

**המרכז האקדמי למשפט ולעסקים**  
**הפקולטה למערכות מידע ומדעי המחשב**

# **ניתוח וניטור מידע כוזב ברשת החברתית Twitter**

## **בשעת חירום**

**הצעת מחקר מורחבת לפרוייקט מחקרי**

**יקיר אטיאס**

**ת.ז- 318602208**

**מנגשה אטלאי**

**ת.ז- 324453737**

**מנחה: ד"ר אבי יוסיפוף**

**שם הקורס: סדנה לפרוייקטים**

**מס' הקורס: 6172**

**גורם מאשר: ד"ר אבי יוסיפוף**

**05.03.2023**

**י"ב באדר, ה'תשפ"ג**

## תוכן עניינים

2	מסמך אישור
3	תודות
3	איורים, טבלאות וגרפים
3	תקציר
4	Abstract
4	מבוא
6	מטרות ויעדים
8	סקירת ספרות
11	שיטות
14	תוצאות
21	דיון
23	סיכום ומסקנות
24	ביבליוגרפיה
25	פוסטר הפרוייקט

## מסמך אישור

קבלת אישור לביצוע פרוייקט גמר יקיר ומנגשה

Yakir Attias

yakirattias7@gmail.com

Avi Yosipof

avi.yosipof

אחלן אבי, אני ומנגשה נשמח לקבל ממך אישור רשמי לביצוע פרוייקט הגמר " ניתוח וניסור מידע כוזב ברשת החברתית טוויטר בפנימי חירום".

יוס, ג' 27 בדצמ, 11:52 (לפני 23 שעות)

Avi Yosipof

אני

שלום,  
מאשר.  
ד"ר אבי יוסיפוף  
...

Dr. Avi Yosipof

Senior Lecturer

Dean: The Faculty of Information Systems and Computer Science

Head of the Information Systems Program (BSc)

College of Law & Business, Ramat-Gan, Israel

2

## תודות

אנו רוצים להביע את תודתנו לד"ר אבי יוסיפוף, דיקן הפקולטה למערכות מידע ומדעי המחשב והמנחה של פרויקט גמר זה. ההנחיה, התמיכה והמומחיות שלו תרמו לנו המון במחקר. לאורך הפרויקט, אבי היה מקור מתמיד של עידוד וידע, שתמיד דוחף אותנו למצוינות. תודה למרצה אור אלרואי, שסיועו תרם למחקר רבות. התובנות והמומחיות שלו העשירו את ההבנה שלנו בנושאים שונים ועזרו לעצב את הניתוח שלנו.

## איורים, טבלאות וגרפים

11	איור 1 : תיאור גרפי של המחקר
11	איור 2 : דוגמא לניסוח שאילתה באנגלית
12	איור 3 : הצצה למערך הנתונים של הפרויקט
13	איור 4 : תרשים קווים- מספר הציוצים ביום עבור כל שפה
13	איור 5 : מספר הציוצים עבור כל שפה : ההאנגלית הינה בעלת נתח הציוצים הגדול ביותר (78.7%)
13	איור 6 : אלגוריתמים ומודלים שהשתמשנו בפרויקט
14	איור 7 : סרוק על מנת להכנס לאתר
15	איור 8 : מפת חום עבור מתאם הקורלציה בין השפות
17	איור 9 : ניתוח רגישות ממוצע לכל השפות ביחד לפי תאריכים
18	איור 10 : פלט ארימה עבור השפה האנגלית
21	איור 11 : ענן המילים המדוברות ביותר בשפה האנגלית
26	איור 12 : פוסטר הפרויקט

## תקציר

מחקר זה חקר את התפשטות המידע השגוי של COVID-19 בטוויטר בהתמקדות בתיאוריית הקונספירציה 5G. מערך הנתונים שלנו כלל כ-441,000 ציוצים ברוב השפות המדוברות באירופה.

השתמשנו בכריית נתונים בעזרת שימוש בapi של טוויטר כדי לאחזר ציוצים, ספריות של Python למניפולציה של נתונים וטכניקות עיבוד שפה טבעית (NLP) לניקוי טקסט וניתוח סנטימנטים. יישמנו את מודל BERT לניתוח סנטימנטים ואת המודלים של ARIMA ו-SARIMAX לחיזוי סדרות זמן.

הממצאים שלנו חשפו מתאמים משמעותיים בנפחי הציוצים בשפות שונות, מה שמצביע על קשר עולמי בתגובה למגיפה. ניתוח סנטימנטים הראה שונות בין שפות, וניתוח אוטו קורלציה הצביע על התמדה בדיונים בתיאוריות קונספירציה לאורך זמן. שיא עוצמת התדירות של הציוצים התרחשו לעתים קרובות בקנה אחד עם אירועים או הודעות מרכזיות הקשורות למגפה בשפות שונות.

המחקר תורם להבנתנו כיצד מידע שגוי מתפשט במהלך משברים גלובליים, ומוצא דרכים ליצירת אסטרטגיות לניטור ומניעת מידע שגוי. הוא מדגיש את תפקידו של מדעי הנתונים בהתמודדות עם אתגרים חברתיים ומספק כיוונים למחקר עתידי כדי לחדד את השיטות ולהרחיב את מקורות הנתונים.

## Abstract

This research investigated the spread of COVID-19 misinformation on Twitter. Our dataset comprised of approximately 441,000 tweets in multiple languages, focusing on conspiracy theories related to the pandemic.

We used data mining to retrieve tweets, Python's pandas library for data manipulation, and Natural Language Processing (NLP) techniques for text cleaning and sentiment analysis. We applied the BERT model for sentiment analysis and the ARIMA and SARIMAX models for time series forecasting.

Our findings revealed significant correlations in tweet volumes across different languages, suggesting a global interconnectedness in response to the pandemic. Sentiment analysis showed variations across languages, and autocorrelation analysis indicated persistence in conspiracy theory discussions over time. Peaks in tweet volume often coincided with key pandemic-related events or announcements.

The research contributes to our understanding of how misinformation spreads during global crises, informing strategies for monitoring and counteracting misinformation. It highlights the role of data science in addressing societal challenges and provides directions for future research to refine the methods and expand data sources.

## מבוא

### נושא המחקר

מצבי חירום, כמו אסונות טבע ומשברי בריאות הציבור, עלולים להוביל לעתים קרובות להפצת מידע כוזב ברשתות החברתיות. לכך יכולות להיות השלכות חמורות, שכן אנשים מסתמכים על מידע זה כדי לקבל החלטות שיכולות להשפיע על בטיחותם ועל בטיחותם של אחרים.

משמעות המושג 'מידע כוזב' הוא הפצה מכוונת של מידע שקרי ומטעה שנועדה ליצור תמונת מציאות מעוותת.

'תאוריית קונספירציה' היא ניסיון להסביר אירוע או תופעה באופן חתרני ובניגוד להסבר המקובל, מתוך תפיסה שגורמים בעלי כוח מסתירים מהציבור הרחב את האמת, ומונעים את חשיפתה בשל אינטרסים שונים.

טוויטר היא פלטפורמת מדיה חברתית פופולרית המשמשת לעתים קרובות לשיתוף מידע בשעת חירום, והיא הפכה למקור מידע מרכזי עבור אנשים רבים, הפלטפורמה מאפשרת העברת מידע ממוקד אל האוכלוסייה בתחומי מגוונים, ממידע על סיכונים והתמודדות עם מצבי חירום ועד מידע הקשור לרשות המקומית ולשירותים חשובים אחרים לאזרח בשגרה ובחירום.

עם זאת, התפשטות מהירה של מידע כוזב בטוויטר בזמן חירום עלולה להוביל לבלבול, פאניקה ומידע מוטעה, אחת מהבעיות המרכזיות הקיימות בעידן הרשתות החברתיות זה מידע כוזב ותיאוריות קונספירציה שמתפשטות בשיח ברשת בשניות ובמקרה חירום יכולות לגרום אף להחמרה של המצב.

במהלך מגפת הקורונה המדינות השונות לא רק נלחמו במגפה אך גם במידע הכוזב שהופץ ברשת. יושב ראש ארגון הבריאות העולמי העניק במהלך מגפת הקורונה אף את השם INFODEMIC, כלומר מגפה של מידע כוזב, לכן חשוב לנתח ולנטר את התפשטות המידע הכוזב בטוויטר בזמן חירום על מנת לזהות ולמתן את השפעתו. (Zarocostas, 2020).

הצעת מחקר זו מתארת תוכנית לנתח ולנטר מידע כוזב בטוויטר בזמן חירום, במטרה להבין את עומק הדיון והפצה של מידע כוזב בנושא הקורונה במדינות אירופה וישראל.

### מהות הבעיה

אחת הבעיות העיקריות שנגרמות מהפצת תיאוריות קונספירציה בטוויטר בזמן חירום היא שהן עלולות לערער את האמינות של מקורות מידע רשמיים.

כאשר אנשים מתמודדים עם משבר, הם מסתמכים על מידע מדויק ממקורות מהימנים כדי לקבל החלטות שיכולות להשפיע על ביטחונם ועל ביטחונם של אחרים. עם זאת, ריבוי תיאוריות הקונספירציה בטוויטר עשוי להקשות על אנשים להבחין בין מידע מדויק ולא מדויק, מה שעלול להוביל לבלבול וחוסר אמון במקורות מידע רשמיים.

בעיה נוספת הנגרמת מהפצת תיאוריות קונספירציה בטוויטר בזמן חירום היא שהן יכולות להסיט את תשומת הלב והמשאבים הרחק מבעיות דחופות יותר. לדוגמה, אם אנשים מתמקדים בהפצת תיאוריות קונספירציה, הם עשויים להיות פחות מסוגלים להתמקד במשימות המיידיות והמעשיות יותר של תגובה למצב החירום.

לכך עלולות להיות השלכות חמורות, מכיוון שהוא עלול להפריע ליכולתם של מגיבים חירום וארגונים אחרים לנהל את המשבר בצורה יעילה.

