לגבולות פינלנד זמין בתחום הבריאות, על בסיס פרויקט הגנום הפיני "FinnGen" ובשיתוף עם IBM	אין פירוט בנושא	פינלנד

אינה משקפת את כלל ההשקעות בנושא).	ווטסון בריאות. נישה נוספת היא פיתוח פרוטוקול להחלפת נתונים בינלאומית (IHAN), בדומה לפרוטוקול IBAN להעברות בנקאיות.		
כ-1.7 מיליארד אירו עד 2022.	אין פירוט בנושא	אוניברסיטאות מעניקות לסטודנטים תקופות היעדרות של עד ארבעה סמסטרים כדי לטפל בסטארטאפים שלהם.	דרום קוריאה

טבלה 1: אסטרטגיות לאומיות לקידום בינה מלאכותית ברחבי העולם¹⁷

לצד התוכניות הלאומיות הגלויות, קיימים בעולם מספר מאמצים מקבילים ומסווגים לפיתוח יישומים שנוגעים לביטחון לאומי. במאי 2020 הושקה שותפות גלובלית בבינה מלאכותית (GPAI, Global Partnership on AI) על ידי אוסטרליה, קנדה, צרפת, גרמניה, הודו, איטליה, יפן, מקסיקו, ניו זילנד, דרום קוריאה, סינגפור, סלובניה, אנגליה, ארה"ב והאיחוד האירופי. השותפות שמה לה למטרה לפתח בינה מלאכותית בדרך התואמת זכויות אדם, חירות, וערכים דמוקרטיים, ושמה דגש על מציאת דרכים בהן בינה מלאכותית יכולה לעזור לעולם להתאושש ממגפת ה-COVID-19¹⁸.

במחקר שנערך על ידי חברת Tortoise בשנת 2019 נוצר אינדקס המדרג 54 מדינות לפי יכולות בינה מלאכותית, תוך התבססות על 7 מאפיינים: כישרונות, תשתית, סביבת פעילות, מחקר, פיתוח, אסטרטגיית ממשלה ומסחרי. ישראל דורגה במקום 12, כאשר הדירוג שלה בסעיפים השונים היה: כישרונות 13, תשתית 36, סביבת הפעלה 50, מחקר 5, פיתוח 5, אסטרטגיה ממשלתית 46, מסחרי 3. כפי שניתן לראות, היה פער משמעותי בין הדירוג הגבוה של ישראל במחקר ופיתוח ובמסחרי, למול הדירוג הנמוך בתשתיות הנדרשות בסביבת פעילות ובאסטרטגיית ממשלה¹⁹.

¹⁷ <https://www.kas.de/documents/252038/4521287/Comparison+of+National+Strategies+to+Promote+Artificial+Intelligence+Part+1.pdf/397fb700-0c6f-88b6-46be-2d50d7942b83?version=1.1&t=1560500570070>

¹⁸ <https://www.diplomatie.gouv.fr/en/french-foreign-policy/digital-diplomacy/news/article/launch-of-the-global-partnership-on-artificial-intelligence-by-15-founding>

¹⁹ <https://www.tortoisemedia.com/intelligence/ai/>

Country	Talent	Infrastructure	Operating Environment	Research	Development	Government Strategy	Commercial	Total rank
United Kingdom	5	8	1	3	11	7	4	3
Canada	4	23	5	8	10	4	5	4
Germany	9	12	7	4	12	5	9	5
France	8	30	2	12	9	6	7	6
Singapore	2	4	39	16	15	30	6	7
South Korea	28	5	30	22	3	31	25	8
Japan	26	16	17	6	7	12	8	9
Ireland	6	2	28	28	6	42	20	10
Australia	11	32	8	18	4	39	12	11
Israel	13	36	50	5	5	46	3	12
Switzerland	10	14	47	7	14	28	10	13
Finland	20	9	38	11	17	16	15	14
Spain	19	10	21	31	24	9	28	15
The Netherlands	7	17	13	34	8	22	22	16
Denmark	16	19	25	13	20	23	23	17
India	3	46	20	24	13	25	13	18
Luxembourg	25	29	26	10	25	11	17	19
Sweden	15	26	19	9	22	44	16	20

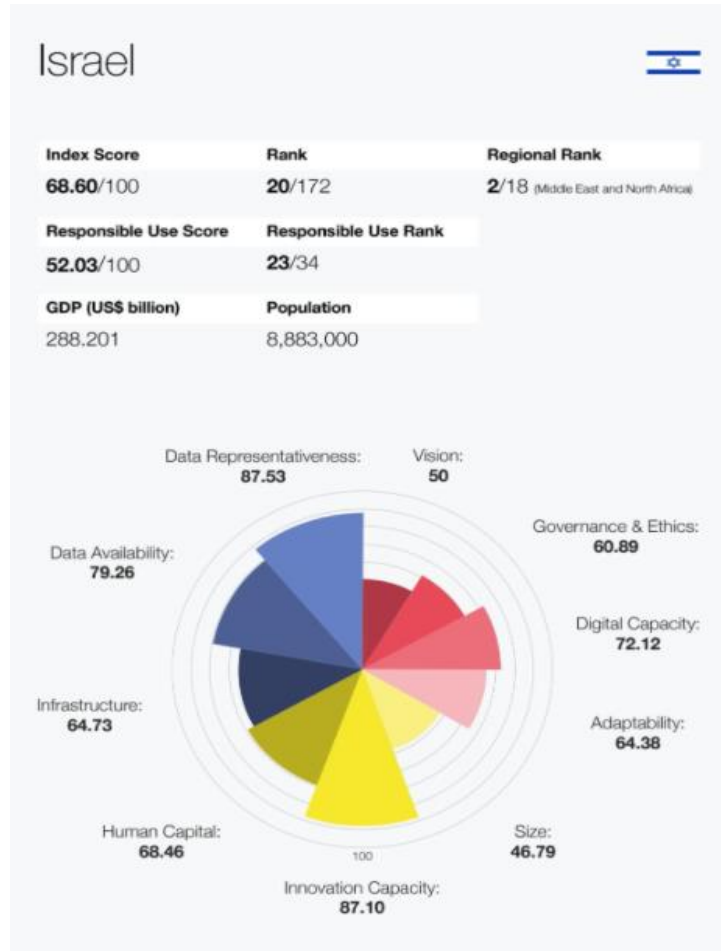
איור 2: 20 המדינות המובילות ב-AI לפי מדד tortoise בשנת 2019

בשנת 2020 ערכה Tortoise אינדקס מעודכן, בו ישראל דורגה במקום 5, כאשר הדירוג שלה בסעיפים השונים היה: כישרונות 5, תשתית 27, סביבת הפעלה 38, מחקר 7, פיתוח 7, אסטרטגיה ממשלתית 51 ומסחרי 3. כפי שניתן לראות, הדירוג הכללי של ישראל השתפר אך התשתיות, ההטמעה והמדיניות הממשלתית נותרו בדירוג נמוך מאוד. כמו כן, ניתן לראות שבמחקר ופיתוח, אשר מעידים על חדשנות, ישראל ירדה בדירוג בשנה שחלפה. בגין אופן החישוב של האינדקס, והפיטורין בענף ההייטק ישנה זמינות גבוהה יותר של הון אנושי. מדיניות הממשלה, כפי שבאה לידי ביטוי לנוכח הקורונה, בתעסוקה ובחינוך – ירדה בדירוג, ולעומת זאת ההשקעות הפרטיות עלו. המלצות הוועדה בנושאי התשתית והרגולציה אמורות לפתור את הפערים הללו, ולעודד מצטיינים שיצאו משוק העבודה ללמוד לתארים מתקדמים.

Country	Talent	Infrastructure	Operating Environment	Research	Development	Government Strategy	Commercial	Total rank
United States of America	1	1	44	1	2	10	1	1
China	19	2	6	2	1	3	2	2
United Kingdom	4	17	23	3	16	13	4	3
Canada	8	15	34	5	19	4	5	4
Israel	5	27	38	7	7	51	3	5
Germany	13	10	21	4	18	8	12	6
The Netherlands	6	7	13	15	10	29	21	7
South Korea	20	5	55	28	3	7	22	8
France	10	13	16	19	17	6	8	9
Singapore	3	3	56	21	21	17	6	10
Finland	16	21	11	9	22	5	19	11
Ireland	7	6	7	36	6	46	24	12
Australia	12	34	28	14	5	18	14	13
Japan	28	8	35	6	8	19	7	14
Sweden	9	14	27	16	20	38	20	15
Denmark	11	16	17	24	29	23	23	16
Luxembourg	23	19	5	30	13	27	15	17
Switzerland	15	11	53	10	4	54	11	18
New Zealand	17	20	10	13	39	33	31	19
India	2	59	33	27	11	36	10	20

איור 2א: 20 המדינות המובילות ב-AI לפי מדד *tortoise* בשנת 2020

במחקר שנערך על ידי מרכז IDRC (International Research Development Centre) הקנדי ו-Oxford Insights, נוצר אינדקס המדרג את כלל מדינות העולם לפי מוכנות ממשלתית בתחום הבינה המלאכותית. הדוח כולל ביקורת נוקבת על הפער והפיגור של ישראל, ומדרג אותנו במקום ה-20 בעולם. הדוח גם מצין שישראל צריכה להשתפר בהיבטי זמינות נתונים, בדגש על נתונים ממשלתיים.



איור 3 : פירוט הדירוג של ישראל לפי מדד IDRC

Annex 1: Full Rankings

Rank	Country	Score
1	United States of America	85.479
2	United Kingdom	81.124
3	Finland	79.238
4	Germany	78.974
5	Sweden	78.772
6	Singapore	78.704
7	Republic of Korea	77.695
8	Denmark	75.618
9	Netherlands	75.297
10	Norway	74.430
11	France	73.767
12	Australia	73.577
13	Japan	73.303
14	Canada	73.158
15	Luxembourg	72.616
16	United Arab Emirates	72.395
17	Estonia	69.922
18	Switzerland	69.219
19	China	69.080
20	Israel	68.825
21	Ireland	68.244
22	Austria	68.230
23	New Zealand	68.146
24	Spain	68.041

איור 4: פירוט 24 המדינות המובילות לפי מדד IDRC

2.3. סיכום

בינה מלאכותית מבטיחה להעלות את הפיריון ולעודד צמיחה, הן עבור חברות והן עבור כלכלות שלמות. בעוד מספר תעשיות כבר הטמיעו פיתוחים משמעותיים המבוססים על יכולות בינה מלאכותית, תעשיות אחרות נותרות מחוץ לטכנולוגיה והיכולות שהיא מביאה. בטחון המדינה, מחייב מדיניות סדורה לשילוב בינה מלאכותית במערכת הביטחון.²⁰

הסיבות המרכזיות שגורמות לכך שלא נעשה שימוש מספק או בכלל בבינה מלאכותית הן: חוסר תמרוץ מצד הדרג הניהולי, התנגדות לשינוי ותרבות ארגונית לא מקבלת, חוסר מוכנות של הנתונים, מחסור באנשי מקצוע העוסקים בבינה מלאכותית, בדגש על כאלו שמבינים בתרגומה למטרות הארגון²¹, ובעיות רגולטוריות ואתיות. הבעיות הללו באות לידי ביטוי בכלל התעשיות בארץ, והדרך הטובה ביותר לפתור אותן היא דרך השקעה באימוץ סביבת בינה מלאכותית כמדיניות ממשלתית, בדומה למה שנעשה בתוכניות הלאומיות שפורטו כאן. השקעה בתשתיות ובהון אנושי, תוך יצירת הסכמות רגולטוריות תסייע להביא את ישראל למעמד תחרותי.

²⁰ <https://www.themarket.com/technation/premium-1.9216588>

²¹ <https://hbr.org/2018/02/you-dont-have-to-be-a-data-scientist-to-fill-this-must-have-analytics-role>

3. תשתיות רלוונטיות

3.1.1. רקע

כשלי השוק המרכזיים בעולמות הבינה המלאכותית ולמידת המכונה הם: השגת הון אנושי מנוסה, נגישות למאגרי מידע (בעיקר מאגרים ממשלתיים ורפואיים), נגישות לכוח מחשוב רב בזמינות גבוהה ובמחיר סביר והתחברות לרשת חישוב ציבורית בינלאומית כדוגמת PRACE²².

כפי שראינו בשיח, הן עם האקדמיה והן עם החברות בתעשייה, ככל שהטכנולוגיה מתקדמת וסך המידע הדיגיטלי הזמין והנדרש להוצאת תובנות עולה, כך עולות העלויות והצרכים הנדרשים על מנת לשמור על יתרון תחרותי. תשתית חישוב משמעותית יכולה לסייע בכל ההיבטים ולהאיץ את הערך שבעיסוק בבינה מלאכותית במדינת ישראל במספר דרכים. ראשית, התשתית תשמש להכשרת כ"א ייעודי הנדרש בתעשייה ובאקדמיה להקמת ותחזוקת תשתיות חישוב. שנית, היא תהווה פלטפורמה לאומית לגופים ציבוריים ותדרבן אותם להנגיש מאגרי מידע בסמיכות לכוח העיבוד ובגבולות ישראל. בנוסף, הקמת תשתית תאפשר סבסוד שירותי אחסון מידע ועיבודו כבסיס לפתרון בעיות אופטימיזציה לכלל שותפי התשתית. כמו כן, יש לבחון חיבור חלק מהתשתית המקומית לרשת האירופאית.

מודל ההקמה עליו הוחלט בנוגע לתשתית מחשוב העל הוא של תשתית דואלית. תוקם תשתית מידע לא מסווג לכלל התעשייה והאקדמיה, תשתית זו תהיה מחויבת ברגולציה מוגדרת לצידה של תשתית עם מידע מסווג לשימוש מערכת הבטחון והתעשיות הביטחוניות. לצד תשתיות המידע תוקם תשתית חישוב ניסויי שתאפשר בחינת חומרה ואלגוריתמים. תשתית כזו אינה קיימת בארץ עבור כל העוסקים במחקר ופיתוח באקדמיה, בתעשייה ובמערכת הבטחון. כיום לא קיימת להם נגישות לתשתית הנדרשת למחשוב על לצורך ניסוי ותהייה. תיעדוף הגישה למשאבי חישוב העל ייתן באופן יחסי לתרומה בפועל להקמת התשתית ולמימון פעילותה השוטפת. יהיה ניתן להרחיב תשתית זו על פי צריכת המשאבים השונים בפועל מתוקף היותה תשתית משותפת וגמישה. התשתית תהיה דינמית ומשודרגת, כאשר יוגדר תקציב המשך קבוע לצורך חידושים עתידיים.

רכישת הציוד, התחזוקה והייעוץ, יתבצע בשיתוף חברות תעשייה שונות, והעלות המוערכת נכון למועד כתיבת שורות אלו תהיה של כ-50 מיליון דולר ארה"ב מידי שנה.

פרק התשתית יעסוק בשלושה תחומים מהותיים לנושא התשתית הנדרשת בהם דנה הוועדה: מחשוב על, מחשוב ענן ונתונים.

3.2. מחשוב על (HPC)

3.2.1. רקע

לשם ביצוע חישובים מורכבים בתחום הבינה המלאכותית נדרש כוח חישוב רב, המסוגל לטפל גם במידע מסוגים שונים בנפחים גדול מאוד. למעשה, צרכי החישוב בבינה מלאכותית והצורך בפתרון בעיות חישוב מורכבות יותר ויותר, דוחפים קדימה את שוק המחשוב באופן עקבי, בין היתר במעבר ליחידות עיבוד מסוג GPU ותכנון מאיצים חדשים אשר יהיו מותאמים במיוחד לחישובי בינה מלאכותית. מעל

²² יוזמת מחשוב העל האירופאית: <https://prace-ri.eu/>

לחצי מהנתונים של העולם נוצרו בשנתיים האחרונות, אך רק פחות משני אחוזים מתוכם נותרו. הדאטה גדל בצורה מעריכית, ואיתו גדלה הדרישה לחישוב, אחסון ותשתית להעברתו ברשת²³. כמו כן, לפני 2012, תוצאות בינה מלאכותית תאמו את חוק מור, עם הכפלת מחשוב אחת לשנתיים. החל מ-2012, המחשוב הוכפל כל 3.4 חודשים.²⁴

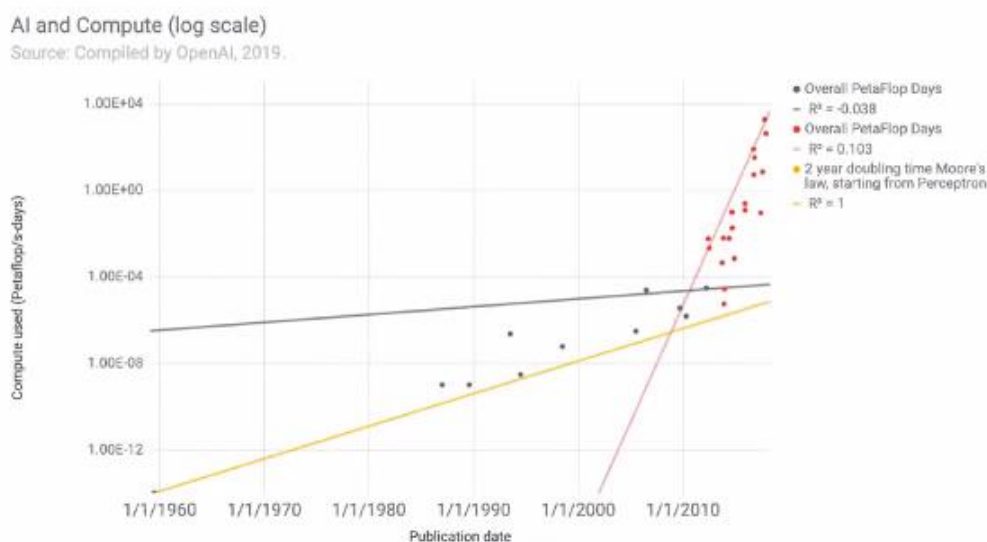


Fig. 3.14b.

איור 5: קצב הכפלת יכולות המחשוב בבינה מלאכותית

השינויים הטכנולוגיים מתפתחים בקצב מסחרר. בנושא של עיבוד תמונה למשל, זמן אימון לאבחון הנתונים בה הצטמצם מ-22 ms בנובמבר 2018 ל-0.82 ms ביולי, 2019, ועלות העיבוד הפכה לנמוכה מאוד (בעוד שכמות התמונות לאימון וכמות התובנות שמחולצות עלתה.²⁵

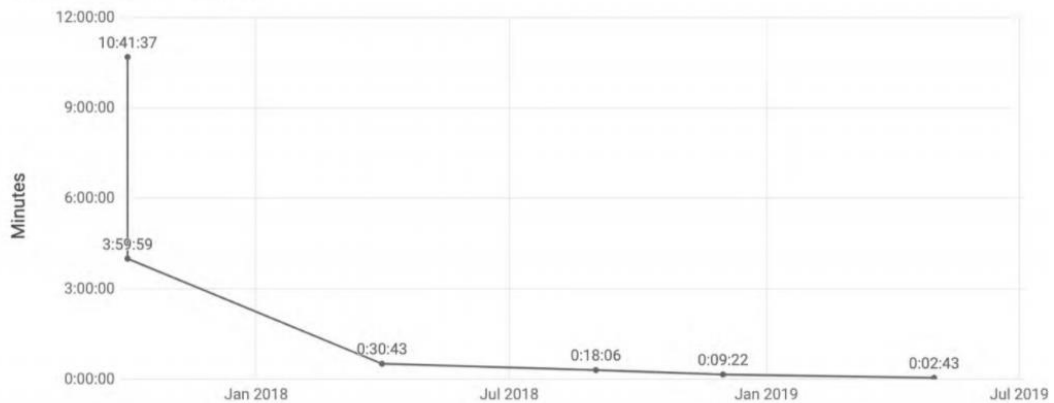
²³ מתוך מצגת של אבנר גורן, חבר הוועדה, שהועברה במסגרת דיוני הוועדה

²⁴ <https://hai.stanford.edu/research/ai-index-2019>

²⁵ <https://hai.stanford.edu/research/ai-index-2019>

ImageNet training time (October 2017 – November 2019)

Source: Stanford DAWN Project, 2019.



איור 6 : זמני האימון בעיבוד תמונה ב- *ImageNet* בין אוקטובר 2017 לנובמבר 2019.

תשתית עיבוד תספק יכולת לפתח ידע חדש בצורה נגישה ובעלות סבירה אך לשם כך עליה להישאר מעודכנת ורלוונטית להתפתחויות הטכנולוגיות כדי לשמר את היכולת לחקור ולחדש באופן עצמאי. חשוב להדגיש כי הצורך בעדכון הדינמי אינם רק בכוח חישוב עצמו, אלא גם במבנה האחסון, תקשורת, תמיכה ועוד.

מחשוב על במדינות אחרות ובארץ

ניתן ללמוד רבות מהגישות בארה"ב, אירופה, יפן, סין, ספרד, סינגפור, פינלנד, צ'כיה ומדינות נוספות. הגישה השלטת הינה לרוב בניית מרכזי חישוב במודל ממשלתי, ציבורי או איזורי, אשר מאפשרים חישובים מורכבים לאקדמיה ולתעשייה, לצד בניית מרכזי מחקר בנושאי חישוב אשר יעודדו חדשנות וינגישו יכולות חישוב מתקדמות לכלל המשתמשים. בארץ אין מרכז חישוב על לכן נוצר פער גדול בין ישראל לעולם. רבים מהחוקרים באקדמיה ובתעשייה נסמכים על יכולות חישוב במקומות אחרים בעולם. שיתופי פעולה אלה מוגבלים במציאת חוקרים וגופים החולקים את העניין במחקר ובתוצאותיו ואינם נגישים אחרת.

לפי הנתונים שעולים מאתר TOP500 המרכז את הידע על מחשבי העל בעולם, סין וארה"ב מובילות בתחום, ובין 10 המחשבים החזקים ביותר מצויים מחשבים שממוקמים גם בשוויץ, יפן וגרמניה. בנוסף למדינות המובילות, גם מדינות קטנות (כדוגמת פינלנד, צ'כיה וסינגפור) הן בעלות מחשבי על²⁶.

בטבלה מס' 2 מוצגים עשרת מחשבי העל המובילים לפי אתר TOP500, נכון ליוני 2020:

²⁶ לרשימת ה-500TOP העדכנית ליוני 2020 : <https://www.top500.org/lists/top500/2020/06/>

מדינה	דירוג	שם	יצרן	מס' ליבות	זיכרון (GB)	מעבד	חיבור	מיקום
יפן	1	Supercomputer Fugaku	Fujitsu	7,299,072	4,866,048	A64FX 48C 2.2GHz	Tofu interconnect D	RIKEN Center for Computational Science
ארה"ב	2	SUMMIT	IBM	2,414,592	2,801,664	IBM POWER9 22C 3.07GHz	Dual-rail Mellanox EDR Infiniband	DOE/SC/Oak Ridge National Laboratory
ארה"ב	3	SIERRA	IBM	1,572,480	1,382,400	IBM POWER9 22C 3.1GHz	Dual-rail Mellanox EDR Infiniband	DOE/ANNSA/LLNL
סין	4	SUNWAY TAIHULIGHT	NRCP C	10,649,600	1,310,720	Sunway SW26010 260C 1.45GHz	Sunway	NATIONAL SUPERCOMPUTING CENTER IN WUXI
סין	5	TIANHE-2A	NUDT	4,981,760	2,277,376	Intel Xeon E5-2692v2 12C 2.2GHz	TH Express-2	NATIONAL SUPERCOMPUTER CENTER IN GUANGZHOU
איטליה	6	HPC5	Dell EMC	669,760	349,440	Xeon Gold 6252 24C 2.1GHz	Mellanox HDR Infiniband	Eni S.p.A
ארה"ב	7	SELENE	Nvidia	272,800	560,000	AMD EPYC 7742 64C 2.25GHz	Mellanox HDR Infiniband	NVIDIA Corporation
ארה"ב	8	FROTERA	Dell EMC	448,448	1,537,536	Xeon Platinum 8280 28C 2.7GHz	Mellanox InfiniBand HDR	Texas Advanced Computing

Center/Univ. of Texas								
CINECA	Dual-rail Mellanox EDR Infiniband	IBM POWER9 16C 3GHz	252,928	347,776	IBM	MARCONI	9	איטליה
Swiss National Supercomputing Centre (CSCS)	Aries interconnect	Xeon E5-2690v3 12C 2.6GHz	365,056	387,872	Cray/HPE	Piz Daint	10	שוויץ

טבלה 2: עשרת מחשבי העל המובילים לפי אתר TOP500

בבואנו להגדיר את צרכי מרכז החישוב לשימושי בינה מלאכותית בישראל עלינו להידרש לשתי שאלות: הראשונה היא מיהם המשתמשים שלנו. השנייה היא מה הצרכים שלהם בנוגע לנושאים הבאים: גישה למידע משותף, חיסכון בעלויות, יכולת פתרון בעיות גדולות בקנה מידה רחב, אילוצי סיווג, חיסיון וביטחון, ביצוע בדיקות test bed וחדשנות, תמיכה בקהילה, חינוך ואימון.

על מנת למקם את דרישות הסף לקידום מחקר ופיתוח תחרותי, בחרנו לבחון שלוש תשתיות למחשוב על, על בסיס הפרמטרים הבאים: מודל הצטיידות ושדרוג, רמת נצילות, סוגי המשתמשים והתפלגותם, מודל חיובי של כל סוג לקוחות, תחזוקה טכנית, צוות התמיכה בחוקרים, כמות GPU, כוח עיבוד, אנרגיה ומבנה החומרה.

בטבלה זו מוצגים הנתונים אודות מחשבים בגרמניה, אנגליה, ומדינה נוספת במפרץ הפרסי, כפי שהתקבלו בשיחות עם מרכזי החישוב עצמם:

מדינה	מודל הצטיידות ושדרוג	רמת נצילות	סוגי המשתמשים והתפלגות	מודל חיובי של כל סוג לקוחות	תחזוקה טכנית	צוות התמיכה בחוקרים	כמות GPU	כוח עיבוד אנרגיה ומבנה החומרה	מבנה החומרה
גרמניה ²⁷	הממשלה הקציבה 5 מיליארד יורו ל-5 שנים	100% מבצעים רכש נוסף בימים אלו	1100 חוקרים ו- PHD (מתכננים לגדול ל-2000 חוקרים)	כל גורם קיבל הקצאה ממחשב העל בהתאם להשקעה	0.5 FTE from the facility DC	2 FTE	600	0.5MV	DGX
אנגליה	התקצוב גדל בהתאם לתקציב הממשלה. לדוגמא, כרגע הוקם מיזם בריאות עם kings college לחקר הסרטן על פלטפורמות DGX חדשות	100% מבצעים רכש נוסף בימים אלו	אין מידע בנושא	כל גורם קיבל הקצאה ממחשב העל בהתאם להשקעה	0.5 FTE from the facility	3 FTE	800	0.5MV	DGX
מדינה במפרץ הפרסי	אין מידע בנושא	אין מידע בנושא	אין מידע בנושא	אין מידע בנושא	אין מידע בנושא	אין מידע בנושא	3200	1.7MV	DGX

<https://www.dfki.de/>²⁷

גם לאחר עשרות פניות למרכזים בעולם התקשנו לקבל מספרים מדויקים אודות הנתונים. דוגמא למרכז עליו הצלחנו לקבל מידע היא מרכז בברצלונה המאופיין ב-165 אלף ליבות, PF 11.15, TB 390 של זיכרון ראשי, PB 14 של אחסון, בנוסף גישה למרכז נוסף שנותן ²⁸PB 24.6.

היות ועבודת הוועדה עוסקת בסוגי משתמשים שונים (ביטחון, הכשרות, תעשייה, משרדי ממשלה ואקדמיה) כאשר לכל אחד מהם ישנם שימושים וצרכים שונים, קיים צורך להתעמק בצרכים של כל משתמש. להלן טבלה המתארת את צרכי המחשוב של המשתמשים השונים:

<https://www.bsc.es/>²⁸

מוטיבציה עיקרית למחשוב-על לאומי	סוג הכוח החישובי	פורמט המידע	סוג המשימות הנדרשות	סוג המשתמש
הפחתת עלויות בהשוואה לענן; הרצת מודלים גדולים שלא ניתן להריצם על חומרה פרטית; מינוף צוות התמיכה של מחשוב העל; מינוף פרויקטי תוכנה קהילתיים	מספר <i>CPUs</i> ומאות או אלפי <i>GPUs</i>	<i>Double precision floating points</i> ; רמה גבוהה של מיקבול מידע	מודלים פיזיקליים; סימולציה	אקדמיה – מחשוב על
הפחתת עלויות בהשוואה לענן; תמיכה ברשתות גדולות יותר, מעבר ליכולת חומרה פרטית	מנועי אימון (מאות <i>GPUs</i> או מאיצים)	סוגי מידע לאימון <i>32/16 float</i>	אימון	אקדמיה – בינה מלאכותית
בדיקת מוצרים בסביבת <i>Data center</i> מלאה	חיבור לציוד של <i>Data centers</i>	סוגי מידע <i>AI</i>	מאגרי מידע הסקה/אימון	בינה מלאכותית – סטארטאפים <i>Silicon/Servers</i>
הפחתת עלויות בהשוואה לענן; במרבית המקרים מסתפקים ברשתות בגודל סביר	מנועי אימון (מאות <i>GPUs</i> או מאיצים)	סוגי מידע לאימון <i>32/16 float</i>	אימון	סטארטאפים – אפליקציות בינה מלאכותית
מעדיפים ענן ציבורי לטובת הגמישות וה- <i>Scale</i> שדרושים להם לטובת רשתות הגדולות במיוחד	מנועי אימון (מאות רבות של <i>GPUs</i> או מאיצים)	<i>Training data types 32/16 float</i>	אימון	אפליקציות בינה מלאכותית לחברות (למשל מובילאי)
הפחתת עלויות בהשוואה לענן; תמיכה ברשתות גדולות יותר, מעבר ליכולת חומרה פרטית; מינוף תשתית אזרחית	מנועי אימון (מאות רבות של <i>GPUs</i> או מאיצים להסקה <i>(Inference)</i>)	סוגי מידע לאימון <i>32/16 float</i> סוגי מידע להסקה <i>(Inference) INT8</i>	אימון	צה"ל
הפחתת עלויות בהשוואה לענן; מינוף ה- <i>Ecosystem</i> המסחרי	מנועי אימון (עשרות <i>GPUs</i> או מאיצים להסקה <i>(Inference)</i>)	סוגי מידע לאימון <i>32/16 float</i> סוגי מידע להסקה <i>(Inference) INT8</i>	הסקה ואימון	משרדי ממשלה

אימון והכשרות לאנשי מקצוע בתחומי ה-AI				הכשרות

טבלה 4: טבלה המפרטת את הצרכים של המשתמשים השונים

3.3. ענן ציבורי

3.3.1. רקע

מחשוב ענן הוא מודל המאפשר גישה נוחה, מכל מיקום לפי דרישה למאגר משותף של משאבי מחשוב ונתונים הניתנים להגדרה (למשל רשתות, שרתים, אחסון, יישומים ושירותים) שניתן לספק במהירות ולשחררם במאמץ ניהול מינימלי או באינטראקציה של ספק שירות. בתחילה, שימוש בענן היה משני, אך כיום הוא נמצא בליבת מערכי המחשוב, ומאפשר לנצל יכולות של בינה מלאכותית ולמידת מכונה, רובוטיקה, ניתוח מעמיק והגירת בסיסי נתונים.

היתרונות המשמעותיים של שימוש בענן הם: שימוש לפי צריכה בלבד, שקיפות ובקרה טובים יותר של הצריכה, התאמה גמישה יותר, ייתור הצורך לבנות ולהריץ בסיס נתונים וחסכון בחלק מההון האנושי הנדרש. חסרונות לשימוש בענן הם: כאשר מדובר בכמות של 8 מאיצים או יותר ובכמויות נתונים גדולות הוא הופך לפחות אטרקטיבי בעלות ובביצועים בטכנולוגיות של היום. בינה מלאכותית מניעים נתונים רבים, העברת נתונים לצורך עיבוד מהענן ואליו יקרה ויוצרת בעיות עלות, ביצועים וסוגיות אבטחה ורגולציה.

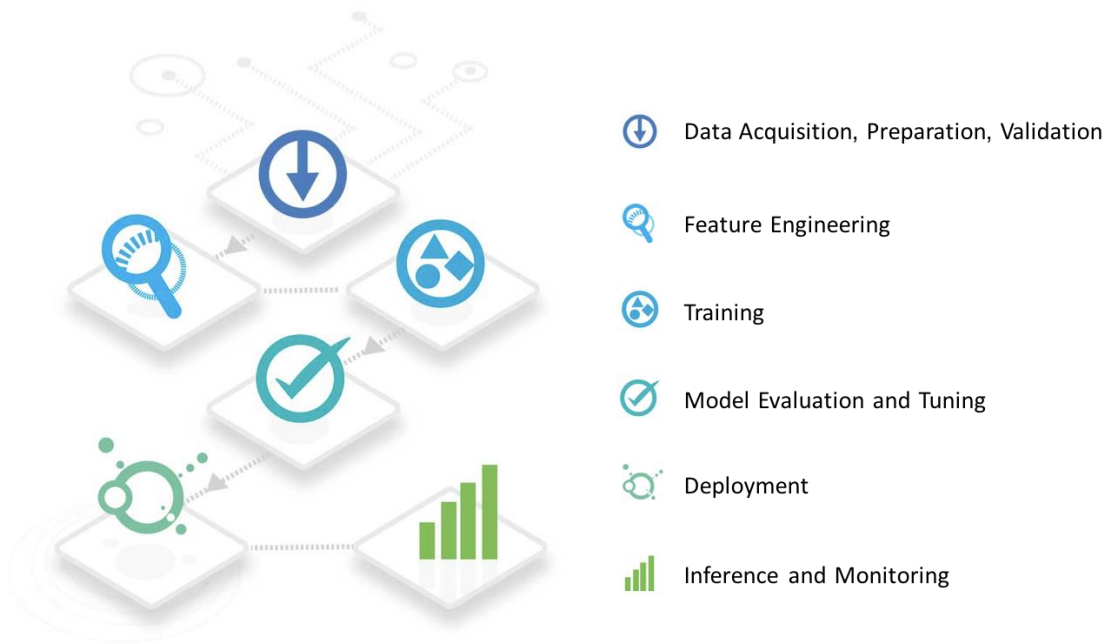
מודל נוסף שניתן לשקול שיאפר גמישות מוגברת הוא מודל היברידי הכולל נגישות לפלטפורמת ענן בשלב אפיון צרכי המחשוב, טרם הרחבתם, תוך שמירה על מגבלות צריכה. מודל זה מאפשר גמישות למשתמשים להרחבת משאבי החישוב והאחסון לענן, טיפול בפערים המובנים בין גרסאות התכנה בענן ו-On-Premise (בענן ציבורי קיימות, בדרך כלל, הגרסאות המתקדמות יותר), ומודל אלטרנטיבי המציג מערכת פיתוח קטנה לצד מערכת הפיתוח הגדולה²⁹.

