Yapılacaklar:

1. Templatelere göre textlerden veri çekilecek.
2. Template verileri ile infobox verilere karşılaştırılacak.
   1. Sorunlar
      1. Metin içinde geçmiyor bazıları?
         1. (http://tr.wikipedia.org/wiki/Bruce\_Cohen)
         2. (<http://tr.wikipedia.org/wiki/Carlos_S%C3%A1nchez_Moreno>)
      2. Yanlış yazılanlar var. Bunlar başarımı dşrecektir?
         1. (Ahmet Çakar (d. 1962, Çankırı, Türkiye)sonuç-->(İstanbul))
3. Hatalara bakılacak.
4. Gerekiyorsa templatelerde düzenleme yapılacak.
5. Gerekiyorsa yeni templateler tanımlanacak.
6. DB deki aşağıdaki alanlar için tum işlemler tekrarlı yapılacak.
   1. isim
   2. doğum\_
   3. doğum\_yeri
   4. ölüm\_tarihi
   5. ölüm\_yeri
   6. diğer\_adları
   7. meslek
7. İl ilçe gibi bilgilerde templatelerde çıkartılacak.

Yapılanlar:

1. DB de doğum yeri alanı ayrıştırıldı.
2. Templateler için örnek cümleler çıkartıldı.( TemplateOrnekDoğumYeri.docx)(02.05.2012)
3. Templateler çıkartılacak .( TemplateOrnekDoğumYeri.docx (02.05.2012)
4. Tum templateler implement edildi. (13.05.2012)
5. Bulunan keliemlerin ekleri çıkartıldı. (13.05.2012)
6. Bulunan kelimelerin sayı olamaması gerek bu kontrol edildi. (13.05.2012)
7. Eğer bulunan yer ismi “ili” içeriyorsa atılıyordu. Ancak yanlış sonuçlar rettiğinden burası değiştirildi.”ilini” yapıldı.
8. Bulunan değerlerin veritabanına performanslı bir şekilde yazdırılması yapıldı.
9. Sonuçlar:
   1. Toplam 2400 kayıt var.
   2. 619 tanesinin infobox da doğum yeri yok

( select count(\*) from infobox ib where ib.p\_dogum\_yer = "null" or ib.p\_dogum\_yer is null)

* 1. Kalan kayıt 1780
  2. 905 doğru bulunan
  3. 885 yanlış bulunan
     1. 935 tanesini boş (Bunların 388 tanesi zaten yok infodada boş)
     2. İnfoda dolu olup benim bulamaddığım yada yanlış bulduğum=550 tane

1. 550 adet hiç bulunamayana bakılacak.

Başaramadıkları:

jan legindork

12353069

12424575

12594422