לבסוף, להפצה של תיאוריות קונספירציה בטוויטר בזמן חירום יכולה להיות גם השפעות חברתיות ופסיכולוגיות שליליות. למשל, זה יכול להוביל להגברת החרדה והלחץ בקרב אנשים שאולי כבר מרגישים פגיעים ואי ודאות לנוכח משבר. זה גם יכול לתרום לשחיקת הלכידות החברתית, מכיוון שאנשים עלולים להפוך לחוסר אמון זה בזה ומפולגים יותר באמונותיהם ובמעשיהם.

להפצה של תיאוריות קונספירציה בטוויטר בזמן חירום יכולה להיות השלכות חמורות, הן עלולות לערער את האמינות של מקורות מידע רשמיים, להסיט את תשומת הלב והמשאבים מהנושאים הדחופים יותר, ולהיות בעלות השפעות חברתיות ופסיכולוגיות שליליות. לכן חשוב להבין את הגורמים התורמים להפצת תיאוריות קונספירציה בפלטפורמה, ולפתח אסטרטגיות לצמצום השפעתן.

## רציונל המחקר

הרציונל של המחקר הוא לשפר את ההבנה שלנו לגבי התפתחות הדיון בתיאוריות קונספירציה שונות הקשורות ל-COVID-19, ולפתח אסטרטגיות למניעת או הגבלת התפשטותו על מנת להבטיח את הדיוק והאמינות של המידע המשותף בפלטפורמה.

מחקר זה חשוב מכיוון שהוא יכול להשפיע לטובה על הבטיחות והרווחה של אנשים וקהילות, והוא יכול לתת מידע לפיתוח מדיניות ופרקטיקות הקשורות לשימוש אחראי במדיה חברתית בזמן חירום, ובכך, להציל חיים. מחקר זה מתמקד באיסוף מידע משפות שונות כגון מדינות אירופה וניתוחם.

## חשיבות המחקר

מחקר זה לראשונה בודק את השיח בנושא מידע כוזב וקונספירציה בשפות שונות בטוויטר. מחקר זה חשוב מנקודת מבט חברתית ומדיניות, תוצאותיו יכולות להיות שימושיות עבור קובעי מדיניות, סוכנויות לניהול חירום ופלטפורמות מדיה חברתית בפיתוח אסטרטגיות להבטחת הדיוק והאמינות של המידע המשותף בטוויטר בזמן חירום.

על ידי ניתוח וניטור מידע כוזב בפלטפורמה, ניתן לפתח אסטרטגיות למניעת או הגבלת התפשטותו, שיכולות לסייע להבטיח את הדיוק והאמינות של המידע המשותף בטוויטר בשעת חירום, מה שיכול לעזור לקדם את השימוש האחראי במדיה החברתית ולהגן על האינטרס הציבורי. כיון שהעולם עובר גלובליזציה, כלומר אירוע במדינה מסוימת משפיעים כמעט בטווח המיידי במדינה אחרת, ניתן לראות את הקורונה כדוגמה מצוינת לעניין זה, חשוב לנתח את השיח ברשתות ובמיוחד בטוויטר שמהווה כר פורה להבעת דעה גם בשפות שונות על מנת להבין איך אוכלוסייה עם תרבויות וסט אמונות שונות מתנהגות במצבי חירום.

## **מטרות ויעדים**

### מטרות המחקר

מטרת המחקר היא לנתח את השיח בנושא מידע כוזב והגורמים התורמים להפצתו בפלטפורמה, להבין את עומק הדיון בזמן המשבר, ולפתח אסטרטגיות לניטור והפחתת השפעתו בכמה שיותר סוגי אוכלוסיות ומדינות.

על ידי כך, ניתן לשפר את הדיוק והאמינות של המידע המשותף בטוויטר בשעת חירום, מה שיכול להיות בעל השפעה חיובית על בטיחותם ורווחתם של אנשים וקהילות, ואפילו להציל חיים בזמן משברים עולמיים.

המטרות העיקריות של הצעת מחקר זו הן:

- א. לבצע השוואה בין שפות שונות בנוגע למידע כוזב שמשותף בטוויטר בשעת חירום.
- ב. לזהות את עומק השיח החברתי בטוויטר על תאוריות קונספירציה שונות בזמן מגפת הקורונה כדי להבין את ההשפעה שיש להן בזמן מצבי חירום דומים אחרים.
- ג. לנתח את תדירות השיח לאורך המגפה בכל שפה.

1. שלב ניסוח השאלות

מטרת השלב הזה היא לנסח שאלות בשפות: אנגלית, גרמנית, איטלקית, ספרדית, פולנית, רומנית, הולנדית, יוונית, צ'כית.  
כדי לאסוף ציוצים הקשורים לתיאוריות קונספירציה שונות על COVID-19. שלב זה יעזור לחוקרים לשלוף ציוצים רלוונטיים המכילים מילות מפתח ספציפיות הקשורות לתיאוריות הקונספירציה הנחקרות, ועל בסיס זה לבנות מסד נתונים של הציוצים הרלוונטיים למחקר.

2. שלב הכנת מסד הנתונים

המטרה של שלב זה בתהליך כריית הנתונים היא להכין מסד נתונים שיכלול את הציוצים שנאספו בשלב הקודם, בצירוף תאריך, שפת הציוצ, סוג קונספירציה.  
מטרת שלב זה הוא לאסוף ולאחסן כמות גדולה של נתונים המתייחסים לתיאוריות קונספירציה בזמן משבר הקורונה. הנתונים שנאספו ישמשו בשלבי הניתוח והמודלים הבאים.

3. שלב הניתוח

המטרה של שלב זה היא לנתח את הציוצים שנאספו לפי שפה ומדינה, תוך שימוש בטכניקות ניתוח רשת וניתוח טקסט. הניתוח יספק תובנות לגבי דפוסי התפשטות המידע הכוזב במהלך מגפת ה-COVID-19, כמו גם את המאפיינים של משתמשי הרשת החברתית המפיצים מידע כוזב זה.

4. שלב בניית המודל

מטרת שלב זה היא לבנות מודלים של חיזוי המבוססים על הנתונים המנותחים. המודלים יסייעו לזהות את גורמי המפתח התורמים להפצת מידע כוזב בעת חירום ולהבין את התפשטות הציוצים על ציר הזמן ואת הקשר בין התפשטות הקונספירציה למקרי חירום שקרו באותו הזמן.  
מטרה נוספת לשלב זה היא לחזות את הסבירות להתפשטות מידע כוזב בהקשרים שונים.

5. שלב הטמעת המערכת

מטרת שלב זה היא ליישם מערכת מבוססת ענן להצגת תוצאות המחקר. המערכת תספק ממשק ידידותי לגישה והדמיה של ממצאי המחקר, שיסייע לקובעי מדיניות, לחוקרים ולציבור לקבל החלטות מושכלות לגבי ניהול מידע כוזב בשעת חירום.

## סקירת ספרות

כיום הרשתות החברתיות ובייחוד טוויטר ממלאות תפקיד חשוב בהפצת מידע במגוון רב של תחומים, עד כדי כך שגם גופים ציבוריים ואפילו מדינות מעבירים מידע דרך אותן רשתות על מנת להנחות את האזרחים (בניגוד לעבר, שכדי להעביר מסרים לאוכלוסייה היה מקובל להשתמש בעיתונות המסורתית ובדרכים אחרות).

היתרון הגדול בשימוש ברשתות החברתיות ובטכנולוגיה העכשווית הוא הזמינות ויכולת ההנגשה לקהל רחב בזמן קצר, אך החיסרון הוא היעדר פיקוח והיכולת של כל אדם להשתמש בכלים הללו לפי רצונו, מה שמהווה כר פורה להפצת מידע כוזב ותאוריות קונספירציה במהירות גבוהה ולקהל רב. הפצת מידע כוזב היה נפוץ במיוחד בתקופת התפרצות מגפת הקורונה, עד כדי כך שארגון הבריאות העולמי נתן לתופעה שם משלה - "info-demic", ואף הקימו מחלקה שכל תפקידה הוא להילחם בפייק ניוז שרץ ברשתות החברתיות. (Zarocostas, 2020).

על מנת להמחיש את גודל ההתפשטות של מידע כוזב ברשתות החברתיות בזמן מצבי חירום, ניתן להסתכל לדוגמה על התפשטות הקונספירציה "#FilmYourHospital" - ההאשטאג, שתפס תאוצה אדירה בטוויטר בזמן תחילת משבר הקורונה, עודד אנשים לבקר בבתי חולים מקומיים כדי לצלם תמונות וסרטונים של מחלקות קורונה בבתי חולים ריקים כדי לעזור "להוכיח" שמגפת - COVID 19 היא בעצם מתיחה. הקונספירציה התפרסמה ברשת החברתית טוויטר וקיבלה חשיפה עצומה בתחילת המגיפה - על פי נתוני מחקר שנערך עליו ניתן לראות כי בתאריכים 28 במרץ עד 9 באפריל, נאספו 99,039 ציוצים על שכללו את האשטאג, שצויצו על ידי לא פחות מ-43,461 משתמשים שונים.

לפי המאמר של Gruz and mai, שבו ניתחו את התפשטות תאוריית הקונספירציה "#FilmYourHospital" ברשת הטוויטר, באמצעות הפיכת נתוני הטוויטר לרשת קשרים, שבה הצמתים מייצגים משתמשי טוויטר, והקשרים ביניהם מייצגים אינטראקציות בין משתמשים, ובעזרת ייצוג זה הומחש איך כל ציוץ התפשט ברחבי רשת הטוויטר.

במאמר, הוסק כי רוב החשבונות הפעילים והמשפיעים מאחורי מסע פרסום האשטאג הינם חשבונות אמיתיים ולא בוטים (חשבונות אוטומטיים), אך עם זאת, נמצאו סימנים של תיאום בקרב אנשי פוליטיקה וקבוצות ימין קיצוני המנסים להשתמש בתאוריית הקונספירציה הזו כנגד יריביהם הפוליטיים.

