

# **המחלקה להנדסת תוכנה**

## **פרויקט גמר – תשע"ח**

### **סוציומטרי מרובה קורסים**

### **Multi Courses Sociometric**

**מאת**

**ספיר אליהו, אביחי והב**

<b>תאריך:</b>	<b>מנחה אקדמי: דר' מרים אללוף, מר שי תבור אישור:</b>
<b>תאריך:</b>	<b>רכז הפרויקטים: מר אסף שפיינר אישור:</b>

מערכות ניהול הפרויקט:

#	מערכת	מיקום
1	מאגר קוד	<a href="https://github.com/AvihaiV/Multi-Courses-Sociometric">https://github.com/AvihaiV/Multi-Courses-Sociometric</a>
2	יומן	<a href="https://github.com/AvihaiV/Multi-Courses-Sociometric/wiki/Project-Diary">https://github.com/AvihaiV/Multi-Courses-Sociometric/wiki/Project-Diary</a>
5	סרטון גרסאת אלפא	<a href="https://youtu.be/xkFo32a7QP0">https://youtu.be/xkFo32a7QP0</a>

תקציר

- תוכן העניינים
- מילון מונחים, סימנים וקיצורים

## 1. מבוא

סוציומטריה או "הערכת עמיתים", היא שיטה כמותית למדידת יחסים חברתיים. כוחה העיקרי של השיטה נובעת מהעבודה שהציון שמקבל אדם מורכב משקלול מספר הערכות. בשיטת "דירוג עמיתים" מתבקש כל אדם לדרג את חבריו לקבוצה על פי מדדים מסוימים. שיטה זו קלה ומהירה לניתוח ומסייעת לאתר את התכונות החיוביות של האדם.

מטרת פרויקט זה הינה לספק כלי לשמירת נתונים סוציומטריים על חילים במהלך שירותם הצבאי, לפי הקורסים אותם ביצעו, ובעזרת נתונים אלה לאפשר התאמה מרבית לתפקיד הצבאי. לצורך בניית כלי זה נשתמש ונחקור את מסד הנתונים Neo4j. מסד הנתונים מציג את המידע בצורת גרף של הקשרים עבור כל ישות (חייל) ובנוסף מציג מידע על כל ישות. המערכת יוצרת פרופיל עבור כל משתמש (חייל) והקורסים שהוא ביצע במהלך שירותו הצבאי, עבור כל קורס יתקבלו הציונים הסוציומטריים שקיבל החייל ע"י חבריו לקורס. במהלך עבודתנו נחקור לעומק את מסד הנתונים Neo4j, נבין כיצד הוא עובד ודרך פעולתו. על מנת להנגיש את המערכת ניצור אפליקציה שאליה ניתן להזין את הנתונים הסוציומטריים, האפליקציה תתממשק עם מסד הנתונים Neo4j.

עד כה, עסקנו בהיכרות עם מסד הנתונים Neo4j. השפה איתה עובד המסד נקראת cypher. את הישויות והקשרים בגרף יוצרים בעזרת שאילתות מתאימות בשפה זו והמסד יוצר גרף ויזואלי שמסייע להבין בצורה טובה יותר את היחסים בין הישויות.

בנוסף בעזרת שאילתות מורכבות יותר ניתן להפיק מידע רלוונטי מהגרף בהתאם לרצוי. יצרנו במסד הנתונים גרף לדוגמא שמדמה חילים, קורסים אותם ביצעו והציונים אותם קיבלו מחבריהם לקורס (ראה פרק מסכים).

בנינו גרסה ראשונית לאפליקציית המשתמש. באפליקציה ישנו מסך input בו ניתן להזין עבור חייל בקורס מסוים את הציונים אותם קיבל מחבריו. האפליקציה משקללת את הציון של החייל באותו קורס. באפליקציה ישנו גם מסך output בו ניתן יהיה לשלוח מידע ממסד הנתונים בהתאם לסינון הרצוי (לפי קורס, חייל או ציון).

