



ג' בכסלו תש"פ  
1 בדצמבר, 2019

## "מידע וטקסט"

עלון קבוצת עניין אחזור מידע וטקסט - SIGTRTS  
חוברת 2 כרך כו' – דצמבר 2019

### חדשות מקבוצת העניין

#### שלום לכולם!

אנו נפגשים שוב בעלון חדש נוסף. עלון זה פותח את כרך כו' של הקבוצה. אני מאחל לכולנו המשך הנאה, עניין ופעילות ברוכה בתחום אחזור המידע.

ראוי לציין שנושאים רבים הנידונים בקבוצה נקבעים ע"י החברים עצמם בדפי המשוב בסיום המפגשים. אתם מוזמנים להמשיך ולהציע נושאים לדיון וכמובן להציע את עצמכם להרצאה.

אני מזכיר לכם את כתובת האתר של הקבוצה [www.sigtrts.org](http://www.sigtrts.org). באתר חומרים רבים שנאספו במשך שנים רבות בתחום. אנה הפיצו בין חבריכם את כתובתנו ועודדו אותם להצטרף לרשימת התפוצה (הרשימה מפוקחת על ידי והתעבורה בה נועדה לעדכונים בלבד).

#### 1. קשר

הקבוצה מקימת קשר עם חבריה באמצעות קבוצת דיוור אלקטרונית (Mailing-list) המאפשרת בצורה חופשית לכל אדם להירשם לקבוצה או למחוק עצמו ממנה. הדרכה כיצד להצטרף או לעדכן כתובת מופיעה באתר תחת התפריט "רשימת תפוצה של הקבוצה".

ספריות וגופים מרכזיים אחרים המעוניינים לקבל עותק מודפס של העלון צריכים לפנות בבקשה מיוחדת ליו"ר הקבוצה.

הרוצים להיות חברים בקבוצה צריכים להירשם לרשימת התפוצה האלקטרונית שלה באתר הקבוצה.

באתר הקבוצה מפה שנועדה להקל על ההגעה של החברים מחוץ לעיר. הדפיסו אותה לפני היציאה. אם אתם עושים שימוש בניווט לוויני (GPS) כונו אותו לשע"ם, רחוב פועלי צדק 4, ירושלים. נסיעה טובה!

מאז הוצאת החוברת האחרונה (כרך כו' חוברת מס. 1) ביוני 2019 נפגשה הקבוצה פעמיים.

המפגש הראשון התקיים ביולי 2019

- א הרצאה ראשונה: שיחה, שתיקה ומה שביניהן: כריית מידע והפקת ידע מאודיו – ע"י דר' ורד זילבר-ורוד  
דר' ורד היא בלשנית בהכשרתה ובהתמחותה ומנהלת המעבדה למדיה ומידע באו"פ. הוצגו דוגמאות של הקלטות מהתקשורת שהגישה אליהם מוגבלת שלא כמו טקסטים כתובים שניתן לבצע חיפוש עליהם. הוצגה ההתפתחות של התחום בעולם ואת התחלת התמלול לעברית רק לפני כ- 3 שנים. הוצגו ניתוחי שיחות מהתקשורת הישראלית בהווה.

המפגש השני התקיים באוקטובר 2019

- א. הרצאה ראשונה: "תיקון סופרים", שימוש בחכמת ההמון לתיקון ושיפור יכולת זיהוי אוטומטי של כתב יד – ע"י צבי קופליק ומשה לביא.  
הוצגו מטרות הפרויקט ביניהן אימון וזיהוי של כתב יד ע"י מנוע הזיהוי, פענוח של כתב היד ושכלול היכולת של קריאת כתב יד אוטומטית.  
הוצגו תכונות מנוע הזיהוי ויכולותיו. הוצגה שיטת העבודה והודגמו ביצועי המשתמשים.
- ב. הרצאה שנייה: השילוב של למידה עמוקה בעולם עיבודי הטקסט – האם מהפיכה ב-NLP? – ע"י אורי חנני.  
היעד המרכזי לעמוד במבחן טורינג, לאפשר לכלי למידה עמוקה לייצר הבנה עמוקה של טקסט שתשווה ליכולות של אדם. לצד זיהוי הטקסט קיימת יכולת רבה בזיהוי תמונות הודות לפרויקט נרחב של גוגל. ניתנה סקירה של עולם ה-NLP כאשר היעד הוא להחליף את היכולות האישיות של בני אדם בהכנת חומרים והבנתם ע"י אלגוריתמים ממוחשבים.

א. האם התוצאות של גוגל אמת אבסולוטית? פורסם ב-BBC

ב. פרס מרכוס כ"ץ לפרופ' יעקב שוויקה

קריאה מהנה.

**חסות** .4

נכון להיום הקבוצה ללא חסות.  
נשמח לקבל הצעות לחסות מארגונים בתחום העיסוק של הקבוצה.

**כללי** .5

קבוצת הענין "אחזור מידע וטקסט" (SIGTRS) מהווה פורום לאנשי מקצוע העוסקים בתחום אחזור טקסט, אחזור מידע וטכנולוגיות קשורות. אנשי המקצוע הם מפתחים (מנתחי מערכות ותוכניתנים) ו/או משתמשים.  
הקבוצה הוקמה בשנת 1994 ע"י החתום מטה ומאז פועלת ברציפות הן במפגשי הרצאות והן בהפצת מידע בין היתר באמצעות עלון הקבוצה.  
הקבוצה נפגשת ארבע פעמים בשנה להרצאות ולהחלפת רעיונות.  
עלון הקבוצה יוצא פעמים בשנה בקביעות משנת 1994. כל חומר העלון בטקסט מלא נמצא באתר הקבוצה.  
הצעות להרצאות, רעיונות או חומר כתוב אחר ניתן לשלוח ליו"ר הקבוצה :

עפר דרורי  
שע"ם

[offerd@gmail.com](mailto:offerd@gmail.com)

טל: 02-5688439

ב ב ר כ ה

עפר דרורי  
יו"ר הקבוצה

בקרו אותנו באתר הבית של הקבוצה : <http://www.sigtrs.org>