שירותי פיתוח בענן הם שירותים או יישומים אשר ניתנים על ידי ספק הענן, המאפשר נגישות לשירותיו תוך שקיפות לתשתיות החישוב, הרישות והאחסון. כך מתאפשר לצוותי המחקר, הפיתוח והיישום למנף מודלים בבינה מלאכותית דרך API (Application Programming Interface) בלי צורך במומחיות בתחומי תשתיות הענן או במדעי הנתונים. מודלים אלו מספקים שירותים עם יכולות שפה, תמונה, ולמידת מכונה אוטומטית. שירותים אלה זמינים לרוב באמצעות גישה לממשק API ספציפי והם בדרך כלל מתומחרים על פי מספר שיחות API (בדומה לשיחות טלפון). במקרים מסוימים, ניתן להשתמש בשירותים באמצעות כלי תצורה משולבים. דוגמאות לשירותים אלה כוללים הבנת שפה טבעית, זיהוי תמונה ויצירת מודל למידת מכונה³⁰.

²⁹ מתוך המצגת של ארווין לייבוביץ שהוצגה במסגרת דיוני הוועדה
³⁰ Gartner Magic Quadrant for Cloud AI Developer Services, 24 February 2020

תשתית ענן עבור בינה מלאכותית ולמידת מכונה:

בניית תשתית ענן עבור פעילויות בינה מלאכותית ולמידת מכונה באופן מיטבי דורשת הקפדה על סדר זרימה של תהליכים מרובים. נדרש לנקות ולתייג נתונים, לחלץ את התכונות המתאימות עבור ביסוס החיזוי ולבסוף נדרש לאמן את המודל, לבצע ולידציה (כולל מחזוריות לשיפור המודל) ולאחר מכן לבצע שימוש במודל תוך שיפור מתמיד ביכולותיו. רצף זה הינו מורכב ביותר גם עבור הארגונים המיומנים ביותר מבחינה טכנית.



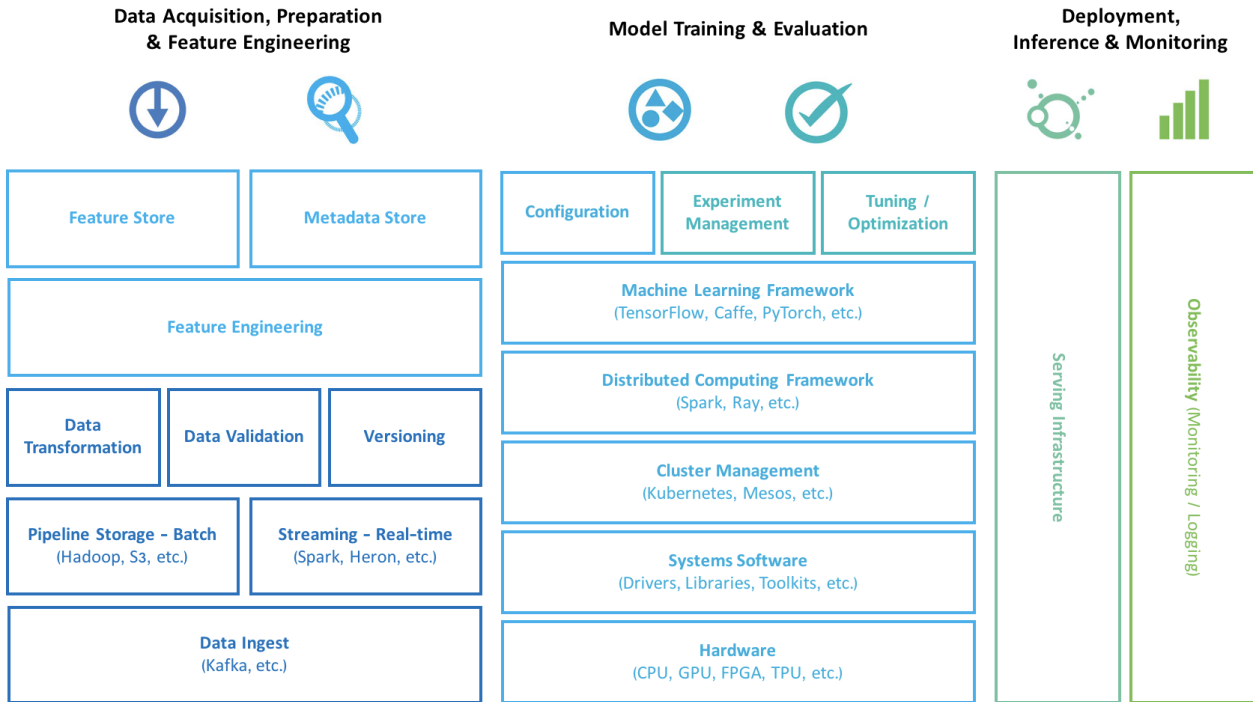
איור 5: תרשים המציג את רצף העבודה בבניית תשתית ענן

על מנת להביא את היכולות הללו למלוא הפוטנציאל נדרש לבצע הפשטה של ממשקים, מערכות וכלי עבודה חדשים כדי לפשט את הפיתוח והפריסה של יישומים עבור המפתח, ולכן בכל שכבה בארכיטקטורה נדרש טיוב לטובת התהליך. על הענן לאפשר נגישות אל:

- חומרה ייעודית המכילה ליבות עיבוד רבות, זיכרון רב ובעל רוחב פס גבוה.
- שבבים אשר מותאמים לחישוב מקבילי בקצב גבוה ולביצוע פעולות מתמטיות.
- תשתיות למחשוב מבוזר לאימונים וליישום האלגוריתם גם על פני צמתי חישוב מרובים.
- ממשק המאפשר הבניית זרימת מידע אמינה (Pipeline) ואחידה על מנת לאפשר יצירה, שחזור וניהול נתוני אימון וחיזוי.
- תשתית אחסון המאפשרת ביצועים בקצב גבוה והשהייה נמוכה על מנת לאפשר פעולות במהירות על בסיס בנתונים ובזמן אמת.

• נגישות לביצוע פעולות על מודלים מבצעיים בסדרי גודל רחבים של:

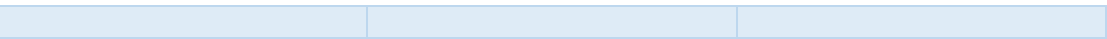
- בקרת איכות
- איתור ותיקון באגים
- כוונון ואופטימיזציה



איור 6: תרשים המציג את רצף העבודה בבניית תשתית ענן

בתרשים זה ניתן לראות את אבני הבניין המרכזיות הנדרשות לטובת הקמת תשתית לחישוב עבור בינה מלאכותית ולמידת מכונה, משלב הכנסת המידע ועד לפריסת המודל והפקת תובנות. להלן הסבר עבור כל אחת מאבני הבניין בשרטוט:

- Metadata Store - מידע שמתאר את משתני המידע עליו המודל עובד.
- Feature Store - ספרייה של פונקציות שיוצרות פיצורים מהמידע.
- Feature Engineering - שימוש בידע על התחום הטכנולוגי אליו שייך הדאטה בכדי ליצור יכולות נוספות.
- Data Transformation - תהליך המרת הנתונים מתבנית אחד לאחרת.
- Data Validation - בדיקה שהמודל רץ על מידע נקי, נכון ושימושי.
- Versioning - יצירה וניהול של מספר גרסאות של המודל.



- Pipeline Storage – Batch - אחסון מידע בנפחים גדולים. גישה ישירות לדיסק.
- Streaming – Real-time - אחסון מידע בנפחים גדולים. גישה בזמן אמת.
- Data Ingest - תהליך קבלה וגישה למידע לשימוש ואחסון מידי.
- Configuration - בחירת מודל התחלתי ופרמטרים ראשוניים.
- Experiment Management - ניהול התוצאות, הקוד, והמודלים המתקבלים ושמירה שלהם.
- Tuning / Optimizations - שיפור המודל בהתאם לתוצאות על סט הבדיקה בעזרת שינוי פרמטרי על.
- Machine Learning Framework - סביבות עבודה שעוזרות להריץ ולבחון את המודלים בצורה יעילה, קלה ומהירה.
- Distributed computing Framework - מסגרת מחשוב מבוזרת לניהול משאבי המחשוב.
- Cluster Management - ניהול היכולת לבצע מקבול תהליכים לכינוס האלגוריתם.
- System Software - תוכנות מערכת ההפעלה שמפעילות את החומרה.
- Hardware - החומרה שמבצעת את החישובים והפעולות של המודל.
- Serving Infrastructure - התשתית שמאפשרת את יישום האלגוריתם הנבנה בשטח.
- Observability - שמירה והצגה בצורה שקופה של המידע הרלוונטי לתפעול כלל המערכות.

ענן ציבורי במדינות שונות

בארה"ב נחתם לאחרונה הסכם בין אוניברסיטאות מובילות לחברות טכנולוגיות שתאפשר לאקדמאים ומדענים אחרים גישה למשאבים שעד כה היו זמינים לענקי הייטק בודדים בלבד. היוזמה, שנקראת ענן המחקר הלאומי, קיבלה את תמיכת שתי המפלגות והוגשו הצעות חוק להקמת צוות משימה שיכלול מדענים ממשלתיים, אקדמאים ונציגי תעשייה שיבנו ויממנו את המהלך³¹. כמו כן, בבריטניה לאחרונה נעשה שימוש בריבוי עננים³².

3.3.2. ענן ציבורי בישראל

קיימים כעת בארץ מספר פרויקטים ממשלתיים הנוגעים לענן. הפרויקט הראשון מכונה נימבוס³³. נימבוס היא תוכנית לרכישת ענן ציבורי למשרדי הממשלה, היות וחלק ממעצמות הענן הבינלאומיות

³¹<https://www.nytimes.com/2020/06/30/technology/national-cloud-computing-project.html>

³²<https://ukcloud.com/>

³³<https://govextra.gov.il/nimbus-mr-gov-il>

מקדמות בימים אלו תשתית ענן בישראל. מתוכננים גם שלבי ביניים בהן חלק משרותי הענן עשויים להינתן ע"י הזוכה כשירותים מחו"ל והשאר מישראל. ביקשנו ששירותי הבינה המלאכותית הרשומים לעיל יכללו בהיצע לנימבוס כדי לעמוד בדרישות רגולטוריות לגבי הנתונים וכדי לספק ביצועים כנדרש. שלב מעבר משירותים הניתנים מחו"ל לשירותים הניתנים מתשתיות מלאות בארץ לא אמור לקחת מעל שנה, כאשר הכוונה היא לייצר רגולציה ותמריצים שיחייבו את משרדי הממשלה לעבור לענן הישראלי ובכך להבטיח הנגשה העומדת בכללי הרגולציה לנתונים שלהם. כך ניתן יהיה לקצץ בעלויות של הפעלת רשתות מחשוב פרטיות ועוד.

פרויקט נוסף הוא תמנע. תמנע הינו פרויקט של משרד הבריאות לאיסוף נתונים רפואיים ומתן נגישות אליהם דרך ענן ציבורי, תוך שמירה על פרטיות המשתמשים ומתוך מטרה לאשר מו"פ רפואי לצורך קביעת מדיניות, פיתוח טיפולים רפואיים ומעקב אחר מגמות. 15 גופי בריאות חתומים על העברת מידע קליני למערכת ובהם בתי חולים וקופות חולים. פרויקט זה לא קודם בקצב סביר מסיבות ניהול ורגולציה.

הפרויקט השלישי הוא פסיפס. פסיפס הוא פרויקט לאומי משותף למשרד הבריאות, רשות החדשנות, ישראל דיגיטלית, ות"ת, משרד המדע ומפא"ת, שמטרתו הקמה של מאגר מידע גנומי וקליני שיאפשר מו"פ להוצאת תובנות לצד שיפור יכולות האבחון והטיפול בחולים. גם פרויקט זה נמצא בתחילת דרכו ויוכל להנות מהתשתיות המפורטות מעלה.

בימים אלה הוקם חל"צ בבעלות האוניברסיטאות בישראל שיארח את פסיפס. עד כה נחתם הסכם עם קופות חולים מאוחדת ולאומית בהן חברים למעלה מרבע מאוכלוסיית ישראל, כדי לאפשר העברת מידע קליני מהקופות. כמו כן בוצע תכנון של התשתית התפעולית, המאפשרת לקופות גיוס תורמים והעברת המידע הקליני הנדרש בהתאם לפרוטוקול שגובש וסוכם. בנוסף מתגבש הסכם התקשרות להקמת מרכז מצוינות לאומי לריצוף גנומי. לאחר הוקמה וועדה מדעית מייעצת בראשותו של פרופ' קרל סקורצקי. עבודה מאומצת של בכירי החוקרים והגנטיקאים במדינה, הגישה הוועדה המלצה למיפוי תמהיל תת האוכלוסיות במדינת ישראל והחלה הקמתה של תשתית טכנולוגית ומערך המידע.

למרות כל היוזמות הללו, כעת אין עדיין תוכנית ענן בישראל. עובדה זו מאלצת גופים אשר מעוניינים להזיל עלויות ולהשתמש בענן להעביר את הפעילות בענן שלהן לחו"ל.³⁴ לאור העובדה שאין בישראל מרכז חישוב לאומי בעל כוח רב שיכול להוות חלופה, והחשיבות של שימוש בענן לפיתוח בינה מלאכותית, עלינו לנצל את העובדה שישנן חברות שמקדמות הקמת תשתית ענן בישראל בימים אלו ולממש יוזמה לשימוש בענן. השימוש בשירותי ענן שנמצאים בישראל יאפשרו אחסון נתונים אשר מטעמי פרטיות ורגולציה צריכים להישאר בתוך המדינה. בנוסף, קיומו של ענן מקומי אשר יעמוד בדרישות האבטחה, עמידות ורציפות תפעולית יאפשר סגירת פער מול מדינות מתקדמות יותר בהיבטי תשתיות מחשוב.³⁵

³⁴ https://www.themarker.com/technation/.premium-1.9278577?utm_source=App_Share&utm_medium=iOS_Native
³⁵ <https://www.globes.co.il/news/article.aspx?did=1001315732>

3.4. נתונים

3.4.1. רקע

אחד החסמים להתקדמות מהירה בתחומי בינה מלאכותית הוא העדר נתונים נגישים. פער משקף מידע רב המצוי שאינו מתוקן במאגרים שונים שאינם מתקשרים ביניהם. נתונים אלה אינם תואמים, מתויגים או מוכנים לשימוש. אין ספק כי לבעלות על מידע קיימת חשיבות רבה לעתידם של ארגונים מסחריים רבים ולכלכלת המדינה עצמה. יצירת יכולת מתוחכמת לאיסוף, ארגון, בקרה ומסחור של מאגרי נתונים וקניין רוחני, ויכולות עיבוד וניתוח הן מהותיות לשימוש בכלי בינה מלאכותית. קיומם של מאגרי מידע רחבים, שלמים ואיכותיים יאפשר גם יצירת תשתית שלמה לשיתופי פעולה בין האקדמיה, התעשייה ומשרדי הממשלה, וליצירת שיתופי פעולה עם גופים ממשלתיים זרים ועם חברות פרטיות זרות. המידע הקיים כיום נוצר בגופים ממשלתיים, בגופים אשר נתמכים על ידי המדינה ובגופים פרטיים רבים. אחד החסמים בתחום שיתוף המידע כיום הוא ההגנה על המידע ועל פרטיותו ולכן חשוב לרכז מאמץ בנושאים רגולטוריים אלו כדי לאפשר הן את השמירה על המידע המוחזק בבעלות המדינה וגופים ממשלתיים והן את הקמת פלטפורמות לשיתוף מידע בין חברות, אקדמיה וגופים זרים.

טרנספורמציה דיגיטלית

נושא נוסף אשר רלוונטי מאד לעולם הנתונים הוא הטרנספורמציה הדיגיטלית. הטרנספורמציה הדיגיטלית היא תהליך שמטרתו לעזור לארגונים למצוא דרכים חדשות ליצירת תוצר בעל ערך עסקי על בסיס נתונים ונתונים ציבוריים כדי להגדיל את יתרונם התחרותי. השינוי הדיגיטלי השפיע על ארגונים שונים ויצר הזדמנויות עסקיות לשירותים ולתעשיות שונות, תוך מיצוי נתונים ודיגיטציה של תהליכים עסקיים ותפעוליים. שינוי זה השפיע על האינטראקציה עם הלקוחות, על אופן ניהול המידע וניתוח הנתונים על מנת לחלץ תובנות רלבנטיות עסקיות ותפעוליות. שינויים אלו כוללים מעבר ממילוי טפסים וכתובת מסמכים בכתב יד למילואים במחשב, ביצוע פעולות שונות וקבלת שירותים מרחוק, ועוד.

לביצוע מוצלח של טרנספורמציה דיגיטלית 5 גורמי מפתח:

1. נוכחותם של מנהיגים ומנהלים נכונים, בעלי ניסיון דיגיטלי.
2. התאמת ההון האנושי בארגון - הכשרת הון אנושי עתידי ופיתוח כישרונות, הגדרת תפקידים מחדש כך שיהיו מיושרים עם מטרות הטרנספורמציה, גיוס אינטגרטורים ומנהלי חדשנות טכנולוגית שיגשרו על הפערים בין התחומים המסורתיים לתחומים הדיגיטליים של העסק
3. עידוד העובדים לאמץ את הטכנולוגיות החדשות - תרבות ארגונית אשר מחזקת התנהגויות ודרכי עבודה באמצעות מנגנונים רשמיים. נתינת מקום לעובדים להציע ולקדם רעיונות חדשים.
4. הטמעה חיובית של השינוי, תוך תקשורו התמידי לעובדים הן בדרכים המסורתיות והן בדרכים הדיגיטליות, ויצירת מחויבות בקרב העובדים לדרכי העבודה החדשות.
5. שדרוג העבודה היום-יומית והפיכתה לדיגיטלית בכלל פעולות החברה. הפיכת המידע לנגיש ברחבי הארגון, יישום טכנולוגיות לשירות עצמי לשימוש העובדים, הלקוחות והשותפים העסקיים. בנוסף, ביצוע החלטות מבוססות מידע יכול גם הוא להגדיל את סיכויי ההצלחה

בביצוע הטרנספורמציה. מחקרים מראים שאחוזי ההצלחה גבוהים יותר בקרב חברות המיישמות שימוש נרחב בטכנולוגיות, כאשר בינה מלאכותית עומדת בראשן³⁶.

שמירה, אגירה ויכולת שיתוף ושימוש נרחב בנתונים הן כלי חיוני ביישום יכולות בינה מלאכותית. כמו כן, תהליכי ניתוח מידע מתקדמים תוך שימוש בבינה מלאכותית ולמידת מכונה גם הן משמעותיות במיוחד לביצוע תהליכי טרנספורמציה דיגיטלית, היות והן מאפשרות לחברות למנף את הנתונים שברשותן על מנת לקבל תובנות עסקיות ולייצר מודלי התנהגות של השוק והלקוחות³⁷.

תקינה נכונה ואחידה של הנתונים גם היא חשובה מאד. בעולם התעסוקה למשל, קיים תקן אירופאי בשם ESCO (European Skills, Competences and Occupations) עבור סיווג נתונים בנושאי מקצועות, מיומנויות ומקצועות. תקן זה משמש כמילון, אשר מזהה, מסווג, ומתאר את המקצועות השונים הרלוונטיים לשוק התעסוקה האירופאי³⁸.

דוגמה עדכנית נוספת לצורך בנתונים ובכלים לעיבוד נתוני טקסט חופשי בעת עיבוד תיקים רפואיים, ניתן לראות בעבודה עדכנית שנערכה על ידי צוות רב תחומי בניהול מפא"ת לניתוח נתונים של חולי COVID-19 שאושפזו בבית החולים תל השומר. במסגרת עבודה זו נבחנו תיקים של כ-400 חולי COVID-19 מאושפזים, בניסיון לחזות התדרדרות להנשמה מלאכותית חודרנית של החולים. בעוד התיק הרפואי בכללותו סופק לצוות המחקר על ידי בית החולים, הצוות נאלץ לעשות שימוש רק במידע שהיה קיים בשדות מובנים (כגון בדיקות מעבדה, נתוני הנשמה, מדידות מדדים חיוניים ותרופות שנופקו) ולא במידע העשיר שהיה קיים בטקסטים חופשיים בעברית שתיעד את מצבם הקליני של החולים - וזאת בשל הזמן הרב שהצוות היה צריך להשקיע על מנת לפתח כמעט מאפס כלים לעיבוד מידע זה. דוגמה ראשונית מאוד לפוטנציאל הלא ממומש של המידע הטקסטואלי, ניתן לראות במודל שתוכן על ידי אחד הצוותים - הוספה של מידע בסיסי מאוד הנמצא בטקסט חופשי הכולל בעיקר אבחנות כרוניות, הוביל ליצירת 200 שדות מידע חדשים, ולמודלים בעלי הביצועים הטובים ביותר מבין המודלים של הצוותים שהשתתפו בפרויקט. במידה וכלים מאומנים לניתוח מידע חופשי היו קיימים עבור הטקסט בעברית באותו אופן חופשי שהם קיימים עבור חוקרים העובדים עם טקסטים באנגלית - הצוות היה יכול לעשות שימוש במידע, ולייצר בזמן דומה תחזיות אמינות ומדויקות בהרבה, ולהכניס אותן לשימוש קליני³⁹.

3.4.2. אסטרטגיות דאטה שונות

דוגמה לאסטרטגיית דאטה מוצלחת היא של האיחוד האירופי⁴⁰. עד כה האיחוד כבר נקט מספר צעדים לקידום האסטרטגיה מאז 2014. עם תיקון ה-GDPR (General Data Protection Regulation) האיחוד האירופי יצר מסגרת איתנה לאמון דיגיטלי. יוזמות אחרות אשר טיפחו את התפתחות כלכלת הנתונים הן התקנה על זרימה חופשית של נתונים לא אישיים (FFD, free flow of non-personal)

³⁶ <https://www.mckinsey.com/business-functions/organization/our-insights/unlocking-success-in-digital-transformations>

³⁷ <https://www.e-zigurat.com/innovation-school/blog/5-key-factors-successful-digital-transformation-2019/>

³⁸ <https://ec.europa.eu/esco/portal/howtouse/21da6a9a-02d1-4533-8057-dea0a824a17a>

³⁹ מתוך מאמר "עיבוד שפה טבעית (NLP) בעברית - הצורך הלאומי וניתוח use case בתחום הרפואי", סא"ל דהן, מרכז הוועדה וד"ר צביקה סגל, רופא ומדען נתונים, מטריקס.

⁴⁰ Communication from the Commission to the European Parliament, the Council, the Economic and Social Committee and the Committee of the Regions, "A European strategy for data", Brussels, 19.2.2020 COM(2020)

(data), חוק אבטחת הסייבר (CSA, Cybersecurity Act), ההנחיה על נתונים פתוחים. הנציבות עסקה גם בדיפלומטיה דיגיטלית שהכירה ב-13 מדינות כמספק הגנה נאותה על נתונים אישיים. בתחומים מסוימים אומצה חקיקה ייעודית למגזרים רלוונטיים בנושא גישה לנתונים בכדי להתמודד עם כשלים בשווקים מסוימים, כגון שוק הרכב, ספקי שירותי תשלום, מדדי נתוני רשת חשמל, תעסוקה או מערכות תחבורה חכמות.

החזון הוא שהאיחוד האירופי ייצור סביבת מדיניות אטרקטיבית, כך שעד 2030, חלקו של האיחוד האירופי בכלכלת הנתונים - נתונים המאוחסנים, מעובדים ומועילים לשימוש חשוב באירופה - לפחות יתאם את המשקל הכלכלי שלה. המטרה היא ליצור מרחב נתונים אירופאי אחד הפתוח לנתונים מרחבי העולם - בו נתונים אישיים ותעשייתיים באיכות גבוהה, כולל נתונים עסקיים רגישים, הם מאובטחים וגם לעסקים יש גישה קלה אליהם. מרחב זה יהיה אמון על: זרימת נתונים בתוך האיחוד ובכל המגזרים, הקפדה על חוקים וערכים אירופיים, בפרט הגנה על מידע אישי, חקיקה להגנת הצרכן וחוק התחרות, הכללים לגישה ושימוש בנתונים יהיו הוגנים, מעשיים וברורים ויהיו מנגנוני ניהול נתונים ברורים ואמינים. תהיה קיימת גישה פתוחה אך אסרטיבית לזרימת נתונים בינלאומית. דרך להתמודד עם מאגרים מידע שונים שלא מתקשרים ביניהם היא יצירת אגם נתונים, אשר מאפשר לאסוף את כל המידע הקיים בארגון במאגר מרכזי אחד. הדגש הוא על העברה של כל הנתונים ללא אבחנה בצורתם או בתבניתם. כל הנתונים נאספים במאגר הנתונים בצורתם המקורית, ללא שינוי או התאמות. כאשר הנתונים נגישים ניתן להפעיל תהליכי עיבוד על מנת לאפשר פעילויות ומשימות שונות כמו הפקת דוחות, החזיה (visualization) ופעילויות מתקדמות יותר כגון למידת מכונה. השם אגם נתונים מסמל את העובדה שהמידע צריך להופיע בצורתו הטבעית ולזרום בצורה חופשית ממקום למקום על פי הצורך. מכרז פומבי לאספקה ותפעול מערכת אחסון מרכזית בתצורת DATA-LAKE הנגשת מידע וניהול תהליכי עבודה יצא באוקטובר 2020.⁴¹

https://mr.gov.il/ilgstorefront/he/p/4000521102?fbclid=IwAR2_R91ze3NmW5gDmRS5H-focI-u4fzbNOh3P5AFrU_sylgVOeLHxY3ICzc⁴¹

מדינת ישראל יכולה ללמוד ממאמצים אלו ולהישען עליהם בניסוח כללים ועקרונות לתחום, תוך התאמתם לחברה והערכים הישראליים כדי להבטיח פיתוח ושימוש אתי ואמין בבינה מלאכותית.

4.2.1 ארגון ה-OECD

הארגון לשיתוף פעולה ולפיתוח כלכלי (Organisation for Economic Co-operation and Development, OECD) פיתח עקרונות לבינה מלאכותית חדשנית ואמינה שאומצו על ידי המדינות החברות בארגון במאי 2019, וכן על ידי מדינות נוספות שאינן חברות כמו ארגנטינה, ברזיל, קולומביה, קוסטה ריקה, פרו ורומניה.⁴⁴ העקרונות מציבים תקנים לבינה מלאכותית שהינם מעשיים וגמישים מספיק כדי לעמוד במבחן בזמן בתחום המשתנה תכופות, ומשלימים תקנים קיימים של ה-OECD לגבי פרטיות, ביטחון דיגיטלי וניהול סיכונים, ופעילות עסקית אחראית. המלצות הארגון כוללות התייחסות לחמישה עקרונות שחיוניים לשימוש אחראי בבינה מלאכותית אמינה: *צמיחה כוללת, פיתוח בר-קיימא ורווחה; ערכים והוגנות ממוקדים באדם; שקיפות והסברתיות; איתנות (רובסטיות), ביטחון ובטיחות; אחריותיות*.⁴⁵ כמו כן, הארגון מספק חמש המלצות לממשלות, שעומדות בקנה אחד עם העקרונות:

- סייעו בהשקעות ציבוריות ופרטיות במחקר ופיתוח כדי לעודד חדשנות בבינה מלאכותית אמינה.
 - טיפול אקו-סיסטם לבינה מלאכותית נגיש באמצעות תשתית דיגיטלית וטכנולוגית, ומנגנונים לשיתוף דאטה וידע.
 - הבטחת סביבת מדיניות שתפתח את הדרך לפריסת מערכות בינה מלאכותית אמינות.
 - העצמת אנשים בעלי הכישורים הנחוצים לבינה מלאכותית ותמיכה בעובדים למעבר הוגן.
 - שיתופי פעולה בין גבולות ובין מגזרים כדי לקדם את השימוש האחראי בבינה מלאכותית אמינה.
- המלצות הארגון אינן מחייבות חוקית, אך עשויות להיות בעלות השפעה ולהציב רף בינלאומי למדיניות בתחום.

4.2.2 האיחוד האירופי

לצד המדיניות הלאומית שניסחו מספר מדינות אירופאיות לתחום, קבוצת מומחים מטעם האיחוד האירופי (AI HLEG, High-Level Expert Group) ניסחה קווים אתיים מנחים נוספים עבור בינה מלאכותית אמינה, כהמשך לאסטרטגיית הבינה המלאכותית האירופית שהוכרזה ב-2018. המומחים מגדירים בינה מלאכותית אמינה ככזו הפועלת באופן חוקי, המבטיחה שמירה על עקרונות וערכים אתיים, ואיתנה במובן הטכני והחברתי.⁴⁶ בהתבסס על זכויות בסיסיות ועקרונות אתיים, הקווים המנחים מונחים שבע דרישות מרכזיות בהן מערכות מבוססות בינה מלאכותית צריכות לעמוד כדי להיות אמינות: *סוכנות ופיקוח אנושיים; איתנות ובטיחות טכניים; פרטיות ונתוני ממשל; שקיפות; גיוון, מניעת אפליה והוגנות; רווחה חברתית וסביבתית; אחריותיות*.⁴⁷

⁴⁴ <http://www.oecd.org/going-digital/ai/principles/>

⁴⁵ <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>

⁴⁶ High-Level Expert Group on Artificial Intelligence. (2019). "Independent Set Up by The European Commission: Ethics Guidelines for Trustworthy AI." Artificial Intelligence

Brussels: European Commission. <https://ec.europa.eu/futurium/en/ai-alliance-consultation>

⁴⁷ <https://ec.europa.eu/futurium/en/ai-alliance-consultation>

כמו כן, בוועדה המשויכת למועצת אירופה, אשר הוסמכה להתמודד עם האתגרים השונים העולים מפיתוח טכנולוגיות בינה מלאכותית כונס פורום CAHAI⁴⁸ (Ad hoc Committee on Artificial Intelligence). מטרת פורום CAHAI היא לתת מענה לסוגיות העולות מטכנולוגיות של בינה מלאכותית באמצעות חקיקה מתאימה. המכנה המשותף למדינות החברות הוא אירופי, ואילו ארצות הברית, קנדה וישראל הן רק מדינות משקיפות בפורום זה. פורום CAHAI מעוניין ליצור מצב עניינים בו החקיקה תקבע מראש כיצד יכול השוק לפעול. הפורום שוקל יצירת אמנה חדשה בנושא בינה מלאכותית, או לכל הפחות ניסוח המלצה רשמית של מועצת אירופה.

4.2.3. ארצות הברית

ארצות הברית היא מעצמה בתחום הבינה המלאכותית ופועלת בשנים האחרונות לקידום התחום ולשימור יתרונה הטכנולוגי, כאשר התייחסות לאתגרים האתיים נתפשת בקרב גורמי הממשל כאמצעי חשוב לשימור יתרון זה. תפישה זו באה לידי ביטוי במספר תוכניות ויוזמות ממשלתיות.

- הוועדה המייעצת לחדשנות בביטחון (Defense Innovation Board) פרסמה באוקטובר 2019 המלצות לשימוש אתי בבינה מלאכותית על ידי מחלקת ההגנה. המסמך מונה חמישה עקרונות אתיים לבינה מלאכותית: *אחריות, שוויון, יכולת מעקב והבנה של הטכנולוגיה, אמינות ויכולת משילות*. המסמך מספק גם שורת המלצות לתחום בהן ביסוס וועדת היגוי רחבה במחלקת ההגנה; חיזוק תוכניות האימון וההון האנושי במחלקת ההגנה; השקעה במחקר על היבטי אבטחה חדשים של בינה מלאכותית; השקעה במחקר לחיזוק יכולת שחזור והתייחסות לסוגיית הקופסה השחורה; הגדרת ספים או מדדים לאמינות המערכות; חיזוק טכניקות לבחינה והערכה של בינה מלאכותית; פיתוח מתודולוגיה לניהול סיכון; הבטחת הטמעה ראויה של עקרונות אתיים של בינה מלאכותית; והרחבת המחקר להבנת האופן הנכון להטמעת הערכים הללו.⁴⁹
- כמו כן, בינואר 2020 הבית הלבן הציע קווים מנחים לרגולציה של יישומי בינה מלאכותית⁵⁰ שיסייעו לסוכנויות פדרליות מסדירות לפתח ולשמר גישות ליצירה ואימוץ בטוחים ואמינים של טכנולוגיות בינה מלאכותית מבלי לפגוע בזימות. טיוטת המסמך כוללת שורת עקרונות שיש לקחת בחשבון בניסוח רגולציה לעיצוב, פיתוח ושימוש ביישומי בינה מלאכותית: *אמון ציבורי בבינה מלאכותית; שיתוף הציבור; יושרה מדעית ואיכות מידע; הערכת וניהול סיכונים; עלות ותועלת; גמישות; הוגנות ואי-אפליה; גילוי נאות ושקיפות להגברת אמון הציבור*⁵¹; *בטיחות וביטחון; תיאום בין סוכנויות*. הטיוטה מונה גם אמצעים לא-רגולטורים לשימוש הסוכנויות הממשלתיות כמו יצירת תוכניות וניסויים כמו אתגרים והאקטונים לעידוד פיתוח יישומים שונים שיאפשרו הקלה רגולטורית, ופיתוח סטנדרטים וולונטריים ומוסכמים ליישומי בינה מלאכותית שיאפשרו לנהל את הסיכונים הקשורים בפיתוח באופן אדפטיבי יותר לצרכי הטכנולוגיה המתפתחת. ב-3 לדצמבר 2020 פורסמו הנחיות לקידום שימוש אמין בכלי בינה מלאכותית בממשל הפדרלי. בהנחיות הוגדר כי תוך 180 יום תפורסם בפומבי מפת דרכים להנחיות המדיניות, בהתאם לצו זה.

