**כלי לאחזור מידע טקסטואלי מאתרים**

1. מטרת הכלי: איסוף כתבות, התגובות להן, ופרופיל המשתמשים
2. תהליך החיפוש:
   1. רושמים מילות מפתח לחיפוש. אפשרות לבחור שמילות המפתח יופיעו בכותרת או בטקסט עצמו או בתגובות, או במשולב.
   2. אפשרות חיפוש לפי תקופות: ואז התוכנה תחלק את מספר הכתבות המבוקש ליחידות זמן מבוקשות (יום/חודש/שנה)
   3. מגדירים את האתרים מהם יתבצע החיפוש מתוך רשימה שנבנתה מראש
   4. ניתן להגדיר מספר מקסימלי של תוצאות
   5. ואז התוכנה נכנסת לכל אחת מהתוצאות, שולפת את האמור מטה, ומכניסה לשורות באקסל בהתאמה.
3. **אתרי חדשות:**

10 אתרי החדשות באנגלית הראשונים ברשימה שמופיעה כאן:

<http://www.ebizmba.com/articles/news-websites>

wall street journal ו- financial times.

5 אתרים בעברית:

www.themarker.com

Ynet.co.il

Walla.co.il

[http://www.globes.co.il](http://www.globes.co.il/)

<http://www.maariv.co.il/>

models.com

מבנה הקבצים באקסל:

* 1. גיליון 1 – כתבות – 7 עמודות של נתונים בהתאם לפירוט הבא (במילים אחרות – כל כתבה היא מקרה נפרד ומייצרת שורה):
     1. מספר הכתבה
     2. שם הכתבה
     3. תאריך פרסום הכתבה
     4. שם הכתב
     5. אתר שבו התפרסמה הכתבה
     6. הכתבה עצמה במלואה
     7. סך הטוקבקים לכתבה (מוכללים כטקסט אחד)
     8. מספר צפיות בכתבה
     9. מספר התגובות
     10. תקציר הכתבה
     11. מספר מילים בכתבה
  2. גיליון 2 – תגובות – 6 עמודות של נתונים בהתאם לפירוט הבא (כל תגובה היא מקרה נפרד ומייצרת שורה):
     1. מספר הכתבה – בהתאמה לגיליון 1
     2. אתר ממנו הורדו הכתבות
     3. שם המגיב/ כינוי
     4. נושא התגובה (למשל בynet לכל תגובה יש נושא)
     5. התגובה עצמה
     6. מספר השיחה: תגובות עצמאיות זוכות למספר רץ ואילו תגובות ששייכות לאותה שיחה – מקבלות אותו מספר
     7. מספר מזהה (0 /1)- האם זו תגובה לתגובה
     8. תאריך תגובה
  3. גיליון תמונות (בשלב הבא, לא כרגע)
     1. תמונות בתיקייה נפרדת, עמודה עם לינק באקסל לתמונה בתיקייה
     2. תאור תמונה (אם אפשר בעמודות נפרדות שוללות את בעלי המקצוע)

1. **אתר** [**WWW.QUORA.COM**](http://WWW.QUORA.COM)

חיפוש לפי מילות מפתח ב: כותרת / בטקסט עצמו/ תאור הכתב

אפשרות לסמן מינימום צופים (views), upvotes, downvotes

אפשרות לבחור קטגוריות (חיפוש חופשי)

כמו לאתרי החדשות, יש להכניס לאקסל.

1. פרטי כתבה (שאלה):

כותרת הכתבה

הטקסט של הכתבה

מספר מגיבים

מספר צופים (views)

מספר Upvotes

מספר Downvotes

תאור הכתב בכתבה (סלוגן)

קטגוריות של הכתבה (תגיות למעלה)

כל התגובות יחד כטקסט אחד בעמודה מימין

1. פרופיל הכתב:

שם הכותב

תאור (סלוגן) של הכותב

מספר שיש לו של: Answers, Questions, Activity, Posts, Blogs, Followers , Following , Topics , Edits, answer views

מקום מגורים

מקום לימודים

1. תגובות (בלשונית נפרדת באקסל, עם ID של הכתבה אליה התגובה שייכת):

תאריך

טקסט הכתבה

שם כתב

סלוגן כתב

תאור הכתב בכתבה (סלוגן)

מספר צופים (views)

מספר Upvotes

מספר Downvotes

1. הערות לתגובה (לשונית נפרדת, עם ID של התגובה לשיוך)

שם מגיב,

טקסט תגובה

1. **קיקסטרטר**

אפרט בהמשך. כרגע לא בטוח.