



Annotációs eszközök

Excel

- Táblázatos adatok
- Néhány annotálendő kategória
- Szótárszerkesztés, szemantikai jegyek, kulcsszavazás...
- + Egyszerű kezelni
- + Könnyen elérhető
- + Szűrés, statisztika elérhető
- Nem mindig átlátható
- Nincs GUI



TextAnnotator

- Nyers szöveg (txt)
- Többféle (hierarchikus) annotációs réteg
- Token és frázis szint is
- Bizonytalanság, többszavas kifejezések, tulajdonnevek, véleménydetekció, CV-k...
- + Egyszerű kezelni
- + Ingyen elérhető
- + Áttekinthető annotáció (GUI)
- Testre szabás kell (minimális kódolás)
- Az annotáció kigyűjtése (kódolás)



MMAX

- Token, frázis és szöveg szintű annotáció
- Egy elemhez többféle választási lehetőség
- Hierarchikus jelölés
- Szavak közti relációk jelölhetők
- Szófaji egyértelműsítés, szintaxis, koreferencia...
- + Sokrétű feladatra egy felület
- + Ingyen elérhető
- XML-formátum (előelemzés, konvertálás)
- Az annotáció kigyűjtése (kódolás)



VisDic

- Synset szintű annotáció
- Szavak hálónba kapcsolása
- Wordnet építése, ontológiák kezelése...
- + **Egyszerű kezelni (GUI)**
- + **Ingyen elérhető**
- + **Célfeladat sajátosságaira van szabva**
- + **Keresést is támogat**
- + **Többnyelvű keresés / adatszerkesztés**
- **XML-formátum (konvertálás)**
- **Kisebb működési problémák (bugok)**



SenseTagger

- Token, frázis és szöveg szintű annotáció
- Egy elemhez többféle választási lehetőség
- Kontextus megjeleníthető
- Jelentés-egyértelműsítés, véleménykinyerés...
- + Egyszerűen kezelhető (GUI)
- + Ingyen elérhető
- XML-formátum (előelemzés, konvertálás)
- Az annotáció kigyűjtése (kódolás)



TrEd

- Mondatszintű annotáció
- Perl
- Függőségi fák (dependencia)
- + Egyszerűen kezelhető (GUI)
- + Átlátható ágrajzok
- + Könnyen mozgatható csomópontok
- + Ingyen elérhető
- fs formátum (konvertálás, fs2conll, conll2fs)



Saját projektek

- Milyen szövegek?
- Mennyi szöveg?
- Szövegek forrása (web, MEK...)
- Mi az annotációs feladat?
- Mik az annotálandó kategóriák?
- Mi az annotálás szintje?
- Milyen eszköz lenne jó a célra?