⁴⁸ <https://www.coe.int/en/web/artificial-intelligence/cahai>

⁴⁹ <https://media.defense.gov/2019/Oct/31/2002204458/-1/>

[1/0/DIB_AI_PRINCIPLES_PRIMARY_DOCUMENT.PDF](https://www.whitehouse.gov/wp-content/uploads/2020/01/Draft-OMB-Memo-on-Regulation-of-AI-1-7-19.pdf)

⁵⁰ <https://www.whitehouse.gov/wp-content/uploads/2020/01/Draft-OMB-Memo-on-Regulation-of-AI-1-7-19.pdf>

⁵¹ למשל, התייחסות לשאלות לגבי האופן בו הטכנולוגיה תשפיע על משתמשי הקצה האנושיים.

מפת דרכים זו תכלול, במידת הצורך, לוח זמנים ליצירת קשר עם הציבור ולוחות זמנים לסיום הנחיות המדיניות הרלוונטיות.⁵²

• במקביל פועלות מספר תוכניות מחקר שמסייעות להתמודדות עם ההיבט הערכי בפיתוח ובשימוש במערכות מבוססות בינה מלאכותית. החל מ-2018 דארפ"א מנסה להתמודד עם אתגר ההסברתיות ולסייע למשתמש האנושי להבין את המכניזמים ואופן קבלת ההחלטות של המכונה. במסגרת תוכנית בינה מלאכותית הסברית⁵³ הסוכנות מבקשת ליצור טכניקות ללמידת מכונה שיפיקו מודלים הסבריים יותר מבלי לפגוע ברמת הביצוע או ברמת האמון של המערכת.⁵⁴ תוכנית נוספת של הקרן הלאומית למדע (The National Science Foundation - NSF) בשיתוף עם אמזון מבקשת לתרום לאמינות והבטיחות של מערכות מבוססות בינה מלאכותית באמצעות מימון ותמיכה במחקר הממוקד בהוגנות בבינה מלאכותית. בין הסוגיות שהתוכנית מבקשת להתייחס אליהן, ניתן למנות שקיפות, הסברתיות, אחריותיות, פוטנציאל להטיה והשפעות שליליות, התקדמות אלגוריתמית, יעדי הגינות ותוקף של הגינות.⁵⁵

4.2.4 . צרפת

בצרפת פורסמה ב-2018 תוכנית אסטרטגית לבינה מלאכותית שביקשה ליצור מודל ייחודי צרפתי, מבלי להסתמך על מסמכים לאומיים דומים ממדינות אחרות. בתוכנית ישנה התייחסות לסוגיית האתיקה, כאשר עיקר ההתמקדות הוא בחשש מהטיה של מערכות בינה מלאכותית ובאתגר הקופסה השחורה – הקושי להבין את תהליך קבלת ההחלטות של המכונה ולסמוך על שיקול דעתה בסוגיות רגישות. הכותבים רואים ב*שקיפות* כערך משמעותי להתמודדות עם האתגר, ולביקורת על המערכת. התוכנית לא מציגה רשימת עקרונות נוספים, אך ממליצה על הקמת וועדה לאתיקה של בינה מלאכותית וטכנולוגיה שתוביל דיון ציבורי בנושא. הוועדה תבחן את ההשפעות האפשריות של בינה מלאכותית בטווח הארוך והקצר, ותיצור תקן לפיתוחה האתי ולבקרה שלה. כמו כן, התוכנית גורסת שיש לקחת בחשבון שיקולים אתיים הקשורים לבינה מלאכותית כבר בתהליך הפיתוח של האלגוריתמים וקוראת להכשרה מתאימה לאנשי המקצוע הרלוונטיים.⁵⁶

יתר על כן, במסגרת החוק לחירויות דיגיטליות⁵⁷ הוטל על הוועדה הצרפתית הלאומית לטכנולוגיות מידע וחירויות (CNIL, National Data Protection Commission) לדון בהשפעה האתית והחברתית של בינה מלאכותית, ולהבטיח את קיומן של שורת חירויות בפיתוח מודל צרפתי לבינה מלאכותית – *הגנה על פרטיות, הגנה על מידע אישי, שקיפות, אחריותיות של השחקנים ותרומה לרווחה משותפת*.⁵⁸ מסקנות הדיון הוצגו במסמך שפורסם ב-2018 ומתמקדות בשני עקרונות מרכזיים:⁵⁹

• *הוגנות*, כלומר בינה מלאכותית שאינה פוגעת בקהילת המשתמשים והאזרחים, ושוקלת את התוצאות הקולקטיביות.

⁵² <https://www.whitehouse.gov/presidential-actions/executive-order-promoting-use-trustworthy-artificial-intelligence-federal-government/>

⁵³ Explainable Artificial Intelligence (XAI)

⁵⁴ <https://www.darpa.mil/program/explainable-artificial-intelligence>

⁵⁵ https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505651

⁵⁶ https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf
Digital Republic Bill⁵⁷

⁵⁸ <https://futureoflife.org/ai-policy-france/>

⁵⁹ https://www.cnil.fr/sites/default/files/atoms/files/cnil_rapport_ai_gb_web.pdf

- תשומת לב וכוונות מתמשכות באמצעות תהליכים ומדדים לבקרה כדי להתמודד עם אופיין הלא צפוי והלא יציב של בינה מלאכותית ולמידת מכונה, וכדי להבטיח הימנעות מהסתמכות רבה מדי על האלגוריתמים.

המסמך גם מספק שש המלצות למדיניות בנושא חינוך לאתיקה, שיפור יכולת ההבנה והביקורת של מערכות בינה מלאכותית וקידום חירויות אנושיות.

4.2.5 בריטניה

ב-2018 הכריזה ראש הממשלה בריטניה דאז, תרזה מיי, שהקידום של בינה מלאכותית וניצול ההזדמנויות הטמונות בה יעשה בצורה אחראית שתבטיח שהשינוי עובד עבור כולם.⁶⁰ מספר יוזמות ממשלתיות חקרו ותכננו את האופן הנכון לרתום את טכנולוגיית הבינה המלאכותית.

וועדת המדע והטכנולוגיה של בית הנבחרים הבריטי פרסמה ב-2016 מסמך על רובוטיקה ובינה מלאכותית שהתייחס למגוון סוגיות אתיות וחוקיות הקשורות בבינה מלאכותית. דגש מרכזי במסמך הושם על בטיחות ושליטה, כאשר הכותבים התייחסו למספר סוגיות הנגזרות מכך בהן: הצורך באמצעים לאימות ותיקוף כדי להבטיח שהמערכות פועלות האופן הרצוי; שקיפות בתהליך קבלת ההחלטות במיוחד בתחומים רגישים; מזעור הטיה במערכות בינה מלאכותית ואפליה הנגרמת מכך; פרטיות והסכמה; אחריותיות וחבות משפטית.⁶¹

וועדה נבחרת לבינה מלאכותית פרסמה ב-2018 דוח שבוחן את ההתפתחות והמדיניות בתחום בבריטניה. הדוח מכיר בכך שבריטניה תתקשה להתחרות בענקיות כמו ארצות הברית וסין בכל הקשור למימון והון אנושי, אך מציע שאתיקה של בינה מלאכותית תוכל להקנות למדינה יתרון תחרותי. הדוח קורא לביסוס אסטרטגיה לאומית לבינה מלאכותית ומציע "כללים לבינה מלאכותית":⁶² יש לפתח בינה מלאכותית לטובת הכלל ולרווחת האנושות; בינה מלאכותית צריכה לפעול על פי עקרונות של הגינות ומובנות; אין להשתמש בבינה מלאכותית כדי לפגוע בזכויות בנתונים או בזכות לפרטיות של אינדיבידואלים, משפחות וקהילות; לכל האזרחים הזכות לרכוש השכלה ולהתפתח מנטלית, כלכלית ורגשית לצד בינה מלאכותית; לעולם אין לתת בידי בינה מלאכותית את הכוח לפגוע, להרוס או להונות בני אדם.^{63 64}

ההתייחסות להיבטים האתיים ביחס לבינה מלאכותית באה לידי ביטוי גם במסמך משותף של החברה המלכותית והאקדמיה הבריטית. אלו סקרו את הצרכים בתחום ניהול ושימוש בדאטה, כאשר תוצאות הסקירה הרחבה פורסמו במספר מסמכים. במסמך המרכזי מוצגת שורת עקרונות ל-Data Governance, שמהווים קווים מנחים ליצירת מודלים וכלים יותר מאשר הנחיות לפעולה. העקרון המנחה הכללי הוא שהנתונים צריכים לקדם גשוג אנושי. לעקרון זה מתווספים ארבעה נוספים: הגנה על אינטרסים וזכויות של אינדיבידואלים וקבוצות; הבטחת שקיפות ואחריותיות בשקלול-התמורות

<https://futureoflife.org/ai-policy-united-kingdom/>⁶⁰

<https://publications.parliament.uk/pa/cm201617/cmselect/cmsctech/145/145.pdf>⁶¹

AI Code⁶²

<https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf>⁶³

<https://futureoflife.org/ai-policy-united-kingdom/>⁶⁴

הנגרם כתוצאה מניהול הנתונים ושימוש בהם; למידה מהצלחות וכישלונות וחיפוש אחר שיטות עבודה מוצלחות; חיזוק הממשל הדמוקרטי הקיים.⁶⁵

4.2.6 דרום קוריאה

בדרום קוריאה פורסמו מספר תוכניות לאומיות לבינה מלאכותית שהציגו את הדרכים והפעולות הנחוצות לקידומה במדינה. ב-2016 פורסמה תוכנית-על שנועדה להכין את דרום קוריאה למהפכה התעשייתית הרביעית ומתייחסת למספר טכנולוגיות כמו בינה מלאכותית, האינטרנט של הדברים, נתוני-עתק ומחשוב ענן, המכונות במסמך טכנולוגיות מידע תבוניות. התוכנית בוחנת את ההשפעה האפשרית של הטכנולוגיות הללו על החברה, שוק העבודה והתיעוש וקוראת ליצירת מסגרת נורמטיבית וחוקית מעודכנת שתוכל להתמודד עם האתגרים העולים מהשימוש הגובר בהן. אמנם אין בתוכנית התייחסות לעקרונות או ערכים מסוימים שנחוצים כדי להבטיח שימוש אתי בבינה מלאכותית, אך מפורטים אמצעים וצעדים שונים שיש לקדם כדי להגן על רווחת האזרחים. למשל, ביסוס אמנה אתית מכוונת-אדם שתנחה את תהליך איסוף הנתונים ויצירת האלגוריתמים של בינה מלאכותית, ותמזער את הסכנות החברתיות והסיכוי לשימוש פוגעני בטכנולוגיה. האמנה תהווה מדריך אתי ברור עבור מפתחים ומשתמשים, ותכלול פרוטוקולים לפיתוח בינה מלאכותית ואיסוף נתונים כדי להבטיח הוגנות ואמינות.⁶⁶

תוכנית אסטרטגית לאומית נוספת פורסמה במרץ 2020. גם בה מוזכר הצורך בביסוס קוד אתי לבינה מלאכותית, ונסקרות בה בקצרה יוזמות דומות מהעולם ומצבן בעת פרסום הדוח. דגש מיוחד מושם בתוכנית על המלצות ה-OECD, שפורטו לעיל, כנקודת ייחוס למדיניות לאומית ולביסוס תקנים לבינה מלאכותית אתית שיעמדו בקנה אחד עם נורמות בינלאומיות.⁶⁷

חשוב לציין שגם בישראל החל מהלך משמעותי בתחום. ב-2018 פעלה וועדה לגיבוש תוכנית לקידום תחום הבינה המלאכותית.⁶⁸ אחת מתתי-הוועדות מהן הורכבה הוועדה הממשלתית עסקה בהרחבה באתיקה ורגולציה של בינה מלאכותית, ופרסמה בנובמבר 2019 שורת המלצות נרחבת לתחום, שכללו בין היתר עקרונות אתיים מנחים עבור מקבלי ההחלטות למחקר, פיתוח ושימוש בבינה מלאכותית.⁶⁹ העקרונות שנכללו בהמלצות הוועדה הם: שקיפות, הסברתיות, אחריות וניהול סיכונים, כיבוד זכויות אדם והגנה עליהן (שמירה על שלמות הגוף, פרטיות, שמירה על אוטונומיה, זכויות אזרחיות ופוליטיות), הגנת סייבר ואבטחת מידע, בטיחות, קיום שוק תחרותי. דוח נוסף הוגש על ידי מוסד שמואל נאמן עבור המועצה הלאומית למחקר ופיתוח במשרד המדע והטכנולוגיה ב-2018, עוסק גם כן באתיקה, משפט ופרטיות בתחום הבינה המלאכותית. הדוח מרכז מספר עקרונות משותפים עבור מקבלי ההחלטות הישראלי אליהן יש לתת דגש בניסוח מסגרת אתית לאיסוף מידע ועיבודו: מניעת הטיה, אחריות

⁶⁵ <https://royalsociety.org/-/media/policy/projects/data-governance/data-management-governance.pdf>

⁶⁶ https://english.msit.go.kr/cms/english/pl/policies2/_icsFiles/afieldfile/2017/07/20/Master%20Plan%20for%20the%20intelligent%20information%20society.pdf

⁶⁷ https://english.msit.go.kr/cms/english/pl/policies2/_icsFiles/afieldfile/2020/03/23/National%20Strategy%20for%20Artificial%20Intelligence_200323.pdf

⁶⁸ <https://www.globes.co.il/news/article.aspx?did=1001307714>

⁶⁹ <http://ekarine.org/heb/wp-content/pubs/AIEthicsRegulationReport-hebrew.pdf>

ציבורית, פרטיות מרבית, קידום הגישה למידע מצד ממשלות, ניתוח סיכונים אחראי, השכלה בתחום האתיקה, הצטרפות לדיון הבינלאומי בנושאי האתיקה⁷⁰.

4.3. מדיניות, אסדרה ותקינה

יצירת רגולציה לתחומים טכנולוגיים היא מאתגרת, שכן היא נדרשת לאזן בין הכוונה לשמור על אתיקה ומוסר בתהליכי המחקר, הפיתוח והשימוש הטכנולוגיים ובין הרצון לעודד חדשנות ויזמות ולהימנע מפגיעה בצמיחה הטכנולוגית. יתר על כן, קיים אתגר בניסוח רגולציה או תקנים שיעמדו במבחן הזמן לאור קצב ההתפתחות המהיר של התחום, במיוחד בכל הקשור בבינה מלאכותית. בנוסף, העובדה שתחום הבינה המלאכותית הוא גלובלי ומונע על ידי חברות ותאגידים בינלאומיים מעוררת קושי נוסף ביצירת אסדרה משפטית.

בשנים האחרונות יותר ויותר מדינות נדרשות לאתגר זה, ופועלות ליצירת אסדרה ותקינה של תחום הבינה המלאכותית כדי לקדם את המחקר, הפיתוח והיישום של בינה מלאכותית ברמה הלאומית, תוך עמידה בעקרונות אתיים. גם כאן, ההסתכלות המשווה על הנעשה בעולם עשויה לשרת את מקבלי ההחלטות בישראל וללמד על גישות שונות להתמודדות עם האתגר הרגולטורי, שיוכלו להוות בסיס או מודל לאימוץ בהקשר הישראלי. יש לציין שגם בישראל החלו מהלכים לניסוח כללים ותקנות לבינה מלאכותית, בין היתר במסגרת הוועדה הממשלתית שהוזכרה לעיל.

דגש מיוחד יושם על אסדרה ותקינה של דאטה לאור חשיבותה ותפקידה המרכזי באימון אלגוריתמים מבוססי בינה מלאכותית, וכן על התייחסות לאסדרה של תשתיות רלוונטיות, לאור עיסוקה של הוועדה בכך.

4.3.1. האיחוד האירופי

האסדרה הכללית להגנה על מידע (GDPR) נכנסה לתוקף ב-2018 ומשחקת תפקיד מכריע באסדרה של דאטה באירופה. האסדרה מתייחסת לאגירת נתונים פרטיים והשימוש בהם על ידי חברות, ולמעשה יוצרת תשתית לאסדרה ותקינה של תחום הנתונים עליה ניתן יהיה להוסיף ולשנות בהתאם למקרה הייחודי של בינה מלאכותית.

ה-GDPR מבחין בין מידע אישי שיכול לשמש לזיהוי אינדיבידואלים ובין מידע אישי רגיש כמו מידע גנטי, נטיות מיניות והשקפות פוליטיות של אינדיבידואלים. ארגונים מורשים לעבד מידע אישי בתנאי שקיבלו אישור מפורש מהאינדיבידואל שהמידע מתייחס אליו, או כאשר קיים צורך חוקי, הגנתי או חיוני לעיבוד המידע. במידה והארגון מעבד מידע אישי לפי צורך אך ללא אישור מפורש מטעם האינדיבידואל, עליו ליידע אותו. כמו כן, ארגונים חייבים לספק לאינדיבידואלים מידע אודות הדרך בה הנתונים שלהם יעובדו ולספק הסבר לבקשה למידע ולהשלכות של סירוב להעבירו. לפי ה-GDPR כל אדם רשאי להגיש בחינם בקשה מיוחדת לחברה או ארגון שאוספים מידע אודותיו, ואלו יאלצו לספק לו את המידע הנאסף תוך חודש, כאשר במקרים מסוימים יוכל האינדיבידואל לדרוש את מחיקת המידע. החלטה משמעותית של ה-GDPR היא שלאנשים תהא הזכות לדרוש שלא יהיו כפופים להחלטות אוטומטיות של המכונות אודותיהם, ושהם זכאים להסבר לגבי כל החלטה שהתקבלה עבורם. במסגרת התקנות, חברות מחויבות להגן על המידע, להעריך את השפעת המידע על

https://www.neaman.org.il/Files/Artificial%20Intelligence,%20Data%20Science,%20and%20Smart%20Robotics-%20A%20report%20on%20Ethics,%20Law%20and%20Privacy_20181230121054.072.pdf

האינדיבידואל במידה וישתחרר ולתעד את דרך עיבודו. עליהן ליישם את העקרונות הבסיסיים של הגנת מידע (למשל, אנונימיזציה במידת האפשר), ולדווח על כל מקרה של הרס, אובדן, שינוי או גישה לא מאושרים של המידע אם עשויה להיות השפעה שלילית על האנשים המצוינים במידע.^{71 72}

דוח של הנציבות האירופית שפורסם ב-2020 עוסק בהצגת מדיניות להשגת שתי מטרות – קידום השימוש בבינה מלאכותית והתייחסות לסיכונים הקשורים בשימוש בה – באמצעות יצירת "אקו-סיסטם של מצוינות ואמון". בין הפעולות שמציגה התוכנית יש לציין את: הקמתם של מרכזי יזמות דיגיטלית בכל אחת מהמדינות החברות באיחוד כדי להבטיח רמת מומחיות גבוהה; הקמת שותפות פרטית-ציבורית בתחומי בינה מלאכותית, נתונים ורובוטיקה כדי לשלב את המאמצים בתחום באופן מתואם; קידום דיאלוג שקוף ופתוח בין מגזרים שמטרתו קידום האימוץ של בינה מלאכותית על ידי המגזר הציבורי, בדגש על תחומי הבריאות והשירותים הציבוריים; קידום פרקטיקות אחראיות לניהול נתונים בהתאם לעקרונות FAIR⁷³ לצד השקעה של הנציבות במחשוב קוונטי, מחשוב קצה ובינה מלאכותית, ותשתיות נתונים וענן.⁷⁴

המסמך המשלים לדוח מתמקד באסטרטגיה לתחום הנתונים ומבקש להטיב עם אזרחי אירופה באמצעות חיזוק ושיפור תהליכי קבלת החלטות מבוססות נתונים במגזר הציבורי והפרטי. מטרת המסמך היא ליצור שוק אירופי יחיד ופתוח לנתונים מכל העולם, שיאפשר העברת נתונים אישיים ולא אישיים באופן מאובטח ונגיש לשימוש כולם, תוך הפחתת טביעת הרגל האקולוגית. לתוכנית ארבע בסיסים עיקריים:⁷⁵

- א. יצירת מסגרת עבודה ממשלתית וחוצת מגזרים לגישה לנתונים ושימוש בהם. הנציבות מכירה בקושי בניסוח מסגרת מקיפה לכלכלה הטכנולוגית המשתנה ולכן נמנעת מראש מפירוט יתר ורגולציה מקדימה, ומעדיפה גישה גמישה יותר שתאפשר ניסוי, כמו "ארגזי-חול" (Sandboxes) רגולטוריים. מוזכר הצורך במנגנונים לתקינה של פעילויות הקשורות בנתונים, וכן צורך בחתירה לעבר תיאור הרמוני ורחב של מאגרי נתונים, אובייקטים ומזהים כדי לאפשר שימוש טכני בנתונים (interoperability). כמו כן, הנציבות תפעל להפיכת נתונים ציבוריים באיכות גבוהה זמינים לשימוש חוזר ותבחן את הצורך בפעילות חקיקתית שתשפיע על היחסים בין השחקנים בכלכלת הנתונים כדי לספק תמריצים לשיתוף נתונים אופקי.
- ב. השקעה תקציבית בפרויקט "High Impact project on European Data Spaces" שיעסוק בארכיטקטורות לשיתוף נתונים (כולל תקינה של העברת נתונים, שיטות עבודה מומלצות וכלים) ומנגנונים ממשלתיים, לצד הפדרציה האירופית של תשתיות ענן אמינות וחסכוניות באנרגיה (תשתיות-כשירות, פלטפורמות-כשירות, ותוכנה-כשירות) ותתייחס לצורך של תעשיות מסוימות במודלים המאפשרים העברת נתונים ואיבודם ללא זמן שהייה (latency).
- ג. הכשרת יחידים והשקעה בכישורים ובחברות קטנות ובינוניות.

⁷¹ https://www.neaman.org.il/Files/Artificial%20Intelligence,%20Data%20Science,%20and%20Smart%20Robotics-%20A%20report%20on%20Ethics,%20Law%20and%20Privacy_20181230121054.072.pdf
<https://gdpr-info.eu/art-5-gdpr/>⁷²

⁷³ עקרונות מנחים להפיכת הנתונים נגישים למציאה, מובנים, ניתנים לשימוש, וניתנים לשימוש חוזר (Findable, Accessible, Interoperable and Reusable).

⁷⁴ https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

⁷⁵ https://ec.europa.eu/info/sites/info/files/communication-european-strategy-data-19feb2020_en.pdf

ד. יצירת מרחבי נתונים אירופיים נפוצים במגזרים ומרחבים אסטרטגיים עבור הציבור (תעשייה וייצור, קיימות, תעבורה, בריאות, פיננסיים, אנרגיה, חקלאות, ומנהל ציבורי). מרחבים אלו יובילו לנגישות למאגרי נתונים גדולים עבור התעשיות והמרחבים הללו, לצד כלים טכניים ותשתיות הנחוצים לשימוש בנתונים ושיתוף שלהם כאשר ישנם מנגנוני פיקוח ממשלתיים. המסגרת האופקית המוצעת תושלם במידת הצורך בחקיקה מגזרית לגישה ושימוש בנתונים ועל ידי מנגנונים להבטחת יכולת השימוש בהם ברמה הטכנית.

4.3.2. ארצות הברית

כתגובה לזינוק בדיווחים על דליפות מידע צרכני ולחשש הגובר מפני אובדן הפרטיות, נחקק בחודש יוני 2018 במדינת קליפורניה החוק להגנת פרטיות הצרכן⁷⁶ (California Consumer Privacy Act (CCPA) אשר נכנס לתוקפו בשנת 2020. תחולת ה-CCPA היא על תושבי מדינת קליפורניה, גם כשהם נמצאים זמנית במקום אחר, בשונה מה- GDPR אשר חל על אנשים הממוקמים פיזית בתחומי האיחוד האירופי. כך, שני החוקים יכולים לחול בו זמנית על אותה פעילות במידע אישי, למשל, כאשר תייר תושב קליפורניה מבקר במדינה באיחוד האירופי. ישנם מספר הבדלים בין שני החוקים כגון דרישות שונות לתוכן מדיניות הפרטיות, זכויות שונות והגדרה שונה של 'מידע אישי' כך שהיא כוללת גם מידע אישי הקשור למשקי בית ולמכשירי וכו'. בין היתר החוק מאפשר: מתן זכות לצרכן לדעת אילו קטגוריות של מידע אישי נאספו עליו ועל משפחתו, מתן זכות לצרכן לדעת האם המידע האישי שלו נמכר או הועבר למטרות עסקיות לצדדים שלישיים ומה זהותם, הטלת חובה על עסקים לחשוף בפני הצרכן אם הם מוכרים את המידע האישי שלו ומתן אפשרות לצרכן להורות לעסק לחדול מכך או למנוע זאת מבעוד מועד. הטלת איסור על עסקים למנוע שירות, לשנותו או לגבות יותר עבורו אם הצרכן מבקש לממש את זכויותיו, דרישה מעסקים לאבטח את המידע האישי של הצרכנים והטלת אחריות עליהם במקרה של דליפת מידע אישי.

בשנת 2008 חוקקה מדינת אילינוי את חוק הפרטיות של מידע ביומטרי (Biometric Information Privacy Act, BIPA). חוק זה מחייב חברות לבקש רשות לפני איסוף נתונים ביומטריים, לציין את הסיבה ואת משך הזמן שהנתונים יאוחסנו, ולכלול פרטים אלה במדיניות פרטיות כתובה. החוק גם מאפשר לצרכנים לתבוע חברות שאינן עומדות בדרישות אלו⁷⁷.

ב-2019 משרד הבית הלבן לניהול ותקצוב פרסם את אסטרטגיית הנתונים הפדרלית כמסגרת לעקרונות פעולה ושיטות עבודה מומלצות לשימוש בנתונים וניהולם על ידי הסוכנויות הפדרליות⁷⁸. מטרתה היא למנף את השימוש בנתונים על ידי הסוכנויות הפדרליות כמשאב אסטרטגי. ניתן לאגד את הפרקטיקות המוצעות באסטרטגיה תחת שלוש כותרות עיקריות:

- **בניית תרבות שמעריכה נתונים ומקדמת שימוש ציבורי בהם:** התמקדות בהבנת הצרכים של הגורמים השונים בתחום הדאטה והאופן בו ניתן להשתמש בדאטה באופן משותף ככלי ליעול קבלת ההחלטות.

⁷⁶ <https://oag.ca.gov/privacy/ccpa>

⁷⁷ <https://www.ilga.gov/legislation/ilcs/ilcs3.asp?ActID=3004&ChapterID=57>

⁷⁸ <https://strategy.data.gov/assets/docs/2020-federal-data-strategy-action-plan.pdf>

- **משילות, ניהול נתונים והגנה על נתונים**: התמקדות באמצעים להגנה על נתונים איכותיים ואמינים, לצד הגנה על זכויות פרטיות וחסינון הקשורות בנתונים. בין היתר באמצעות יצירת סטנדרטיזציה לתחום, בחינת היכולות והבשלות בתחום הנתונים בכל משרד ממשלתי, שימור Inventory Data Assets ושיתוף נתונים ברמה המדינית והמקומית.
- **קידום שימוש יעיל וראוי בנתונים**: התמקדות בשימוש משותף ויעיל בדאטה בדרכים שונות, בהן רתימת פלטפורמות מחשוב משותפות, חיזוק הקשרים עם התעשייה והאקדמיה והרחבת הגישה לנתונים באמצעות מתודות מגוונות.

כדי לקדם את יישום האסטרטגיה, מפורטת במסמך תוכנית פעולה לשנת 2020 שמתמקדת בפעולות של הסוכנויות הפדרליות, פעולות קהילות ופעולות לפתרון משותף. **הפעולות של הסוכנויות** מבוצעות על ידי כל סוכנות בנפרד ומיועדות לקדם את יכולתה של כל תוכנית למנף את הנתונים שברשותה כנכס אסטרטגי. פעולות אלו כוללות זיהוי של צרכים בתחום הנתונים כדי לענות על שאלות דחופות, יצירת גוף מגוון למשילות בנתונים (Data Governance), הערכת בשלות הנתונים והתשתית, זיהוי הזדמנויות להגדלת הכישורים בתחום של חברי הצוות, ופרסום ועדכון של מאגרי הנתונים. **פעולות קהילתיות** נעשות על ידי קבוצת סוכנויות באמצעות וועדה משותפת או מנגנון לתיאום במטרה לשלב ולתאם מאמצים קיימים בחקיקה ואסדרה של התחום. פעולות אלו כוללות הקמת וועדה פדרלית לנתונים (Federal Chief Data Officer Council), שיפור הנתונים והמשאבים ויצירת מודלים למשאבים עבור מחקר ופיתוח של בינה מלאכותית, ושיפור תקנים לנתוני ניהול פיננסיים. **הפעולות לפתרון משותף** הן פרויקטים ומאמצים שמובלים על ידי סוכנות יחידה או וועדה משותפת ויטיבו עם כל הסוכנויות. אלו כוללות רשימת פעולות ארוכה, בהן: פיתוח מאגר נתונים פדרלי, פיתוח רשימת מיומנויות לנתונים, פיתוח מסגרת עבודה אתית לנתונים, פיתוח ערכת כלים להגנה על נתונים, פיתוח קווים מנחים למדידת איכות ודיווח על דאטה ופיתוח מאגר תקנים לדאטה⁷⁹.

כמו כן, המכון הלאומי לתקנים וטכנולוגיה (NIST) פרסם אסטרטגיה ראשונה מסוגה למעורבות פדרלית בפיתוח סטנדרטים טכניים לפיתוח בינה מלאכותית.⁸⁰ על פי התוכנית, על הממשל הפדרלי לחזק ידע, מנהיגות ותיאום בכל הקשור לסטנדרטיזציה של בינה מלאכותית כדי להגביר את היעילות. לשם כך, יש למנות מתאם תקנים (Standards Coordinator) שיהיה אחראי על איסוף הצרכים והמידע הקשורים לסטנדרטיזציה של בינה מלאכותית כדי להעריך את התקנים הקיימים ולהתאימם לבינה מלאכותית. כמו כן, על הממשל לפעול לקידום מחקר ממוקד כדי לקדם הבנה רחבה יותר של האופן בו ניתן לשלב היבטים שונים של אמינות בתוך תקנים וכלים לתקינה. בנוסף, על הממשל לתמוך ולהרחיב שותפות ציבורית ופרטית לפיתוח ולשימוש בתקנים ובכלי תקינה לבינה מלאכותית, כדי לקדם בינה מלאכותית אמינה ואיתנה. לבסוף, על הממשל לפעול ברמה הבינלאומית לצד שחקנים נוספים לקידום תקנים לבינה מלאכותית עבור הצרכים הלאומיים והכלכליים האמריקאיים. המסמך מתייחס גם לכלים שנחוצים לתקינה וכוללים בין היתר פורמטים מתוקננים למאגרי נתונים, כלים להצגת תהליך

⁷⁹<https://strategy.data.gov/assets/docs/2020-federal-data-strategy-action-plan.pdf>
⁸⁰<https://www.whitehouse.gov/wp-content/uploads/2020/02/American-AI-Initiative-One-Year-Annual-Report.pdf>

ההנמקה של מערכות בינה מלאכותית ומתודולוגיות לבדיקה שיאפשרו לתקף ולהעריך את הביצוע של טכנולוגיות אלן⁸¹.

באתר ממשלתי ייעודי לתחום הבינה המלאכותית הודגשה חשיבותם של משאבי מחשוב ותשתיות הדרושים למחקר ופיתוח של בינה מלאכותית. צוין בחלק זה כי ארצות הברית מובילה בפיתוח תשתיות מחשוב והוזכר מחשב-העל Summit שמשמש את משרד האנרגיה למחקר מדעי רחב תחומים, ומספק יכולות ייחודיות להטמעת בינה מלאכותית. ב-2021 משרד האנרגיה הודיע שבכוונתו לבנות את מחשב-העל Frontier ולשמר את ההובלה האמריקאית בתחום.⁸² גם הקרן הלאומית למדע (NSF) השקיעה בפיתוח של מגוון רחב של תשתיות סייבר שיוכלו לשמש למחקר ופיתוח בינה מלאכותית, כולל מחשב-על. ב-2018 הקרן מימנה את מחשב העל הגדול והחזק ביותר שאי פעם נתמך על ידה לצרכי מחקר אקדמי, Frontera⁸³.

בעדכון לתוכנית האסטרטגית הלאומית למחקר ופיתוח בינה מלאכותית הוצעו ארכיטקטורות אפשריות לעיצוב בינה מלאכותית אתית. למשל, ארכיטקטורת צג דו-שכבתית (two-tier monitor architecture) שתפריד בין הבינה המלאכותית הפועלת ובין הסוכן שאחראי להערכה אתית וחוקית של כל פעולה מבצעת. גישה נוספת היא מתן עדיפות להנדסה בטוחה, בה המסגרת הקונספטואלית לארכיטקטורת סוכן הבינה המלאכותית משמשת להבטיח שהתנהגות הבינה המלאכותית תהיה בטוחה ולא פוגענית לאדם⁸⁴.

4.3.3 צרפת

בתוכנית הלאומית הצרפתית מוצעים קווים מנחים למקבלי ההחלטות כדי להתמודד עם הסוגיות האתיות של בינה מלאכותית. בין המלצותיהם:⁸⁵

- הקמת גוף מומחים לבחינת האלגוריתמים ומאגרי המידע תוך התייחסות לקוד המקור ולתיעוד הפיתוח כדי להבטיח שהבינה המלאכותית מספקת שירותים הוגנים.
- הומלץ לפעול לפי תוכנית המחקר של דארפ"א לפיתוח בינה מלאכותית ניתנת להסברה. זאת באמצעות הפקת מודלים שניתנים להבנה יותר בקלות, יצירת ממשקי משתמש מובנים יותר והבנת התהליכים הקוגניטיביים הכרוכים בהפקת הסברים הולמים.
- פיתוח מנגנון הערכה להשפעת אפליה כדי לאפשר לרשויות לפעול במהירות וביעילות אך בו בזמן להבטיח מניעת אפליה ולשקול את ההשפעות החברתיות של האלגוריתם המפותח.