בשלב הבא נממשק את האפליקציה עם מסד הנתונים. הנתונים שיוכנסו לאפליקציה יועברו למסד הנתונים ושם ייווצר הגרף שיכיל את כל החיילים, הקורסים שעשו והציונים שקיבלו. לאחר מכן בעזרת שאילתות מתאימות יישלף המידע הרצוי לפי דרישות הסינון במסך ה output ויוצגו התוצאות.

## 2. תיאור הבעיה

### דרישות ואפיון הבעיה

במהלך השירות הצבאי חייל עובר מגוון קורסים אשר בסיומם ניתנים לו ציונים סוציומטריים על ידי חבריו ומפקדיו. בהתאם לציונים אלה נקבע התפקיד אותו יבצע במהלך השירות הצבאי. נכון להיום לא קיים כלי לאחסון ושמירה של נתונים וציונים מקורסים שונים, כך שניתן יהיה להצליב את המידע ביניהם ולהוציא נתונים רלוונטיים. דבר זה גורם לחוסר יכולת של התאמה מרבית של חייל לתפקיד אותו יבצע כיוון שתפקיד מסוים עלול לדרוש יכולות אשר נמדדות בקורסים שונים והצלבת המידע אינה פשוטה ומידית כיום.

מטרת הפרויקט היא לבנות מודל גרפי בעזרת מסד הנתונים Neo4j שבעזרתו ניתן יהיה לשמור ולעבד את הציונים הסוציומטריים עבור כל חייל. בעזרת מסד הנתונים Neo4j ניצור מערכת המאגדת עבור כל חייל את כל הקורסים אותם ביצע ואת הציונים שקיבל וכך נוכל בעזרת שאילתות מתאימות להצליב מידע ולהוציא נתונים רלוונטיים. בעזרת נתונים אלה תתאפשר התאמת תפקיד אופטימלית לחייל.

### הבעיה מבחינת הנדסת תוכנה

מסד הנתונים Neo4j הינו מסד נתונים מסוג NoSQL גרפי. מסד NoSQL נותן פתרון אחסון וגישה למידע שאינו במבנה של טבלאות.

פתרון לבעיית הציונים הסוציומטריים במסד נתונים רלציוני כלומר מבנה של טבלאות הינו מסורבל ולא ברור שכן הצלבת מידע בין מספר רב של טבלאות על מנת להוציא נתונים על חייל הינה מסורבלת וארוכה. על מנת לתת פתרון לבעיה הנ"ל במסד נתונים מסוג זה נצטרך ליצור K טבלאות של  $N \times N$  כאשר N מייצג את מספר החיילים ו K את מספר הקורסים. על מנת לנתח את התוצאות ולהגיע למסקנות הרצויות נצטרך לעבור על  $N \times K$  וקטורים (ראה תרשים בעמוד הבא). בהתבסס על כך שמספר החיילים הינו עצום, מעבר על מספר כזה של וקטורים הוא בלתי אפשרי.

דרך זו הינה מסובכת ולא אידאלית, והצלבת המידע בין הטבלאות הינה מסורבלת ולא יעילה.

K2

	N1	N2	...	Nn
N1				
N2				
...				
Nn				

K1

	N1	N2	...	Nn
N1				
N2				
...				
Nn				

Kn

	N1	N2	...	Nn
N1				
N2				
...				
Nn				

.....

K3

	N1	N2	...	Nn
N1				
N2				
...				
Nn				

ההצלבה בין הטבלאות במטרה לשלוף נתונים על חייל מסוים היא מסובכת ולא יעילה. לעומת זאת נראה כי בשימוש ב-Neo4j מסד נתונים גרפי הפתרון יהיה פשוט יותר שכן הקשרים בין החיילים (הקדקודים בגרף) יותר מוחשיים והוצאת הנתונים עתידה להיות יותר פשוטה.

בפרויקט זה נרצה לבחון אם אכן השימוש ב-Neo4j אכן יפשט את הפתרון של בעיה זו.

