

# ARAMA MOTORU UYGULAMASI

*M.Furkan Tufan ve Muhammed Şara*

150201111 ve 150201123

Bilgisayar Mühendisliği Bölümü  
Kocaeli Üniversitesi

furkantufan0127@gmail.com ve muhammed.sara271@gmail.com

## Özet

Bu projede bizden istenilen web ortamında arama motoru geliştirmektir. Bu işlemi gerçekleştirirken kullanıcıdan site ve kelime bilgisi girmesini istedik. Girilen kelimeyi sitede arama işlemi gerçekleştirip eğer kelime sitede varsa kaç defa geçtiğini bulup buna göre bir skor tablosu geliştirdik. Bu skor tablosuna göre siteleri sıralama işlemi gerçekleştirdik. Projedeki bir diğer amaç ise kullanıcı tarafından girilen kelimenin eş anlamlısına göre de sitede arama yapabilmektir. Bunun için kelimenin eş anlamlısını veren bir site bulmamız gerekti. Kullanıcı tarafından girilen kelimeyi bu siteye yollayıp siteden gelen cevabı yani eş anlamlı kelimemizi bir değişkene atayıp, hem kullanıcının girdiği kelimeyi hem de kullanıcı tarafından girilen kelimenin eş anlamlısını kullanıcının girdiği sitelerde arama işlemi yaptık. Bu adımda da geçen kelime sayısına göre bir skor algoritması oluşturup bu skorlara göre siteleri sıralama işlemi gerçekleştirdik.

## 1.Problem tanımı

Projede bize verilen problem kullanıcının gireceği kelime ve site bilgisine göre girilen kelimenin girilen url de kaç defa geçtiğini bulma, birden fazla kelime ve birden fazla url için bu işlemleri gerçekleştirebilme, sitede yapılan arama sonucunda girilen kelimenin kaç defa geçtiğine göre bir skor tablosu oluşturup bu skorlara göre siteleri sıralama işlemi, girilen kelimenin eş anlamlısını bulup eş anlamlısına göre de arama ve sıralama işlemi yapabilme ve son olarak da bu uygulamanın çevrimiçi olarak yayınlanması işlemidir.

## 2.Proje ile ilgili Temel Bilgiler

Projemizi php dili kullanarak geliştirdik. Uygulamamızın site tasarımını html ve css kodları ile gerçekleştirdik. Derleyici olarak phpstorm derleyicisini kullandık.

### 2.1 Yapılan Araştırmalar

Projemize başlamadan önce hangi dil ile uygulamamızı geliştireceğimize karar verdik. Php dili ile geliştirmeye karar verdikten sonra kullanıcı tarafından girilen url ve kelime bilgisine göre url deki tüm bilgileri çekmemiz gerekiyordu. Bunun için php simple html dom parser kütüphanesini kullandık. Bu kütüphanenin bize sağladığı avantaj girilen sitenin tüm html kodlarını çekip bir string değişkene atayabilmektir. Bu işlemi gerçekleştirdikten sonra string arama fonksiyonları ile site içerisinde kelime arama işlemi gerçekleştirdik. Bu arama sonucuna göre girilen kelime site içerisinde geçiyorsa bu kelimenin kaç defa bu sitede gerçekleştiğini bulduk. Buna göre bir site sıralama algoritması geliştirdik. Bundan sonraki aşama ise kullanıcı tarafından girilen kelimenin eş anlamlısına erişmek oldu. Bunun için kelimenin eş anlamlısını veren bir site bulduk.[11] <https://www.dersimiz.com/yakin-ve-es-anlamli-anlamdas-kelimeler-sozlugu.asp?> Bu siteden faydalanarak kelimenin eş anlamlısını çektik. Bu kelimeyi de sitede arama yapma fonksiyonumuza gönderdik. Bu sayede bizden istenilen aşamaları gerçekleştirmiş olduk.