עוד גילו, כי הדלק שהניע את הקונספירציה הזו בימיה הראשונים הגיע מקומץ של פוליטיקאים שמרנים בולטים מימין הקיצוני, אשר השתמשו בכוחם על מנת ליצור מודעות לקמפיין ולעודד אתה עוקבים שלהם לצלם את המתרחש בבתי החולים. תאוריית הקונספירציה עלתה למודעות האוכלוסייה בעיקר על ידי תומכים פרו-טראמפ, ולאחר מכן נוצר גל משני של התפשטות התאוריה מחוץ לארה"ב. (Gruzd & Mai, 2020).

לאור האמור לעיל, אחת הבעיות המרכזיות שיש לטפל בהן על מנת לבלום את התופעה - היא נפח הנתונים העצום.



זרימת המידע עם מגוון רחב של נרטיבים, נכונים ושקריים, לא ניתנת לניתוח במהירות מדויקת ויעילה אפילו בסיוע ארגונים שנועדו לכך. נפח הנתונים העצום הופך את הגישות המסורתיות ללא מספיק מועילות על מנת לבלום את התפשטות המידע הכוזב.

לכן, על מנת לבלום את קצב התפשטות המידע השגוי בזמני חירום, נוכל להשתמש בטכנולוגיות עדכניות, לדוגמא שימוש בבינה מלאכותית, מדעי נתונים ולמידת מכונה, כל זאת על מנת לסייע לממשלות ולשירותי הבריאות.

כמו שנעשה במאמר של Gradon et al., שבו סיפקו סקירה כללית של האפשרויות ליישום טכנולוגיה ללימוד ולהילחם בדיסאינפורמציה, ותיארו מספר גישות הנמצאות בשימוש כיום להבנה, תיאור ומודל של תופעות של מידע מוטעה ודיסאינפורמציה, בעזרת טכנולוגיות עדכניות כגון למידת מכונה, שיטות כריית מידע וטקסט באיתור מידע מוטעה ועוד. במאמר, תומכים ביישום ניתוח ביג דאטה וטכנולוגיות מידע קשורות כדי למנוע, לשבש ולאתר מידע מוטעה ודיסאינפורמציה ביעילות.

על פי ממצאי המאמר, חשוב להעריך את ההזדמנויות לשימוש בבינה מלאכותית, מדעי הנתונים ולמידת מכונה כדי לסייע לסוכנויות ממשלתיות, לספקי שירותי הבריאות, למדיה החדשותית מכל הסוגים וארגוני החברה האזרחית לעבד ולנתח מידע כדי לספק מידע נכון ויעיל לאזרחים.

הגישה, ששמה דגש רב על שימוש בטכנולוגיה העדכנית למניעת שיתוף מידע כוזב וטיפול בזמן אמת במשברי חירום, דורשות פעולה משותפת בקנה מידה עולמי, שכן שיתוף תאוריות קונספירציה ומידע שגוי ברשת גורמים לנזקים בלתי הפיכים בקני מידה עולמי. (Gradon et al. 2021).

לדוגמא, במאמרו של Micallef, ששם דגש על תפקיד האזרחים בבדיקת עובדות ונכונות המידע שמסתובב ברשת, ותגובת ההמונים להטעיות הפופולריות בתקופת הקורונה, חקרו כיצד והאם מנסים להפריך מידע שגוי ומצאו כי 96% מכל ההפרכות נעשות על ידי אזרחים מודאגים, ויתרה מכך, פרסום בדיקת הפרכה מקצועית למידע שגוי לא רק קטנה באופן משמעותי מבחינת נפח, אלא גם זוכה בדרך כלל להרבה פחות חשיפה בניגוד לפרסום המידע השגוי. (Micallef et al. 2020).

בנוסף ניתן לנתח התנהגויות של משתמשים על פני טווח זמן מסוים וכך להבין באילו זמנים או אילו אירועים גורמים לעלייה בציוצים שמכילים תאוריות קונספירציה כפי שנעשה במחקר Conspiracy theories on Twitter.

הנתונים שאותם חקרו הם כ-142 אלף ציוצים שנחשבים כתומכים בתאוריות הקונספירציה וכ-10,000 ציוצים שאינם נחשבים כתומכים בתאוריות קונספירציה.

המחקר נעשה על ידי שימוש באלגוריתמים שהופכים טקסט לוקטורים, ובעזרתם ניתן להשוות מרחק בין מילה/תו למילה/תו אחרים ואז לנסות למצוא דמיון בניהם בעזרת טכנולוגיות שונות כגון האלגוריתם GloVe.

אנו רואים שמהניתוח של ציוצים שמכילים תאוריות קונספירציה על פני ציר הזמן ניתן ללמוד כי רוב הציוצים על תאוריות קונספירציה הינם ציוצים אימפולסיביים כתגובה על מקרה שקרה באותו היום ולא משהו מתוכנן מראש על מנת להפיץ בשיטתיות. (Batzdorfer et al. 2022).

לאחר ניתוח אילו נסיבות/אירועים גורמים לעלייה בפרסום תאוריות קונספירציה, על ידי שימוש ב-NLP ניתן ללמוד גם על השיח בטוויטר בנושא קונספירציה מסוימת, כמו שנעשה במחקרם של Elroy & Yosipof.

במחקר, השתמשו בדאטה של 331,448 אלף ציוצים מטוויטר (על ידי שימוש ב-API למטרות מחקר) שנוגעים בתאוריית הקונספירציה g5 על מנת להוכיח כי ניתן לסווג בצורה טובה על ידי שימוש בכלים של NLP כמו TF-IDF, GloVe, BERT, Word2Vec אלו ציוצים נחשבים תומכים בתאוריית הקונספירציה, מה היא כמות התומכים בקונספירציה, כמות המתנגדים והניטרליים, מי שלט בשיח הציבורי, באיזה זמן ועוד.

מצאו כי בתאוריה זו, כמות המשתמשים שתמכו בקונספירציה הייתה נמוכה מכמות המשתמשים שהתנגדו לה אך השליטה בחודשים הראשונים הייתה בידי התומכים, ולאחר תקופה קצרה המתנגדים הם אלה ששלטו בשיח הציבורי. (Elroy & Yosipof, 2022).

מטרת השימוש בטכנולוגיות המתקדמות ביותר במדעי הנתונים הינה שתהיה אפשרות לנתח בזמן אמת או בעבר מה גרם לאנשים לציץ בצורה כזאת או אחרת, ציוצים שמכילים תאוריות קונספירציה.

הניתוח לפי תאריכים מאפשר לנו להבין אילו תאוריות קונספירציה הן זמניות, כלומר מתפרצות ועולות ללב הדיון הציבורי אך לאחר תקופה קצרה הם כמעט נעלמות, ואילו תאוריות קונספירציה נשארות בלב הדיון הציבורי באופן תמידי.

ניתן להסיק כי תאוריות קונספירציה שמשתייכות לסוג השני, הינם תאוריות שיש כמות גבוהה של אנשים שמשוכנעים שהם אמת לאמיתה, מה שמשפיע עליהם לקבל את ההחלטה שלא להתחסן. ניתוח כזה, מאפשר לנו לתת כלים למקבלי ההחלטות באילו תאוריות להשקיע משאבים גבוהים כגון פרסום נגד, הפרכות התאוריה ושכנוע של התומכים נגד התאוריה. (Erokhin & Yosipof, 2022).

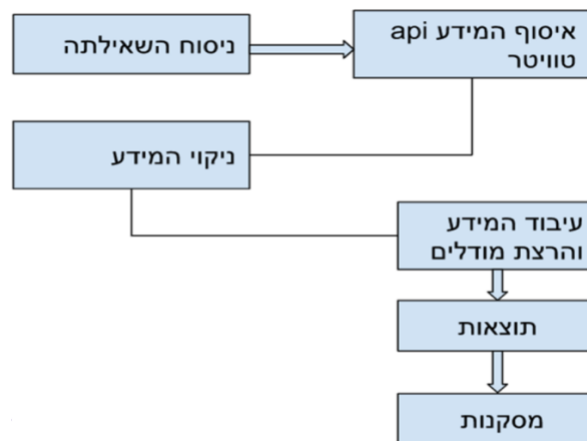
הקורונה ותאוריות הקונספירציה התפשטו והשפיעו על כל העולם ולכן חשוב לחקור את הנושא גם בשפות אחרות על מנת להבין האם אוכלוסיות עם מנהגים, סט אמונות ותרבויות שונות מתנהגים אחרת או מאמינים אחרת.

בצורה זו, נוכל גם להבין האם הטכנולוגיות המתקדמות של מדעי הנתונים ועיבוד שפה טבעית מאפשרות לנו להסיק מסקנות גם עבור שפות ומדינות אחרות כגון ערבית וכו'. (Al-Hashedi & Ali, 2022).

## שיטות

במחקר זה, הנבדקים הם הציוצים עצמם. הציוצים שנבחרו אינם מיוחסים למשתמשים ספציפיים, אלא נחשבים יחד בשל התוכן והרגשות שלהם. הם מייצגים מערך מגוון של שפות, גיאוגרפיות ונקודות מבט, ותורמים למערך נתונים עשיר לניתוח.

השפות המיוצגות במערך הנתונים שלנו כוללות אנגלית, גרמנית, איטלקית, ספרדית, פולנית, רומנית, הולנדית, יוונית וצ'כית. מגוון זה של שפות מאפשר לנו לבצע ניתוח השוואתי בין קהילות לשוניות שונות ולקבל פרספקטיבה גלובלית על התפשטות המידע השגוי.



איור 1: תיאור גרפי של המחקר

1. **ניסוח שאלות** לתאוריית ה-5G לפי רשימת השפות עבור שליפת הציוצים מטוויטר.

דוגמא לניסוח שאלתה בשפה האנגלית-

{covid OR coronavirus OR corona} AND {5G}

איור 2: דוגמא לניסוח שאלתה באנגלית

2. **איסוף הציוצים** מטוויטר בעזרת API - שליפת הציוצים בעזרת רישיון המחקרי של מנחה הפרוייקט, המאפשר לנו לשלוח מידע מלא מטוויטר לצורך ניתוח מחקרי.