התוכנית קוראת להבטיח נגישות לכמות רבה של נתונים לכמה שיותר משתמשים באופן מאוזן – גופים ממשלתיים לצד שחקנים כלכליים קטנים יותר ומחקר ציבורי. רשויות ציבוריות חייבות להציג דרכים חדשות להפקת נתונים, חלוקתם ומשילה בהם על ידי הפיכתם לטובין משותפים⁸⁶. לשם כך יש לעודד

⁸¹ https://www.nist.gov/system/files/documents/2019/08/10/ai_standards_fedengagement_plan_9aug2019.pdf

⁸² <https://www.whitehouse.gov/ai-american-innovation/>

⁸³ https://nsf.gov/news/news_summ.jsp?cntn_id=296431

⁸⁴ <https://www.nitrd.gov/pubs/National-AI-RD-Strategy-2019.pdf>

⁸⁵ https://www.neaman.org.il/Files/Artificial%20Intelligence,%20Data%20Science,%20and%20Smart%20Robotics-%20A%20report%20on%20Ethics,%20Law%20and%20Privacy_20181230121054.072.pdf

⁸⁶ ניתן לראות דוגמה באסטרטגית הבינה המלאכותית הצרפתית: <https://towardsdatascience.com/the-french-national-strategy-on-artificial-intelligence-c8c8fcfdace1>

שחקנים כלכליים לחלוק ולאגד את הדאטה שלהם, כאשר המדינה תפעל כצד שלישי אמין. בחלק מהמקרים המדינה תוכל לכפות פתיחות לגבי דאטה שיש בו עניין ציבורי⁸⁷.

בתחום התשתיות, הכותבים ממליצים על בניית מחשב-על המעוצב במיוחד לבניה מלאכותית כדי לאפשר לחוקרים לעמוד בתחרות מול משאבי המחשוב של המגזר הפרטי. כמו כן, נחוצה חבילת גישה לענן פרטי שיפותח ברחבי אירופה ויותאם לצרכים והמאפיינים של בינה מלאכותית מבחינת זמן עיבוד וכוח אחסון⁸⁸.

בתוכנית ישנה גם התייחסות מעניינת לאקולוגיה וקיימות. המסמך מדגיש את הצורך בפיתוח בינה מלאכותית ירוקה יותר; כלומר, לבחון את טביעת הרגל האקולוגית של מוצרים, שירותים, תוכנה וחומרה הקשורים לבניה מלאכותית בשאיפה לצמצם אותה. בין היתר, התוכנית קוראת לרשויות ציבוריות לפעול כדי להפוך את שרשרת הערך לירוקה יותר על ידי תמיכה בתעשיית הענן האירופאית וקריאה לאופטימיזציה בתחום האנרגיה באמצעות שימוש בשיטות עבודה מומלצות קיימות בכל הסקטור האירופי⁸⁹.

המסמך של הוועדה הצרפתית הלאומית לטכנולוגיות מידע וחירויות (CNIL) מספק גם כן מספר המלצות למדיניות שעולות בקנה אחד עם המלצות התוכנית הלאומית. בין היתר ההמלצות כוללות יצירת אלגוריתמים מובנים יותר וממשקים ברורים למשתמש, והקמת פלטפורמה לאומית לבקרת אלגוריתמים ובחינת פעילותם בהתאם לחוק⁹⁰.

4.3.4 בריטניה

הצורך באסדרה של תחום הבינה המלאכותית ומדעי הנתונים עולה במספר רחב של פרסומים לאומיים, שמספקים המלצות לפעולה ולחקיקה.

- במסמך וועדת המדע והטכנולוגיה של בית הנבחרים הבריטי על רובוטיקה ובינה מלאכותית התייחסו הכותבים לצורך במסגרת עבודה לאסדרה וחקיקה של התחום. הומלץ על הקמת וועדה לבניה מלאכותית שתסקור את ההשפעות החברתיות, האתיות והחוקיות של התפתחות התחום ותנסח עקרונות מנחים לאסדרה⁹¹. כתוצאה מכך, בינואר 2017 הוקמה קבוצת מחוקקים (APPG, All-Party Parliamentary Groups AI) לבחינת הנושא. הוועדה התמקדה בשישה תחומי מדיניות: דאטה, תשתיות, מיומנויות, יזמות וחדשנות, מסחר ואחריותיות. המלצות מדיניות מעניינות לגבי תחום הדאטה קוראות להפקדת תחום ה-Data Governance וה-Data Governance בידי מרכז חדש לאתיקה ויזמות בנתונים, והקמת Data Trusts⁹² שיפקחו על העברת מידע אתית בין ארגונים. לגבי תשתיות, המסמך קורא להבטיח תשתיות אוניברסליות בסיסיות כמו רשתות דור חמישי בכל חלקי המדינה, כך שבכל אזור תהא גישה לתשתית פיזית ודיגיטלית אמינה ובטוחה⁹³.

https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf⁸⁷

שם⁸⁸

https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf⁸⁹

https://www.cnil.fr/sites/default/files/atoms/files/cnil_rapport_ai_gb_web.pdf⁹⁰

<https://publications.parliament.uk/pa/cm201617/cmselect/cmsctech/145/145.pdf>⁹¹

⁹² גישה להתמודדות עם תהליך קבלת החלטות בדומה לאופן שקרנות נאמנות מסייעות לקבל החלטות לגבי נכסים אחרים על ידי העברת הסמכות לקבלת החלטות לידי הנאמן לרווחת קבוצה גדולה יותר של בעלי עניין.

http://www.appg-ai.org/wp-content/uploads/2017/12/appgai_2017_findings.pdf⁹³

- מסמך נוסף שמספק המלצות למדיניות בתחום הבינה המלאכותית והנתונים הוא "בינה מלאכותית בממלכה המאוחדת: מוכנה, נכונה ומסוגלת?" שפורסם ב-2018. המסמך מתייחס לחשיבות דאטה נגיש וזמין, ולקושי בהפיכת הנתונים לאנונימיים מבלי לפגוע בשימושיות שלהם ומבלי לאפשר זיהוי-מחדש של המידע המזהה שהוסר. הומלץ במסמך על הנגשת נתונים ציבוריים באופן אתי עבור חוקרים ומפתחים, כדי לאפשר להם להתחרות בכמויות הנתונים האדירות שברשות החברות האמריקאיות. לתפיסתם, נתונים ציבוריים יהיו מגוונים יותר לרוב, וכך יוכלו לסייע גם בהתמודדות עם הטיה אלגוריתמית. הדבר ידרוש מעורבות של משרדי הממשלה וארגונים ציבוריים להפוך את המסמכים שברשותם לדיגיטליים ותואמים לפורמט אחיד. שחרור הנתונים הללו ייעשה לאחר אנונימיזציה מספקת, ויחוזקו מנגנונים חוקיים וטכניים שיאפשרו שליטה של פרטים על הנתונים שלהם והגנה על פרטיותם.⁹⁴

- בספטמבר 2018 ממשלת בריטניה פתחה בניסוי עם הפורום הכלכלי הבינלאומי לפיתוח מדיניות רכש לבינה מלאכותית,⁹⁵ ומאז 2019 היא המדינה הראשונה שמתנסה בשימוש בקווים המנחים האלו במשרדי המדינה השונים.⁹⁶ מדובר בעשרה קווים מנחים ושיקולים שיש לקחת בחשבון בתהליך הרכש, בכתיבת הצעה לרכישה (request for proposal – RFP) ובהערכת התגובות להצעה:⁹⁷

- השתמשו בתהליכי רכש שמתמקדים בהצגת בעיות והזדמנויות ומותירים מקום לאיטרציה, במקום בכאלו שמציעים פתרון מסוים.
- הגדירו את התועלת הציבורית בשימוש בבינה מלאכותית, תוך הערכת הסיכונים.
- התאימו את הרכישה לאסטרטגיות ממשלתיות קיימות ותרמו לשיפורן.
- התייחסו לחקיקה פוטנציאלית וכללי התנהגות רלוונטיים בהצעת הרכש שלכם.
- נסחו את רמת הגישה אל הנתונים ברמה טכנית ומנהלית.
- הדגישו את המגבלות האתיות והטכניות של השימושים המיועדים בנתונים כדי להימנע מבעיות הטיה.
- עבדו בצוותים מגוונים ומולטי-דיסציפלינריים.
- התמקדו במהלך תהליך הרכש במנגנונים לאחריותיות אלגוריתמית ולנורמות של שקיפות.
- הטמיעו תהליך להמשך התקשרות של הגורם הרוכש עם ספק הבינה המלאכותית להעברת ידע והערכת סיכונים בטווח הרחוק.
- צרו תנאים הוגנים ומאוזנים עבור ספקי הבינה המלאכותית השונים.

- יש לציין את ה-AI Sector Deal שמדגיש את המחויבות הכלכלית הבריטית לרתום את הבינה המלאכותית ומדעי הנתונים לפיתוח כלכלי ותעשייתי. התוכנית כוללת השקעה של 0.95 מיליארד ליש"ט לתמיכה במגזר הבינה המלאכותית שכולל את הממשלה, התעשייה והאקדמיה. התוכנית מבקשת לקדם את האימוץ והשימוש בבינה מלאכותית במדינה. בין המטרות שמציבה התוכנית ישנה התייחסות לחיזוק התשתיות הדיגיטליות והפיזיות, ותשתיות בתחום הנתונים באמצעות

⁹⁴<https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf>

⁹⁵<https://www.eurasiareview.com/21092018-uk-partners-with-world-economic-forum-to-develop-first-artificial-intelligence-procurement-policy/>

⁹⁶<https://www.weforum.org/press/2019/09/uk-government-first-to-pilot-ai-procurement-guidelines-co-designed-with-world-economic-forum/>

⁹⁷http://www3.weforum.org/docs/WEF_Guidelines_for_AI_Procurement.pdf

פריסת רשתות דור חמישי וסיבים אופטיים ברחבי המדינה. כמו כן, התוכנית בוחנת מסגרות-עבודה לשיתוף מידע כמו ה-Data Trusts⁹⁸.

Guideline	Principles
1. Use procurement processes that focus not on prescribing a specific solution but rather on outlining problems and opportunities, and allow room for iteration.	<ul style="list-style-type: none"> a. Make use of innovative procurement processes to acquire AI systems. b. Focus on developing a clear problem statement, rather than on detailing specifications of a solution. c. Support an iterative approach to product development.
2. Define the public benefit of using AI while assessing risks.	<ul style="list-style-type: none"> a. Set out clearly in your RFP why you consider AI to be relevant to the problem and be open to alternative technical solutions. b. Explain in your RFP that public benefit is a main driver of your decision-making process when assessing proposals. c. Conduct an initial AI risk and impact assessment before starting the procurement process, ensure that your interim findings inform the RFP, and revisit the assessment at decision points.
3. Align your procurement with relevant existing governmental strategies and contribute to their further improvement.	<ul style="list-style-type: none"> a. Consult relevant governmental initiatives such as AI national strategies, innovation and/or industrial strategies, and guidance documents informing public policy about emerging technologies. b. Collaborate with other relevant government bodies and institutions to share insights and learn from each other.
4. Incorporate potentially relevant legislation and codes of practice in your RFP.	<ul style="list-style-type: none"> a. Conduct a review of relevant legislation, rights, administrative rules and other relevant norms that govern the types of data and kinds of applications in scope for the project and reference them in the RFP. b. Take into consideration the appropriate confidentiality, trade-secret protection, and data-privacy best practices that may be relevant to the deployment of the AI systems.
5. Articulate the technical and administrative feasibility of accessing relevant data.	<ul style="list-style-type: none"> a. Ensure that you have proper data governance mechanisms in place from the start of the procurement process. b. Assess whether relevant data will be available for the project. c. Define if and how you will share data with the vendor(s) for the procurement initiative and the subsequent project.
6. Highlight the technical and ethical limitations of intended uses of data to avoid issues such as historical data bias.	<ul style="list-style-type: none"> a. Consider the susceptibility of data that could be in scope and if usage of the data is fair. b. Highlight known limitations (e.g. quality) of the data in the RFP and require tenderers to describe their strategies on how to address these shortcomings. Have a plan for addressing relevant limitations that you may have missed.
7. Work with a diverse, multidisciplinary team.	<ul style="list-style-type: none"> a. Develop ideas and make decisions throughout the procurement process in a multidisciplinary team. b. Require the successful bidder(s) to assemble a team with the right skill set.
8. Focus throughout the procurement process on mechanisms of algorithmic accountability and of transparency norms.	<ul style="list-style-type: none"> a. Promote a culture of accountability across AI-powered solutions. b. Ensure that AI decision-making is as transparent as possible. c. Explore mechanisms to enable interpretability of the algorithms internally and externally as a means of establishing accountability and contestability.
9. Implement a process for the continued engagement of the AI provider with the acquiring entity for knowledge transfer and long-term risk assessment.	<ul style="list-style-type: none"> a. Consider during the procurement process that acquiring a tool that includes AI is not a one-time decision; testing the application over its lifespan is crucial. b. Ask the AI provider to ensure that knowledge transfer and training are part of the engagement. c. Ask the AI provider for insights on how to manage the appropriate use of the application by non-specialists.
10. Create the conditions for a level and fair playing field among AI solution providers.	<ul style="list-style-type: none"> a. Reach out in various ways to a wide variety of AI solution providers. b. Engage vendors early and frequently throughout the process. c. Ensure interoperability of AI solutions and require open licensing terms to avoid vendor lock-in.

טבלה 6: קווים מנחים לרכש בינה מלאכותית, הפורום הכלכלי הבינלאומי, 2019

4.3.5 דרום קוריאנה

המסמכים הלאומיים הקיימים בתחום מצויים רשימה מפורטת של צעדי מדיניות שיש לנקוט במספר תחומים כדי למצב את דרום קוריאנה כאחת המדינות המובילות בתחום.

המשרד למדע, טכנולוגיות מידע ותקשורת ותכנון עתידי (Ministry of Science, ICT and MSIP) פרסם אסטרטגיה לפיתוח תעשיית בינה מלאכותית ומידע. האסטרטגיה מתייחסת להשפעות אפשריות של בינה מלאכותית על מגוון תחומים כמו שוק העבודה וכלכלה. שלוש המטרות העיקריות המפורטות בתוכנית הן הקמת תעשיית ברמה עולמית לטכנולוגיות תבוניות, קידום

⁹⁸<https://www.gov.uk/government/publications/artificial-intelligence-sector-deal/ai-sector-deal#executive-summary>

היישום של טכנולוגיות אלו בכל התעשיות ונקיטת צעדים יזומים לחיזוק ושיפור מערכת התמיכה החברתית. כל מטרה כזו מחייבת שורת משימות שמפורטות לאורך המסמך. מספר המלצות מעניינות עולות במסמך בכל הקשור לתשתיות ונתונים:⁹⁹

- **הקמת מערכת לאומית לניהול נתונים שתאפשר לפתח תשתית לנתוני-עתק עבור למידת מכונה.** לשם כך יש להמיר מידע ציבורי ממשלתי לפורמטים פתוחים שמותאמים ללמידת מכונה, ועל סוכנויות וארגונים ציבוריים לזהות ולחשוף את הנתונים שברשותם. כמו כן, יש לפתח מערכת פתוחה המחוברת לפלטפורמות נתונים ציבוריות ופרטיות עבור אחסון ושימוש סיסטמתי בנתונים שמקורם סנסורים מערים חכמות. בנוסף, יש לאפשר אימוץ של שירותי ענן ופלטפורמות מחשוב פרטיים על ידי ארגונים גדולים שברשותם כמויות נתונים גדולות על ידי שינוי הרגולציה הקיימת. יש לפעול לקידום שירותים נוספים בתחום מכירת הנתונים. למשל, שירות שמחבר בין חברות בעלות נתונים וחברות שעוסקות בניתוחם.
- **מציאת מגוון אמצעים לתמוך בהפצת נתונים ובשימושם בהתאם לסוג הנתונים.** יש להבחין בין מידע כללי, מידע לא-פרטי, מידע לא-מזהה (כלומר מידע פרטי שמידע מזהה הוסר ממנו) ומידע פרטי אישי, וליצור אסטרטגיה מתאימה לניהולם ולשימוש בהם.
- **עידוד שימוש בטכנולוגיית בלוקצ'יין לחיזוק ניהול נתונים ואבטחתם.** הטכנולוגיה משמשת היום בעיקר בתחומי הפיננסיים, ולפי התוכנית יש להרחיב את השימוש בה לתחומים נוספים, בין היתר באמצעות יוזמות בינלאומיות לפיתוח טכנולוגיות בלוקצ'יין יישומיות ועל ידי רפורמה ברגולציה הקיימת.
- **טיפוח האקו-סיסטם של תעשיית הבינה המלאכותית על ידי שיתוף נתונים ואלגוריתמים** שפותחו על ידי המגזר הציבורי בצורה של תוכנת קוד-פתוח ו-APIs עם עסקים, אוניברסיטאות וגופי מחקר.
- **יש לאמץ תקנים בינלאומיים מעודכנים** לגבי הבנה לשונית והבנה חזותית (linguistic understanding and visual understanding) ולהבטיח עמידה בתקנים הללו, ולפתח קריטריונים להערכה להגברת התחרותיות שלהם.
- **יש לפרוס שירותי רשת מהדור החמישי** כדי לאפשר רוחב פס משופר, תקשורת אמינה ובעלת זמן שהייה נמוך, שתוכל להתמודד עם תקשורת מסיבית (למשל בערים חכמות).
- **יש למקד את משאבי המדיניות כדי לחזק את האיכות והאמינות של רשתות, ולפעול לקידום קריפטוגרפיה מבוססת מחשוב-קוונטי** כדי למזער את הסכנה של מתקפות סייבר ברמת השורש.
- **יש להפעיל "אתרי-מבחן" (Test-beds) ביעוץ עם משרדי ממשלה רלוונטיים,** כדי לתמוך בחדשנות בשירותי ערים חכמות, רובוטיקה חכמה, רכבים אוטונומיים וכו'. יש לבסס חיבורים בין אתרי המבחן הללו, ואזורים חופשיים מרגולציה כדי לאפשר לתעשיות אסטרטגיות מקומיות להתפתח, ולשפר תשתית תעשייתית קיימת.
- **יש לאמץ פרדיגמה רגולטורית חדשה ולהסיר מכשולים רגולטוריים ליזמות.** יש לעדכן את תוכנית הרישוי הקיימת כדי למנוע עיכוב בפיתוח והגשמה של טכנולוגיות תבוניות לאור מגבלות מוסדיות

⁹⁹

https://english.msit.go.kr/cms/english/pl/policies2/_icsFiles/afieldfile/2017/07/20/Master%20Plan%20for%20the%20intelligent%20information%20society.pdf

וחוקיות, וליצור "ארגז-חול" (Sand-box) רגולטורי – תוכנית מדיניות שמאפשרת בדיקה של טכנולוגיות ושירותים חדשים למטרות ותקופות מוגדרות מבלי להיות מוגבלות על ידי הרגולציה הקיימת.

במסמך ממשלתי נוסף, "Toward AI World Leader: Beyond IT", שפורסם ב-2019, שמטרתו לקדם את היוזמה הנשיאותית עבור בינה מלאכותית, פורטו המלצות נוספות למדיניות, בין היתר לגבי אתיקה, תשתיות ונתונים¹⁰⁰.

המסמך קורא לביסוס אתיקה לבינה מלאכותית ברמה הגלובלית וצעדי המשך להמלצות ה-OECD בנושא, ומבקש לבסס סטנדרטים לבינה מלאכותית אתית שעומדים בקנה אחד עם נורמות גלובליות. ישנה התייחסות במסמך גם לצורך בעדכון חקיקה קיימת להגנת מידע ופרטיות ולהרחבת פרויקטים קיימים שדורשים אישור והסכמה מצד בעלי הנתונים לשימוש בהם על ידי גופים שונים. גם תוכנית זו מתייחסת לצורך ביכולת הסברתיות והבנת תהליך קבלת ההחלטות של אלגוריתמים מבוססי בינה מלאכותית.

גם במסמך זה התייחסות לצורך בפתיחת נתונים ציבוריים שמצויים ברשות מוסדות ציבור ולקידום השימוש בהם בתעשיות חדשות כמו ערים חכמות ומכוניות אוטונומיות, והתייחסות לחשיבות חיזוק הקשר בין הנתונים הקיימים במגזר הציבורי והפרטי. במסמך צוינה הקמת "AI Hub", פלטפורמות מאובטחות לחיבור בין הנתונים הקיימים במגזר הציבורי ואלו הקיימים במגזר הפרטי שישמשו למחקר ופיתוח. הצעה מעניינת נוספת שעולה במסמך היא מערכת ואוצרים לבינה מלאכותית, שתחבר בין חברות בעלות צורך בפתרונות בינה מלאכותית לספקים שיפתחו את הפתרון עבורן.

ישנה התייחסות גם לצורך בהגברת התחרותיות בתחום המוליכים-למחצה, באמצעות פיתוח יזום של הדור הבא של מוליכים-למחצה תבוניים. לשם כך יושקע תקציב לעיצוב טכנולוגיה בעלת מהירות עיבוד משופרת וצריכת חשמל מופחתת, לצד יצירת שבבי בינה מלאכותית חדשים שמשלבים זיכרון ועיבוד (Processing-In-Memory). הושם דגש גם על ביטחון סייבר, ופיתוח מערכות לגילוי, ניתוח ותגובה לאיומי סייבר המבוססת על בינה מלאכותית, וטכנולוגיה מבוססת בינה מלאכותית להגנה אוטומטית על מידע באמצעות ניתוח חולשות של מכשירים ורשתות ובטיחות סיסמאות.

התוכנית קוראת ליצירת סביבה רגולטורית מאפשרת בה חברות ומפתחים יוכלו "לדמיין ללא גבולות" ולהתמודד ביחד עם אתגרים חדשים. התוכנית קוראת לשינוי הפרדיגמה הרגולטורית הקיימת, כך שקודם יאושר השימוש בטכנולוגיה, והאסדרה המתאימה תיווצר לאחר מכן (Approval first and "Regulate later"). לשם כך יעשה שימוש בארגזי חול רגולטוריים שיאפשרו ניסוי של טכנולוגיות חדשות ללא המגבלות הרגולטוריות הקיימות ותחת תנאים מסוימים, וכן שימוש בחקיקה גמישה ומקיפה בכל הקשור להגדרת מושגים ומונחים, סיווג של מערכות וכו'.

4.4. התייחסות מיוחדת: תחומי מפתח

ישנם מספר תחומי מפתח שיתרמו משמעותית מבינה מלאכותית, אך מעוררים אתגר בכל הקשור לתשתיות משותפות ושיתוף נתונים. חלק מהתוכניות הלאומיות שנסקרו מתייחסות לתחומים אלו ומציגות את הפוטנציאל הטמון באימוץ הטכנולוגיות על ידי מגזרים ותעשיות אלו, אך גם את האתגרים

¹⁰⁰https://english.msit.go.kr/cms/english/pl/policies2/_icsFiles/afieldfile/2020/03/23/National%20Strategy%20for%20Artificial%20Intelligence_200323.pdf

שעשויים לעלות מהטמעת בינה מלאכותית לתוכם, והאפשרויות להתמודדות איתם. שלושה תחומים נבחרו כדוגמה לפעילות בתחום – בריאות, פיננסיים ותחבורה.

4.4.1 בריאות

מדינות שונות התייחסו לפוטנציאל הטמון ברתימת בינה מלאכותית ומדעי הנתונים ככלים עבור מערכת הבריאות הלאומית והציגו את הצורך והקושי העולים מכך. מספר תוכניות מתייחסות לאפשרות להשתמש בכלים טכנולוגיים אלו לאבחון מחלות וטיפול בהן, ואת התשתיות הנחוצות לשם כך.

בדרום קוריאה מודגשת החשיבות של קידום השימוש בבינה מלאכותית במגוון תעשיות (פרויקט AI+X) באמצעות שימוש בנתונים ציבוריים רחבי היקף, כאשר אחד התחומים המוזכרים הוא רפואה. מוזכרות בתוכניות השונות מספר מטרות חשובות לצד מועד השלמתן הרצוי. אחת המטרות החשובות, שמהווה בסיס לאחרות במובנים מסוימים, היא ביסוס סביבת נתונים שתתמוך במתן שירותי בריאות יעילים ואיכותיים. לפי התוכנית, לשם כך נחוצה רשת שתאפשר לכל המוסדות הרפואיים במדינה לשתף את התיקים הרפואיים האלקטרוניים שברשותם, ונחוצה סטנדרטיזציה שתגדיר פורמט אחיד לתיקים רפואיים אלו (כרגע בכל בית חולים פורמט אחר). התוכנית מבקשת לאגד את הנתונים הרלוונטיים המפוזרים בארגונים ציבוריים ופרטיים בפלטפורמה משותפת כדי לבסס תשתית מקיפה עבור יצירת שירותי בריאות מותאמים אישית ופיתוח תרופות מהיר. הפלטפורמה שתפותח תכלול שש קטגוריות נתונים עיקריות, שיאפשרו דיוק רפואי רב יותר: תיקים רפואיים, מרשמים, מידע בריאותי אישי, מידע גנטי, מידע על אורח חיים, מידע סביבתי. בין היתר, צוין בתוכנית גם פיתוח מאגר ידע מבוסס בינה מלאכותית, שישתמש בנתונים הקיימים בבתי החולים ובשירותי הביטוח הלאומיים, ופיתוח מערכת שתשלב בין מידע גנטי והנתונים הקיימים בתיקים רפואיים דיגיטליים. מטרה תומכת נוספת היא ביסוס מקבץ מבחן (test cluster) לבחינת רובוטיקה ושירותים תבוניים בתחום הבריאות, לצד תמיכה בפיתוח ומסחור של ציוד ומכשור רפואי תבוני מתהליך הניסוי הקליני ועד הרישוי^{101 102}.

מגפת הקורונה (COVID-19) שפקדה את העולם בשנה האחרונה הדגישה את האופן בו כלים טכנולוגיים שונים יכולים לסייע ולתמוך במערכת הבריאות במגוון אופנים, ומדינות רבות אכן ניצלו את הטכנולוגיה ברשותם במאבק. דרום קוריאה פרסמה באפריל 2020 מסמך מקיף שמציג את התגובה הלאומית למגיפה בשימוש בטכנולוגיות מידע ותקשורת. מוצג השימוש בטכנולוגיות אלו לשמירה על ריחוק חברתי, איתור ומעקב מהיר אחר הפצת והתפשטות הנגיף וטיפול בנגיף¹⁰³.

במסמכים לאומיים שונים של **בריטניה** ישנה התייחסות לכך שהשימוש בבינה מלאכותית ומדעי הנתונים יכול להוביל למהפכה של ממש בתחום הבריאות, אך דגש רב הושם על אבטחת הנתונים הרפואיים והגנה על המטופלים. למשל, נרשם כי הנתונים המצויים בידי שירות הבריאות הלאומי (NHS, National Health Service) יוכלו להוות מקור בעל ערך רב, אך שיתופם צריך להיעשות במשנה זהירות תוך הבטחת אנונימיזציה ויידוע של בעל הנתונים על השימוש בהם, כולל אפשרות לביטול

¹⁰¹https://english.msit.go.kr/cms/english/pl/policies2/_icsFiles/afieldfile/2020/03/23/National%20Strategy%20for%20Artificial%20Intelligence_200323.pdf

¹⁰²https://english.msit.go.kr/cms/english/pl/policies2/_icsFiles/afieldfile/2017/07/20/Master%20Plan%20for%20the%20intelligent%20information%20society.pdf

¹⁰³https://english.msit.go.kr/cms/english/pl/policies2/_icsFiles/afieldfile/2020/04/17/Flattening%20the%20curve%20on%20COVID-19.pdf

הסכמתו (opt-out). לכן החלו לפתח מדיניות לאופן בו ניתן לקבל גישה עבור נתוני המטופלים למטרות מחקר וחדשנות, ולטובת המטופל והציבור¹⁰⁴.

ב-2019 פורסם קוד התנהגות לטכנולוגיות נתונים בתחום הבריאות, שמתייחס לאתגר האתי בשימוש בנתונים רפואיים ורגישים בטכנולוגיות נתונים. הקוד כולל עשרה עקרונות. בין העקרונות ניתן לציין שימוש בנתונים בהתאם לקווים מנחים ראויים ובהתאם למטרה לשמה משמשים תוך שימוש מינימלי בנתונים אישיים נחוצים, ושימוש בסטנדרטים פתוחים להבטחת תקשורת קלה עם מערכות קיימות וביצוע הערכות לאיכות הנתונים¹⁰⁵. שירות הבריאות הלאומי אף פרסם ארגז כלים מתעדכן להגנה וביטחון בנתונים – כלי מקוון להערכה עצמית שמאפשר לארגונים למדוד ולפרסם את ביצועיהם בתחום הנתונים. כל ארגון שיש לו גישה לנתוני מטופלים חייב להשתמש בארגז הכלים כדי להבטיח שימוש תקין בהם¹⁰⁶.

ב-2019 פורסם נייר מדיניות בשם "The Grand Challenge missions" שמציג את האתגרים הגדולים העומדים בפני המדינה בהובלת התעשיות שלה לחזית בעתיד, והצגת פתרונות לאתגרים האלו. אחת המשימות המצוינות שם היא שימוש בנתונים, בינה מלאכותית וחדשנות במניעה, אבחון מוקדם וטיפול במחלות כרוניות עד 2030 והצבת בריטניה כמובילה גלובלית בתחום. מספר פעולות לקידום המטרה הושגו, כמו קוד ההתנהגות המוזכר לעיל. בין הפעולות העתידיות מתוכננת יצירת "Local Health and Care Record Exemplars" – פלטפורמה לשיתוף ושימוש בנתוני מטופלים מכלל חלקי מערכת הבריאות הלאומית, כולל ברמה המקומית, בהתאם לכללי הגנת מידע ופרטיות והבטחת שקיפות¹⁰⁷.

גם ארצות הברית מכירה בערך של בינה מלאכותית עבור מגוון תעשיות, כאשר בריאות היא אחת מהן. נרשמו מספר יוזמות לקידום התחום. כדי להסיר חסמים ליזמות טכנולוגית, מנהל המזון והתרופות האמריקאי (FDA, Food and Drug Administration) אפשר לראשונה שיווק של טכנולוגיית בינה מלאכותית לזיהוי בעיות עיניים הנגרמות מסוכרת, ואפשר שיווק של תוכנה מבוססת בינה מלאכותית המסייעת לספקי טיפול רפואי לזהות שברים בשורש כף היד יותר בקלות. כמו כן, המנהל חבר למרכז לבקרת ומניעת מחלות (CDC, Centers for Disease Control) לקידום מחקר בלמידת מכונה ועיבוד שפה טבעית ליצירת כלים חינוכיים לשיפור איסוף נתונים קליניים. ב-2019 החל המנהל בפיתוח מסגרת עבודה רגולטורית ייחודית לתוכנות חכמות בכלים רפואיים המבוססת על מעגל-החיים הכולל של המוצר^{108, 109}.

כמו כן ה-FDA פרסם תוכנית בשם "The Software Precertification Program" שמהווה נתיב וולונטרי בדרך למודל רגולטורי מותאם אישית להערכת הבטיחות והיעילות של טכנולוגיות תוכנה (למשל אפליקציות רפואיות) מבלי למנוע את הגישה של מטופלים לטכנולוגיות האלו. מטרת התוכנית היא לספק פיקוח רגולטורי יעיל יותר על מכשירים וכלים רפואיים מבוססי תוכנה, לאור העובדה

¹⁰⁴ <https://www.parliament.uk/documents/lords-committees/Artificial-Intelligence/AI-Government-Response.pdf>

¹⁰⁵ <https://www.gov.uk/government/publications/code-of-conduct-for-data-driven-health-and-care-technology/initial-code-of-conduct-for-data-driven-health-and-care-technology>

¹⁰⁶ <https://digital.nhs.uk/data-and-information/looking-after-information/data-security-and-information-governance/data-security-and-protection-toolkit>

¹⁰⁷ <https://www.gov.uk/government/publications/industrial-strategy-the-grand-challenges/missions>

¹⁰⁸ <https://www.whitehouse.gov/ai/ai-american-industry/#sector-healthcare>

¹⁰⁹ <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device>

שהרגולציה הקיימת עבור חומרה רפואית אינה תואמת את הצרכים העולים משימוש ביישומי תוכנה רפואיים שונים ואת קצב השינוי המהיר שלהם¹¹⁰.

4.4.2. פיננסים

ההתקדמות הטכנולוגית המהירה והדיגיטציה של נתונים כלכליים נותנת את אותותיה גם על המגזר הפיננסי, כאשר מדינות שונות בוחנות כיצד להתאים את הרגולציה והמדיניות הקיימת לפיתוחים טכנולוגיים בתחום וכיצד לרתום אותם לשימושה מבלי לעורר סיכון למשתמשים.

ארצות הברית פועלת לקידום הטמעה של בינה מלאכותית בתחומי הפיננסים. למשל, פעילות משרד האוצר לאימוץ והטמעת כלים מבוססי בינה מלאכותית ולמידת מכונה באמצעות הסרת חסמים רגולטוריים פוטנציאליים. המסמך שפורסם על ידי המשרד מכיר בכך שפיתוחים טכנולוגיים שונים יוכלו לשפר את השימוש בשירותים פיננסיים שונים עבור אינדיבידואלים ועסקים, ולהפוך אותו למהיר, יעיל וחסכוני יותר. המסמך מכיל ארבע המלצות עיקריות: אימוץ גישות רגולטוריות לשינויים באגירת, שיתוף ושימוש במידע פיננסי של צרכנים ותמיכה בפיתוח טכנולוגיות מפתח תחרותיות; התאמת המסגרת הרגולטורית למאבק בפיצול רגולטורי מיותר, והתייחסות למודלים עסקיים חדשים המתאפשרים על ידי טכנולוגיות פיננסיות; עדכון רגולציות של פעילויות ספציפיות במגוון שירותים ומוצרים המוצעים על ידי מוסדות פיננסיים, שרובן אינן מעודכנות או תואמות את ההתקדמות הטכנולוגית; קידום גישה לרגולציה שמאפשרת ניסוי אחראי במגזר הפיננסי, משפרת גמישות רגולטורית ומקדמת את האינטרסים האמריקאים בעולם¹¹¹. בנוסף, ב-2019 הרשות להגנת צרכנים פיננסיים (Consumer Financial Protection Bureau) פרסמה מדיניות חדשה שמאפשרת שימוש רב יותר בנתונים ובאלגוריתמי למידת מכונה בשירותים ומוצרים פיננסיים. בין היתר, המדיניות מאפשרת בחינת מוצר או שירות פיננסי היכן שיש אי וודאות רגולטורית באמצעות ארגזי חול רגולטוריים¹¹².