## 2.2 Yazılım Mimarisi

*ara(\$stringUrl,\$stringKelime) fonksiyonu:* Bu fonksiyon ile kullanıcı tarafından girilen sitede kelime arama işlemini gerçekleştiriyoruz. Fonksiyonumuz iki parametre içermektedir. Bu parametreler kullanıcı tarafından girilen url ve kelime değişkenleridir. Bu işlemi gerçekleştirirken string fonksiyonlarından strpos fonksiyonunu kullandık. Bu fonksiyonun bize avantajı site içerisinde kelimeyi ararken eğer kelimeyi bulursa kelimenin indexini bize vermesidir. Kullanıcının girdiği kelimeyi arayacağımız için örneğin kullanıcı arı kelimesi aradığında başarı kelimesindeki arıyı da alıyordu. Strpos fonksiyonu sayesinde kelimenin indexine erişebildiğimiz için kelimenin bir önceki indisine de erişebiliyoruz. Eğer kelimenin bir önceki veya bir sonraki indisi herhangi bir harfe eşit ise aradığımız kelime olamayacaktır. Arama fonksiyonumuz bu şekilde çalışmaktadır.[13]

*urlara(\$stringUrl,\$stringKelime) fonksiyonu:* Bu fonksiyon ile kullanıcı tarafından girilen sitedeki alt siteleri arıyoruz. Bunu yaparken kullanıcı sitenin urlini girdiği için url adı elimizde olacaktır. Strpos fonksiyonu ile urlden sonraki indexe erişebildiğimiz için urlin sonraki değeri “\” karakterine eşit olduğu durumda o sitenin alt linkini bulmuş oluruz.[13]

*replace\_tr(\$text) fonksiyonu:* Bu fonksiyon ile kelimelerde Türkçe karakter düzeltme işlemi gerçekleştiriyoruz. Eğer kullanıcı tarafından girilen kelime grubu Türkçe karakter içeriyorsa aranan sitede ise aynı kelime Türkçe karakterler olmadan geçiyorsa bunun kontrolünü yapıyoruz. Örneğin güvenlik kelimesini aradık ve sitede bu haliyle değil de guvenlik olarak olarak geçiyorsa bu kelimeyi de saydırmamız gereklidir. Bunun düzeltmesini bu fonksiyon ile gerçekleştirmektediriz.

*selectionSort(\$array) fonksiyonu:* Siteleri sıralama işlemi gerçekleştirir. Bu algoritma ile skoru en fazla olan siteye göre sıralama işlemi gerçekleştirir.

*skor(\$kelimeler,\$kelimesayi) fonksiyonu:* Bu fonksiyon ile kullanıcı tarafından girilen kelimeyi

ve kelimenin kaç defa sitede geçtiği parametrelerini içermektedir. Parametreler kullanıcı birden fazla kelime girdiyse kelimeler dizisi ve kelimelerin kaç defa geçtiği bilgisini içeren diziye dönüşürler. Bulunan kelimelerin sayıları karşılaştırılır. Hangi kelime en çok geçerse onun puanı daha yüksektir. Örneğin 2 sitede 3 farklı kelime arıyoruz. 1.sitede kelimelerimiz 50-40-50 defa geçsin diğerinde ise 10-20-10 geçsin. Burada en çok geçen kelimeleri ve kendi kelime grupları arasındaki farkı en az olan sitenin puanı daha yüksek olacaktır.

```
$toplampuan=0;
for($i=0;$i<$kelimesayi;$i++){
    for($j=$i+1;$j<$kelimesayi;$j++){
        $puan=((($kelimeler[$i]-
$kelimeler[$j])/1)*10);
        $toplampuan=$toplampuan+$puan;
    }
}
```

**return** \$toplampuan;

*esanlami(\$kelime) fonksiyonu:* Bu fonksiyon ile linkinin kaynakça da belirttiğim siteden eş anlamlı kelimeyi bulmamızı sağlıyor. Eş anlamlı kelimeyi bulduğumuz siteye erişmek için curl\_setopt fonksiyonunu kullandık. Örnek kod aşağıda verilmiştir.[12]