### 3. תיאור הפתרון

מטרת הפרויקט הינה לתת מענה לבעיית ניתוח הציונים הסוציומטריים בצורה הפשוטה והיעילה ביותר. מענה זה יינתן על ידי הקניית יכולות חיפוש והצלבת מידע המתבססים על מסד נתונים גרפי – Neo4j. מערכת זו תאפשר חיפוש פשוט וויזואלי. המערכת תציג קשרים בין הישויות השונות ותתמקד בתוצאות החיפוש הרצויות.

#### מהי המערכת:

המערכת תכלול אפליקציית WEB בה המשתמש יוכל להזין נתונים על חייל לפי כל הקורסים אותם ביצע והציונים אותם קיבל על ידי עמיתיו לקורס. אפליקציית ה-WEB תכלול 5 עמודים מרכזיים: עמוד הבית בו יינתן הסבר אודות האפליקציה, דף הזנת נתונים בו תינתן האפשרות להזין חייל למערכת וציונים עבור קורס מסוים, דף שליפת נתונים בו ניתן לקבל מידע לפי קריטריונים רצויים (חייל, קורס או ציון), דף צור קשר ודף אודות מפתחי האפליקציה. באמצעות אפליקציית WEB שכזו אנו ניצור חיבור בין האפליקציה למסד הנתונים של Neo4j שם ייווצר הגרף ובו הנתונים על כל החיילים והקורסים, ומשם תתבצע גם שליפת המידע.

#### תהליכים ונתוני המערכת:

למערכת ישנם שני תהליכי עבודה:

1. הכנסת הנתונים ויצירת הגרף – המשתמש יכניס את הנתונים דרך אפליקציית ה-WEB הנתונים יועברו למסד הנתונים Neo4j ושם ייווצר גרף אשר יכיל את כל החיילים, הקורסים אותם ביצעו והציונים שאותם קבלו על ידי העמיתים לקורס. הקדקודים בגרף יהיו החיילים והקשרים בניהם יהיו הקורסים בהם השתתפו הציונים אותם קיבלו.
2. שליפת נתונים והחזרת תשובה – המשתמש יזין באפליקציית ה-WEB את הנתונים אותם רוצה לראות לפי סינון של: חייל, קורס או ציון, בהתאם לסינון תישלח שאילתה מתאימה למסד הנתונים ותישלף תשובה מהגרף, תשובה זו תוחזר למשתמש דרך האפליקציה.

### תיאור הפתרון המוצע:

בפתרון לבעיה זו השתמשנו במסד נתונים Neo4j. ההשערה היא שהשימוש במסד נתונים מסוג NoSQL יהיה פשוט יותר מבחינת הכנסת ושליפת הנתונים. השערה זו מתבססת על העובדה שעל מנת לתת פתרון לבעיה זו במסד נתונים רלציוני דורש מעבר על מספר כמעט אינסופי של וקטורים. נניח שמספר החיילים הוא  $N$  עבור כל קורס ניצור טבלה של  $N \times N$  כאשר כל חייל מדרג את כל חבריו לקורס על פי קריטריון מסוים. בהנחה שמספר הקורסים הינו  $K$  נצטרך לעבור על  $K \times N$  וקטורים על מנת להשיג תוצאה רצויה. מעבר על כמות כזו של וקטורים הוא ארוך ולא יעיל.

### תיאור הכלים המשמשים לפתרון:

- על מנת לפתור בעיה זו השתמשנו בשני כלים עיקריים:
1. אפליקציית WEB – האפליקציה כתובה בשפת Java Script
  2. מסד נתונים גרפי
  3. Neo4j – מסד נתונים מסוג NoSQL כלומר מסד נתונים שאינו מתבסס על טבלאות אלא על גרף.

#### 4. סקירת עבודות דומות בספרות והשוואה

בשוק כיום קיימות מערכות שונות לניתוח מערכות שאלונים. מערכות אלה רלציוניות כלומר בנויות מבסיס נתונים הבנוי מטבלאות כאשר כל טבלה מכילה מידע על רשות מסויימת. הקשרים בין הרשומות בטבלאות נעשה באמצעות שדה מיוחד הנקרא מפתח. למרות שקיימות מערכות כאלה נרצה לבדוק האם ניתן לבדוק היתכנות שלהם במסד נתונים אחר שהוא לא רלציוני כלומר מסד נתונים NoSQL. בפרויקט שלנו נתמקד במסד הנתונים Neo4j.