3. **אחסון הנתונים** בבסיס נתונים ייעודי והוספת מאפיינים לנתונים כגון - תאריך, שפה, מדינה, תאוריית קונספירציה. בפרויקט שלנו השתמשנו בעיקר בספריית פנדס מבית פייתון, בה אחסנו את הנתונים בדאטה פריים:

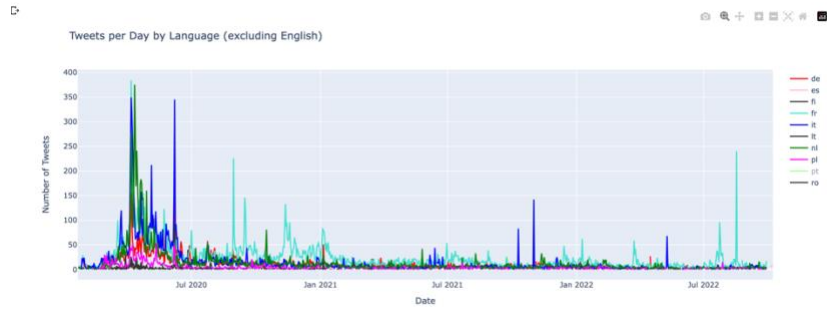
created_at	id_b36	author_idb36	id	original_text	metric_retweets	metric_likes	metric_replies
2020-01-09 11:22:20+00:00	98DNWPVP1N28	1CUYV3J	1215232331054796800	@ElijahSchaffer Corona Premier is only 2.5g carts	0	0	0
2020-01-19 22:49:16+00:00	99F1QU7EDCN9	554FV9U61QH	1219029084220555269	#CHINA - #disease in Wuhan City of 8.5 million...	0	1	1
2020-01-20 17:00:03+00:00	99HR1OQX01DU	7PPELOYJCAO	1219303586358889922	Coronavirus in China: Virus spreads like Huawei...	0	0	0
2020-01-21 05:12:22+00:00	99JKDIFJL134	CWJMR	121948788072341760	Claims that 5G is causing Coronavirus starting...	3	10	2
2020-01-21 07:25:24+00:00	99JW8QDMO35S	IHQ72	1219521359568498688	@garthbiggs @juhasaninen Coronavirus is the g...	1	2	2
2022-10-08 12:42:30+00:00	BZSS7CP2BNST	BYEW9Q06HIWW	1578727533717360641	@Kamash_Dev @Anaxadition @EmbajadaRusiaES Mi ...	0	2	0

איור 3: הצעה למערך הנתונים של הפרוייקט

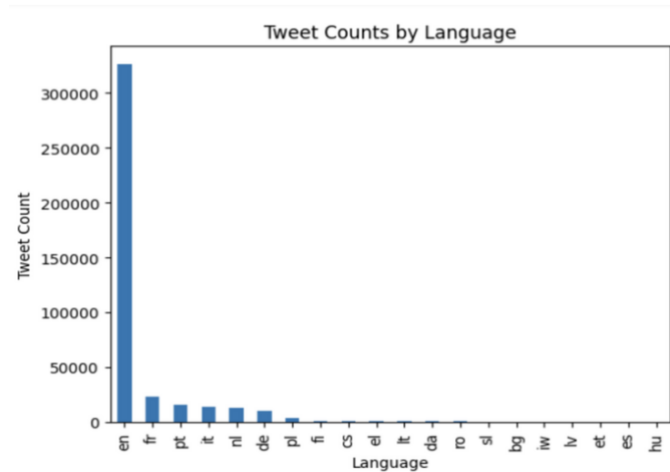
4. **ניקוי הנתונים והכנתם עבור ניתוח ועיבוד הנתונים**  
תהליך זה מבטיח שהנתונים איתם אנו עובדים מדויקים, עקביים ומוכנים לניתוח.

- הסרת נתונים לא רצויים: הצעד הראשון היה להסיר כל מידע לא רלוונטי ממערך הנתונים שלנו. זה כלל כל מידע שאינו רלוונטי לשאלת המחקר שלנו, כמו ציוצים שלא הכילו אף אחת ממילות המפתח הקשורות לתיאוריות הקונספירציה שאנו חוקרים.
- טיפול בערכים חסרים: בשלב הבא, נאלצנו להתמודד עם נתונים חסרים. לפעמים, ייתכן שלציוץ אין את כל המידע שאנחנו צריכים, כגון השפה או המיקום של הציוץ. במקרים אלה, השתמשנו בשיטות סטטיסטיות כדי להשלים את הפערים במידת האפשר, או השמטנו את הציוצים האלה ממערך הנתונים שלנו.
- תיקון אי עקביות: היינו צריכים גם לוודא שהנתונים עקביים. לדוגמה, לוודא שתמיד מתייחסים לאותה שפה באותו קוד, או שהתאריך והשעה של כל ציוץ הם באותו פורמט.
- עיבוד מוקדם של טקסט: בזמן שאנו עובדים עם נתוני טקסט (תוכן הציוצים), היינו צריכים גם להכין את הטקסט לניתוח. זה כלל הסרת אלמנטים מיותרים כמו כתובות URL ותווים מיוחדים, והמרת כל הטקסט לאותיות קטנות כדי להבטיח אחידות.

5. **סטטיסטיקה תיאורית** - סטטיסטיקה תיאורית מציעה סיכום כמותי של הנתונים. לדוגמה, חישובנו את הממוצע, החציון והמצב של ציוצים ליום עבור כל שפה, בדקנו כמה ציוצים יש בכל שפה לאורך מגפת הקורונה.



איור 4: תרשים קווים- מספר הציוצים ביום עבור כל שפה.



איור 5: מספר הציוצים עבור כל שפה: ההאנגלית הינה בעלת נתח הציוצים הגדול ביותר (78.7%)

6. אפיון ידני של אחוז מסויים מהנתונים על מנת ליצור בסיס אמין ובקרה למודלים- תייגנו בצורה ידנית 5000 ציוצים בשפה האנגלית כ-מתנגדי/תומכי/ניטרליים לתאוריית הקונספירציה 5G.

7. שימוש בטכניקות ניתוח רשת וניתוח טקסט ומודלי סיווג על מנת לספק תובנות לגבי דפוסי התפשטות המידע הכוזב.

שם המודל	תיאור
cross corolation	מתאם קורולציה הוא מדד סטטיסטי המראה את המידה שבה שני משתנים (במקרה שלנו- 2 שפות), נעים זה ביחס לזה
auto correlation	קורלציה אוטומטית היא בדיקה של פיגור בסדרת הזמן בשפה עצמה- כלומר ההשוואה היא לא בין 2 שפות, אלא בין אותה שפה בזמנים שונים
auto arima	מודל סטטיסטי המיועד לניבוי עתיד על פני ציר זמן-בעזרתו ניתן לחזות פעילות ציוצים עתידית עבור כל שפה.
Word Cloud	הצגה טקסטואלית של המילים הנפוצות ביותר בדאטה
senmtiment analysis	ניתוח רגישות- טכניקה המשמשת לקביעת הסנטימנט (הרגש) מאחורי קטע טקסט נתון, במקרה שלנו ציוץ
Bert NLP	ברט הוא מודל עיבוד שפה, המשמש כדי לסווג את הסנטימנט של הציוצים כמתנגדי/תומכי/ ניטרלים לקונספירציה

איור 6: אלגוריתמים ומודלים שהשתמשנו בפרויקט

8. מרכיבי תוכנה וממשק- פיתוח מערכת מבוססת ענן להצגת ממצאי המחקר. המערכת בעלת ממשק ידידותי למשתמש היכולה לשמש קובעי מדיניות, חוקרים, עיתונאים וכלל הציבור.

- דאש: השתמשנו ב-Dash, מסגרת Python לבניית יישומי אינטרנט אנליטיים, כדי ליצור את הממשק של האתר שלנו. Dash אפשרה לנו לשלב אלמנטים אינטראקטיביים באתר, כגון תרשימים וגרפים שמשתמשים יכולים לתמרן כדי להמחיש את הנתונים בדרכים שונות.

- רנדר: כדי להנגיש את האתר שלנו לציבור, פרסנו אותו ב-Render, פלטפורמת ענן לאירוח שירותי אינטרנט. Render מטפל בתשתית של האתר שלנו, מאפשר לנו להתמקד בתוכן ולהבטיח שהאתר שלנו יישאר פעיל.

באמצעות אתר אינטרנט אינטראקטיבי זה, אנו שואפים לספק לקובעי מדיניות, לחוקרים ולציבור כלי ידידותי למשתמש לחקור ולהבין את התפשטות המידע השגוי בטוויטר במהלך מגפת COVID-19.

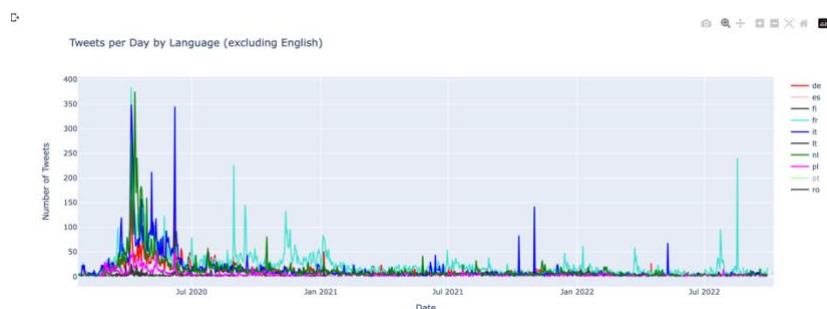


איור 7: סרוק על מנת להכנס לאתר

## תוצאות

### ניתוח סדרות זמן-

גרף כמות הציוצים בכל יום בתקופת המגיפה:



בעזרת ניתוח זה, ניתן לבצע הסבר וחקר של עליות חדות "פיקים" של ציוצים הקשורים לתאוריית הקונספריציה 5G בתאריכים מסויימים בשפות מסויימות:

בתאריך 7.6.2020 בגרמניה, תועדה עלייה חדה במספר הציוצים (318), אירוע משמעותי אחד בתקופה זו היה המחלוקת המתמשכת הקשורה לתגובת COVID-19 בראשות הקנצלרית אנגלה מרקל ושר הבריאות ג'נס ספאן. אסטרטגיות הפתיחה המחודשת והוויכוח סביב השימוש במסכות פנים ואמצעי ריחוק חברתי יכולים להיות סיבה אפשרית לעלייה במספר הציוצים.