בצרפת פורסם ב-2020 מסמך שמבקש להאיר על הסוגיות של הסברתיות ומשילות בבינה מלאכותית בהקשר הפיננסי. במסמך ניתן דגש על הערכה של אלגוריתמים מבוססי בינה מלאכותית ומשילות שלהם. הוצגו ארבעה קריטריונים עיקריים להערכת אלגוריתמים וכלים פיננסיים מבוססי בינה מלאכותית: (1) ניהול נתונים ראוי שמתייחס לשיקולים אתיים כמו הוגנות והעדר הטיה; (2) ביצוע של האלגוריתם, אותו ניתן להעריך בכמה מדדים, כשמגוון המדדים הזמין מספק להערכת הדיוק של האלגוריתם הן מבחינה טכנית והן מבחינת פונקציונליות. לעתים יש לאזן בין הקריטריונים שנבחרו למידת ההסברתיות הרצויה; (3) יציבות מתארת עד כמה התנהגות האלגוריתם היא איתנה ועמידה לאור מחזור החיים שלו; (4) הסברתיות נחוצה להבנת תוצאה מסוימת של האלגוריתם שעשויה להיות בעלת ערך למשתמש הקצה, שתאפשר לתקף את החלטות האלגוריתם וגם ליידע את המשתמש על התהליך כך שיוכל לקבל החלטה מושכלת. בכל הקשור למשילות בהטמעת בינה מלאכותית לארגון פיננסי, הושם דגש על כמה היבטים שצריכים להילקח בחשבון: בחינת ההטמעה לתוך תהליכי הארגון – מה התפקיד שממלא האלגוריתם? מה המתודולוגיה לאורך מחזור חייו?; אינטראקציות אדם/מכונה; בטיחות ומיקור חוץ להתמודדות עם מתקפות ולהערכת סיכון; תהליך תיקוף פנימי;

¹¹⁰<https://www.fda.gov/media/113802/download>

¹¹¹<https://home.treasury.gov/sites/default/files/2018-07/A-Financial-System-that-Creates-Economic-Opportunities---Nonbank-Financi...pdf>

¹¹²<https://www.consumerfinance.gov/about-us/newsroom/bureau-issues-policies-facilitate-compliance-promote-innovation/>

תהליך תיקוף מתמשך לאחר הצבת האלגוריתם לשימוש; יכולת בקרה ובדיקה, הן של קוד המקוד, הנתונים והשיטות, והן של ההסברים והתנהגות האלגוריתם.¹¹³

4.4.3 תחבורה

התפתחות בתחום הרכבים האוטונומיים הביאה מדינות רבות לעסוק בתשתיות והרגולציה הדרושות להיערכות לפריסה רחבה של טכנולוגיה זו ולקידומה.

בדרום קוריאה התוכנית הלאומית העדכנית מגדירה כמטרה את יצירתה של סביבה לנהיגה אוטונומית בהתבסס על טכנולוגיית בינה מלאכותית. התוכנית קוראת לפיתוח טכנולוגיה שיתופית של רכבים אוטונומיים, ולפיתוח תחבורה ציבורית אוטונומית שתדע לבחור את הנתבים וזמני הנסיעה האופטימליים לפי דרישה ועומס.¹¹⁴ במסמך אחר צוין ביסוסה של תשתית לתחבורה חכמה, שתחבר את כל כלי הרכב ומתקני התנועה כדי לשפר בזמן אמת את הפעילות בכביש ולספק עדכוני תנועה מותאמים אישית. לשם כך נחוץ פיתוח מערכת תחבורה חכמה שתגובה על ידי טכנולוגיה לאיסוף נתוני-עתק בזמן אמת ובדיקת רב, כולל נתוני GPS של רכבים חכמים ונתונים מסנסורים אחרים; פיתוח מערכת לצמצום עומסי תנועה, שתדע לזהות את הגורמים להיווצרותם ולהפנות את התנועה בהתאם; ביסוס דרישות בטיחות לרכבים אוטונומיים וביצוע רפורמה בחקיקה ובמוסדות רלוונטיים, לצד פיתוח תשתית למסחור רכבים אוטונומיים ויזמות במגזר הפרטי (ערי מבחן, מפות מדויקות וכו'); פיתוח מערכת תחבורה שיתופית, שתאגד מידע על מכוניות אוטונומיות ושרותי רכב שיתופי, כך שאזרחים יוכלו לשכור ולנהוג ברכבים ותצומצם הבעלות על רכבים.¹¹⁵

משרד התחבורה בבריטניה פרסם הצהרה ב-2018 לגבי שימוש בבינה מלאכותית בתחום התחבורה. בהצהרה נכתב שמערכי נתונים בתחום המוחזקים על ידי גופי ציבור יבחנו עבור פתיחתם והנגשתם, כדי לאפשר לחברות בריטיות ולאקדמיה גישה לנתונים בעלי ערך. כמו כן, צוין שכדי ליצור שיתוף נתונים בין ארגונים, מספר גופים ומערכות קיימים יפעלו ליצירת חיבור בין הגורמים המחזיקים בנתונים והגורמים המעוניינים להשתמש בהם לפיתוח בינה מלאכותית, ויחליטו על תנאים לשימוש שיתאמו את צורכיהם.¹¹⁶

מסמך לאומי אחר מתייחס לאבטחת סייבר של רכבים אוטונומיים ומגדיר 8 עקרונות עיקריים (שמפורקים ומפורטים לאורך המסמך): הביטחון הארגוני נתון בידי רמת ההנהלה; סיכונים לביטחון מוערכים ומנוהלים כראוי ובמידתיות, כולל סיכונים לשרשרת האספקה; כדי להבטיח שהמערכות בטוחות בכל מחזור חייהן, על ארגונים לספק שירותי טיפול למוצר ותגובה לתאונות; על כל הארגונים והגורמים הרלוונטיים המעורבים לפעול יחד להגברת בטיחות המערכת באמצעות תיקוף, אימות וחיבור בטוח למכשירים חיצוניים; המערכות מעוצבות תוך שימוש בגישת defence-in-depth, כדי לא להסתמך על נקודת כשל יחידה; הביטחון של התוכנה כולה מנוהל לכל אורך חייה, בין היתר באמצעות עדכונים מאובטחים ופרקטיקות עיצוב פתוח; האחסון והתקשורת של הנתונים מאובטחת וניתנת

¹¹³ [https://acpr.banque-](https://acpr.banque-france.fr/sites/default/files/medias/documents/20200612_ai_governance_finance.pdf)

[france.fr/sites/default/files/medias/documents/20200612_ai_governance_finance.pdf](https://english.msit.go.kr/cms/english/pl/policies2/_icsFiles/afieldfile/2020/03/23/National%20Strategy%20for%20Artificial%20Intelligence_200323.pdf)

¹¹⁴ [https://english.msit.go.kr/cms/english/pl/policies2/_icsFiles/afieldfile/2020/03/23/National%20Strat-](https://english.msit.go.kr/cms/english/pl/policies2/_icsFiles/afieldfile/2020/03/23/National%20Strategy%20for%20Artificial%20Intelligence_200323.pdf)

[egy%20for%20Artificial%20Intelligence_200323.pdf](https://english.msit.go.kr/cms/english/pl/policies2/_icsFiles/afieldfile/2017/07/20/Master%20Plan%20for%20the%20intelligent%20information%20society.pdf)

¹¹⁵ [https://english.msit.go.kr/cms/english/pl/policies2/_icsFiles/afieldfile/2017/07/20/Master%20Plan%20-](https://english.msit.go.kr/cms/english/pl/policies2/_icsFiles/afieldfile/2017/07/20/Master%20Plan%20for%20the%20intelligent%20information%20society.pdf)

[20for%20the%20intelligent%20information%20society.pdf](https://www.gov.uk/government/publications/review-of-artificial-intelligence-in-transport-2017/position-statement-on-artificial-intelligence-in-transport#recommendations)

¹¹⁶ [https://www.gov.uk/government/publications/review-of-artificial-intelligence-in-transport-](https://www.gov.uk/government/publications/review-of-artificial-intelligence-in-transport-2017/position-statement-on-artificial-intelligence-in-transport#recommendations)

[2017/position-statement-on-artificial-intelligence-in-transport#recommendations](https://www.gov.uk/government/publications/review-of-artificial-intelligence-in-transport-2017/position-statement-on-artificial-intelligence-in-transport#recommendations)

לשליטה, כך שמשתמשים יוכלו למחוק נתונים רגישים; המערכת מעוצבת להיות עמידה בפני תקיפות ולהגיב בהתאם אם יש כשל בסנסורים או בהגנה¹¹⁷.

בנוסף, ננקטו מספר צעדים ויוזמות לשימוש בבינה מלאכותית עבור תחום התעבורה. למשל, שימוש משרד התחבורה בבינה מלאכותית כדי לשפר את תנאי הדרך בדרכים מקומיות ב-2019. הטכנולוגיה תאפשר למפות את מצע הדרכים הקיים ולשפר את איכותן עבור רכבים ואופניים¹¹⁸.

משרד התחבורה בארצות הברית פעל לפיתוח קווים מנחים להטמעה של רכבים אוטונומיים ומערכות נוהגות לתוך מערכת התעבורה הקיימת. בינואר 2020 פרסם המשרד תוכנית המפרטת 10 עקרונות ממשלתיים להגנה על משתמשים וקהילות, לקידום שווקים יעילים וליצירת מאמצי תיאום שיבטיחו סטנדרטיזציה אחידה¹¹⁹. מסמך זה נבנה על מסמכים קודמים של המשרד, שהתמקדו בחדשנות באמצעי הובלה על פני השטח ובבדיקה בטוחה והטמעה של מערכות נהיגה אוטונומיות¹²⁰.

4.5. המלצות למדיניות

מטרת חלק זה להציע את התשתית הערכית למדיניות וכן שורת המלצות לרגולציה של תחום הבינה המלאכותית.

4.5.1. אתיקה

על בסיס הסקירה האתית נבחרו מספר ערכים בולטים, שיוכלו להוות בסיס אתי לרגולציה ישראלית בתחום הבינה המלאכותית.

1. שקיפות

הנגשה של מידע אודות תהליך הפיתוח והעיבוד של האלגוריתמים עבור המשתמשים בהם, כדי לבסס אמון בטכנולוגיה ולהבטיח אחריותיות רבה יותר מצד העוסקים בפיתוח טכנולוגי. הנגשת המידע תוכל להתבצע באופן מדורג בהתאם לסוג המשתמש, המידע הנחוצ לו והבנתו הטכנולוגית ובאמצעות שיתופיות בנתונים ובמודלים.

2. הסברתיות

היכולת של משתמשים אנושיים להבין ולהסביר את תהליך קבלת ההחלטות של המערכת, כדי לפתור את סוגיית הקופסה השחורה ולאפשר מידת הסתמכות רבה יותר על מסקנותיה, במיוחד בתחומים בהם יש סיכון לחיי אדם. הסברתיות תאפשר לבקר את תהליך העיבוד של המערכת, ותבטיח יישום אחראי יותר של מסקנותיה. כמו כן, הסברתיות תאפשר להבין בשלבים מאוחרים האם המערכת חוותה "התעסקות" חיצונית אשר השפיעה על תכנונה המקורי.

3. הגנה על פרטיות ונתונים אישיים

כדי לאמן מערכות בינה מלאכותית ולהשתמש בהן, נחוצים נתונים רבים, בהם גם נתונים אישיים. לפיכך, קיים צורך בהגנה על נתונים אישיים ומידע פרטי, שתקיף את איסוף הנתונים, עיבודם והשימוש

¹¹⁷https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/61135/cyber-security-connected-automated-vehicles-key-principles.pdf

¹¹⁸<https://www.gov.uk/government/news/dft-to-embrace-artificial-intelligence-technology-in-plans-for-local-roads-health-check>

¹¹⁹<https://www.transportation.gov/sites/dot.gov/files/docs/policy-initiatives/automated-vehicles/360956/ensuringamericanleadershipav4.pdf>

¹²⁰<https://www.whitehouse.gov/ai/ai-american-industry/#sector-healthcare>

בהם על ידי גורמים שונים. המטרה היא לצמצם את הפגיעה בפרטיותם של אזרחים, אך בה בעת לאפשר שימוש בנתונים שמהווים חומר גלם חיוני למערכות בינה מלאכותית.

4. אבטחת סייבר

כאמור, בינה מלאכותית זקוקה לנתונים רבים בשלב האימון ובשלב הפעילות, ובמידה רבה מתבססת על טיבם. מאגרי נתונים ורשתות תקשורת המשמשים לפיתוח והטמעה של בינה מלאכותית עשויים להפוך מטרה נכספת לתוקפים, שיבקשו לגשת למידע רגיש או אישי, לבצע שימוש לרעה בנתונים, ואף לבצע שינויים בנתונים כדי לפגוע באיכות האלגוריתם המתבסס עליהם. לפיכך, נחוצה הגנת סייבר ואבטחה ראויה ומספקת לתשתיות הדרושות לפיתוח ושימוש בבינה מלאכותית.

4.5.2 רגולציה

בהתבסס על סקירת המדיניות לעיל והתייחסות לצורך הישראלי, מצורפות מספר המלצות לרגולציה בתחום הבינה המלאכותית. המלצות אלו מבקשות להתמודד עם הסוגייה האתית ובה בעת לספק מענה לקידום ופיתוח של בינה מלאכותית ומדעי הנתונים ברמה לאומית.

1. רגולציה מאפשרת

כדי למנוע סטגנציה טכנולוגית נחוצה רגולציה מאפשרת, במיוחד בכל הקשור לנתונים. כלי מדיניות בולטים בהקשר זה הם "ארגזי חול" ו"אתרי מבחן" רגולטוריים. הכוונה היא ליצירת סביבות מבוקרות לניסוי טכנולוגי, שיאפשרו לבחון את היכולות, ההשפעות והכשלים האפשריים הטמונים בטכנולוגיה טרם הטמעתה הרחבה בחברה או בארגון. ניסויים אלו תחומים לזמן ומקום מוגדרים, ומאפשרים לרגולטור לבחון את ההתפתחות הטכנולוגית ואת הצורך הרגולטורי שהיא מעלה, ובו בזמן מאפשרים התפתחות טכנולוגית מהירה שאינה מואטת על ידי רגולציה קיימת ולא מעודכנת.

2. תקינה וחקיקה

קיים צורך בפיתוח תקינה של תחום הבינה המלאכותית – הן מבחינת אלגוריתמיקה ומודלים והן מבחינת הנתונים – כך שתובטח הטמעה של העקרונות האתיים והבטיחותיים בשלבי המחקר, הפיתוח השימוש, וכדי להבטיח יכולת בקרה ומדידה של התוצרים ועמידתם בתקנים השונים. כמו כן, ייתכן ונחוץ עדכון של חקיקה קיימת בתחומי הנתונים והגנת המידע כך שהחקיקה תאפשר התייחסות לטכנולוגיות חדשות באופן רחב וגמיש.

3. מדיניות רכש ומכירה

נחוצה מדיניות לתחום הרכש והמכירה של אלגוריתמים ומוצרים בתחום הבינה המלאכותית. מדיניות זו תצטרך להתייחס לרכש ומכירה בעולם הביטחוני והממשלתי, כדי להבטיח את הביטחון הלאומי בכל הקשור למכירת תוצרים של מערכת הביטחון ושלוחותיה ורכישת מוצרי מדף. כמו כן, המדיניות תוכל להתייחס גם למגזר האזרחי, ולצורך בהסרת חסמים רגולטוריים מסוימים במנגנוני הרכש כדי לאפשר התפתחות טכנולוגית מהירה.

4. הקמת מרכזי נתונים ופלטפורמות לשיתוף נתונים ומודלים

קיים צורך בהקמת מרכזי נתונים בהם ירוכזו הנתונים ויוגשו לשימוש, וכן פלטפורמות לשיתוף נתונים ומודלים בין מגזרים וגופים. כך יחוזק הקשר בין חלקי האקו-סיסטם ויתאפשר ניצול יעיל יותר של כלל הנתונים למטרות מגוונות. כמו כן, השיתופיות תסייע לקדם שקיפות. לשם כך, יש להגדיר פורמטים קבועים ומסודרים לאגירת הנתונים על ידי גורמים במגזר המסחרי והציבורי, ולשיתופם בין גורמים אלו.

5. ניהול הנתונים, אבטחת סייבר והגנה על מידע

טמון ערך רב בשימוש בנתונים ציבוריים במחקר ופיתוח, לאור היקפם ולאור התפיסה שיהיו מגוונים ויוכלו לסייע במניעת הטיה אלגוריתמית. כדי להשתמש בנתונים אלו עולה הצורך באבחנה בין סוגי נתונים (אישיים, לא מזהים, ציבוריים), ועולה החשיבות בפיתוח מנגנוני אבטחה ואנונימיזציה מספקים שיאפשרו שימוש בנתונים פרטיים במידת הצורך. כמו כן, נחוצה הגנת סייבר מספקת עבור התשתיות שישמשו את פיתוח והטמעת הבינה המלאכותית, ועבור מרכזי הנתונים המשותפים. כאמור, אלו יהיו מטרה למתקפות סייבר ויש להבטיח אמצעים בטוחים להגנתם ולהתמודדות עם מתקפות סייבר. במסמכי המדיניות הוצעו כמה טכנולוגיות אפשריות שיוכלו לסייע לאבטחת סייבר והגנת מידע, כמו בלוקצ'יין ומחשוב קוונטי, לצד קריאה לשימוש באלגוריתמים מבוססי בינה מלאכותית כאמצעי לזיהוי והתמודדות עם איומי סייבר.

5. הון אנושי

5.1. רקע

מחקר ופיתוח בתחומי בינה מלאכותית מציב דרישות גבוהות בהיבטי הון אנושי מוכשר ומנוסה. ליבת המומחיות והידע בתחום זה מצויה באקדמיה המאפשרת הכשרה של הון אנושי מתאים בתחום, החל במחקר ופיתוח גישות אלגוריתמיות חדשות, עבור דרך יישום אלגוריתמים חדשים לבעיות בעולמות תוכן שונים, טיפול בהכנת המידע ותיוגו, בחינת מודלים שונים ובחירת המודלים המתאימים ביותר, טיפול בהטמעה ועוד. במצב השוק הנוכחי קיים פעם משמעותי בהיקפי ההון האנושי בין צד הביקוש לצד ההיצע וכתוצאה מכך יש לחברות רבות קושי לגייס אנשי מקצוע טובים המתאימים בתחום. ראוי גם לציין כי במצב עניינים זה מתחוללת תחרות עזה על העובדים המיומנים ביותר ולאלה מוצעות פעמים רבות, תנאים מפתחים ביותר אותם יכולים להציע רק מרכזי הפיתוח של ענקיות האינטרנט הפועלים בישראל. חברות אלו ביניהן אמזון, גוגל, פייסבוק ועוד, משופעות לא רק באמצעים כספיים אלא גם בנתוני אמת שהם מרכיב יקר המציאות לאימון של בינה מלאכותית. על מנת לאפשר כניסה לעולמות הבינה המלאכותית גם לחברות רבות נוספות, נדרש להגדיל באופן משמעותי את כמות כ"א בתחום.

5.2. מחקר אקדמי

מחקר אקדמי בבינה מלאכותית הוא מרכיב מרכזי בפיתוח התחום בארץ. המחקר מסייע בהכשרת הון אנושי לתעשייה ולמשרדי הממשלה, החל בבוגרים ועד דוקטורנטים, מעביר ידע לתעשייה, ויזמות מתוך האקדמיה, ומחזק את המעמד הבינלאומי של ישראל בתחום. סקירה של פרסומי מאמרים בתחומי בינה מלאכותית¹²¹ מראה כי תחומי הליבה בבינה מלאכותית מודרנית הם: למידה חישובית, ראייה חישובית, עיבוד שפה טבעית, סוכנים אוטונומיים ורובוטיקה. בתחומים אלו מתקיימים כנסים תחרותיים מאד, שנוכחות בהן משקפת נאמנה את היקף ואיכות פעילות המחקר. כנסים אלו הם:

- NeurIPS, ICML, ICLR, COLT : למידה חישובית
- CVPR, ICCV, ECCV : ראייה חישובית
- ACL, EMNLP : עיבוד שפה טבעית
- IJCAI, AAI : מגוון תחומים ב-AI
- RSS, ICRA : רובוטיקה

¹²¹ <https://hai.stanford.edu/research/ai-index-2019>

Number of AI papers on arXiv, 2010-2019

Source: arXiv, 2019.

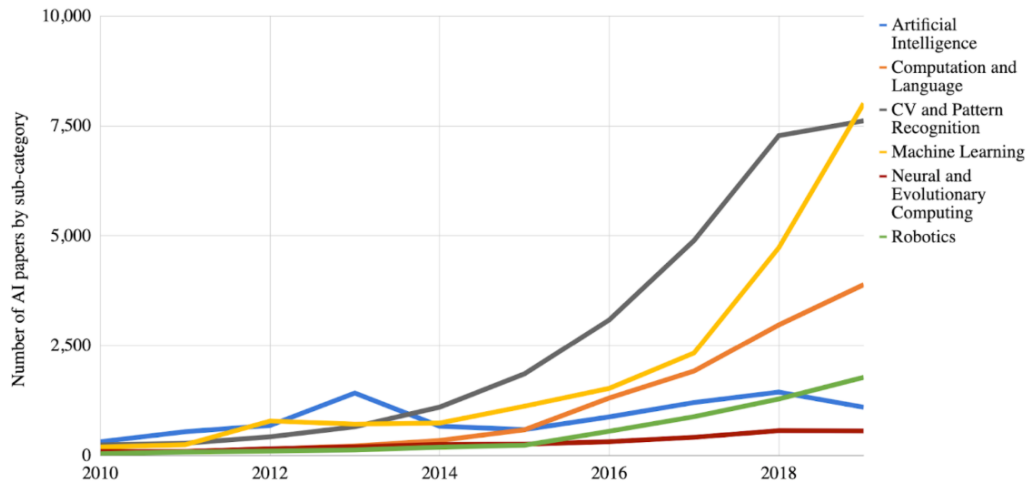


Fig. 1.6.

איור 7 : פרסומי מאמרים העוסקים בבנייה מלאכותית בחלוקה לתחומים, בין השנים 2010-2019

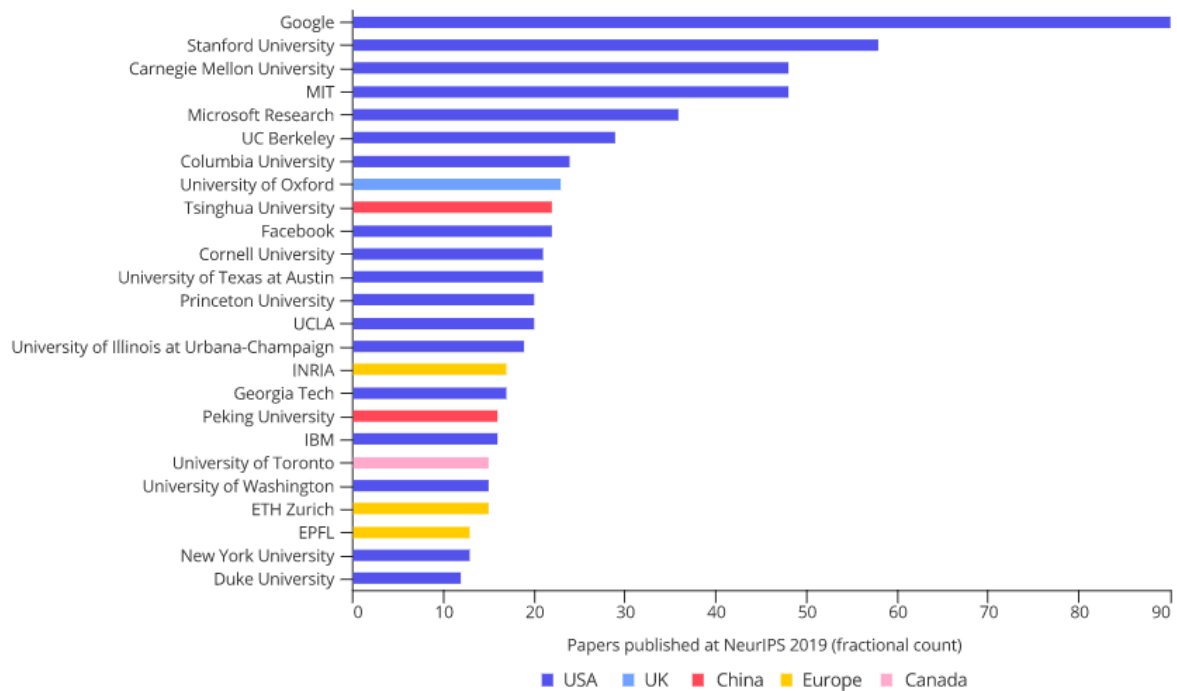
כיום, לישראל ישנם מספר יתרונויות בתחום : איכות גבוהה של ההכשרה האקדמית והסטודנטים, מספר חוקרים מובילים בינלאומיים בתחום שונים של בינה מלאכותית (על פי פרסומים בכנסים מובילים, מענקי ERC -- European Research Council ומאמרים משפיעים), היקף גדול של תעשיה בתחום שמגדיל את הביקוש להכשרה אקדמית בכל הרמות והכשרות רלוונטיות במערכת הביטחון אשר נותנת נקודת התחלה טובה ללימודים אקדמיים.

האתגרים והחסמים באקדמיה הישראלית הם : כמות קטנה של אנשי סגל בתחום (פחות מ-70 פעילים), זליגה של אנשי סגל מהאקדמיה למרכזי מחקר של חברות, דבר המקשה על הכשרת סטודנטים, תחרות משמעותית שקיימת בעולם על הובלה בתחום, קצב התקדמות מרשים מאד במדינות אחרות (כגון קנדה, פירוט בהמשך), מדינות מובילות הופכות למובילות בחדשנות, ומוקד הגירה להון אנושי איכותי. אי עמידה בקצב המחקר האקדמי פוגעת ביכולות המדינתיות בתחום.

על מנת לשמור על תחרותיות יש לקיים כמה תשתיות יסוד : תשתית של כוח חישוב שמאפשרת בנייה של מודלים מתקדמים (למשל בעיבוד שפה ותמונה), הגדלת מספר אנשי הסגל, הגדלת האפשרות להכשיר מספר רב של תלמידי מחקר, בקבוצות של 10 תלמידים לאיש סגל, יצירת שיתופי פעולה אפקטיביים עם התעשייה, יצירת ושימור קבוצות עם מובילות בינלאומית בתחומיהן והפיכת מוקד משיכה לחוקרים מחו"ל.

מחקר שבוצע על ידי מכון פולסון, ומבוסס על עבודות שהוגשו ל- NeurIPS 2019 עוקב אחר תנועתם של חוקרים מעולם הבינה המלאכותית בעולם לאחר סיום לימודיהם. במחקר זה ניתן לראות כי 59% מהחוקרים בתחום עובדים בחברות ואוניברסיטאות בארה"ב, אשר מצטיינת במשיכת חוקרים ממדינות אחרות. זאת ניתן לראות מכך שמעל לשני שליש מהחוקרים המובילים בתחום סיימו את התארים הראשונים שלהם במדינות אחרות. סין, לעומת זאת, היא המקור הגדול ביותר של חוקרים מובילים, כאשר 29% מהם הגיעו ממנה, אך רובם (56%) עובדים, לומדים וחיים בארה"ב. 7%

מהחוקרים המובילים בתחום סיימו את התואר הראשון בישראל, ו-5% מהם עובדים בישראל¹²². עובדה המעידה על פוטנציאל גבוה לקליטת חוקרים מובילים נוספים בארץ. קנדה היא המדינה הבולטת ביותר מבחינת אסטרטגיה אפקטיבית בבינה מלאכותית, ובבריטניה ישנה תוכנית דומה אשר זוכה להצלחות גם כן. הגישות המובילות במדינות אלו מבוססות על מענקי מחקר גדולים לחוקרים מובילים, תמיכה בתוכניות דוקטורט ותלמידים באופן ישיר, קשר הדוק לתעשייה ותשתיות חישוב מתקדמות¹²³.



איור 8 : 25 הגופים המובילים במחקר AI לפי מדד מכון פולסון

בחינה של המתחולל בכנסים עולמיים בנושא בינה מלאכותית יכולה לשקף את מעמדה של ישראל ואת בעיית הגודל מול האיכות בקהילת המחקר בבינה מלאכותית ולמידת מכונה בארץ.

להלן טבלה המסכמת נתונים על מחקר בבינה מלאכותית מכנסים בינלאומיים שהתקיימו בשנתיים האחרונות:

¹²² <https://macropolo.org/digital-projects/the-global-ai-talent-tracker/> ; ¹²³ <https://www.ukri.org/news/200m-to-create-a-new-generation-of-artificial-intelligence-leaders/> ; <https://www.cifar.ca/cifarnews/2019/04/08/cifar-expands-canada-cifar-ai-chairs-program-to-46>

כנס	שנה	מאמרים שהוגשו	מאמרים שנבחרו	פרס המאמר הטוב ביותר
ICML(International Conference on Machine Learning)	יולי 2020 ¹²⁴	4,990	1,088 : ישראל-42 בריטניה-123 קנדה - 82 צרפת -59 גרמניה- 51 שוויץ -46 MIT – 69 סטנפורד-62 ברקלי-58	On Learning Sets of Symmetric Elements ¹²⁵
IJCAI (International Joint Conference on Artificial Intelligence)	אוגוסט 2019 ¹²⁶	4,752	850 ישראל-8	
COLT (Conference on Learning Theory)	יולי 2020 ¹²⁷			Proper Learning, Helly Number, and an Optimal SVM Bound ¹²⁸
ICCV (International Conference on Computer Vision)	נובמבר 2019 ¹²⁹			SinGAN: Learning a generative model from a single natural image ¹³⁰

<https://icml.cc/virtual/2020>¹²⁴

Maron, H., Litany, O., Chechik, G., & Fetaya, E. (2020). On Learning Sets of Symmetric Elements. *arXiv preprint arXiv: 2002.08599*.
<https://arxiv.org/pdf/2002.08599.pdf>

<https://www.ijcai19.org/>¹²⁶

<https://www.learningtheory.org/colt2020/>¹²⁷

Bousquet, O., Hanneke, S., Moran, S., & Zhivotovskiy, N. (2020). Proper Learning, Helly Number, and an Optimal SVM Bound. *arXiv preprint arXiv: 2005.11818*.

<http://iccv2019.thecvf.com/>¹²⁹

Shaham, T. R., Dekel, T., & Michaeli, T. (2019). Singan: Learning a generative model from a single natural image. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 4570-4580).

An Equivalence Between Private Classification and Online Prediction ¹³²	<p>1,903 ארה"ב – 1178 ; סין – 259 ; בריטניה – 195 ; קנדה – 95 ; שוויץ – 88 ; גרמניה – 66 ; ישראל – 54 ; אוסטרליה – 53 ; צרפת – 52 ; דרום קוריאה – 51 ; סינגפור – 43 ; יפן – 34 ; איטליה – 22 ; הודו – 19 ; הולנד – 17 ; שבדיה – 14 ; פינלנד – 14 ; דנמרק – 11 ; ערב הסעודית – 10.</p>	<p>8,186</p>	<p>נובמבר 2020¹³¹</p>	<p>FOCS (Foundations of Computer Science)</p>
--	---	--------------	----------------------------------	---

טבלה 5: נתונים על מעורבות ישראלית בכנסים מובילים

אם נבחן את הכנסים אלו לאורך ציר הזמן, לפני 10-15 שנים הייתה לישראל נוכחות הרבה יותר בולטת בכנסים המובילים, הן במספר המאמרים והן בזכייה בפרס המאמר הטוב ביותר. בחינה של הנתונים מכנס NIPS (Neural Information Processing Systems) שהתקיים בשנת 2013¹³³ ומהנתונים מהכנסים שהתקיימו ב-2019¹³⁴ וב-2020¹³⁵ מראה שמוסדות רבים בארה"ב, סין, דרום קוריאה וכו' 'עקפו' את ישראל בדירוג המחקר בבינה מלאכותית. העובדה שמאמרים פרי עטם של חוקרים ישראלים עדיין זוכים בפרס המאמר הטוב ביותר בכנסים בינלאומיים, אך כמות המאמרים יורדת, מלמדת שההידרדרות של ישראל לא נובעת מכך שהמחקר בישראל היום פחות טוב ממה שהיה פעם, אלא מכך שאנחנו התקדמנו מעט, בעוד המתחרים שלנו התקדמו מאוד. ארגונים מובילים צריכים לפרסם עוד ועוד מאמרים מדי שנה על מנת לשמר את מעמדם בתחום, והם עושים זאת בדרך של קבלת חוקרים ומשאבים נוספים בכל שנה ושנה. אפילו על מנת לשמר את דירוגה הנוכחי, שכאמור הוא נמוך יותר משנים עברו, ישראל צריכה להגדיל את כמות החוקרים והמשאבים שלה. לאור זאת, המטרה שלנו צריכה למקסם את תוצרי האקדמיה, לא רק משום שהפרסום כשלעצמו הוא המטרה, אלא משום שהוא משקף את איכות וכמות פעילות המחקר בתחום הבינה המלאכותית בארץ. עם השקעה נכונה במחקר

¹³¹<http://ieeefocs.org>

Bun, M., Livni, R., & Moran, S. (2020). An equivalence between private classification and online prediction. *arXiv preprint arXiv: 2003.00563*.

¹³³<https://www.quora.com/How-can-I-see-the-list-of-institutions-who-have-submitted-papers-at-NIPS>

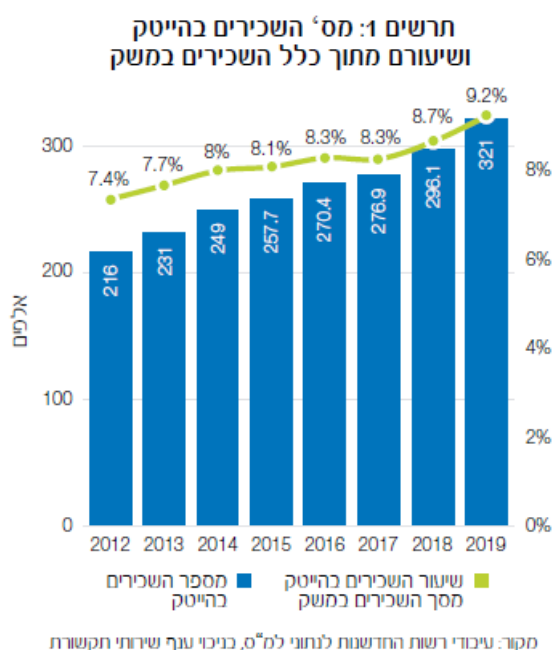
¹³⁴<https://medium.com/@chuvpilo/ai-research-rankings-2019-insights-from-neurips-and-icml-leading-ai-conferences-ee6953152c1a>

¹³⁵<https://chuvpilo.medium.com/whos-ahead-in-ai-research-at-neurips-2020-bf2a40a54325>

אקדמי בבינה מלאכותית, יש לישראל פוטנציאל לגדול ולייצר מסה קריטית ועצמה לאומית בבינה מלאכותית.

