**סוגי ישויות (עם רמות וודאות שונות)**

1. NAME
   1. Person
   2. Company
   3. Location
2. Numbers
   1. Date
   2. Currency
   3. Other numbers

**אינדיקטורים- ממושקלים, אופציונלים**

* ניתוח תחבירי- שע"צ
* פייסבוק- פרופיל, page, places
* google maps- מקומות
* linkedin- חברות, אנשים
* שם domain- חברה
* אותיות גדולות
* מרכאות
* מילים מרמזות- mr. dr., from, to,
* חיפושים קודמים- אולי במקום לחפש
* לבדוק מילה, ולהרחיב למילים לפני / אחרי (או- להסתכל על זוגות, שלשות)- פרמטר?
* מילון

**פעולות הכנה**

* למצוא מאגרים
* חלוקה למשפטים
* להוריד upper case ?
* tokenization
* להסיר stop words

**ניסויים**

* בדיקה מול קורפוס (אם יש)
* בדיקה מול סטנפורד / כלי קיים
* בדיקה מול בדיקה טריוויאלית של אות גדולה

**פינצ'רים נוספים**

* מידע נוסף שנמצא בגוגל- למשל על אדם- מין, גיל, חברה- במה מתעסקת...