קיום בשוק ישנם 4 סוגים של מסדי נתונים מסוג NoSQL:

1. Document Databases – בסיס נתונים זה מצמיד מפתח עם מבנה מורכב של מידע הנקרא מסמך. מסמך יכול להכיל צמדים מורכבים של מפתח-מערך, מפתח-ערך, מפתח-מסמך.  
לדוג' – MongoDB.
  2. Key-Value Stores – משמש לבסיס נתונים פשוט ומהיר התומך בצמדי מפתח-ערך בלבד.  
לדוג' – Riak , Redis , Voldemort.
  3. Wide-Column Stores – בסיס נתונים שמבוסס על עמודות במקום שורות. משתמשים בו בעיקר לשאילתות על מערכי מידע גדולים במיוחד.  
לדוג' – Hbase , Cassandra.
  4. Graph Stores – משמש לשמירת מידע הקשור לרשתות וקשרים חברתיים.  
לדוג' – Hypergraph DB , Neo4j.
- סוגי מסדים אלה נבדלים ביכולת להתמודד עם כמות מידע שגדלה וביכולת להתמודד עם המורכבות של המידע. Data size Vs. Data Complexity.

במקרה של מסד נתונים מסוג Graph Stores חשוב לציין שכמות המידע שאיתו הוא יכול להתמודד היא הקטנה ביותר מבין כל הסוגים, אך הוא עדיין יכול ליישם מיליוני קדקודים וקשרים. מצד שני ההתמודדות של מסד נתונים מסוג זה עם מידע מורכב במהירות רבה רלוונטי ביותר לפרויקט זה מכיוון שכמות המידע שעלינו להתמודד איתו במקרה זה הינו סופי – מספר חיילי צה"ל.



## 5. נספחים

ספרות, תרשימים נוספים, תכנון הפרויקט, טבלת ניהול סיכונים, טבלת דרישות (URD),

א. רשימת ספרות \ ביבליוגרפיה

<https://www.slideshare.net/jexp/intro-to-neo4j-presentation>

<https://www.infoq.com/articles/graph-nosql-neo4j>

<https://neo4j.com/developer/graph-db-vs-rdbms/>

<https://dl.acm.org/citation.cfm?id=1900067>

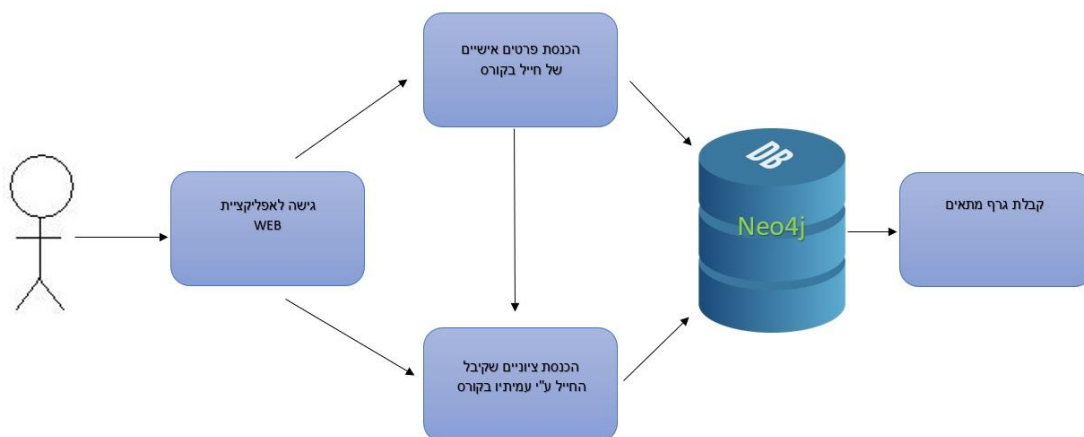
[https://books.google.co.il/books?hl=en&lr=&id=vetaBAAQBAJ&oi=fnd&pg=PT11&dq=neo4j&ots=dsljn143-9&sig=vBz2n1T5Q-xA3OJL8AR-QpZgp7U&redir\\_esc=y#v=onepage&q=neo4j&f=false](https://books.google.co.il/books?hl=en&lr=&id=vetaBAAQBAJ&oi=fnd&pg=PT11&dq=neo4j&ots=dsljn143-9&sig=vBz2n1T5Q-xA3OJL8AR-QpZgp7U&redir_esc=y#v=onepage&q=neo4j&f=false)