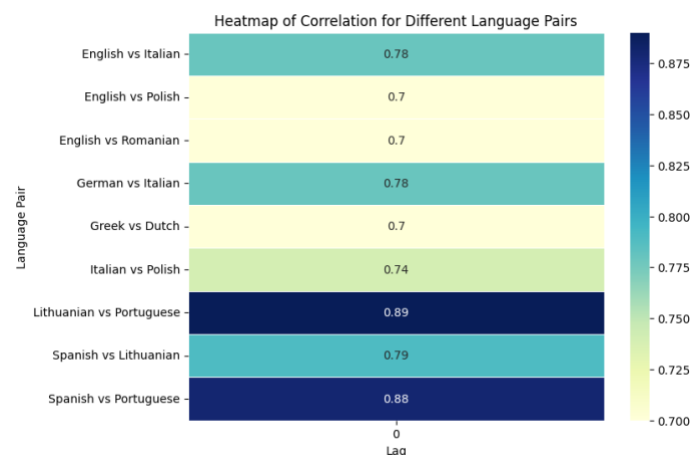
בתאריך 6 באפריל 2020, בשפה האיטלקית תועדה עלייה חדה גם כן (349). איטליה הייתה בשלב חמור של מגפת COVID-19. באותו היום המסוים, דווח שמספר ההרוגים היומי ב-COVID-19 באיטליה החל לרדת. הירידה הייתה ידיעה חיונית שכן איטליה הייתה המוקד האירופי של המגיפה עם שיעור תמותה גבוה משמעותית.

### מתאם קורלציה-

מתאם קורלציה הוא מדד סטטיסטי המראה את המידה שבה שני משתנים נעים זה ביחס לזה. מתאם של 1 אומר ששני המשתנים מתואמים בצורה מושלמת, אם משתנה אחד גדל, גם המשתנה השני גדל. מתאם של 0-1 (מתאם שלילי) אומר ששני המשתנים נמצאים בקורלציה הפוכה לחלוטין. אם אחד גדל, השני יורד. מתאם של 0 אומר שאין קשר ליניארי בין המשתנים (מדד פירסון).

"פיגור" בניתוח סדרות זמן מתייחס לפרק זמן מסוים שתכונה אחת (כמו מספר הציוצים בשפה מסוימת) נגררת אחרי אחר. לדוגמה, אם משווים ספירת ציוצים יומית בשתי שפות ונמצא פיגור של 1, פירוש הדבר שמספר הציוצים של היום בשפה אחת נמצא בקורלציה החזקה ביותר עם ספירת הציוצים של אתמול בשפה השנייה. במחקר שלנו, נמצא כי כמעט כל הפיגורים הם 0, מה שאומר שההשוואות הן בין מדדים בו-זמניים בשפות שונות.

הסיבה לכך שאנו רואים מתאמים מקסימליים בפיגור 0 עשויה להיות בגלל שהציוצים בשפות שונות מגיבים לאותם אירועים גלובליים הקשורים ל-COVID-19, מה שיגרום לעלויות ציוצים בכל השפות בו זמנית.



איור 8: מפת חום עבור מתאם הקורלציה בין השפות

הציוצים באיטלקית ('lt') ופורטוגזית ('pt') מראים את המתאם החזק ביותר (0.89), כלומר דפוסי הציוץ על COVID-19 בשפות אלו נעים בצורה דומה מאוד לאורך זמן. גם ציוצים בספרדית ('es') ופורטוגזית ('pt') מראים מתאם חזק מאוד (0.88), וכך גם ציוצים ספרדית ואיטלקית (0.79). ציוצים באנגלית ('en') ואיטלקית ('it'), וגרמנית ('de') ואיטלקית מראים קורלציות חזקות (0.78). מעניין שגם ציוצים גרמניים ופולניים ('pl') מראים מתאם של 0.70, אבל עם פיגור של 1.

זה אומר שהציוצים הגרמניים היום קשורים הכי חזק לציוצים הפולניים מהיום הקודם. הניתוח שלנו הראה שציוצים בשפות שונות נוטים להראות דפוסים דומים לאורך זמן, במיוחד בין ליטאית לפורטוגזית, ולבין ספרדית ופורטוגזית.

### **קורלציה אוטומטית-**

בשונה מקורלציה רגילה, קורלציה אוטומטית היא בדיקה בין פיגור בסדרת הזמן בשפה עצמה.

באופן כללי, יש קורלציה חיובית לפיגור של 1 בכל השפות, מה שמצביע על כך שהציוצים שהדנים בתיאוריית הקונספירציה G5 יש רמה מסוימת של תלות בעוצמת הציוצים של היום הקודם. מידת הקורלציה הנ"ל משתנה בין השפות. לדוגמה, ציוצים באנגלית (en) מציגים את המתאם האוטומטי הגבוה ביותר בפיגור 1 (0.9163), בעוד שציוצים בעברית (iw) מציגים את הנמוך ביותר (0.1464).

ברוב השפות, ערכי המתאם האוטומטי יורדים ככל שהפיגור גדל, מה שצפוי בנתוני סדרות זמן. עם זאת, שפות מסוימות, כמו הולנדית (nl), מראות עלייה בערכי קורלציה אוטומטית לאחר פיגור של 2 (0.7410), כשהיא מגיעה לשיא בפיגור של 5 (0.7656).

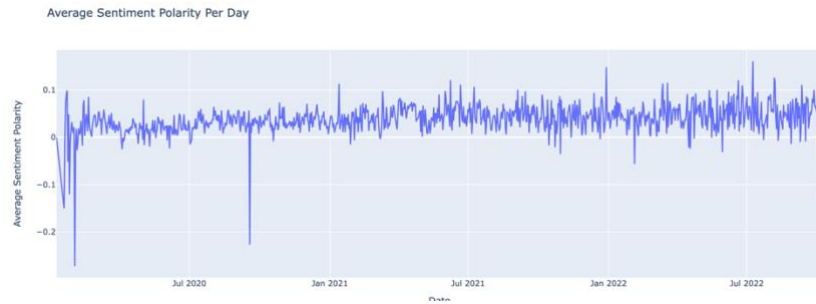
### **ניתוח רגישות-**

ניתוח רגישות היא טכניקה המשמשת לקביעת הסנטימנט (הרגש) מאחורי קטע טקסט נתון, במקרה שלנו ציוץ.

ניתוח כזה יכול להיות מועיל בהבנת דעת הקהל, תגובות לאירועים או מגמות בדיונים מקוונים. במחקר שלנו, אנו משתמשים בניתוח סנטימנטים כדי לנתח את 441,000 הציוצים הקשורים לתיאוריית הקונספירציה של G5 בתקופת הקורונה.

השתמשנו ב-text blob כדי לבצע את ניתוח הסנטימנט.





איור 9: ניתוח רגישות ממוצע לכל השפות ביחד לפי תאריכים

קוטביות הסנטימנט הממוצעת משתנה בין השפות, כאשר לחלק מהשפות יש קוטביות סנטימנט ממוצעת מעט חיובית (כגון אנגלית, צ'כית והונגרית) כלומר מדינות עם יותר חיוביות, בעוד שלאחרות יש קוטביות סנטימנט ממוצעת מעט שלילית (כגון יונית, איטלקית וליטאית). לרוב השפות יש ציוני קוטביות סנטימנט ממוצעים קרובים לאפס, מה שמצביע על סנטימנט כללי ניטרלי. ברוב השפות, רוב הציוצים מסווגים כניטרליים. עם זאת, עבור השפה האנגלית, יש שיעור גבוה יותר של ציוצים חיוביים (115,441) בהשוואה לציוצים שליליים (81,122). בשפות מסוימות יש שיעור גבוה יותר של ציוצים שליליים בהשוואה לציוצים חיוביים, כמו גרמנית (שלילית: 324, חיובית: 702) ואיטלקית (שלילית: 914, חיובית: 774). לעומת זאת, בשפות אחרות כמו הולנדית (שלילי: 593, חיובי: 988) וצרפתית (שלילית: 1,300, חיובית: 1,820) יש שיעור גבוה יותר של ציוצים חיוביים בהשוואה לציוצים שליליים. ממצאים אלה מצביעים על כך שהדיונים סביב תיאוריית הקונספירציה של G5 בתקופת הקורונה השתנו בין השפות במונחים של קוטביות הסנטימנט. נראה שהסנטימנט הכללי ברוב השפות הוא ניטרלי, כאשר חלק מהשפות נוטות יותר לכיוון חיובי או שלילי.

### מודל ארימה-

ARIMA, ראשי תיבות של Autoregressive Integrated Moving Average, היא טכניקת חיזוי המשמשת להבנה ולניבוי ערכים עתידיים בסדרת נתונים. המודל "SARIMAX" שהשתמשנו בו מייצג ממוצע נע עם משתנים.

במילים פשוטות יותר:

"אוטו רגרסיב" פירושו שהמודל משתמש בקשר בין תצפית למספר מסוים של תצפיות בפיגור (כלומר, תצפיות קודמות).

"משולב" פירושו שהמודל משתמש בהבדלים בין תצפיות כדי להפוך את סדרת הזמן לנייחת (כלומר, נתונים עם ממוצע ושונות קבועים לאורך זמן).

"ממוצע נע" אומר שהמודל משתמש בתלות בין תצפית לבין שגיאה שיורית ממודל ממוצע נע המיושם על תצפיות בפיגור.