5.3. תעשייה ומגזר ציבורי

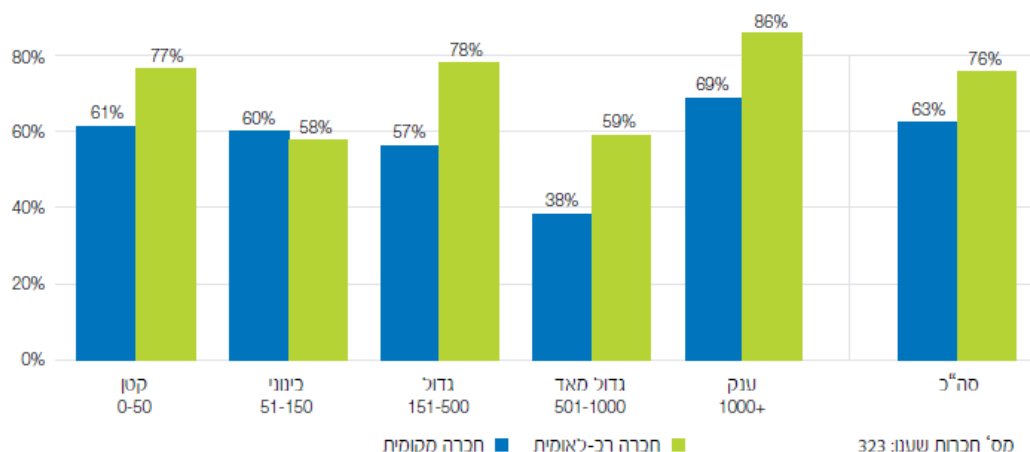
ההייטק הישראלי ממשיך לצמוח ללא הרף בשנים האחרונות. דוח שנכתב על ידי רשות החדשנות ו-start-up nation central, ומבוסס על נתונים ונתוני ות"ת, הלמ"ס, וסקר שנשלח לחברות הייטק ביולי 2019, מציג כי ישנן אינדיקציות ברורות למחסור בעובדים טכנולוגיים וקושי בגיוסם¹³⁶. הפתרון המרכזי למחסור הוא הגדלת היצע העובדים, וראשית כל הכשרת עובדים טכנולוגיים נוספים. האקדמיה עודנה שער הכניסה המרכזי לתעשיית ההייטק. לפי נתונים שהוצגו בדוח כ-75% מהג'וניורים הטכנולוגיים, שהם עובדים עם עד שנתיים ניסיון מעשי, שגויסו במחצית הראשונה של 2019, ומוסד הכשרתם דווח בסקר, הגיעו מאוניברסיטאות או ממכללות. עם זאת, הנתונים מעידים כי על אף המחסור בעובדים טכנולוגיים, רוב החברות שענו על הסקר לא העסיקו עובדים ללא ניסיון במחצית הראשונה של 2019.



איור 9 : מספר השכירים בהייטק ושיעורם מתוך כלל השכירים במשק, מתוך דוח הון אנושי בתעשיית ההייטק

¹³⁶ דוח הון אנושי בתעשיית ההייטק, רשות החדשנות ו-start-up nation central, 2019.

תרשים 3: אחוז הטכנולוגיים מתוך מצבת העובדים ביוני 2019 לפי גודל ובעלות החברה



איור 10 : אחוז הטכנולוגיים מתוך מצבת העובדים ביוני 2019 לפי גודל ובעלות החברה, מתוך דוח הון אנושי בתעשיית ההייטק

מטרת הוועדה בהקשר זה היא לסקור הכשרות מחוץ למסגרות הלימודים לנוער ולמבוגרים, אשר יהיו כלי משלים להכשרות האקדמיות וישמשו כמנגנון מרכזי לשיפור הפריון, להרחיב את מגוון האוכלוסיות הנהנות מכניסת הבינה המלאכותית, מדע הנתונים והדיגיטציה לחיי היומיום. כמו כן, גם להעלות, בהתאם לפריון ולהתייעלות, את השכר של המועסקים בדגש על אלו אשר ממוקמים בחציון התחתון. בנוסף, חלק מהמשתתפים בהכשרות אלו יהיו גם בעלי תארים קודמים המבקשים להרחיב את ידיעותיהם בתחום.

הממצאים העולמיים מראים, בהינתן כישורי הלומדים, השכלה על-תיכונית שאינה אקדמית מאופיינת בתשואות חיוביות. כמו כן, מתוכניות כגון עתידים לעתודה, למדנו גם כי ניתן להרחיב את מספר המועמדים ללימודים מקצועיים ואקדמיים טרום שירותם הצבאי.

המאפיין את ההכשרות המקצועיות האיכותיות הוא קיצור שעות הלימוד ומיקודן תוך עדכון תוכניות הלימודים בשיתוף המעסיקים, ובנייה מחדש של מסלולי הלימוד, בדומה למה שנעשה בקורסים הטכנולוגיים הצבאיים. קהלי היעד העיקריים של ההכשרות המקצועיות הללו הם אנשים שלתועלת המשק ולרווחתם ראוי שיוכשרו בהתאם למקצועות העתיד או יעברו הסבות מקצועיות.

הכשרות מקצועיות לצעירים שטרם נכנסו לשוק העבודה, לצעירים הנכנסים לשוק העבודה ואינם פונים להשכלה אקדמית, ולעובדים שישנו חשש לכך שמקצועם נעלם והם על סף פיטורים, צפויות להביא לגידול משמעותי של החוסן הלאומי, הכלכלי והבטחוני של ישראל, כמובן בכפוף לאיכות ההכשרה וכמות המוכשרים שמוסמכים בהצלחה.

ההכשרות המקצועיות תתבססנה על העקרונות הבאים:

- ראייה הוליסטית של ההכשרות המקצועיות כנקודת פתיחה של מסלול קריירה הפתוח לכולם, עם אפשרות של לימודי או הכשרות המשך.
- כל הכשרה תעמוד בקריטריון איכות: שיפור פריון ותשואה, חיזוק תחומים בהם יש במשק ביקושים גבוהים ושכר עולה, כולל תחומים שלא קיימים במערכת כיום.
- עדכון תוכניות הלימוד על פי הביקוש בשוק ובשיתוף מעסיקים, כך שכל תוכניות הלימוד יעודכנו באופן שוטף.

• עדיפות לבניית רצף אקרדיטציה ובניית שרשראות קורסים בכל תחום לימוד אפשרי, כך שכל תלמיד הנכנס למערכת יוכל לראות את מסלול ההתפתחות המקצועית שלו. בקריטריון התשואה, יורחב היקף המשתתפים בהכשרות אלו ביחס למספר המשתתפים בהן כיום. לאור זאת, בהינתן עמידה ביעדי איכות וביקוש יורחבו ההכשרות כדי למלא את הביקוש וההיצע. כמו כן, ולאור הטרנספורמציה הדיגיטלית וההתפתחויות והשינויים המתמידים בשוק ההייטק, יש לבצע מעקב אחר המגמות בשוק. נדרשת פלטפורמה להנגשת מידע, הכוונה ואבחון להכשרות מקצועיות. יש לפתח כלים פדגוגיים ואחרים להקניית כישורים חדשים ולבחון הטמעה שלהם במערכי ההכשרה, ההשכלה והתעסוקה בהתאם לאפקטיביות. בפרט הוועדה רואה חשיבות בקידום הקניית מיומנויות בסיס לאוכלוסיות בהם מזוהה חוסר משמעותי, כדוגמת עברית לחברה הערבית¹³⁷.

לאור הפריחה הגדולה של ההייטק הישראלי ככלל, וחברות המתעסקות בבינה מלאכותית בפרט, והמחסור בהון אנושי בתחום, נוצרו מספר תוכניות העוסקות בהכשרה והסבה מקצועית של הון אנושי איכותי. בין התוכניות ניתן למצוא Boot Camps, אשר מבצעים הכשרות בתחומי ה-Data science ולמידת מכונה, לבוגרים בעלי פרופיל גבוה (לרוב בוגרי תארים במדעים מדויקים). הכשרות אלו אשר מנוהלות על ידי רשות החדשנות, הן בעלות אחוזי השמה גבוהים. סוג נוסף של הכשרות שנבנו הוא הכשרות הון אנושי אשר מבוצעות במסגרת המעסיק, אשר מאגדות חברות בעלות צרכי הכשרה דומים. נושא חשוב נוסף הוא איתור הון אנושי איכותי לתעשיית ההייטק, אשר גם עבורו נבנתה תוכנית.

הדיון בנושא זה טרם התקיים באופן נרחב במסגרת הוועדה ועל כן נושא זה יורחב בשלבים הבאים של עבודת הוועדה. בימים אלה מתגבשות המלצות לאשכולות קורסים אקדמיים שיפורסמו בנספח לגרסה הסופית של דוח זה.

¹³⁷ מתוך שולחן עגול בנושא תעסוקה שהתקיים במרכז הבינתחומי הרצליה.

6. הצעת ות"ת וצוות המשנה של המיזם הלאומי למערכות נבונות

במקביל לעבודת הוועדה, התנהלו ועדות נוספות בהן התבצעה עבודת מחקר והסקת מסקנות. בין העבודות הללו ישנן התוכנית הלאומית האקדמית שהותוותה על-ידי הוועדה המייעצת לות"ת ומסמך צוות המשנה של המיזם הלאומי למערכות נבונות בנושא כוח מחשוב וקוונטים. להלן תמצית המסקנות וההמלצות אשר נכללו בהצעת ות"ת לתל"ם, וכן בדוח צוות המשנה של המיזם הלאומי למערכות נבונות בנושא כוח מחשוב וקוונטים.

6.1. תוכנית הוועדה המייעצת לות"ת להקמת תוכנית לאומית אקדמית

התוכנית האקדמית של הוועדה נבנתה מתוך הבנה שהיא תהווה חלק מתוכנית לאומית רחבה יותר המשתלבת עם שאר הגופים הרלוונטיים בתעשייה ובביטחון. הוועדה סקרה את תמונת המצב בפן האקדמי בלבד, והיא בחנה את מצב האקדמיה בארץ בתחום מדעי הנתונים¹³⁸.

- הוועדה מנתה את נושאי המחקר של חוקרים בליבת התחום, ביניהם: Machine learning, Statistics, Deep learning, NLP, Computational biology, Network Analysis, Bioinformatics, Computer vision, Image Processing, Big Data, Signal processing, Selective Inference, Econometrics, Optimization, Game Theory, Algorithms, Biostatistics & Health, Robotics, Information Theory, Databases, Casual Inference, AI, Speech.
- הוועדה מנתה את תחומי המעטפת של המחקר במדעי הנתונים, שהם: בינה מלאכותית (החלק שאינו מכוסה בתחומי הליבה), אופטימיזציה, חקר ביצועים, תורת המשחקים, תהליכים סטוכסטיים, אינפורמציה, רובטיקה, בקרה, ניתוח אותות, ביו-אינפורמטיקה, ביולוגיה חישובית, רפואה דיגיטלית/ מותאמת אישית, אפידמיולוגיה, אקונומטריקה, פסיכומטריקה, סייבר, חישוב קוונטי, ניהול ובניית בסיסי נתונים / מידע.
- הוועדה ציינה כי כיום קיימים 158 חוקרים שליבת מחקרם בתחום זה, 230 חוקרים ממעטפת התחום, ועוד 297 בתחומים נוספים. שליש מהחוקרים בליבת התחום עוסקים גם בתחומים יישומיים. החוקרים מנחים 487 תלמידי מחקר, 314 תלמידי דוקטורט, ו-65 משתלמים לפוסט-דוקטורט.
- הוועדה ציינה נכון ל-2018 היו קיימים בכל המוסדות (תל אביב, בר אילן, בן גוריון, מכון ויצמן, הטכניון, אוניברסיטת חיפה, האוניברסיטה הפתוחה, אוניברסיטת אריאל והאוניברסיטה העברית) מרכזים לתת תחומים של מדע הנתונים. קיימים 7 מרכזי מחקר במענקי ות"ת, חלקם נקראים מדעי הנתונים ו-AI. נכון לחודש אפריל 2020, 6 מהם בפעילות ענפה.
 - הוועדה הציגה מספר חסמים לפיתוח התחום באקדמיה, כולל מחסור בתלמידי מחקר, חשיפה מועטה של סטודנטים לנושא, תשתיות מחקר חסרות ועוד.
- הוועדה הגדירה מספר תכולות כחלק מתוכנית לאומית אקדמית לפיתוח התחום במערכת ההשכלה הגבוהה בישראל:

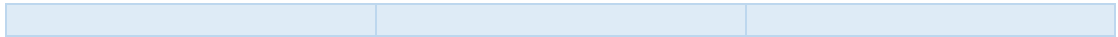
¹³⁸ המידע לקוח מתוך מצגת שהועברה על ידי פרופסור יואב בנימיני במסגרת עבודת הוועדה בתאריך 26.4.20

- תמיכה במרכזי מחקר מוסדיים – עיקר המשאבים יוקצה לפיתוח מרכזים באוניברסיטאות השונות, ובכלל זה קליטת אנשי סגל, הקמת מעבדות ותשתיות מחקר, מלגות לסטודנטים ועוד, וכמו כן לעידוד שיתוף התעשייה והמגזר הציבורי בפעילותם.
- פעילות לאומית, ניהול ועתודה – ערוץ העונה על צרכים לאומיים המשותפים לכלל החוקרים והמוסדות, כגון הוצאות ניהול התוכנית, מלגות לישראלים במוסדות זרים ועריכת כנסים בין-לאומיים.

6.2. דוח צוות המשנה של המיזם הלאומי למערכות נבונות בנושא כוח מחשוב וקוונטים

- צוות ההיגוי המליץ על מספר צעדים עבור הפן התעשייתי בתוכנית חישוב על הלאומית¹³⁹:
 - הקמת מרכז חישוב במסגרת מעבדה לאומית בנושא חישוב על. מעבדה שכזאת צריכה להיות חלק ממעבדה לאומית בנושא בינה מלאכותית. בניתוח הנתונים שהועלו ברור כי יהיה צורך בכוח חישוב רב לטובת נושאי בינה מלאכותית בשנים הקרובות. העלויות הכרוכות בשימוש בכוח חישוב רב הינן נמוכות יותר בשימוש במרכז חישוב עצמאי, כל עוד קיימת מסה קריטית המצדיקה את הקמת התשתית, ותחזוקתה. ניתן לראות מהנתונים כי מסה קריטית כזאת תהיה קיימת, ולכן יש הצדקה להקמת מרכז חישוב מקומי. אין הדבר אומר שלא יהיה שימוש בארץ בחישוב ענן מסחרי, במקומות בהם הדבר יהיה מתאים, בין אם בשל חוסר במשאבים, אלגוריתמיקה חדשנית שלא ניתן להשיג, ועוד.
 - הקמת ענן ישראלי מקומי שיהיה משותף לגופי הממשלה האזרחיים ולמערכת הביטחון, תוך הפרדת החלק האזרחי מהביטחוני (לצורך מידור של פעילות בעלת סיווג בטחוני גבוה).
 - עיסוק רב בנושא הקמת מאגרי מידע בעלי אופי ארצי, וכן יזימה של מאגרים אשר יאפשרו לישראל להיות בחזית הבינלאומית (למשל, מאגרי מידע רפואי אשר בארץ יחסית קל לאסוף). ביצוע עבודת מטה מעמיקה בנוגע לתשתיות הפיסיות הגדולות, על-מנת להעריך באופן מדויק את הצרכים ואת המשמעויות התקציביות.
 - מינוי גורם מרכזי העוסק בסוגיות של מדיניות, חוק, אבטחה, וכלל הסוגיות המשפטיות העשויות לעלות בהקשר של מאגרי מידע (כולל שמירה על פרטיות).
 - מינוי גורם אשר יקבל החלטות לגבי תפוצת המאגרים הנאספים – שיקולים של כדאיות כלכלית, פרטיות, צרכי מחקר ועוד.
 - מינוי גורם המתמחה בשמירת מאגרים, טיוב שלהם, והפצתם.
 - לתעשיות להוציא את ההון האנושי שלהן לתארים מתקדמים רלוונטיים.
 - סה"כ ההשקעה הממשלתית שהומלצה על-ידי הצוות, מעבר לתוכנית האקדמית, מסתכם ב-145 מיליון ₪ על-פני 5 שנים.
 - הדוח כלל גם הצעות נוספות שנבחנו על-ידי הצוות, וכן סיכום הערכת עלויות.

¹³⁹ בן ישראל, י', מתניה, א' ופרידמן, ל' (עורכים). (ספטמבר 2020). המיזם הלאומי למערכות נבונות בטוחות להעצמת הביטחון הלאומי והחוסן המדעי-טכנולוגי: אסטרטגיה לאומית לישראל. דוח מיוחד לראש הממשלה.

- 
- לפרטים נוספים ראו מסמך צוות המשנה של המיזם הלאומי למערכות נבונות בנושא כוח מחשוב וקוונטים

7. רציונל עבודת הוועדה

עד כה, מרבית הפרויקטים של פורום תל"ם¹⁴⁰ עסקו במימוש תשתית פיסיית משמעותית עבור תחומים, שזוהו על-ידי חברי הפורום, כאסטרטגיים לעתיד המחקר והפיתוח במדינת ישראל. בבואנו לעבודת וועדה זו, היה ברור כי יש לבחון את שאלת התשתיות לתחום בינה מלאכותית בראייה לאומית רחבה יותר. גם בעבודת הוועדה האקדמית שעל בסיסה הובאה ההצעה לשולחן תל"ם, נאמר כי יש לראות את התוכנית האקדמית כחלק מתוכנית לאומית כוללת יותר. בנוסף, התוכנית האקדמית, מטבעה, עסקה בזיהוי וגיבוש הצרכים למערכת האקדמית בלבד, וזו לתקופה מוגבלת של חמש שנים. עבודת הצוות שהוקם ברשות החדשנות בשיתוף מפא"ת נועדה להשלים הסתכלות בראי התעשייה, מעבר לצרכים האקדמיים שהוגדרו.

כבר בתחילת עבודת הוועדה התברר כי אופי הפעילות הנדרשת עבור תחום זה מתבטא בתוכנית לאומית רחבה ומקיפה יותר מהמנדט הרגיל של פורום תל"ם, וכי התקציבים הנדרשים למימוש תוכנית לאומית לתחום מדעי זה גדולים מהתקציבים הסטנדרטיים של שותפי תל"ם, הנשענים באופן עקרוני על התקציבים השוטפים של הגופים.

עבודת הוועדה נעשתה לכן מתוך תפיסה כוללת של צרכי האקדמיה, התעשייה הקיימת והעתידית, ומערכת הביטחון, בראייה רחבה מעבר לתשתיות הפיסיות. על-מנת למצב את ישראל למול התוכניות במדינות אחרות, נדרשת התייחסות בראש ובראשונה לפערי ההון האנושי ולא רק לתשתיות הפיסיות. חשוב לציין שתהליכי הכשרת הון אנושי בתחום זה דורשים זמן רב, ולכן הכרחי להתחיל תהליכים אלו מוקדם ככל הניתן.

יצוין עוד כי במקביל לעבודת וועדה זו, התקיים גם תהליך לבחינת המיזם הלאומי לחוסן מדעי-טכנולוגי בנושא 'מערכות נבונות', פרי יוזמת רוח"מ שריכוזה הוטל על פרופ' אלון במילי יצחק בן-ישראל ופרופ' אביתר מתניה. כחלק מעבודה זו, הוקמה תת-וועדה לנושא כוח המחשוב הדרוש למערכות הנבונות העתידיות. תת-וועדה זו עסקה הן בצרכים לתשתיות חישוב חזק 'רגילות' (חישוב-על ותשתיות ענן), והן בנושא הקוונטי. על מנת לסנכרן את עבודת שתי הוועדות, הוסכם על שותפי תל"ם ועל מובילי המיזם כי ד"ר ארנה ברי תעמוד בראש שתי הוועדות, שכן יש ביניהן חפיפה משמעותית¹⁴¹.

על כן, במסמך זה אנו מביאים המלצה לתוכנית לאומית כוללת לתחום מדע ובינה מלאכותית הנשענת על תוצאות בחינת המסמכים המקדימים שהונחו על שולחן תל"ם, אך מקיפה היבטים רבים מעבר לתשתית גרידא. בהמלצות שיובאו בהמשך מסמך זה, אנו כוללים את התכולות הנדרשות, את גובה ואופן החלוקה (ברמת הפעילויות העיקריות) של התקציב הנדרש, מזהים מרכיבים הנמצאים בשלבי מימוש שונים, ומנתחים את ההמלצה בחתכים שונים, כולל חתך גופי הביצוע העיקריים (אקדמיה/ביטחון/תעשייה), חתך הון אנושי לעומת תשתית חומרה, ועוד¹⁴².

¹⁴⁰ פורום (וולונטרי) של ות"ת, הרשות לחדשנות, משרדי הביטחון, מדע, אוצר, האקדמיה הלאומית למדעים, ולעתים גופים נוספים המצטרפים לפי נושא עניין ספציפי, שמנהלים במשותף פרויקטי תשתית לאומית למחקר.

¹⁴¹ בן ישראל, י', מתניה, א' ופרידמן, ל' (עורכים). (ספטמבר 2020). המיזם הלאומי למערכות נבונות בטוחות להעצמת הביטחון הלאומי והחוסן המדעי-טכנולוגי: אסטרטגיה לאומית לישראל. דוח מיוחד לראש הממשלה.

¹⁴² תכנון תקציבי מפורט יבוצע על-ידי הנהלת התוכנית וגופי התקצוב במהלך שלב ההתארגנות.

8. אפיון הצורך והתועלות הצפויות

- 8.1. תחום הבינה המלאכותית משתלב כאמור בכלל תחומי המחקר האקדמי והתעשייה. גישה ליכולות מחשוב על, מאפשרת השגת רמות ביצועים גבוהות ומהירות בהרבה, ובמקרים רבים מאפשרת פתרון בעיות ויכולות שלא ניתן להשיג כלל ללא הטכנולוגיה.
- 8.2. נראה כי תחום הבינה המלאכותית יישאר רלוונטי ופורץ דרך למשך שנים רבות, וההכרה בפוטנציאל העצום הטמון בו הביאה מדינות רבות להשקעות עתק בתחום בשנים האחרונות, כפי שפורט בסעיף 2.1 לעיל.
- 8.3. השקעת המשאבים הרבים בעולם מעידה על ההבנה כי פריצות הדרך המשמעותיות שהתחום צפוי להביא, יובילו לפתרונות לאתגרים גלובליים, לצמיחה כלכלית משמעותית וליתרונות בזירה הביטחונית.
- 8.4. על-מנת להבטיח את חוסנה הלאומי של מדינת ישראל, עליה להשקיע משאבים בתחום מדע הנתונים ובינה מלאכותית. בשלב זה, למדינת ישראל ישנה הזדמנות להשתלב כשחקן מפתח בתחום, ותהיה לכך השלכה כלכלית אדירה. ישראל צריכה לכוון להובלה היכן שידה משגת, שכן טכנולוגיות בינה מלאכותית עתידות להשתלב במנעד רחב של יישומים, המושפעים בחלקם כבר היום מטכנולוגיות אלו.

9. מסקנות

בעקבות המחקר שהתבצע על-ידי הוועדה, הבנת המצב במדינת ישראל כיום, דיונים רבים בין חברי הוועדה וכן דיונים עם גורמים מחוץ לוועדה ואף מחו"ל, התגבשו המסקנות העיקריות המובאות להלן. בהמשך, בפרק 7, נביא את המלצות הוועדה למימוש התוכנית הלאומית.

9.1. תחום בינה מלאכותית ומדע הנתונים הינו תחום אסטרטגי, אשר כבר היום מהווה אחד התחומים המשפיעים ביותר על המדע, הכלכלה והביטחון, השפעה אשר תגבר בשנים הבאות. דבר זה בא לידי ביטוי, בין היתר, בתוכניות הלאומיות הגדולות שממומשות וימומשו בכלל המדינות המובילות בעולם, ובהשקעות הגדולות, כולל על-ידי גורמים תעשייתיים משמעותיים בעולם.

9.2. בישראל זוהה פער מרכזי בנושא התשתיות, הן באקדמיה והן בתעשייה. בשני המקומות הללו קיים כיום ביזור משאבים בקרב המשמשים. מסקנת הוועדה היא כי נכון לרכז את המשאבים במקום אחד, זאת כדי לאפשר גמישות ומיצוי יעיל שלהם, תוך הקצאה בהתאמה לצרכים השונים. הריכוז יאפשר מתן תמיכה משמעותית למשתמשים, יעילות והוזלת עלויות, בהקשרי אחסון נתונים וגמישות בנושא שדרוג החומרה, וכן יקל על אבטחת הסייבר והאבטחה הפיזית. בהמשך לכך נכון להכשיר הון אנושי לנושאי מחשוב על ולהקים מרכז ידע אשר יקשר בין נתונים לטכנולוגיה.

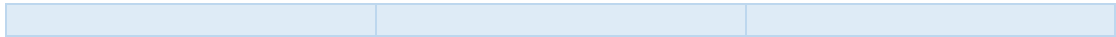
9.3. פער נוסף שזוהה הוא בהון האנושי. זהו משאב הנמצא בחסר באקדמיה וכפועל יוצא מכך גם בתעשייה. חלק מהמסקנות וההמלצות להלן יתייחסו להון האנושי ולדרכים לפיתוחו, אם בהגדלת מספר העוסקים בתחום, בהכשרתם והכוונתם, ואם בפלטפורמות החומרה הנדרשות לפיתוחו בחלק מהמקרים.

9.4. בתחום ביצוע NLP לשפה העברית אין מספיק פעילות באקדמיה ובתעשייה הישראלית. זהו תחום בעל חשיבות אסטרטגית, אשר מהותי לשימוש ביכולות בינה מלאכותית במשרדי הממשלה ובתעשיות נוספות. על כן מסקנת הוועדה היא שגם אם תחום זה איננו במוקד העיסוק הנוכחי, הוא חשוב דיו להרחבה משמעותית של העיסוק בו.

9.5. הוועדה מצביעה על חשיבות השת"פ הבינ"ל הן בתעשייה והן באקדמיה, לצורך שיפור היכולות הישראליות בתחום. חלון ההזדמנויות לשת"פ בינ"ל כיום הינו ייחודי, לאור התוכניות הבינ"ל הנמצאות בתהליך הקמה. מסקנת הוועדה היא כי יש לחזק ולמקד את השת"פ הבינ"ל הייעודי בתחום בינה מלאכותית ומדע הנתונים, ולסייע למנגנוני השת"פ הקיימים בגופי הביצוע השונים, הן ברמה המקצועית הייעודית לתחום הבינה המלאכותית ויכולות מחשוב, והן בתקציב.

9.6. נושא הנתונים והנגישות אליהם מהווה מכשול משמעותי, הן לגורמים בתעשייה והן לגורמים באקדמיה. מסקנת הוועדה היא כי נכון יהיה להגדיר רגולטור ממשלתי אחד כדוגמת הלמ"ס, כבעל אחריות למיפוי, אגירה, זמינות, רישוי ונגישות לנתונים. צעד זה מאפשר פתרון לבעיית קושי בנגישות לנתונים, בכפוף לתקנים הבינלאומיים ולחוק הישראלי להגנה על פרטיות, תוך עמידה בהיבטים של פרטיות ואבטחת מידע.

9.7. כחלק מהראייה הלאומית הכוללת, המביאה בחשבון את פיתוח הידע המדעי, את הפריון הכלכלי ואת החוסן הביטחוני, ועל סמך לקחים מתוכניות שונות, הן בתחום הבינה המלאכותית בעולם והן בתחומים אחרים בישראל בעבר, מסקנת הוועדה היא כי תוכנית



לאומית מעין זו חייבת להיות מנוהלת הן ברמה האופרטיבית היום-יומית, והן ברמה
האסטרטגית .

10. המלצות

לאור המסקנות שהוצגו בפרק 10, הוועדה ממליצה להקים בישראל תוכנית לאומית לתחום בינה מלאכותית ומדע הנתונים. משך התוכנית המומלצת הוא חמש שנים, והיא תחל בשנה נוספת שתוקדש להתארגנות, למעט נושאים שכבר הותנעו בגופי התקצוב השונים, אשר לא יעצרו, ויעודכנו באופן שוטף על-ידי גוף הניהול של התוכנית בשיתוף עימם. התוכנית תתמוך בפעילות באקדמיה, במערכת הביטחון ובתעשייה האזרחית. המשך פרק זה יציג את מטרות התוכנית, האסטרטגיות המומלצות ותכולות התוכנית המומלצת. פרק 11 יציג את הפן התקציבי בחומש הראשון וניתוח שלו בחתכים שונים. פרק 12 יציג חזון להמשך התוכנית, ללא העמקה או הצגת תקציב מפורט. פרק 13 יציג את המשמעויות של אי-מימוש של התוכנית המומלצת, והישארות במצב הקיים.

10.1 תוכנית לאומית לבינה מלאכותית ומדע הנתונים

מטרות העל של התוכנית המוצעת כוללות:

1. יצירת תנאים שיאפשרו מחקר ברמה גבוהה והיקף ראוי של מחקר מדעי בתחומים אלו בישראל (מטרה זו תומכת את מטרות העל הבאות, אך מהווה מטרת על בפני עצמה).
2. יצירת סביבה (אקו-סיסטם) שתאפשר את המשך ההתפתחות והתחרותיות של התעשייה, האקדמיה והמגזר הציבורי בתחומים אלו בישראל.
3. קיום הידע בישראל יאפשר הבנה ופיתוח יישומים קריטיים למערכת הביטחון. ידע זה יאפשר זיהוי העיתוי הנכון ליישומים קריטיים, ושיתוף פעולה עם מדינות אחרות.
4. הטמעת השימוש בבינה מלאכותית במשרדי ממשלה, תוך שיפור השירות הממשלתי לאזרח. כמו כן, שימוש בעולמות הבריאות והתעסוקה כמקרי בוחן.
5. הטמעת השימוש בבינה מלאכותית בתעשייה ובענפי השירות השונים.

10.2 אסטרטגיה להשגת מטרות העל

לאור מטרות העל שגובשו, המלצת הוועדה היא שישראל תאמץ את האסטרטגיה הבאה בכדי לקדם את תחום מדע וטכנולוגיות הקוונטים ברמה הלאומית. חשוב להקדים ולהדגיש כי כל מרכיבי התוכנית צריכים להיות מנוהלים על בסיס מצוינות מדעית ומקצועית. האסטרטגיה תחולק תיושם בחמישה מישורים אלו:

1. הפיכת מדינת ישראל למובילה אקדמית בתחומי הליבה בבינה מלאכותית.
2. יצירת תעשייה משגשגת בתחומי הליבה בבינה מלאכותית.
3. בניית מאגר נתונים ממשלתיים שישמשו הן את משרדי הממשלה והן את התעשייה לביצוע מחקר ופיתוח.
4. שיפור השירות לאזרח בעזרת שימוש בנתונים וכלי בינה מלאכותית.
5. העלאת הפריון בתעשייה בעזרת הטמעת השימוש בבינה מלאכותית.

כפי שנכתב בפרק העוסק במחקר אקדמי, נושאי הליבה בבינה מלאכותית, בהם גם התמקדנו בעבודת הוועדה, הם: למידה חישובית, ראייה חישובית, עיבוד שפה טבעית, סוכנים אוטונומיים ורובוטיקה. תחומים אלו מקובלים כתחומי ליבה בבינה מלאכותית ממספר סיבות. סיבה אחת היא שהביטויים

העיקריים של בינה אנושית הן יכולת למידה, שימוש בשפה, ראייה, ותנועה. סיבה נוספת היא סיבה היסטורית, היות ושיטות שמפותחות בתחומים אלו נוטות להיות בעלות השפעה על תחומי מחקר אחרים (לדוגמה רשתות קונבולוציה בראייה, מודלים כמו Bidirectional Encoder BERT Representations from Transformers, טכניקת אימון NLP בשפה וכו'). הסיבה השלישית היא השימושים בתעשייה, השימושים העיקריים של בינה מלאכותית בתעשייה הם בתחומים אלו.

כפי שאמר ראש תחום הבינה המלאכותית ב-Google:

"Deep neural networks are responsible for some of the greatest advances in modern computer science, helping make substantial progress on long-standing problems in computer vision, speech recognition, and natural language understanding," said Jeff Dean, Google Senior Fellow and SVP, Google AI. "At the heart of this progress are fundamental techniques developed starting more than 30 years ago by this year's Turing Award winners, Yoshua Bengio, Geoffrey Hinton, and Yann LeCun. By dramatically improving the ability of computers to make sense of the world, deep neural networks are changing not just the field of computing, but nearly every field of science and human endeavor."

יכולות המחקר והפיתוח היישומי שיושג בזכות התכולות המבוקשות בדוח זה משותפות לכלל רשויות המחקר, ורלוונטיות גם למוצרים ספציפיים למגזר העסקי והבטחוני. התשתיות המשותפות ישמשו את כלל רשויות המחקר.

10.3 מבנה ניהולי מוצע לתוכנית הלאומית

הדיון בנושא זה טרם התקיים באופן מחייב במסגרת הוועדה עם השותפים בתל"ם. נושא זה יורחב יחד עם חברי הוועדה ותל"ם על מנת להבטיח את חיזוק התחום, ביסוס הסינרגיה הלאומית במחקר ופיתוח, מיקוד הממשלה באימוץ התחום ובתמיכה בו, כולל תמיכה במוסדות לאומיים (למ"ס כמשל). קיים צורך ליצור רגולטור בודד, אשר יאחד את הבעלות על הנתונים, תוך שאיפה להסתכל על נתונים כעל 'טובין ציבוריים' אשר יכולים לשמש לטובת הציבור. הרגולטור ייצר, ישביח וישתף את הנגישות לנתונים תוך הגדרת רגולציה מאפשרת¹⁴³. בשל כל אלה נדרשת תוכנית לאומית הנתמכת ומנוהלת בשיתוף תל"ם וממשלת ישראל.