<https://pdfs.semanticscholar.org/322a/6e1f464330751dea2eb6beecac24466322ad.pdf>

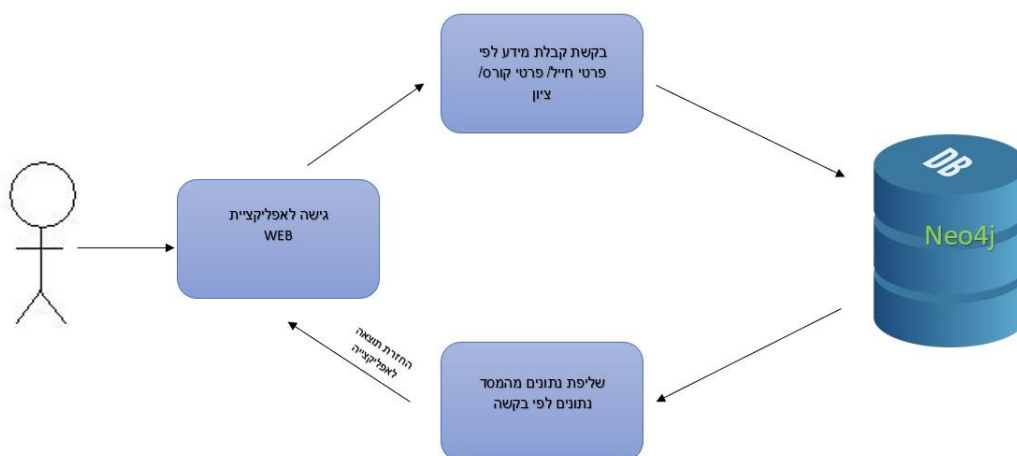
<https://dl.acm.org/citation.cfm?id=2457351>