Language: en

Model Summary:

SARIMAX Results						
=====						
Dep. Variable:	y	No. Observations:	796			
Model:	SARIMAX(0, 1, 2)	Log Likelihood	-5819.085			
Date:	Tue, 25 Jul 2023	AIC	11644.171			
Time:	13:01:37	BIC	11658.206			
Sample:	01-09-2020	HQIC	11649.564			
	- 03-14-2022					
Covariance Type:	opg					
=====						
	coef	std err	z	P> z	[0.025	0.975]
-----						
ma.L1	0.2849	0.010	29.044	0.000	0.266	0.304
ma.L2	-0.2670	0.007	-38.866	0.000	-0.280	-0.253
sigma2	1.352e+05	1160.315	116.492	0.000	1.33e+05	1.37e+05
=====						
Ljung-Box (L1) (Q):			0.00	Jarque-Bera (JB):	288197.57	
Prob(Q):			0.99	Prob(JB):	0.00	
Heteroskedasticity (H):			0.01	Skew:	3.98	
Prob(H) (two-sided):			0.00	Kurtosis:	95.93	
=====						

Warnings:

[1] Covariance matrix calculated using the outer product of gradients (complex-step).

איור 10: פלט ארימה עבור השפה האנגלית

### אנגלית

פרמטרי המודל נשארים זהים לאלו בניתוח הקודם. מונחי הממוצע הנע MA 1 ו MA 2 הם 0.2844 ו-0.2669 בהתאמה, והשונויות של מונח השגיאה,  $\sigma^2$ , היא  $1.08 \times 10^5$ . הנתונים עדיין מראים גבנוניות גבוהה (119.79) והטיה חיובית (4.45), מה שמצביע על התפלגות עם ערכים קיצוניים יותר מהתפלגות נורמלית. התחזית ממודל זה מצביעה על כך שמספר הציוצים של COVID-19 בשפה האנגלית יקטן לסביבות 66 ליום עד ה-1 באוקטובר 2022, וישיאר יציב ברמה זו עד ה-9 באוקטובר 2022. מרווחי הסמך לתחזית זו נעים בין שפל של כ-572 ציוצים לשיא של כ-716 ציוצים ב-30 בספטמבר, 2022. עד ל-9 באוקטובר 2022, טווח זה מתרחב בין כ-2063 ציוצים ל-2196 ציוצים. מרווחים רחבים אלו מרמזים על רמה גבוהה של אי ודאות בתחזית.

### צרפתית

הפלט הזה הוא מניתוח (SARIMAX(1, 1, 1) על ציוצים של COVID-19 בשפה הצרפתית. פרמטרי המודל שונים מהמודל בשפה האנגלית. המונח האוטו רגרסיבי AR הוא 0.4263 והמונח הממוצע הנע MA הוא -0.8879. השונויות של מונח השגיאה,  $\sigma^2$ , היא 347.1915. הנתונים מראים גם גבנוניות גבוהה (89.61) והטיה חיובית (6.48), מה שמצביע על התפלגות עם ערכים קיצוניים יותר מהתפלגות נורמלית. התחזית ממודל זה מצביעה על כך שמספר הציוצים של COVID-19 בשפה הצרפתית יקטן לסביבות 9 ביום עד ה-4 באוקטובר 2022, וישיאר יציב ברמה זו עד ה-9 באוקטובר 2022. רווחי הסמך לתחזית זו נעים בין שפל של כ-24 ציוצים לשיא של כ-49 ציוצים ב-30 בספטמבר, 2022. עד ל-9 באוקטובר 2022, טווח זה מתרחב בין כ-39 ציוצים ל-57 ציוצים. מרווחים רחבים אלו מרמזים על רמה גבוהה של אי ודאות בתחזית. מסקנות:

הניתוח שלנו מצא שמספר הציוצים של COVID-19 בשפה הצרפתית צפוי לרדת לסביבות 9 ליום עד תחילת אוקטובר 2022 ולהישאר ברמה זו לפחות שבוע.

### פורטוגזית

הפלט הזה הוא מניתוח SARIMAX(2, 1, 0) על ציורים של COVID-19 בשפה הפורטוגזית. פרמטרי המודל מראים שהמונח האוטו רגרסיבי AR 1 הוא 0.0271 ו AR 2 הוא -0.4704. השונות של מונח השגיאה,  $\sigma^2$ , היא 14370. אין רכיב ממוצע נע במודל זה. מודל זה גם מצביע על גבנויות גבוהה מאוד (337.54) והטיה חיובית (4.40), מה שמצביע על כך שלנתונים יש ערכים קיצוניים יותר מהתפלגות נורמלית. התחזית ממודל זה מצביעה על כך שמספר הציורים של COVID-19 בשפה הפורטוגזית ירחף סביב 2 ביום 30 בספטמבר 2022 ועד ה-9 באוקטובר 2022. יש עלייה קלה לסביבות 2.25 ב-3 באוקטובר 2022, לפני שהוא יחזור לקדמותו. רווחי הסמך עבור תחזית זו הם רחבים מאוד, ונעים בין שפל של כ-233 ציורים לשיא של כ-237 ציורים ב-30 בספטמבר 2022. עד 9 באוקטובר 2022, טווח זה מתרחב בין כ-558 ציורים ל-562 ציורים. מרווחים רחבים אלו מרמזים על רמה גבוהה של אי ודאות בתחזית. מסקנות: הניתוח שלנו מצא שמספר הציורים של COVID-19 בשפה הפורטוגזית צפוי לנוע סביב 2 ביום מסוף ספטמבר ועד תחילת אוקטובר 2022.

### עברית

הפלט הוא מניתוח SARIMAX(0, 1, 2) על ציורים של COVID-19 בשפה העברית (iw). פרמטרי המודל מראים שמונחי הממוצע הנע MA 1 ו MA 2 הם -0.8947 ו-0.0672, בהתאמה. השונות של מונח השגיאה,  $\sigma^2$ , היא 0.4483. אין רכיב אוטו רגרסיבי בדגם הזה. מודל זה מצביע גם על גוניות גבוהה (12.96) והטיה חיובית (2.43), מה שמצביע על נוכחות של ערכים קיצוניים או חריגים בנתונים. התחזית ממודל זה מצביעה על כך שמספר הציורים של COVID-19 בשפה העברית ינוע סביב 0.16 ליום 27 בספטמבר 2022 ועד ה-6 באוקטובר 2022. מרווחי הסמך לתחזית זו הם די רחבים, ונעים בין שפל של כ-1.09 ציורים לשיא של כ-1.53 ציורים ב-27 בספטמבר, 2022. עד ל-6 באוקטובר 2022, טווח זה עובר מעט בין כ-1.17 ציורים ל-1.49 ציורים. מרווחים רחבים אלו מרמזים על רמה גבוהה של אי ודאות בתחזית. מסקנות: הניתוח שלנו מצא שמספר הציורים של COVID-19 בשפה העברית צפוי לנוע סביב 0.16 ליום מסוף ספטמבר ועד תחילת אוקטובר 2022.

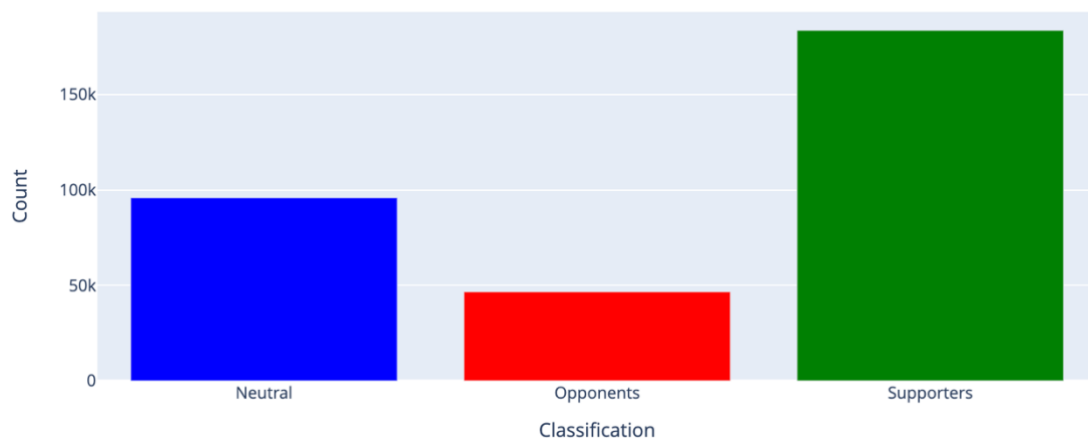
### מודל שפה טבעית ברט-

בניתוח שלנו של נתוני טוויטר הקשורים לתאוריית הקונספריציה 5g, השתמשנו במודלים CT-BERT ו-SBERT, כדי לסווג ציורים ולתרגם מבני משפטים מורכבים לצורות מספריות. מודל ה-CT-BERT, שתוכן במיוחד להבנת שפה בהקשר, סיפק סיווג מדויק של ציורים, תוך שהוא משמר את ניואנס השפה הייחודי שלהם. מודל ה-SBERT סייע להטמעת משפטים בווקטורים מספריים, ובכך פיתח ניתוח יעיל של המשמעויות הבסיסיות בתוך הציורים. המתודולוגיה שלנו

הייתה קפדנית, שכללה חלוקה של הנתונים לדוגמאות אימון ומבחן, ושימוש בשיטת Fold-5 Validation כדי להבטיח הערכה קפדנית ומאוזנת של ביצועי המודלים שלנו. התוצאות היו מלאות תובנות, עם 95.914 אלף ציוצים שסווגו כנייטרליים, 46.627 אלף כמתנגדים לעמדה ספציפית ו-183.577 אלף כתומכים. ממצאים אלה לא רק שופכים אור על הגישה של הציבור לתיאוריית הקונספריציה g5, אלא גם הדגימו את היעילות של הגישה שלנו בהבנת נתונים טקסטואליים מורכבים.