10.4 תכולות מומלצות לתוכנית הלאומית

התכולות שיפורטו להלן הינן פועל יוצא של המסקנות העיקריות, מטרות העל והאסטרטגיה שהוצגו לעיל, ולפיהן:

- בתחום הבינה המלאכותית בתעשייה והאקדמיה יש להציב יעד של ישראל כמדינה מובילה ומתחרה.
- נושא ההון האנושי כמפתח לכלל המרכיבים הנוספים.

¹⁴³ להרחבה ניתן לראות את אסטרטגיית הבינה המלאכותית של צרפת: <https://towardsdatascience.com/the-french-national-strategy-on-artificial-intelligence-c8c8fcfdace1>

- המחקר והפיתוח המדעי-טכנולוגי כבסיס יסודי ורחב לכל העיסוק בתחום.
- שימוש בבינה מלאכותית ככלי לקידום הדיגיטציה במשרדי הממשלה.

10.4.1. הון אנושי :

נושא ההון האנושי זוהה ע"י הוועדה כפער מרכזי לפיתוח התחום בישראל. בסעיף זה כלולים מספר נושאים הקשורים להון אנושי, והם עוקבים אחר ההמלצות שניתנו בדוח הוועדה המייעצת לות"ת.

- הרחבת מספר סטודנטים לתארים מתקדמים במקצועות הליבה.
- הרחבת ההכשרה במדעי הנתונים לכל רוחב האוניברסיטה על מנת לחשוף תלמידים מחוגים שאינם עוסקים בליבת התחום.
- יצירת שיתוף פעולה עם התעשייה ומוסדות הציבור תוך קידום מסגרת אחידה לשיתוף פעולה עם התעשייה.
- קידום שיתוף פעולה עם מוסדות ציבור תוך פתרון בעיות רגולטוריות ואתיות.
- הנגשת נתונים לאומיים וניצול נתוני מערכת הבריאות הישראלית.
- עידוד מחקר אינטרדיסציפלינרי וקידום נושאי המחקר השונים, תוך יצירת הזדמנות לנושאי מחקר חדשים בליבת מדעי הנתונים.
- גיוס סגל נוסף והגדלת מספר החוקרים.
- שמירה על האוניברסיטה כמרכז פעילות גם בעת עבודה חלקית בתעשייה.
- עידוד יציאה ללימודי פוסט דוקטורט בחו"ל.
- איסוף וארגון בסיסי נתונים מונגשים.
- יצירת מרכז תמיכה טכנית: חישובית, מקצועית ניהולית ותשתיתית.
- רכיב זה בתוכנית העבודה יכלול תקצוב לתשתיות פיזיות באקדמיה ולציוד תשתיתי שישרת את כלל האקדמיה והתעשייה.
- הכשרות מחוץ לאקדמיה – Bootcamps בתחומי מדע הנתונים ולמידת מכונה.
- תוכניות להכשרת כ"א במסגרת המעסיק, תוך איגוד חברות בעלות צרכי הכשרה דומים.
- תוכניות לפיתוח הון אנושי בתעשיית ההייטק בעזרת כלים גמישים וייעודיים לצרכים מסוימים – תוכניות לאיתור פוטנציאל מיון, הכוונה, הכשרה, התמחות או השמה של כ"א.

10.4.2. מו"פ תשתיתי

שיתוף מידע ופיתוח כלים חדשים המבוססים על שיתוף פעולה בין האקדמיה והתעשייה ליצירת פלטפורמה אחת, פיתוח כלים חדשים על בסיס רעיונות המבשילים מהאקדמיה וכן ולידציה של תובנות אקדמיות על בסיס מאגרי נתונים גדולים בעולם האמיתי.

- פיתוח מו"פ גנרי מבוסס AI לתעשיית הבטחון ולתעשייה האזרחית.
- מסחור ידע.
- מו"פ רב תחומי – מתן מענה לכשלי שוק ופיתוח כלי עזר הממוקדים בצרכי המגזר הציבורי ומערכת הבטחון.

10.4.3. ייצור מתקדם

פיתוח יכולות ייצור חכמות מבוססות AI. הוצאת תובנות ונתונים ממערכת הייצור היא פעולה נדרשת בעולם הייצור התחרותי, אשר מאפשרת חסכון במשאבים, פרסונליזציה גבוהה של מוצרים, וניבוי צרכים ומגמות צרכניות המשפיעות על הייצור עצמו.

10.4.4. ניהול

הקמת מנהל אזורי ובטחוני לתוכנית אשר יתאפיינו בראיה רוחבית לכלל התוכנית וסנכרון עם גורמים אחרים, תוך התמודדות עם עליה ניכרת בהיקפים.

10.4.5. מחשוב על:

קהל היעד של יכולות מחשוב העל לפי סדר העדיפויות הוא:

- בתעשייה ובאקדמיה: אימון (training) של מודלים בתהליכי מו"פ (עדיפות א), עדכון מודלים (re-training) (בעדיפות ב').
- תעשייה, אקדמיה ומגזר ציבורי: שימוש להסקה (inference) של תובנות על בסיס מידע.
- ביצוע אופטימיזציה של בעיות חישוב מורכבות (HPC) (סימולציה, הדמיה, ניתוח גנומי).

דרישות התשתית הן:

- נגישות בצורה מקוונת כשירות ענן לצד נגישות פיזית למפעילים או בליווי המפעילים לתשתית הניסוי.
- נגישות והרחבה דרך מספר תשתיות ענן שונות (AWS, RED HAT, AZURE) ותוך שימוש בכלי פיתוח המקובלים בשוק.
- מערכת מבצעת המשרתת את המשתמשים לצד מערכת ניסויית/לימודית המאפשרת חיבור של כלי תוכנה ורכיבי חומרה לצרכי הערכה (testing and evaluation)
- מערכת מרובת משתמשים על סמך מנגנון הקצאת משאבים מבוסס הקצאות ע"ב קריטיות, סוג משתמש, צרכים לביצוע המשימה, הזמנה מראש וכו'.
- יכולת לריכוז מקסימלי של משאבים לזמן קצר לטובת משתמש בודד – כדי להתמודד עם מודלים גדולים במיוחד.
- מערכת בקרה של נצילות וחיוב משתמשים לפי סוג המשאב (מעבדים, מאיצים, זיכרון)
- מערכת מודולרית המאפשרת חיבור פיזי של מאיצים ומעבדים מסוגים חדשים, יכולות אחסון נוספות וכדומה – מיצרנים שונים.
- מערכת המאפשרת שדרוג רציף של כלל המרכיבים ותשתית פיזית (למשל שדרוג של 20% מהמרכיבים כל שנה)
- יכולת גידול - מערכת המאפשרת די מקום פיזי ואספקת חשמל בדמות המקבילות למחשב-על המדורג במקום ה 20 ב Top500.
- אבטחת מידע - יכולות הפרדה מלאה בין המשתמשים בהקשר לקוד, למידע ונגישות לתשתית.

- כלים להתממת מידע (אנונימיזציה) על מנת לאפשר שימוש מקביל בין משתמשים שונים באותו מידע.
- כ"א בצוות התפעולי יהיה צוות ישראלי שיהיה בקיא ביכולת להתחבר לתשתית עצמה ויבצע הדרכה לצוותי מו"פ רלוונטיים באקדמיה, תעשייה ומגזר ציבורי כבשגרה.
- יכולת ביצוע וירטואליזציה על כרטיסי ה GPU לצורך עבודה מקבילית של מספר חוקרים על אותם משאבים (לפחות ב 500%)
דרישות החומרה המינימליות הן :
- מספר ליבות מעבדים מבוססי CPU : 88,000 (שקול למקום 50 בעולם נכון ליוני 2020).
- מעבדי GPU : 1,000 בספציפיקציה של A100 (או כוח עיבוד מקביל). תינתן עדיפות למעבדים עם ניסיון בשוק של שנה לפחות.
- Accelerators AI של לפחות שלושה יצרנים שונים כאשר כל אחד מהם מרכז יכולת של 1 PetaFLOPS לפחות.
- סך החומרה לעיל צריכה להסתכם ביכולת ביצוע במקסימום משאבים בלפחות 350 Petaflops לפי FP16 ולפחות 170 Petaflops לפי FP32 .
- זיכרון מהיר של לפחות 1000 TB (בתוך המכונות) ובנפרד מהמכונות נדרש לפחות PB30 של אחסון מהיר.
- רשת תקשורת מהירה בין ה-GPUs שמאפשרת Scale-out והרצת מודלים בפורמט Model parallelism ותקשורת מהירה ל-CPU ולאחסון שתומכת בקצבי ה-GPUs.
דרישות שירותי נגישות ל-HPC :
- יישומים.
- ספריות – ספריות AI לאחר אופטימיזציה (כולל תחזוקה למשך כלל זמן השירות), כגון MXnet, TensorFlow, ו-Caffe2.
- כלי עזר ופיתוח - זיהוי פנים וניתוח ווידאו (Rekognition) , מתן המלצות (Personalize), ניתוח תחזיות (Forecast) , ניתוח טקסטואלי (Comprehend), המרת טקסט לדיבור (Polly), המרת דיבור לטקסט (Transcribe).
- כלי ניתוח ביצועים למעקב אחר שימוש בפועל במשאבים.
- קומפילרים.
- חומרה ייעודית.
- כלי הדרכה ותמיכה – כגון AutoML של גוגל ו-AMT (Automatic Model-Tuning).

רכישת שירותי ענן ציבורי מאחת החברות אשר עתידות להקים תשתית ענן בארץ. גישה זו, רצוי שתכלול גם ליווי מעמיק של התעשייה ממנה רוכשים את השירות. מטרת תכולה זו היא לעודד מחקר ישראלי הן באקדמיה והן בתעשייה ככל שזו תתהווה עם הזמן, בחומרת state-of-the-art. מחקר כזה יתרום לפיתוח ההון האנושי והידע המדעי-טכנולוגי, ויכול לאפשר יתרונות יחסיים ברבדים העיליים של החישוב.

10.4.6.1. השירותים הנדרשים הם: שירותי אחסון מאובטחים, שירותי DB למידע מסוגים שונים, שירותי פרטיות והגנת מידע בפורמטים שונים, שירותי הנגשת מידע בפורמטים שונים, שירותי פיתוח web services ואפליקציית WEB, שירותי רשת מהירה ושירותי עיבוד ואנליטיקה.

הדרישות העיקריות הן:

- על הספק לכלול צוות תמיכה עבור כל אחד מאבני הבניין מעלה וכן להחזיק ברשימת ספקי שירות המסוגלים לבצע פיתוחי השלמה לשירותים אלה.
- על הספק להקים ולתחזק ספריית קוד משותפת למימושי רכיבים עיקריים בעולמות ה AI וה- HPC שמותאמת ונבחנת על גבי התשתית המוצעת.
- תפעול כלל השירותים על ידי צוות תפעול מהארץ המחויב לתפעול ואחזקת כלל השירותים הניתנים ב 24x7.
- רציפות חשמל, פתרון גיבוי מידע
- קישוריות מגובה ורחבת סרט (יש להגדיר את הרחב הנדרש)
- על הספק להבטיח מימוש של 100% מיכולות העיבוד על פי דרישה למשימות קריטיות בנוסף לדרישה הבסיסית לשירותי העיבוד
- אפשרות לשימוש בכלל השירותים העיקריים והרלוונטיים לבינה מלאכותית הקיימים בעננים אחרים (מתן אפשרות לייבוא שירותים).

דרישות כמותיות:

- Virtual machines – כ 40,000
- Vcpus – 100,000
- 300TB - Vram
- שירותי DB רלציוניים ולא רלציוניים – 800
- אחסון - active - 30 Petabytes
- אחסון - archive - 100 Petabytes
- שירותי GPU - לפחות 600 כרטיסים של V100 ויכולת שדרוג ל A100.
- יכולת ויטרואליזציה של משאבי GPU (חלוקת כרטיס GPU למספר חוקרים).

10.4.7. שת"פ בינ"ל:

מרכיב זה יסייע בהגדלת תקציבי השת"פ הבינ"ל הקיימים בגופי התקצוב, למימוש תוכניות שת"פ בינ"ל ייעודי לתחום מדע הנתונים ובינה מלאכותית.

10.4.7.1. דוגמא לשיתוף פעולה בינלאומי אשר ייטיב עם מדינת ישראל רבות במחקר בינה מלאכותית הוא הצטרפות ליוזמת PRACE (Partnership for Advanced Computing in Europe) לרשת מחשוב ציבורית בינלאומית.

11. הפן התקציבי בתוכנית המומלצת

11.1. ניתוח תקציב- עקרונות כלליים

- התקצוב נערך על ידי תתי הצוותים של הוועדה, ונבחן ונבדק במהלך דיוני רוחב על ידי כל הגורמים הרלוונטיים בוועדה ובקרב מי שליווה את עבודתה.
- התקציב במונחי שקלים קבועים.
- הונח כי תחילת הפעילות תהיה ב 2021 ; הונח תקצוב אחיד בעיקר בשנים המאוחרות היכן שלא היה מידע מלא ; לפיכך צפויים שינויים ועדכונים ככל שיתבררו פרטים תפעוליים. שעי"ח מול דולר ארה"ב : 3.5.

11.2. מרכיבי התקציב

- מרכיבי תשתית :

○ HPC

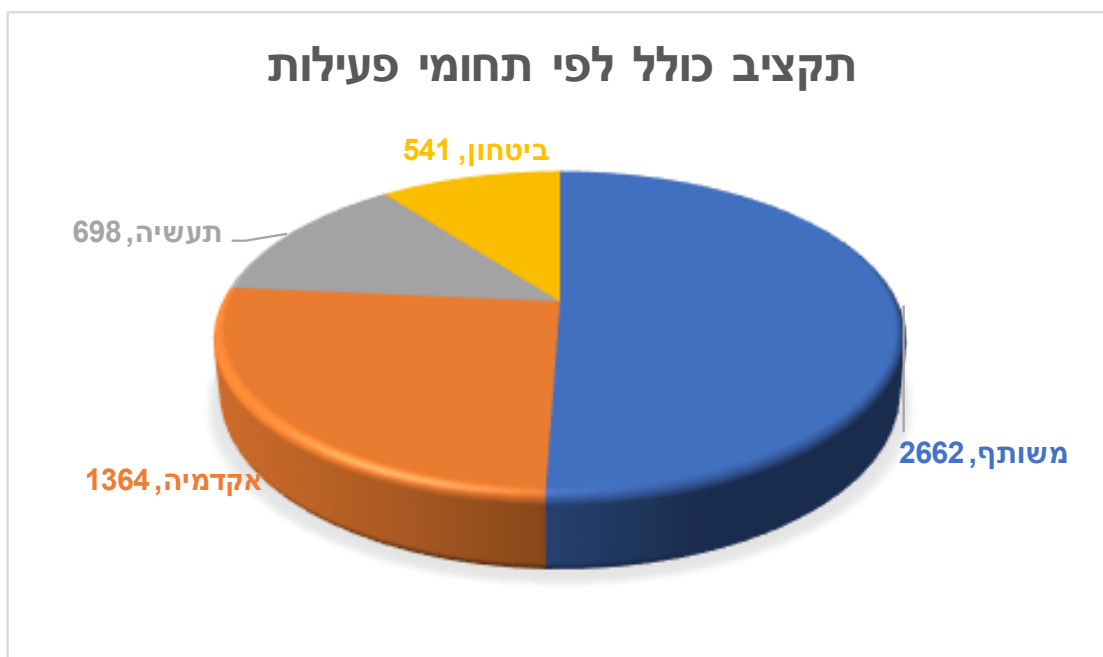
○ ענן

- צרכי אקדמיה

- צרכי תעשייה

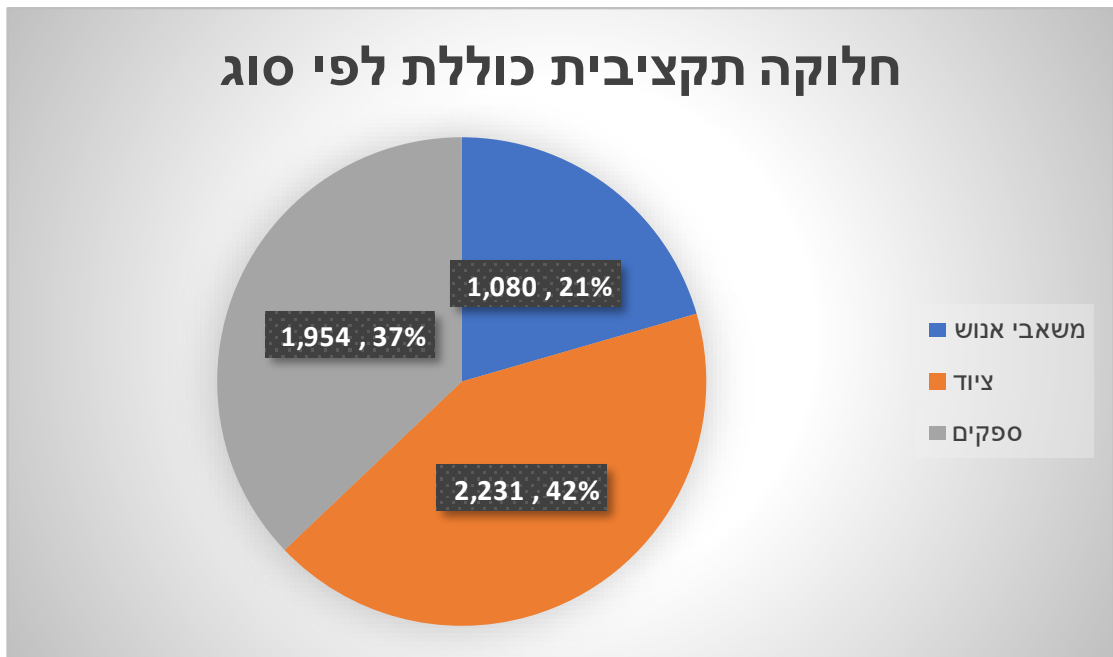
- צרכי מערכת הביטחון

11.2.1. חתך אקדמיה/תעשייה/ביטחון/משותף - האיור הבא מציג את התקציב בחתך גופי התקצוב, כאשר חלק מהפעילויות יבוצעו במנגנונים משותפים וחלקם בגופי התקצוב



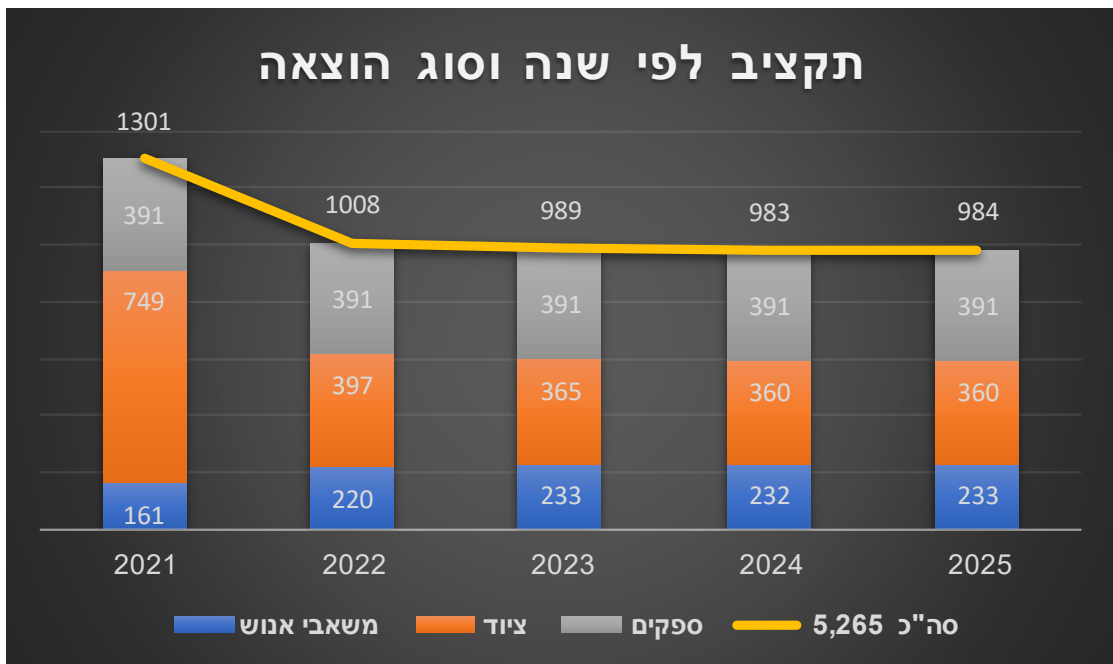
איור 12 : חתך תקציב על פי תחומי פעילות

11.2.2. חתך הון אנושי / ציוד / ספקים/שונות

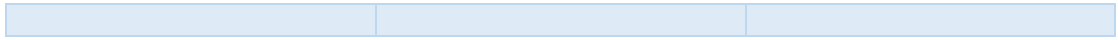


איור 9 : חתך חלוקה תקציבית

11.2.3. חתך התפתחות התקציב על ציר זמן – האיור הבא מציג את נתוני התקציב על ציר הזמן לאורך שנת ההתארגנות והחומש הראשון



איור 10 : התפתחות התקציב על ציר הזמן



12. חזון לתקופה שלאחר תום החומש הראשון

12.1. תוכנית הפעולה המפורטת במסמך זה תחומה לתקופה של 5 שנים ומטרתה לבסס פלטפורמה איתנה בתחום מדע הנתונים ובינה מלאכותית.

12.2. כפי שתואר בפרקים הקודמים, התחום ימשיך להתפתח ולהתרחב בעשורים הקרובים, כך שנדרשת הסתכלות ארוכת טווח והמשך הקצאת משאבים משמעותיים מעבר לחומש הקרוב. עם זאת, כפי שיתואר בפרק זה, מרכז הכובד לאחר החומש הראשון צפוי להתקרב לתעשייה. גם החלוקה בין תתי-התחומים השונים צפויה להשתנות לפי הבשלתם והיתרונות היחסיים שיתפתחו בעקבות החומש הראשון.

12.3. על-פי חזון הוועדה, עם התבגרות התחום, ובתום חמש השנים הראשונות, מדינת ישראל תימצא בנקודת פתיחה שונה משמעותית מזו בה אנו נמצאים כעת. ההבדלים יתבטאו במספר היבטים:

- בשלות לשיתופי פעולה בין הגופים השונים - בתום החומש הראשון, הצפי הוא שכלל הגופים המושפעים מתכולות התוכנית יעברו כברת דרך ויביאו להישגים משמעותיים בתחומם. תוצאה זו, תאפשר שיתופי פעולה עמוקים יותר ביניהם.
- בשלות לשיתופי פעולה בינלאומיים – שיתוף פעולה עם יוזמות בינלאומיות כגון יוזמת PRACE.
- מובילות עולמית בבינה מלאכותית – לאור העובדה כי ישראל כבר מדורגת גבוה בתחומי המחקר והפיתוח המסחרי בבינה מלאכותית, הוועדה צופה כי למדינת ישראל תהיה יכולת התמודדות בשורה הראשונה העולמית. חיזוק התעשייה יאפשר פיתוח מוגבר של הטכנולוגיות השונות, באופן שיוביל לכך שבתום חמש השנים הראשונות, מדינת ישראל תהיה בחזית ותוכל להתחרות מול הטכנולוגיות החדשניות ביותר.
- זיהוי מבוסס יותר של כיווני התפתחות – שימור מיקומה של מדינת ישראל בחוד החנית של טכנולוגיות בינה מלאכותית וזיהוי מוקדם של מגמות טכנולוגיות בתחום.
- ביסוס יתרון יחסי לישראל – בתום החומש הראשון לתוכנית, הוועדה חוזה כי היתרון היחסי בתחומי המיקוד יבוסס יותר, וכן שיזוהו כיוונים ונישות נוספים לפיתוח כיתרון יחסי להמשך, כתוצאה ממחקר בסיסי והנבטות שיבוצעו כבר בחומש הראשון, בכלל תתי-התחומים.
- גידול הקהילה האקדמית – הגדלת מספר אנשי הסגל והסטודנטים המצטיינים בתחומי הליבה, מניעת בריחת מוחות לחו"ל, חיזוק הקשר בין האקדמיה לתעשייה.
- הטמעת השימוש בבינה מלאכותית במשרדי ממשלה ושיפור השירות לאזרח – תוך יצירה וטיוב של מאגרי הנתונים, הטמעת תהליכים של טרנספורמציה דיגיטלית, וגיוס והכשרת הון אנושי מתאים.

13. משמעויות אי-מימוש התוכנית הלאומית המומלצת

מדינת-ישראל, במידה ולא תפעל מידי, לפתוח במרוץ ובמאמץ עליון (כפי שמציגה תוכנית זו שאנו מביאים), תישאר מאחור, **מדעית וטכנולוגית**, כבר בשנים הקרובות ותהיה למעשה כפופה לעליונות של מדינות ושחקנים 'זרים' אחרים, שיאפילו על מדינת-ישראל בטכנולוגיות עוצמתיות וידחקו אותה למצבים קשים בהיבטי ביטחון, כלכלה, בריאות, חינוך, משק ותעשייה. יתרה מזאת, חוסר משימתיות

בתחומי המחקר והטכנולוגיה במדעי הנתונים ובינה מלאכותית יעבירו רבים מבין טובי החוקרים והמפתחים למקומות בהם קצב ההתפתחות בתחום מביא למבחר התעסוקתי שופע. מדינת-ישראל, איננה, יכולה להרשות לעצמה להיות במצב שכזה.

להלן מספר דוגמאות קונקרטיות למשמעויות אי-מימוש התוכנית בחתכים אלו:

13.1. חתך טכנולוגי

- דוגמה א' – בריאות. מגפת ה-COVID-19 המחישה לישראל ולשאר מדינות העולם עד כמה חשובה היכולת לייצר, לאגור מאגר נתונים רפואיים, ולשתף ולנתח אותם כראוי, הן בהיבטי מחקר רפואי, הן בהיבטי טיפול פרטני בחולים, והן בהקשרי חקירות אפידמיולוגיות. מחסור בכלים ותשתיות דיגיטליות מתאימים לנוכח גל נוסף של המגפה עלול להחזיר את המשק הישראלי חודשים אחורה ולהוות מכה קשה על הכלכלה והבריאות במדינה. התמודדות ישראל עם המגיפה דרדרה את מוניטין החדשנות הישראלית ויתרה מכך את מעמדה האסטרטגי של ממשלת ישראל.

- דוגמה ב' – תעסוקה. מגפת ה-COVID-19 הותירה רבבות של ישראלים ללא תעסוקה, ולשכת התעסוקה מתקשה להתמודד עם עומס מבקשי העבודה. אם יעשה שימוש בכלים טכנולוגיים, בדומה לכלים שהוצגו בפרק ההטמעה, הדבר יכול לדאוג למציאת מקומות עבודה רבים, ולהורדה משמעותית של אחוז האבטלה במשק הישראלי.

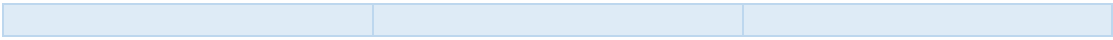
- מעבר לכך, ישנן גם משמעויות ביטחוניות לאי התקדמות בתחומים אלו, כבר בטווח הזמן המידי, שהינן מסווגות.

13.2. חתך סקטוריאלי - חתך הסתכלות אחר הינו החתך הסקטוריאלי ולא הטכנולוגי. כאן ניתן דוגמא למשמעויות של אי מימוש התוכנית על האקדמיה.

- דוגמה ג' - האקדמיה. דוח וועדת ההיגוי של ות"ת לתחום בינה מלאכותית ומדעי הנתונים, בראשותו של פרופ' יואב בנימיני, ביצע ניתוח מעמיק של איכות החוקרים הישראליים בתחום מדעי הנתונים, והראה כי הרמה האקדמית בארץ בתחום זה, על אף שהקהילה קטנה, הינה רמה גבוהה הן ביחס למחקר בתחום הבינה המלאכותית בעולם והן ביחס למחקר בתחומים מדעיים שכנים. שמירה על רמה שכזו מאפשרת פרסומים מדעיים בכתבי העת המובילים, זכייה במענקי מחקר יוקרתיים, וביצוע שת"פ בינ"ל מחקרי עם טובי החוקרים בעולם. אי מימוש התוכנית המומלצת בהקשר ליבת הבינה המלאכותית, תשאיר את הקהילה המחקרית מאחורי ההתפתחות המדעית המואצת בעולם, תקטין את האימפקט המדעי ואת היכולת להמשך שילוב (מחקרי ותקציבי) עם העולם, תגרום לבריחת מוחות של חוקרים שעוסקים בתחום זה לחו"ל, ולמעשה עלולה לנוון את העיסוק בתחום חשוב זה באקדמיה. בנוסף, מכיוון שהידע האקדמי הינו הבסיס ההכרחי ליכולת למעבר לפיתוח טכנולוגיה ומוצרים, פגיעה בהתפתחות הידע האקדמי תשליך הלאה גם על הסקטורים היישומיים, ועל היכולת להפיק תועלת כלכלית לתעשייה, וביטחוניית מתחום זה.

13.3. סינרגיה ושת"פ -

- דוגמה ד' – היות ובכל העולם נוצרות יוזמות ותוכניות לבינה מלאכותית, וישנן תוכניות בינלאומיות כגון PRACE, נדרשת גם שותפות של מדינת ישראל לבל תישאר לבדה מאחור. יש לציין כי באי-מימוש התוכנית, מלבד הישארות מדינת ישראל במסגרת המצומצמת שפורטה מעלה, מדינת ישראל עתידה לייצר פער למול מדינות אחרות בעולם. פער שיתכן ולא ניתן יהיה לגשר עליו



בעתיד. לאור זאת, וועדה זו רואה חשיבות עליונה במימוש התוכנית כפי שהוצעה, וזאת מתוך הבנה שבניית החוסן הלאומי תלויה באופן ישיר ביכולת שלנו לפעול בנושא זה, וכן ביכולת שלנו לפעול באופן נכון ברמה האסטרטגית.

לבקשתכם הוועדה מתכבדת להגיש דוח סופי על עבודתה.

קיימת בארץ קהילייה איכותית של חוקרים מהאקדמיה ומהתעשייה אשר יכולה להיות בסיס מצוין להתרחבות בתחום. לבינה מלאכותית ולמדע הנתונים יכולת לשפר את חיי האזרחים במדינת ישראל, את תעסוקתם, את הפיריון במשק ואת רווחתם הכלכלית של האזרחים באופן משמעותי.

הוועדה גיבשה אסטרטגיה אשר כוללת את הרכיבים שלהלן :

- (1) הרחבת הקהיליה האקדמית המצטיינת אשר עוסקת בליבת תחום הבינה המלאכותית מסדר גודל של 70 חוקרים לכ-100 חוקרים בתוך 5 שנים. חוקרים אלה הם אבן הראשה של כל ההתפתחות בתחום. בחוות דעתם של הפרופסורים יוגיין קנדל ומיכאל גופמן נדרשת הרחבה גדולה יותר כדי להתמודד עם ההגירה לתעשיות עתירות הידע בכלל ולסטארטאפים בפרט.
- (2) הרחבת קהילת הסטארטאפים הישראלית הפעילה כיום בתחום בעשרות אחוזים (כ-1200 חברות).
- (3) זמינות שירותי מחשוב על לחוקרים ומפתחים באקדמיה, בתעשייה ובמערכת הביטחון.
- (4) זמינות שירותי ענן מתקדמים לחוקרים ומפתחים באקדמיה, בתעשייה ובמערכת הביטחון.
- (5) שיפור השירות הממשלתי הניתן לאזרח בעזרת הטמעת בינה מלאכותית ושימוש בבסיסי נתונים זמינים ומובנים.
- (6) שיפור הפיריון בתעשייה הישראלית באמצעות הטמעת בינה מלאכותית.
- (7) העלאת רמת הביטחון בעזרת פיתוח יכולות בינה מלאכותית.

בהיבט הטקטי אנו ממליצים על הצעדים שלהלן :

- (1) מימון תקנים חדשים לקליטה במוסדות להשכלה גבוהה – מלגות אלון משופרות.
- (2) בניית קבוצות מחקר חזקות מסביב לאותם 100 חוקרים בעזרת 10 תלמידי תארים מתקדמים בכול קבוצה.
- (3) תקציבי מחקר תחרותיים בהיקף גדול מהמקובל לבניית קבוצות מחקר משמעותיות.
- (4) תוכנית לתמיכה בתעשייה באמצעות הגדלת היצע ההון האנושי, תמיכה במו"פ מוצרי, מו"פ תשתיתי, והטמעת השימוש בבינה מלאכותית בתעשייה.
- (5) שורת פרויקטי פיתוח לאומיים מסביב לבעיות גנריות.
- (6) שיפור הבנת השפה העברית בתהליכים אוטומטיים.
- (7) הקמת מחשב על ישראלי בהיקף של 88,000 ליבות.
- (8) רכישת שירותי ענן מתקדמים אשר יועמדו לטובת חוקרים מהאקדמיה, התעשייה ומערכת הביטחון.

המלצותינו מסתכמות בסדר גודל של 5.26 מיליארד ₪. מתוכם כ-1.15 יופנו לאקדמיה הישראלית, כ-1.09 יופנו לפעילות תעשייתית, כ-1.08 לפעילות במערכת הביטחון (לשתי האחרונות שת"פ גנרי הדדי ניכר בשיעור של כ-80% בתעשייה וכ-95% בביטחון), כ-1.90 ישמשו להקמת התשתיות

הלאומיות (מהם 1.08 למחשב העל, כולל תשתית תחזוקה ועדכון ל-5 שנים, וכ-0.86 לשירותי הענן). שני מרכיבי התשתית ישמשו הן להטמעת בינה מלאכותית במגזר הממשלתי והן למחקר ופיתוח.

15. תודות

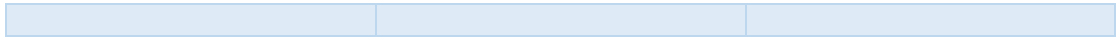
שיתוף הפעולה בין חברי הוועדה, המרכזים והסיוע של פרופ' חיים לוי קידר, ד"ר ניר ברילר וצוות מחברת רבתית הנדסה וייעוץ בע"מ הביאו לעבודה מעמיקה בעלת משמעות חברתית וכלכלית ולתוצר חשוב למדינת ישראל.

תודות נוספות לגב' ענבר דולינקו, חוקרת טכנולוגיה ומדיניות, מאסטרנטית באוניברסיטת ת"א, על הסיוע בכתיבת פרק הרגולציה, ולאלי"מ במיל' רות שוהם, מנכ"לית האוניברסיטה הפתוחה.