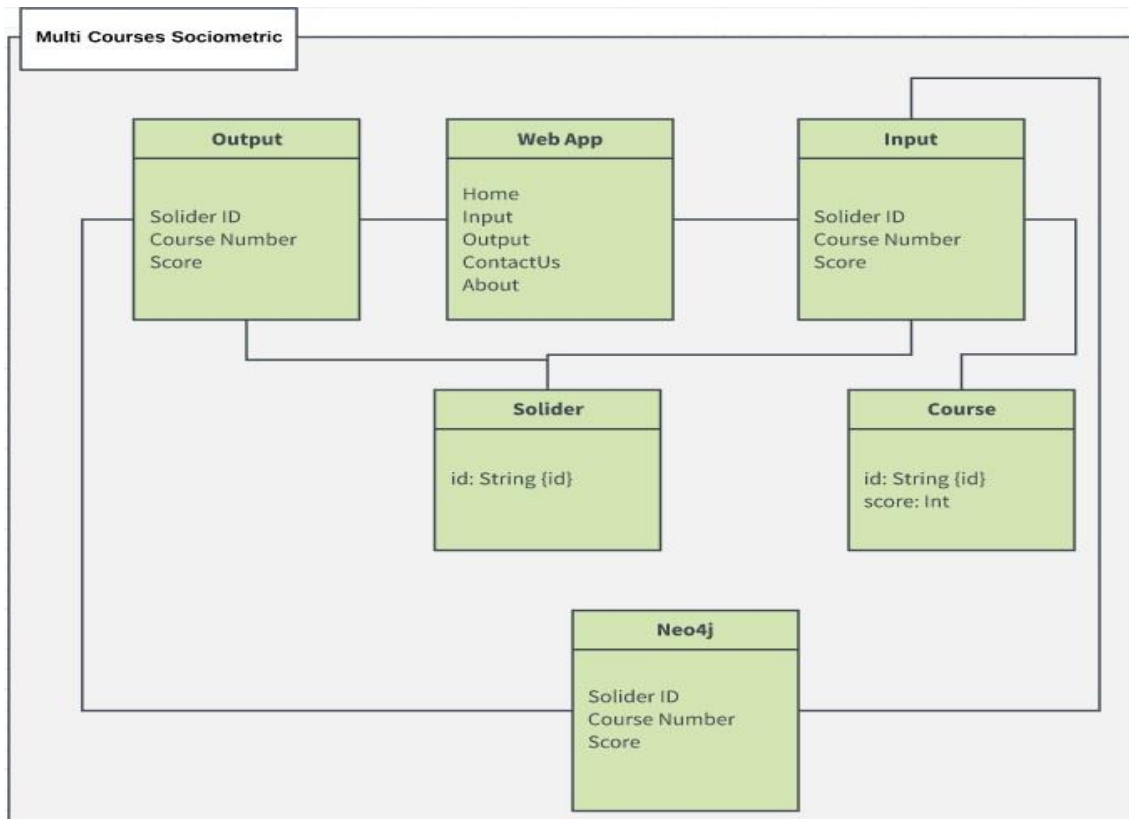
## ב. תרשימים וטבלאות

### תרשים הכנסת מידע



### תרשים שליפת מידע

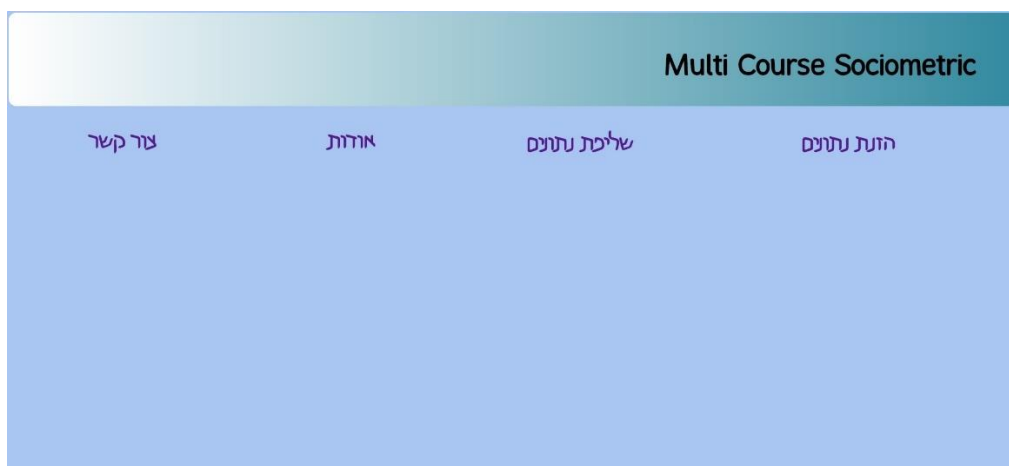




## מסכים:

### מסך 1: דף הבית

גרסה ראשונית של דף הבית לאפליקציית המשתמש



### מסך 2: מסך הכנסת הנתונים על החיילים

גרסה ראשונית. במסך זה יוזנו הציונים הסוציומטריים שניתנו לחייל מסוים ע"י עמיתיו והציון המשוקלל הסופי בקורס. ניתן יהיה להוסיף ציונים לפי מספר המשתתפים בקורס.



### מסך 3: מסך התוצאות:

גרסה ראשונית. המראיין רואה לפניו את הציונים המשוקללים שקיבל חייל מסוים בכל הקורסים אותם ביצע.