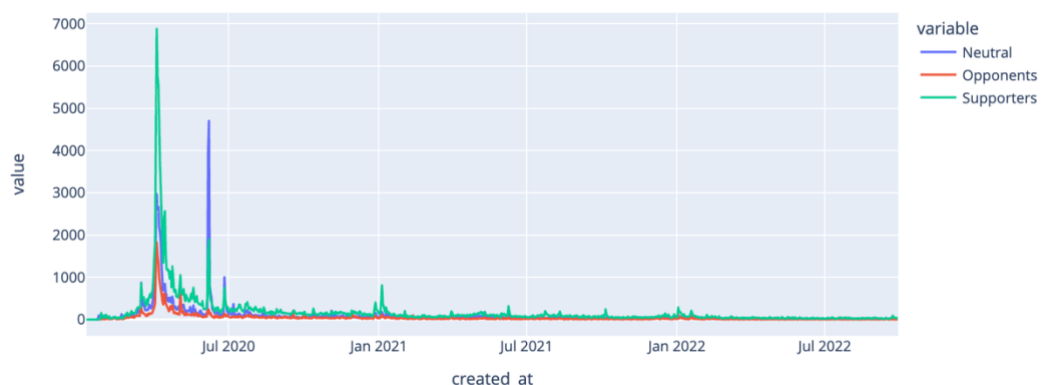
על ידי מינוף מודלים חדישים, הצלחנו להפוך נתוני טוויטר גולמיים לתובנות משמעותיות, לחשוף דפוסים ומגמות שאחרת היו נשארים מעורפלים. הצלחנו של ניתוח זה מדגישה את הערך והפוטנציאל של CT-BERT ו-SBERT בניתוח מדיה חברתית, ומציעה לחוקרים ומקבלי החלטות ערכת כלים רבת עוצמה לאמוד ולהגיב לדעת הקהל. העבודה שלנו עומדת כעדות לקידום טכניקות ניתוח טקסט ויישום רחב יותר של למידת מכונה בהבנת התקשורת והרגשות האנושיים.

Counts of Predicted Classifications



איור 11: התפלגות תוצאות מודל ברט לגבי ציוצים בעד/נגד/נייטרליים לתיאוריית הקונספריציה 5g בארה"ב.

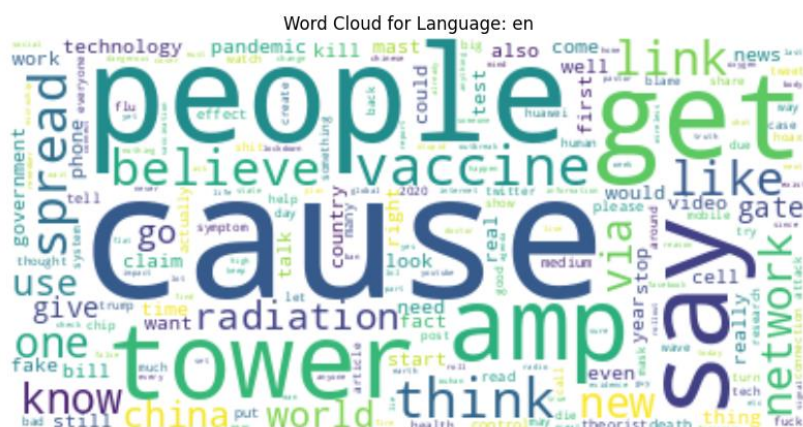
Tweets Over Time by Classification



איור 12- תוצאות מודל ברט לאורך כל תקופת המחקר

## ניתוח טקסט-

ביצענו ניתוח טקסט כדי לזהות את המילים והביטויים הנפוצים ביותר הקשורים ל-5G.



איור 13: ענו המילים המדוברות ביותר בשפה האנגלית

## דיון

המחקר שלנו נועד לחקור את השיח וההפצה של מידע כוזב בטוויטר בזמן חירום, תוך התמקדות במגפת COVID-19. מטרתנו הייתה לשפר את הדיוק והאמינות של המידע המשותף בפלטפורמה ולפתח אסטרטגיות לניטור והפחתת ההשפעה של מידע מוטעה.

## ממצאים

הממצאים שלנו סיפקו כמה תובנות לגבי הדינמיקה של שיחות טוויטר על תיאוריות קונספירציה של COVID-19 בשפות שונות. ראינו שיאים משמעותיים בנפח הציוצים בהתאם לאירועים גדולים או הודעות מדיניות. לדוגמה, בגרמניה, שיא בנפח הציוצים התרחש במקביל לוויכוחים סביב אסטרטגיות פתיחה מחדש ושימוש במסכות פנים.

המחקר שלנו חשף דפוסים משמעותיים של מידע מוטעה שהתפשטו על פני שפות ואזורים גיאוגרפיים שונים. ראינו שהנפח והרגש של ציוצים הדנים בתיאוריות הקונספירציה מושפעים מהדיונים של היום הקודם. ממצא זה מדגיש את הצורך באסטרטגיות התערבות בזמן כדי למנוע הגברת מידע כוזב.

יתר על כן, מצאנו מתאמים חזקים בין נפחי ציודים בשפות מסוימות, מה שמצביע על כך שאירועים גלובליים הפעילו דפוסי תגובה דומים. לדוגמה, דפוסי הציוד על COVID-19 באיטלקית ובפורטוגזית הראו מתאם חזק, מה שמרמז על תגובות בו זמנית לאירועים גלובליים.

באופן בלתי צפוי, גילינו שההשפעה וההתמדה של דיונים על תיאוריות הקונספירציה השתנו בין השפות. זה מצביע על כך שגישה אחת מתאימה לכולם לניהול מידע שגוי עשויה להיות לא יעילה, ואסטרטגיות מותאמות בהתחשב בהבדלים לשוניים ותרבותיים עשויות להיות מוצלחות יותר.

בנוסף, ניתוח הסנטימנט שלנו גילה שהסנטימנט הכללי ברוב השפות נוטה לכיוון ניטרליות, כאשר חלק מהשפות מראות יותר חיוביות או שליליות.

בעוד שהממצאים שלנו מציעים תובנות משמעותיות, לא כל ההשערות אושרו. לדוגמה, שיערנו שלשפות עם סנטימנט חיובי יותר יהיו נפחי ציוץ גבוהים יותר. עם זאת, הניתוח שלנו לא מצא קשר עקבי בין קוטביות הסנטימנט לעוצמת הציוץ. זה עשוי לנבוע מכמה גורמים, כולל הבדלים תרבותיים בהבעת רגשות או האופי המגוון של דיונים בכל שפה תרומות תיאורטיות ויישומיות

המחקר שלנו תורם הן תיאורטית והן מעשית להבנה של התפשטות מידע מוטעה במהלך משבר עולמי.

מנקודת מבט תיאורטית, המחקר שלנו תורם להבנה של האופן שבו מידע שגוי מתפשט במדיה חברתית בזמן משברים ומציע תובנות לגבי הדינמיקה של שיח מדיה חברתית במהלך מגיפה, תוך הדגשת משחק הגומלין בין נפח, סנטימנט ודפוסים זמניים.

מנקודת מבט יישומית, הממצאים שלנו מספקים תובנות ניתנות לפעולה לקובעי מדיניות, פלטפורמות מדיה חברתית וארגוני בריאות הציבור כדי לפתח אסטרטגיות יעילות לניהול מידע שגוי. הבנת הדינמיקה של הפצת מידע מוטעה יכולה לעזור בתכנון התערבויות בזמן ויעילות.

#### מגבלות מתודולוגיות ומחקר עתידי

למחקר שלנו יש כמה מגבלות. ראשית, הסתמכנו על נתוני טוויטר, שאולי אינם מייצגים באופן מלא את הדעות של האוכלוסייה הרחבה יותר. מחקר עתידי יכול לשלב נתונים מפלטפורמות אחרות של מדיה חברתית לניתוח מקיף יותר. שנית, ניתוח הסנטימנט שלנו התבסס על טכניקות אוטומטיות, שעשויות שלא ללכוד את מלוא המורכבות של הרגשות האנושיים.

מחקר עתידי יכול להרחיב את העבודה על ידי שילוב נתונים מפלטפורמות אחרות של מדיה חברתית כמו אינסטגרם, פייסבוק ושימוש בטכניקות ניתוח רגישות מתקדמות יותר. בנוסף, מחקרים עם טווחי זמן ארוכים יכולים לשפוך אור על ההשפעות ארוכות הטווח של מידע מוטעה על התנהגות ומדיניות ציבורית.

מחקרים עתידיים נוספים יכולים לחקור את התוכן הספציפי של הציוצים כדי לספק תובנות מעמיקות יותר לגבי אופי הפצת המידע השגוי. כמו כן, יהיה מעניין לחקור את ההשפעה של אמצעי הנגד על הפצת מידע מוטעה.

המחקר שלנו מדגיש את החשיבות של הבנה והתייחסות להפצה של מידע מוטעה בזמן חירום. ככל שאנו ממשיכים להתמודד עם האתגרים שמציבה מגפת COVID-19 ומשברים עולמיים אחרים, אסטרטגיות להילחם במידע מוטעה ימלאו תפקיד קריטי בהבטחת הבטיחות והרווחה של האוכלוסייה.

### **סיכום ומסקנות**

תהליך המחקר כלל גיבוש שאילתות לכל תיאוריית קונספירציה, אחזור ציוצים באמצעות API של Twitter, אחסון וניקוי הנתונים וביצוע ניתוח סטטיסטי. אימנו גם מודל BERT לניתוח סנטימנטים והשתמשנו במודלים של ARIMA ו-SARIMAX לניתוח סדרות זמן.

התוצאות היו מעוררות מחשבה, זיהינו שיאים משמעותיים בנפח הציוצים בהתאמה לאירועי מפתח או הכרזות ברחבי העולם, מה שמספק הבנה כיצד השיח הציבורי בטוויטר מגיב לאירועים בעולם האמיתי. גילינו גם מתאמים חזקים בין נפחי ציוצים בשפות שונות, מה שמצביע על קשר עולמי בתגובה למגיפה.