16. נספחים

1. נספח תקציבי

נספח זה מפרט מעט יותר לעומק את הרציונל התקציבי לכל מרכיב בתוכנית. המהות של כל מרכיב תוארה כבר בפרק 12 בדוח ולכן לא נחזור עליה כאן. התכנון בוצע בהתאם להערכות תקציביות ובעקבות פגישות רבות עם גורמים שונים בממשלה, באקדמיה, ובתעשיות, ובעקבות דיונים בוועדה עצמה. התכנון התקציבי בוצע על פי חתך צפוי של הפעילויות השונות על ציר הזמן (שנת התארגנות + חומש ראשון של פעילות), כאשר כפי שתואר בדוח חלק מהפעילויות למעשה כבר קיימות/החלו לאחרונה בגופי התקצוב השונים. בנוסף, בוצע חתך לפי הוצאות הון אנושי, רכש ציוד וחומרים, והוצאות ספקים (כאשר במקרים מסוימים גם סעיפים שאינם בדיוק קבלני משנה הוכללו תחת קטגוריה זו).



HPC												
שנים	2020	2021	2022	2023	2024	2025	סה"כ	ביטחון	תעשיה	אקדמיה	משותף	עלות כוללת
משאבי אנוש	0.0	0.0	0.0	0.0	0.0	0.0	0.0	25%	25%	25%	25%	0.0
ציוד	0.0	542.0	135.3	135.3	135.3	135.3	1083.2	25%	25%	25%	25%	1083.2
ספקים	0.0	0.0	0.0	0.0	0.0	0.0	0.0					0.0
סה"כ	0.0	542.0	135.3	135.3	135.3	135.3	1083.2	270.8	270.8	270.8	270.8	1083.2

ענן												
שנים	2020	2021	2022	2023	2024	2025	סה"כ	ביטחון	תעשיה	אקדמיה	משותף	עלות כוללת
משאבי אנוש	0.0	0.0	0.0	0.0	0.0	0.0	0.0					0.0
ציוד	0.0	0.0	0.0	0.0	0.0	0.0	0.0					0.0
ספקים	0.0	172.7	172.7	172.7	172.7	172.7	863.5					863.5
סה"כ	0.0	172.7	172.7	172.7	172.7	172.7	863.5	0.0	0.0	0.0	0.0	863.5

צרכי אקדמיה												
שנים	2020	2021	2022	2023	2024	2025	סה"כ	ביטחון	תעשיה	אקדמיה	משותף	עלות כוללת
משאבי אנוש	0.0	65.9	124.4	137.9	136.7	137.9	602.8	0.0		1.0		602.8
ציוד	0.0	35.0	90.0	57.5	52.5	52.5	287.5	0.0		0.5	0.5	287.5
ספקים	0.0	52.5	52.5	52.5	52.5	52.5	262.5	0.0		0.5	0.5	262.5
סה"כ	0.0	153.4	266.9	247.9	241.7	242.9	1152.8	0.0	0.0	877.8	275.0	1152.8

צרכי תעשיה												
שנים	2020	2021	2022	2023	2024	2025	סה"כ	ביטחון	תעשיה	אקדמיה	משותף	עלות כוללת
משאבי אנוש	0.0	39.2	39.2	39.2	39.2	39.2	196.1		17.0%		83.0%	196.1
ציוד	0.0	91.2	91.2	91.2	91.2	91.2	456.0		20.0%		80.0%	456.0
ספקים	0.0	87.2	87.2	87.2	87.2	87.2	436.0		20.0%		80.0%	436.0
סה"כ	0.0	217.6	217.6	217.6	217.6	217.6	1088.1	0.0	211.7	0.0	876.4	1088.1

מערכת הביטחון												
שנים	2020	2021	2022	2023	2024	2025	סה"כ	ביטחון	תעשיה	אקדמיה	משותף	עלות כוללת
משאבי אנוש	0.0	56.3	56.3	56.3	56.3	56.3	281.6	5.0%			95.0%	281.6
ציוד	0.0	80.9	80.9	80.9	80.9	80.9	404.3	5.0%			95.0%	404.3
ספקים	0.0	78.4	78.4	78.4	78.4	78.4	391.8	5.0%			95.0%	391.8
סה"כ	0.0	215.6	215.6	215.6	215.6	215.6	1077.8	53.9	0.0	0.0	1023.9	1077.8

2. כתב מינוי הוועדה



16 פברואר 2020

תל"מ – פורום לתשתיות לאומיות למחקר ולפיתוח מיסדה של האקדמיה הלאומית הישראלית למדעים

לכבוד:

ד"ר ארנה ברי - יו"ר הוועדה

חברים:

פרופ. יאיר וויס
פרופ. אמיר גלברזון
פרופ. שרית קראוס
ד"ר איה סופר
מר אבנר גורן
מר מוטי בסך

מרכזים:

ד"ר אביב זאבי – מרכז הוועדה
סא"ל ערן דהן - מרכז הוועדה
מר גדי לוי - מרכז הוועדה

שלום רב,

הנדון: מינוי ועדת בדיקה מקצועית לבחינת הצורך בהתערבות ממשלתית לשם האצת התפתחות תחום הבינה המלאכותית ומדע
הנתונים (להלן – הוועדה)

1. על יסוד החלטת פורום תל"מ מיום 26.12.2019, אני מתכבד למנותכם כחברי הוועדה. ד"ר ארנה ברי תהיה יו"ר הוועדה. ד"ר אביב זאבי, סא"ל ערן דהן ומר גדי לוי יישמשו מרכזיה.

2. בהתאם להחלטת הפורום, הוועדה תתמקד בהיבטים הבאים: הון אנושי, תשתיות פיזיות, הנגשת מאגרי מידע והעברת הידע מהאקדמיה לתעשייה. זאת תוך התבססות, בין היתר, על התכניות הקיימות והמתוכננות ברשויות המחקר ומשרדי הממשלה הרלבנטיים.

3. להלן תפקידי הוועדה:

א. לבדוק את הצורך והתועלת הצפויים למדינת ישראל מהתערבות ממשלתית בתחום הבינה המלאכותית ומדע הנתונים, בין היתר על ידי הקמת תשתית נדרשת או העצמת תשתית קיימת, לרבות מפיתוח המחקר האקדמי הבסיסי והטכנולוגי-יישומי. הנכם מתבקשים לבחון את הנושא בהתחשב בצרכי המו"פ של האקדמיה, מערכת הביטחון, וכן התעשייה בישראל, תוך השוואה לתכניות הקיימות בממשלה וכן היקף ומצב התשתיות הקיימות בגופים אלה.

תל"מ – אגף רשות החדשנות, זירת תשתיות טכנולוגיות, רח' הרדן 4, קריית שדה התעופה
טיפון: 052-2899501



תל"מ – פורום לתשתיות לאומיות למחקר ולפיתוח מיסודה של האקדמיה הלאומית הישראלית למדעים

ב. לזהות ולהעריך את הפוטנציאל הגלום בתכניות המתוארות במסמכים המצורפים, לפיתוח תשתית מדעית וטכנולוגיות מתקדמות, לשימוש אזרחי וביטחוני על בסיס התשתית המוצעת, ולהמליץ על סדרי עדיפויות בין התחומים בשטח. לאפיין ולפלא את אוכלוסיית המשתמשים הפוטנציאליים (כולל חברות ומספרן) במגזרים השונים, ולהעריך את השימוש הצפוי בתשתיות המוצעות.

ג. לבחון תכניות ומודלים מקבילים במדינות אחרות ולבדוק את האפשרויות לשיטופי פעולה איתן.

ד. להמליץ על אבני דרך לפיתוח תשתיות המחקר האקדמי והפיתוח היישומי, והדרכים למימוש התוצרים והתועלות הכלכליות והמדדים להשגתם, והיבטים נוספים בהתאם לשיקול דעתה.

ה. בחינת הגדרת קטרים טכנולוגיים ופרייקטליים, שטמרתם הצבת המדינה כמובילה ביני"ל בתחום הבינה המלאכותית והמערכות הנתונות. זאת בשם לב לבחינת עלות-תועלת, פעולות הממשלה הקיימות והמתוכננות בתחום וסדרי עדיפויות.

ו. להמליץ על הגדרת אסטרטגיה לאומית לריכוז המאמצים להשגת מידע הדרוש לצורך אימון בינה מלאכותית בתחומי מפתח, בין היתר באמצעות "רגולציה מאפשרת", או תהליך מוצע לשם השגתה של "רגולציה מאפשרת", שמטרתה להנגיש מאגרי ידע קיימים ועתידיים, תוך הבטחת כללי הסדרה ראויים. זאת תוך התבססות, ככל הניתן, על רגולציה מאפשרת בינלאומית, ובתיאום עם רשת C4IR של הפורום הכלכלי העולמי, לאור החלטת הממשלה להצטרף לרשת.

ז. לבחון את אפשרויות המסחור והייצוא של תוצרי המו"פ לרבות המגבלות שיחולו עליהן.

ח. לאפיין מודל פעילות לתשתית הנדונה תוך תי"חיסות לתחומים הבאים:

(1) היבטים מדעיים / טכנולוגיים –

• תיחום שטח הבינה המלאכותית ומדע הנתונים ושטחים רלוונטיים.

(2) היבטים ארגוניים –

• ריכוז או ביוור התשתית (מרכז או מרכזים, לרבות תשתיות משותפות) ומקומה (באוניברסיטאות המחקר ו/או בתעשייה).

• זמינות ונגישות לחוקרים מהתעשייה, מהמוסדות להשכלה גבוהה וממערכת הביטחון.

• טיפול בהנמכת חסמים ליצירת יישומים חדשים, לרבות כוח אדם מיומן בתחומים כגון אלגוריתמים למיקובל בעיות חישוב מדיסציפלינות מדעיות שונות.

• שיטת הפעולה והגדרת מודל השירות המוצע לאוניברסיטאות, לתעשייה ולמערכת הביטחון, לרבות שיטת הניהול והבקרה (ברמת מרכז/מרכזים וברמת התכנית הכוללת).

• אפשרויות היישום על ידי המשרדים והגופים החברים בפורום תל"מ בכל אחת מאופציות המיקום ובהתחשב בנהלים המחייבים בהם והיקף פעילותם הקיים והמתוכנן.

(3) היבטי תשתיות –

• תשתית פיזית (למעט בינוי) - פירוט הציוד הנדרש לקידום המו"פ לפי קדימויות.



תל"מ - פרום לתשתיות לאומיות למחקר ולפיתוח מיסודה של האקדמיה הלאומית הישראלית למדעים

- תשתית אנושית לתפעול, תחזוקה והטמעה.
- מניעת כפילות (בהתייחס לתשתית וציוד קיימים).

(4) היבטים כלכליים ותקציביים -

- מודל תקצוב - לרבות טבלת מקורות ושימושים, עלות הקמה ועלות תפעול (כולל כ"א טכני).
- מודל מימון - לרבות שיעורי המימון העצמי של הארגון / ארגונים בהם תמוקם התשתית, דמי שימוש למשתמשים חיצוניים ושיטת המימון בתום תקופת התמיכה הממשלתית / ציבורית, לשם פעולה ברת קיימא.
- הבטחת נגישות הידע המפתח באקדמיה לתעשייה והעברתו באופן מהיר ויעיל
- הגדרת סדרי עדיפויות לגבי המלצות הוועדה.

4. כללי:

- א. יו"ר הוועדה, מרכזי הוועדה וכלל חברי הוועדה יחתמו על הסדר ניגוד עניינים, כתנאי להשתתפותם.
- ב. חברי הפרום יוכלו למנות נציג מטעמם שיהיה משקיף בוועדה. משקיפים יוכלו להשתתף בכלל ישיבות הוועדה.
- ג. הוועדה תשים דגש על מיפוי התכניות הקיימות והמתוכננות ברשויות המחקר ומשרדי הממשלה הרלבנטיים, וכן תימנע ככל הניתן מעיסוק בנושאים אשר לגביהם מתקיימת עבודה של ועדה ממשלתית שמונתה בהחלטת ממשלה.
- ד. הוועדה מתבקשת להזמין לשיבותיה את נציגת החשבי"ל בתל"מ, להיפגש עם חברי פרום תל"מ/נציגיהם. הוועדה חופשית להיפגש עם אישים ומומחים נוספים עפ"י שיקול דעתה.
- ה. הוועדה תקבע את דרכי ושיטות עבודתה, במסגרת תקציבה המאושר ותוכל להיעזר בעבודתה בשלושת מרכזיה לצורך איסוף ועיבוד החומרים והנתונים.

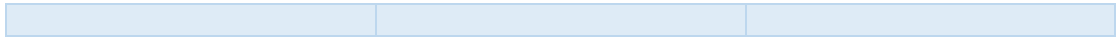
5. אני מודה לכם על הסכמתכם ליטול חלק בעבודת הוועדה ומעריך מאד את התנדבותכם למען פיתוח תשתיות המו"פ בישראל. אהיה אסיר תודה אם דוח ראשוני, לרבות הממצאים, המסקנות והמלצות הוועדה יוגש לנו עד יום 1.6.2019 ויוצג על ידכם למלאות פרום תל"מ. ככל הניתן, אבקש כי יוגש דו"ח ביניים עם הממצאים המרכזיים טרם מועד זה.

לוט:

- מצגת מיום 26 דצמבר 2019.
- דו"ח ועדת היגוי מייצגת לנושא מדעי הנתונים בראשות פרופ' יואב בנימיני.

בברכה,

שמעון אולמן
יו"ר תל"מ



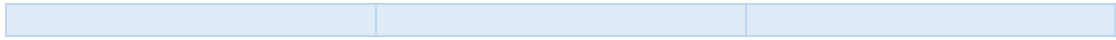
3. פירוט צרכי תעשייה ובטחון

תחום	מטרה	כלי מימון	תקציב	תקציב	תקציב	דוגמאות
			לשנה - רשות	לשנה מפא"ת	שנים	
			מונחי מענק מיליון ₪			
הכשרת הון אנושי	200 מוכשרים בשנה מחוץ לאקדמיה - שת"פ רשות לחדשנות - מפא"ת	סבסוד Boot Camps	1.5	4	27.5	<ul style="list-style-type: none"> סיירת ITC שמבצעת הכשרות מתקדמות בתחום ה- data science לבוגרים בעלי פרופיל גבוה (לרוב תארים במדעים מדויקים). בעלת אחוזי השמה גבוהים ושכר השמה ממוצע מעל K20. סיירת פרימרוז שמבצעת הכשרות בתחומי ה- machine learning לבוגרים בעלי פרופיל גבוה. בעלת אחוזי השמה גבוהים ושכר ממוצע מעל K20.
	הכשרת כ"א במסגרת המעסיק - סבסוד תוכניות הכשרה של איגודי חברות בעלות צרכי הכשרה דומים	3 איגודי הכשרה : איגוד קיים, 6 חברות באיגוד - (לפחות 200 בוגרים ל-3 שנים) כ-70 בשנה. + שני איגודים חדשים, לפחות 5 חברות באיגוד (לפחות 200 ל-3 שנים) כ-70 בשנה	3.5	0	17.5	<p>חברות הייטק בעלות עניין בהכשרה בתחומי הידע החדשים ביותר, שיחברו משאבים של ידע מקצועי, זמן חניכה, קשב ניהולי ומשאבים חומריים על מנת ליצור ולהוביל הכשרה משותפת למהנדסים ומהנדסות. מדובר בעובדי החברות הקיימים ולא בהכשרת עובדים חדשים, במטרה לשמור על רמה טכנולוגית גבוהה בקרב המהנדסים.</p> <p>במסגרת הקונסורציום החברות בנו 2 מסלולי הכשרה עיקריים:</p> <ul style="list-style-type: none"> - מהנדסים בליבת המו"פ בצוותי פיתוח, אנשים שנמצאים בצוותי R&D או שמגייסים במיוחד בשביל להשתלב בצוותים אלה, ויבצעו משימות פיתוח בליבת עולמות הבינה המלאכותית והאלגוריתמיקה - מהנדסים שעוסקים בהטמעה ובכתיבת קוד. מפתחים שיוכלו לקחת את הבסיס מהחוקר ולכתוב קוד לפיו יוטמע היישום במערכת
	פיתוח הון אנושי לתעשיית ההייטק בעזרת כלים גמישים וייעודיים לצרכים מסוימים שת"פ רשות לחדשנות - מפא"ת	קרן הון אנושי - שלושה פרויקטים של תאגידים (עסקיים או מלכ"ר)	10	5	75	<p>קרן תמיכה (על ידי מתן מענקים כספיים) בתוכניות אשר מבקשות לתת מענה לאתגרים בפיתוח הון אנושי מיומן עבור תעשיית ההייטק בישראל בדגש על פתרונות חדשניים.</p> <p>1. ערוצי הכניסה לתעשייה עתירת הידע בישראל, של הון אנושי מיומן בתחומי מחקר ופיתוח, באמצעות שימור ו/או שיפור של הון אנושי איכותי, מיומן בתחומי מחקר ופיתוח, בתעשייה עתירת הידע בישראל.</p>
מו"פ מוצרי	הוכחת ביצוע והתאמה ללקוחות שת"פ רשות לחדשנות - מפא"ת	פיילוטים בישראל ובחור"ל - בריאות, ייצור, תעסוקה וביטחון	75 מיליון ₪ - 50 פרויקטים לשנה במוצע של 3 מיליון ₪ לפרויקט ב 50%	32.5 מיליון ₪ - 5 פרויקטים לשנה בעלות של כ 8 מיליון ₪ לפרויקט, פיילוטים	537.5	<ul style="list-style-type: none"> זיהוי מניפולציות בשוקים באמצעות בינה מלאכותית זיהוי חשד לקורונה ומעקב אחרי התקדמות המחלה ע"י ניתוח אוטומטי של CT חזה מערכת ai לתכנון אוטומטי של תוכניות רמזור אופטימליות

זיהוי אנומליות ברשתות תקשורת		350	במערכת הביטחון	20 - 40 פרויקטים לשנה בממוצע של 4 מיליון ₪ לפרויקט. 50% מימון	פיתוח מוצר – קרן המו"פ (החלק המימושי מתוך התשתית המשותפת) חלק אזורי חלק ביטחוני	שת"פ רשות לחדשנות - מפא"ת	
<ul style="list-style-type: none"> • עיבוד נתונים דיגיטאלי מתקדם לחיזוי של חסינות לסימפטומים Covid19 והצבעה על תרופות אפשריות • מערכת תומכת החלטה מבוססת AI למיון ראשוני של חולים ב-Covid-19 ובמחלות זיהומיות אחרות של דרכי הנשימה לצורך מיקוד תהליך האבחנה המבדלת ומתן טיפול ראשוני מדויק. • re: view - ai av and adas sensor positioning verification • ציפ זיכרון חישובי והאצת אלגוריתמי ai מרובי מידע • אלגוריתמיקת ראייה מתקדמת מבוססת ai לסביבת embedded • advanced bluetooth soc with anc, voice trigger and ai engine • מערכת AI לחקירת רשומות לוג וזיהוי אנומליות בתחום האבטחה • עיבוד נתונים דיגיטאלי מתקדם לחיזוי של חסינות לסימפטומים Covid19 והצבעה על תרופות אפשריות • אפליקציות ביטחוניות מודיעניות, מבצעיות בשלל תחומים על בסיס התאמה למידע מבצעי ולאטגרי מידע מבצעי • שילוב האלגוריתמים וביצוע fine tuning על מידע מבצעי • פיתוח פיצ'רים מתאימים להגעה לביצועים מיטבים • מערכת AI לחקירת חשודים מודיעיניים • מערכת AI למציאת אנומליות וסיווג במגוון ספקטרומים 	<p>פיתוח מוצרים טכנולוגיים (חומרה ותוכנה) על מודלים מאומנים המבוססים על תובנות שפותחו על סמך אנליזה של נתונים</p>	350	30 מיליון ₪ – 5 פרויקטים בשנה – בממוצע של 8 מיליון ₪ לפרויקט	20 - 40 פרויקטים לשנה בממוצע של 4 מיליון ₪ לפרויקט. 50% מימון	פיתוח מוצר – קרן המו"פ (החלק המימושי מתוך התשתית המשותפת) חלק אזורי חלק ביטחוני	שת"פ רשות לחדשנות - מפא"ת	
<ul style="list-style-type: none"> i. גילוי תרפיות חדשות ii. יעול תהליכי המו"פ בשלבי הפיתוח הפרה-קליני והקליני iii. אבחון חכם ורפואה מותאמת אישית iv. טיפול מבוסס דיגיטציה של מידע גינומי וקליני 	<p>פיתוח שיטות וטכנולוגיות חדשניות מבוססות AI והביולוגיה החישובית ע"י חברות ומחקר אקדמי ליצירת פתרונות ביוטכנולוגיים מתקדמים והגעה להוכחת התכנות</p>	32	0	6.4	מעבדת חדשנות/חממה בתחומי הבריאות הדיגיטלית, פארמה ו-AI	סביבת עבודה טכנולוגית לחברות	הזנק

					0	6.4	32	מקום ותשתית מסווגת לפיתוח מודלים מבוססי למידת מכונה על DATA מסווג שנאסף מסנסורים ביטחוניים	i. שימושים ביטחוניים שונים
					5	5	50	בדיקת חסינות AI : א. הטעיית של מערכות AI ו ML בשלבי אימון והסקה ב. הטעייה הנובעת מהטיה לא שוויונית והמלצות לשיפור ג. פרטיות. האם הדלפה צפויה מהנתונים ששימשו לאימון. ד. פיתוח מודלי חסינות AI	ii. תשתית לשירוטי בדיקה של פיתוחי חברות על עמידות בפני איומים iii. תקנון עבור רגולציה שתפתח בקרוב. iv. בדיקת מודלים אופציונליים לשימוש
מו"פ תשתית	פיתוח מו"פ גנרי העברות ידע שת"פ רשות לחדשנות - מפא"ת	מאגדי מו"פ מבוססי AI + ביטחון כ 50% משותף (ביטחון/אזרחי) 50% ביטחוני	24	90	684	שיתוף מידע וכן גורמים שונים באקדמיה והתעשייה המייצגים טכנולוגיות שונות להפקת מידע בפלטפורמה אחת, עשוי להביא ליצירת תובנות חדשות בכלים של בינה מלאכותית	א. הטמעת שיטות בינה מלאכותית עבור ניטור סביבתי של מידע ביולוגי בעזרת שופכין כמעקב לניטור זיהומים ותחלואה ... ב. יכולות ניבוי של הצלחת טיפולים רפואיים ותרופות בעזרת מודלים ביולוגים חוץ גופיים על בסיס מדידה של נתונים ואנליזה שלהם לאורך זמן ג. פיתוח יכולות סנכרון של מידע מחיישנים בטכנולוגיות שונות (lidar, תחום נראה...) להוצאות תובנות חכמות. ד. תשתית ואלגוריתמיקה NLP גנרית בשפות בעלות עניין בישראל ה. תשתית ואלגוריתמיקה לניתוח תקשורת אלחוטית במגוון מתארים ומרחקים. יכולות מעבר משמע לטקסט בעברית ופיתוח מדדים לשילוב בין ניתוח טקסט לניתוח קול ראייה ממוחשבת לטובת הוצאת תובנות ויצירת שכבות וקטוריות מתמונות ח. הבנת סצינה ותמלול וידאו ממגוון תחומי חישה		
	מסחור ידע	מגנטון/יבוא ידע – 25 מיליון ש"ח - 5 פרויקטים בתקציב 1.7 מיליון ש"ח כ"א – 5 מיליון ש"ח	0	0	125	פיתוח כלים חדשים על בסיס רעיונות המבשילים מהאקדמיה וכן וולידציה של תובנות אקדמאיות על בסיס מאגרי נתונים גדולים " בעולם האמיתי"	א. מעקב, זיהוי שניוניים ומניעה מוקדמת של סרטן ב. פיתוח טכנולוגיה לניבוי הצלחה ו- Risk Scoring של השקעות		

				במונחי מענק.		
א. אלגוריתמים להקצאת משאבי זיכרון בצורה מבוזרת ב. זיהוי אנומליות בנתוני מידע גנומי ג. הסברתית של מודלי בינה מלאכותית ד. "הרזיית" רשתות נוירונים לריצה בקצה ה. פיתוח מודלים ושיטות ללמידה ממספר קטן של דוגמאות ו. אוטונומיה מבוססת בינה מלאכותית ז. אלגוריתמים לתקיפה או הגנה של מידע מרשתות נוירונים	הבשלת רעיונות ראשוניים המבוססים על רעיונות חדשניים שטרם הגיעו לידי מימוש ביישום אבל יש בהם יכולות ליצירת חדשנות גבוהה.	200 מיליון ₪ – מזה כ 80 מיליון ₪ במשותף וכ 120 מיליון ₪ לטובת הנושאים הביטחוניים	32	8 בין 10 ל 20 פרויקטים בשנה	הכוונת ידע באקדמיה (פיתוח טכנולוגיות ניבוי, explainable AI, ממד גנרי, מו"פ גנרי בתעשיות שת"פ רשות לחדשנות - מפא"ת	מו"פ רב תחומי
א. בניית סדרה של קורפוסים ואנתולוגיות של תחומים טכנולוגיים בעברית וערבים ושילובם בכלי פיתוח NLP. ב. בניין כלי תיוג חכם ומהיר המותאם לצרכים של מאגרי מידע קיימים במערכות שיש הגיון כלכלי בהוצאת התובנות מהם (תחבורה, בריאות, תעסוקה וכו...) ג. מאגר מידע בנושאי תעסוקה המבוסס על מידע זורם ממספר גורמים (מעסיקים גדולים לדוגמא...)	מתן מענה לכשלי שוק בפיתוח כלי עזר הממוקדים בצרכים של המגזר הציבורי בישראל, מערכת הביטחון והיכולת לנטר מידע רפואי קיים שרובו בעברית ואינו מובנה או מותמם ולכן השימוש בו מוגבל בכלים הקיימים כיום בשוק. מיקוד החברות המפתחות בעולם אינו נותן מענה.	60 מיליון ₪ כ 80 אחוז משותף ועוד כ 10 אחוז לרשות לחדשנות ו 10 אחוז לביטחוני	6	6 : 3-5 איגודים	איגוד תשתיתי שת"פ רשות לחדשנות - מפא"ת	פיתוח כלים גנריים ומאגרי מידע משותפים (דוגמת NLP, תיוג חכם וכו...)
א. בניית ספריות הפעלה למערכות רובוטיות המבוססות על נתוני ייצור על מנת לשפר ולייעל את המערכות ב. הוצאת תובנות מנתוני שימוש בחומרי גלם על מנת לייעל שימוש במשאבי ייצור וכדי לאפשר שקיפות בנתונים לאורך תהליכי הייצור	הוצאת תובנות מנתונים במערכות הייצור היא פעולה נדרשת בעולם הייצור התחרותי, היא מאפשרת חסכון במשאבים, פרסונליזציה גבוהה של מוצרים וניבוי צרכים ומגמות צרכניות המשפיעות על הייצור עצמו	80 מיליון ₪	0	6 - 16 פרויקטים לשנה בממוצע של מיליון ₪ 2-3 פרויקטי דגל של 5 מיליון ₪ לשנה	מו"פ מוצרי בתעשיות ייצור	פיתוח יכולות ייצור חכמות מבוססות AI
	ראיה רוחבית לכלל התוכנית וסנכרון עם גורמים אחרים	2.5	0	500 אלף ₪ בשנה		ניהול מנהל התוכנית - אזרחי
	התמודדות עם עליה ניכרת בהיקפים	3.6	0	720 אלף ₪		5 תקני עזר בזירות הרשות 1/3משרה
	ראיה רוחבית לכלל התוכנית וסנכרון עם גורמים אחרים	13.5	2.7	0		מנהלי תוכנית - ביטחוני



	התמודדות עם עליה ניכרת בהקפים	9.75	1.95	0		תקני מנהלי תחומים	
		2,185.35	215.55	221.52			סה"כ

4. הערכת עלויות של מסדי נתונים בענן עבור חקר קורונה (נכתב על ידי פרופסור אורן קורלנד)

הערכת עלויות של מסדי נתונים בענן עבור חקר קורונה

עורך	תאריך עידכון	סטטוס	הערות
נעם כהן	2020-08-03	טייטה	
נעם כהן	2020-08-05	Added summary	

רקע

במדינת ישראל יש כ 9.1 מיליון אזרחים [1]. ישנם מספר גופים המחזיקים נתוני בריאות עבור חלק מהאוכלוסייה:

- קופות חולים
- מד"א , איחוד הצלה
- משרד הבריאות
- בתי חולים

נניח שכל אחד מהגופים מחזיק נתונים על כל האזרחים (worst case scenario) ונסה להעריך עלות ליצור מסד נתונים מאוחד בענן. הנתונים ישמשו לצורך מלחמה בקורונה.

העננים שיבדקו הם של שלושת החברות הגדולות: מיקרוסופט, אמאזון וגוגל.

המדדים שיבחנו:

1. מחיר חודשי לתחזוקה עם יתירות מקסימלית וזמינות online
2. מחיר לטעינה חד פעמית לענן
3. מחיר ל Query/month
4. אבטחת נתונים
5. נוחות שימוש לצורך ביצוע המחקר

--	--	--

Summary

Cost

- The price to pay for the storage will be very similar between the cloud providers
- For the given numbers (of ~100TB) and described processing, a rough estimate is 10..20 thousand USD per month. This value depends on the actual services that will be used

Methods of ingest

In general, there are two variations that can impact cost and delay until data can be used:

- A one time upload of large file (e.g. the whole database of Rambam)
- A steady stream of small files, such as batch processing .

Given the small amount of data (9M records, less than 1MB each), I expect that the ingestion will be of a single file

Other factors to consider

- Data privacy: I am not an expert in this subject. Data must be anonymized and stored in approved facilities. For example, Azure Ireland might be plausible, and Azure US might not.
- Different data formats from each source. This can be a big problem.
- The format to store the data depends on what we want to do with it. We should have a clear direction before starting the ingest (e.g. will we use Spark / RDBMS / Impala)

The whole process is termed ETL: Extract Transform Load, referring to getting the data from current data sources, processing it and then storing in a new container.

The goal

- build ML models to identify people with higher risk to be infected by the COVID-19
- Perform (SQL) queries on the aggregated data

Data Engineering – ingesting the data

For each of the input databases, we need to export the DB into some format (e.g. parquet), upload to a blob storage or similar non-RDBMS format

From the BLOB storage, the data is available to SPARK / Impala /Hadoop

Relational DB might be inefficient at this scale (due to large latency of queries)

Working assumptions

Amount of data per citizen: 50 textual attributes + 30 BLOBs (such as scanned X-rays and handwritten notes)

Average BLOB size: 100KB

Number of operations required to generate the aggregated database:

Read operations: $N^2 * \text{num_db}$

Where N = number of citizens (e.g. 9.1M)

num_db = number of sources (e.g. 10)

Write operations: N

Computing power (Virtual machines): Not accounted for.

[1]

<https://www.cbs.gov.il/he/mediarelease/pages/2019/%D7%90%D7%95%D7%9B%D7%9C%D7%95%D7%A1%D7%99%D7%99%D7%AA-%D7%99%D7%A9%D7%A8%D7%90%D7%9C-%D7%91%D7%A4%D7%AA%D7%97%D7%94-%D7%A9%D7%9C-%D7%A9%D7%A0%D7%AA-2020.aspx>

תקציב תעסוקה 1

- ▶ שלב ראשון - ניתוח המצב הקיים:
 - ▶ הבנת התשתיות ובסיסי הנתונים הקיימים, מאפייניהם והדרישות לתאימות וקישוריות ביניהם.
 - ▶ הכנת מפת דרכים מפורטת – **תאום עם פרויקט ההכשרות בתב"ת**, אילו חלקים מכוח העבודה צפויים להיות מוחלפים או לעבור שינוי משמעותי בהקשר בינה מלאכותית בכלל ובהקשר לקורונה בפרט.
- ▶ **זמן משוער** - שלושה חודשים.
- ▶ שלב שני - שדרוג חומרות, סידור בסיסי נתונים, ביצוע אנליזה מפורטת לרמת המקצועות והמועצה מקומית/ עירייה:
 - ▶ הבנת המקצועות שידרשו "החלפה" של מרכיב
 - ▶ הבנת המוביליות בין המקצועות
 - ▶ הגדרת תכונות ויכולות מתאימות למעבר בין עבודות (המודל של דיילת והמלצה להסבה לעובדת סוציאלית).
- ▶ בניית יכולת של מתכנתים/ מהנדסי ומדעני נתונים - שתשרת את המהלך למשך שלוש שנים לפחות.

תקציב תעסוקה 2

- ▶ **זמן משוער** 9 חודשים.
- ▶ שלב שלישי – הערכת כמות האנשים שיעברו הסבה (מהניסיון בשבדיה - מדובר בסד"ג של מאות אלפים)
- ▶ שלב חמישי - הערכת כמות החברות ה"קבלניות" שיעסקו בביצוע ההסבות ועלות המאמץ (למשך עשור)

תוכנית עבודה: בתי חולים

▶ תיקון נתונים לפי תקן FHIR

▶ EMR – 65% מהנתונים (Electronic Medical Record)

▶ תיקון נתוני רדיולוגיה לפי תקן DICOM

▶ Digital Imaging and Communications in Medicine - DICOM

מערכת החינוך

▶ **חיבורים:**

▶ נתוני בריאות (תלמידים, מורים, אזורים, גילאים)

▶ קורסי מט"ח (לימודים לכיתות נמוכות)

▶ קורסים באנגלית מחו"ל

▶ לומדות מכון דוידסון והטכניון

▶ **תמיכה במורה – להוראה מרחוק**

▶ הערכת תלמידים

▶ שיעורי עזר

▶ תמיכה מערכתית

מערכת החינוך

חינוך לעולם עתיר בינה מלאכותית וביג דאטה

- טיפוח חשיבה ביקורתית
- שילוב ידע על מדעי הנתונים בתוכניות הלימודים
- טיפוח אוריינטציה טכנולוגית

חשיבה מחדש על תשתיות הנתונים בחינוך

- חיבור למערכות נתונים משיקות
- יצירת מתודות אתיות של איסוף נתונים פדגוגיים וחברתיים-רגשיים

הנעת אקו-סיסטם של פיתוח בינה מלאכותית בחינוך

- התגברות על מגבלות העברית והשוק הקטן
- יצירת מסגרות עידוד ליזמים
- הנעת פיתוחים ראשונים בתמיכה ממשלתית
- בחינת היתכנות וכדאיות פיתוח מערכת למידה אדפטיבית בעברית