מספר הציונים יהיה בהתאם למספר הקורסים אותם ביצע החייל



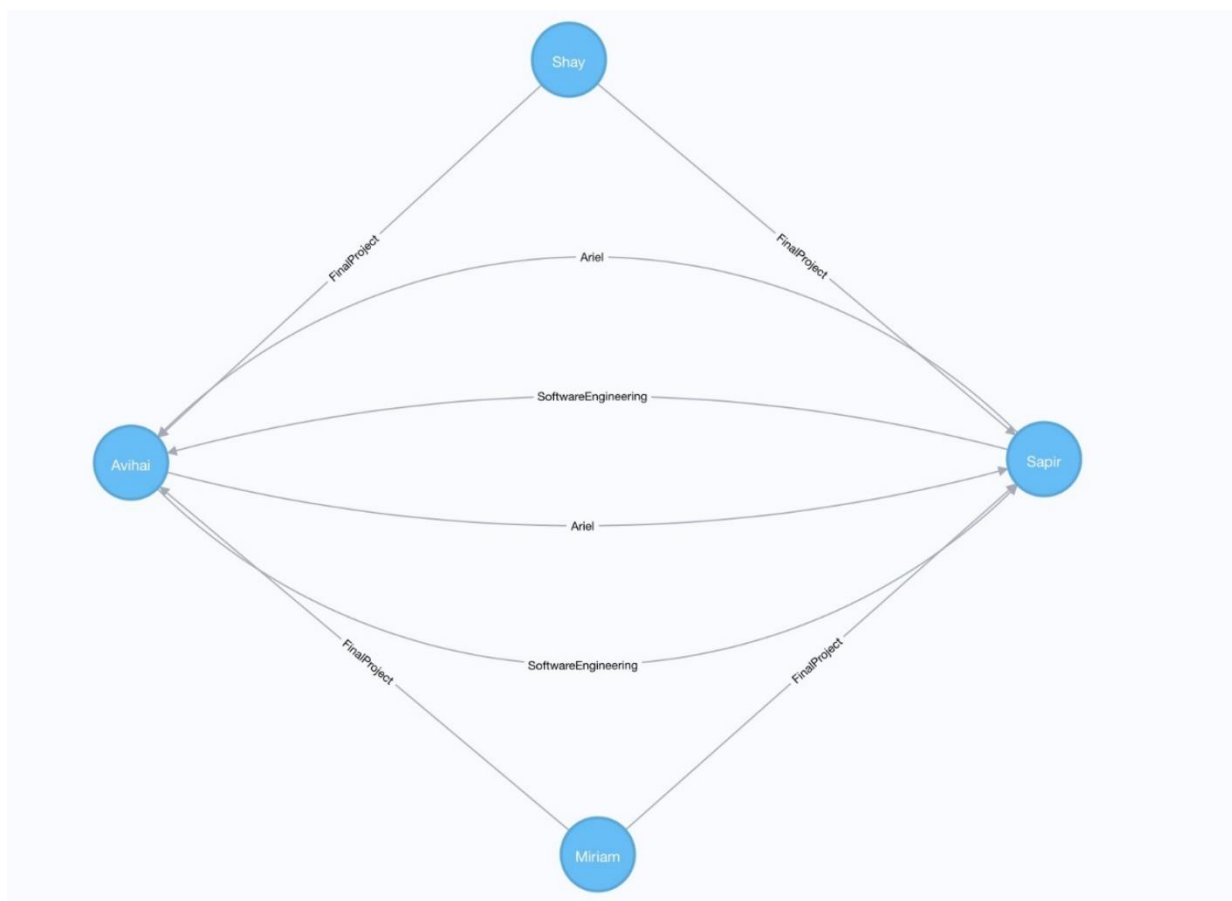
### מסך 4: טבלאות הנתונים ב- Neo4j

הנתונים במסד זה נכנסים לטבלאות ומתורגמים לגרף.

se	av	st	ma
<pre>{   "born": "1991",   "name": "Sapir",   "ID": "308058262" }</pre>	<pre>{   "born": "1989",   "name": "Avihai",   "ID": "201655800" }</pre>	<pre>{   "born": "2000",   "name": "Shay",   "ID": "1234567" }</pre>	<pre>{   "born": "2000",   "name": "Miriam",   "ID": "9876543" }</pre>

מסך 5: הצגה גרפית של נתונים במסד הנתונים Neo4j

דוגמא לנראות גרף ב- Neo4j. כל קדקוד בגרף הינו חייל, הקשרים בין שני קדקודים הם הקורסים אותם בצעו שני החיילים והציון שניתן באותם קורסים (בלחיצה על הקשר מופיע הציון).