יתרה מכך, ניתוח הסנטימנט שלנו חשף שינויים בסנטימנט בין שפות שונות, וניתוח האוטו קורלציה שלנו הצביע על מתאם פיגור חיובי בכל השפות, מה שמרמז על כך שהדיונים על תיאוריות קונספירציה נמשכים לאורך זמן.

למרות ההישגים הללו, היו כמה אתגרים ומגבלות. בתחילה שיערנו קשר בין קוטביות הרגש לנפח הציוץ, אך הניתוח שלנו לא אישר זאת. בנוסף, הנתונים שלנו הוגבלו לטוויטר, שאולי לא מייצגים באופן מלא את דעת הקהל הרחבה יותר. כמו כן, בעוד שניתוח הרגישות הציע תובנות חשובות, ייתכן שהוא לא תופס במלואו את המורכבות של רגשות אנושיים בשל האופי האוטומטי של התהליך.

מתוך התנסויות אלו, למדנו את החשיבות של שילוב מקורות נתונים מגוונים ושכלול השיטות האנליטיות שלנו כדי ללכוד את המורכבות של השיח במדיה החברתית. ניסיון זה גם חיזק את החשיבות של חידוד מתמשך בהשערות ובשיטות המבוססות על תוצאות הניתוח.

בסך הכל, הפרויקט השיג את מטרותיו העיקריות של הבנת הדינמיקה של הפצת מידע מוטעה במהלך משבר עולמי. הממצאים שלנו יכולים לספק אסטרטגיות לניטור ומניעת מידע שגוי, ולהדגיש את התפקיד החשוב של מדעי הנתונים בהתמודדות עם אתגרים חברתיים. אמנם היו כמה מגבלות ותוצאות לא מספיק חד משמעיות, אבל אלה מספקים כיוונים למחקר עתידי.

## ביבליוגרפיה

- Al-Hashedi, Abdullah; Al-Fuhaidi, Belal; Mohsen, Abdulqader M.; Ali, Yousef; Gamal Al-Kaf, Hasan Ali; Al-Sorori, Wedad; Maqtary, Naseebah. (2022). Ensemble Classifiers for Arabic Sentiment Analysis of Social Network (Twitter Data) towards COVID-19-Related Conspiracy Theories. *Applied Computational Intelligence and Soft Computing*, vol. 2022, Article ID 6614730, 10 pages, 2022  
<https://doi.org/10.1155/2022/6614730>
- Batzdorfer, V., Steinmetz, H., Biella, M., & Alizadeh, M. (2022). Conspiracy theories on Twitter: emerging motifs and temporal dynamics during the COVID-19 pandemic. *International journal of data science and analytics*, 13(4), 315–333.  
<https://doi.org/10.1007/s41060-021-00298-6>
- Elroy, O., Yosipof, A. (2022). Analysis of COVID-19 5G Conspiracy Theory Tweets Using SentenceBERT Embedding. In: Pimenidis, E., Angelov, P., Jayne, C., Papaleonidas, A., Aydin, M. (eds) *Artificial Neural Networks and Machine Learning – ICANN 2022*. ICANN 2022. Lecture Notes in Computer Science, vol 13530. Springer, Cham.  
[https://doi.org/10.1007/978-3-031-15931-2\\_16](https://doi.org/10.1007/978-3-031-15931-2_16)
- Erokhin, D., Yosipof, A., & Komendantova, N. (2022). COVID-19 Conspiracy Theories Discussion on Twitter. *Social Media + Society*, 8(4).  
<https://doi.org/10.1177/20563051221126051>
- Gradoń, K. T., Hołyst, J. A., Moy, W. R., Sienkiewicz, J., & Suchecki, K. (2021). Countering misinformation: A multidisciplinary approach. *Big Data & Society*, 8(1).  
<https://doi.org/10.1177/205395172111013848>
- Gruzd, A., & Mai, P. (2020). Going viral: How a single tweet spawned a COVID-19 conspiracy theory on Twitter. *Big Data & Society*, 7(2).  
<https://doi.org/10.1177/2053951720938405>
- Micallef, N., He, B., Kumar, S., Ahamad, M., & Memon, N. (2020). The Role of the Crowd in Countering Misinformation: A Case Study of the COVID-19 Infodemic. In X. Wu, C. Jermaine, L. Xiong, X. T. Hu, O. Kotevska, S. Lu, W. Xu, S. Aluru, C. Zhai, E. Al-Masri, Z. Chen, & J. Saltz (Eds.), *Proceedings - 2020 IEEE International Conference on Big Data, Big Data 2020* (pp. 748-757). [9377956] (Proceedings - 2020



IEEE International Conference on Big Data, Big Data 2020). Institute of Electrical and Electronics Engineers Inc.. <https://doi.org/10.1109/BigData50022.2020.9377956>

Moffitt, J. D., King, C., & Carley, K. M. (2021). Hunting Conspiracy Theories During the COVID-19 Pandemic. *Social Media + Society*, 7(3).  
<https://doi.org/10.1177/20563051211043212>

Pogorelov, K., Schroeder, D.T., Filkuková, P., Brenner, S., & Langguth, J. 2021. WICO Text: A Labeled Dataset of Conspiracy Theory and 5G-Corona Misinformation .Tweets. Association for Computing Machinery, New York, NY, USA, 21–25  
<https://doi.org/10.1145/3472720.3483617>

Zarocostas J. (2020). How to fight an infodemic. *Lancet* (London, England), 395(10225), 676. [https://doi.org/10.1016/S0140-6736\(20\)30461-X](https://doi.org/10.1016/S0140-6736(20)30461-X)

**פוסטר הפרוייקט**

ניתוח וניטור מידע כוזב ברשת החברתית  
טוויטר במהלך משבר הקורונה

**מציגים: יקיר אטיאס ומנגשה אטלאיי**

**מנחה: ד"ר אבי יוסיפוף**

### מחלקה: מערכות מידע

שנה: תשפ"ד

## רקע ומצב נוכחי:

כעידן הרשתות החברתיות, מידע כחב יכול  
להתפשט ברשת בשניות ובמקרי חירום ליצור  
חוקים בלתי הפיכים.

מוקד הפרויקט שלנו הוא מידע שקרי לגבי תיאוריית הקונספירציה ה-5G ברשת החברתית טוויטר במהלך תקופת הקורונה.

## הגדרת הבעיה:

ריבוי תיאוריות הקונספירציה בטוויטר עשוי להקשות על אנשים להבחין בין מידע מדויק ולא מדויק, מה שעלול להוביל לבלבול וחוסר אמון במקורות מידע רשמיים.

בנוסף, להפצה של תיאוריות קונספירציה כטוויטר בזמן חירום יכולה להיות גם השפעות חברתיות ופסיכולוגיות שליליות.

## מטרות הפרויקט:

מחקר זה מתאר תוכנית לפיתוח ולניסוי מידע כחב ותדירותו, בעניין תאוריית ה-5G בטוויטר, בתקופת הקורונה.

מטרת המחקר היא להבין את עומק הדיון והפצה של מידע כחב בנושא ה-5G במדינות אירופה וישראל.

## שיטות ועיבוד נתונים:

הפריקט שלנו משלב שיטות חישוביות וסטטיסטיות שונות כדי לנתח ולפרש נתוני טוויטר הקשורים לתיאוריית ה-5G, ולהשוות את השיח בין שפות שונות באירופה.

להלן הצעדים והכלים העיקריים בהם השתמשנו:

- **ניתוח טקסט:** ביצוע ניתוח טקסט כדי לחלואת המילים וברטיים הנפוצים ביותר הקשורים ל-SCG.
- **ניתוח סטטיסטי (רש):** באמצעות בדיקת מספרית NLTk, קבעת את ההשג (חוב, שלילי או נטרלי) של כל ציץ.
- **ניתוח סדרות זמן:** השתמשו במודל ARIMA כדי לנתח את הפחיצים לאורך זמן, עד כדי להבין את המממזת והדופוסים ביין כל תיאורית.
- **עיבוד שפה:** השתמשו ב-BERT, מודל עיבוד שפה מתוחכם, כדי לטוט את הסנסמיטט של ציצים.
- **ניתוח מתאם:** חישבנת את המתאם בין נפחי הציצים לשפות שונות.

## WORK FLOW

הגדרת יעדי מחקר וניסוח  
שאלות:

הבדלת יעדים ברורים למחקר  
וניסוח שאילות לאחזור ציורים  
הקשורים ל-SG מטוויטר בשמות  
שונות.

57.21

ל מנת להבטיח את מהימנות  
מודל ולאמן אותה תיצ (אפיון)  
ני של כ-5000 ציורים  
תומכי/מתמדי/גטרלים  
תאורית ה-5G.

ג'תות גתונים ובג'ית מודל'ס:

ניתוח ופירוש של הצצים  
באמצעות ויזואליזציות, טכניקות  
ניתוח רשת וטקסט. בבית מודל  
חזוי לזיהוי גורמי מפתח  
התורמים להפצת מידע מוטעה  
ממקורות חיצוניים.

מיון מערכת

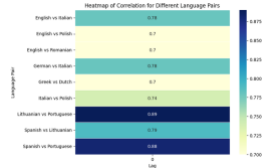
פיתוח מערכת מבוססת ענן להצגת ממצאי המחקר. המערכת בעלת ממשק יידידותי למשתמש היכולה לשמש קבוצי מדיניות, חוקרים,

## ניתוח הנתונים

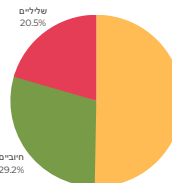
כמות הציוצים בכל יום בתקופת המגיפה:



הקורלציות הגבוהות ביותר בין השפות  
בתקופה אפס:



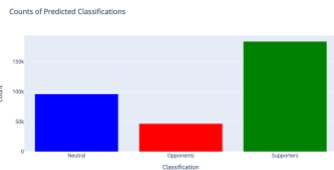
## התפלגות רגישות הציוצים בכל השפות:



ענן המילים המדוברות ביותר באנגלית:



התפלגות עמדות האוכלוסייה לגבי 5G  
בשפה האנגלית (BERTMODEL):



TECHNOLOGY