ג. תכנון הפרויקט

פגישה ראשונית עם המנחים – הצגת הרעיון הכללי.	16.07.17
הכרות ראשונית עם המסד נתונים Neo4j	10.08.17
פגישה עם המנחים – התמקדות ספציפית בנושאי הפרויקט.	27.09.17
יצירת טופס התנעה.	28.09.17
פגישה עם המנחים, בעיות שונות שעלו, קריאת מאמרים וטופס הצעה ראשוני.	08.11.17
יצירת טופס הצעה.	13.11.17
הגשת טופס הצעה.	19.11.17
יצירת קבצי HTML בסיסיים, דף הבית, דף הכנסת ושליפת מידע ודפי אודות וצור קשר	11.12.17
המשך יצירת דפי HTML, יצירת קבצי CSS והתחלת עיצוב האתר.	20.12.17
פגישה עם המנחים לגבי טופס אלפא, שיחה לגבי מורכבות יצירת מסד נתונים שכזה, הכנסה ושליפה של מידע מ-Neo4j	07.02.18
יצירת טופס אלפא, עדכון האתר והקבצים.	11.02.18
המשך עבודה על טופס אלפא	20.02.18
המשך עבודה על טופס אלפא, יצירת קבצי JS עבור האתר	04.03.18
סיום עבודה על טופס אלפא, צילום סרטון הצגה של הפרויקט.	07.03.18

ד. טבלת סיכונים

#	הסיכון	חומרה	מענה אפשרי
1	חוסר ידע מקצועי	גבוהה	חקר באינטרנט, סיוע מצד המנחים ובוגרי המכללה
2	חוסר זמן	גבוהה	תכנון מראש, יצירת יומן וביצוע משימות לפי זמנים.
3	חוסר מידע מצד הצבא	גבוהה	חיפוש נתונים מחיילים שכרגע משרתים בצבא.
4	המוצר אינו מאובטח	גבוהה	יצירת דף כניסה עם מספר אישי של חייל.
5	טעויות בקוד ברמת האפליקציה	בינונית	פנייה לעזרה מהמנחים וחיפוש באינטרנט – Stack overflow. להשתמש בבקרת קוד.
6	מסד הנתונים מתעדכן ולא תומך בגרסה ישנה	בינונית	לא לבצע עדכון גרסה.
7	טעויות בקוד ברמת המסד נתונים	בינונית	פנייה לעזרה מהמנחים וחיפוש באינטרנט – Stack overflow. להשתמש בבקרת קוד
8	התמודדות עם שאלות מסובכות	בינונית	לנסות לפשט או לשנות את השאלתה.
9	התוצאה הסופית אינה מספקת	נמוכה	מעקב תמידי עם המנחים על התקדמות הפרויקט.
10	אחד מחברי הצוות לא יכול להמשיך	נמוכה	הגדרת תפקידים מראש, הצבת מטרות מראש ויכולת עמידה בהן.





ה. רשימת טבלת דרישות  
**טבלת דרישות (User Requirement Document)**

מס' דרישה	תיאור
1	דרישה למסד נתונים שיכול להכיל מידע רב.
2	האפליקציה תיתמך בכל סוגי הדפדפנים.
3	אפשרות לעדכון הגרף ללא צורך יצירה מחדש.
4	אפליקציה נוחה למשתמש.
5	מסד נתונים שמאפשר שליפת מידע פשוטה ומהירה.
6	דרישה למסד נתונים שיכול להכיל מידע מורכב.