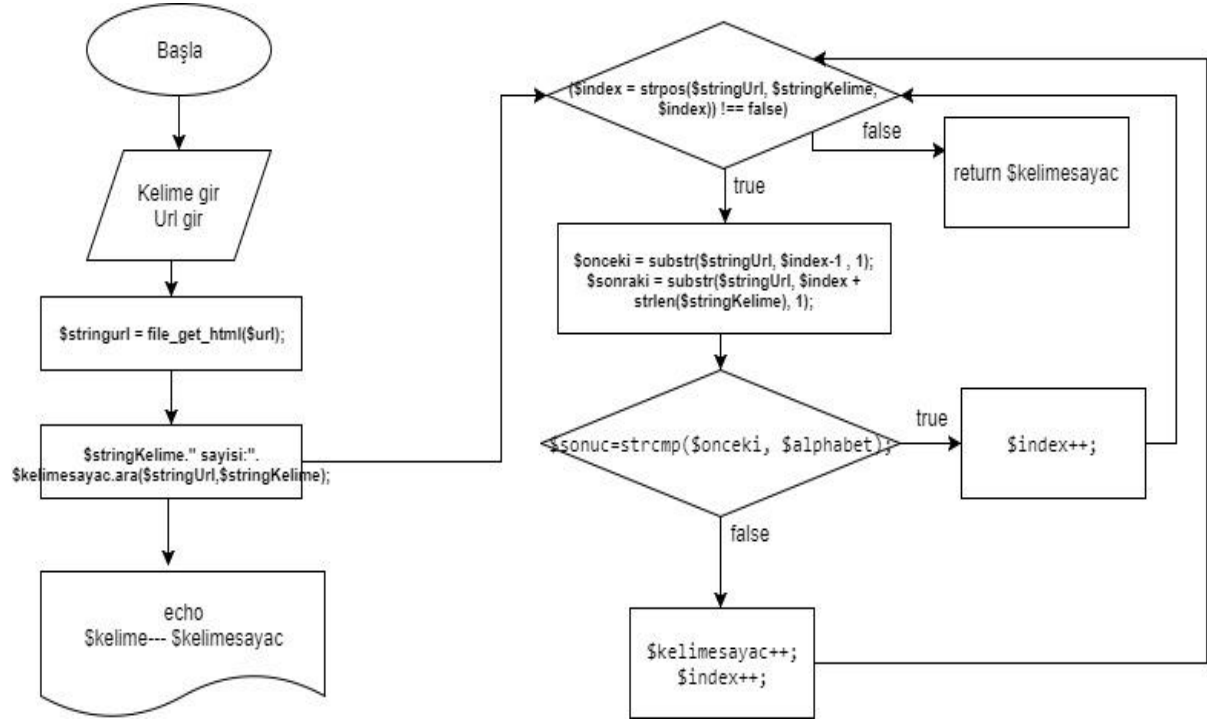
```
curl_setopt($ch, CURLOPT_URL, $url);
curl_setopt($ch, CURLOPT_RETURNTRANSFER, true);
curl_setopt($ch, CURLOPT_BINARYTRANSFER, true);
curl_setopt($ch, CURLOPT_SSL_VERIFYPEER, false);
curl_setopt($ch, CURLOPT_FOLLOWLOCATION, true);
curl_setopt($ch, CURLOPT_POST, true);
curl_setopt($ch, CURLOPT_POSTFIELDS, $data);
curl_setopt($ch, CURLOPT_CONNECTTIMEOUT, $zaman);

$sonuc = curl_exec($ch);
```

\$sonuc değişkeni bize eş anlamlı kelimeyi içermektedir. Bu kelimeyi arama fonksiyonumuza gönderdiğimiz zaman artık eş anlamlı kelimeleri de bulup skorunu üretecektir.

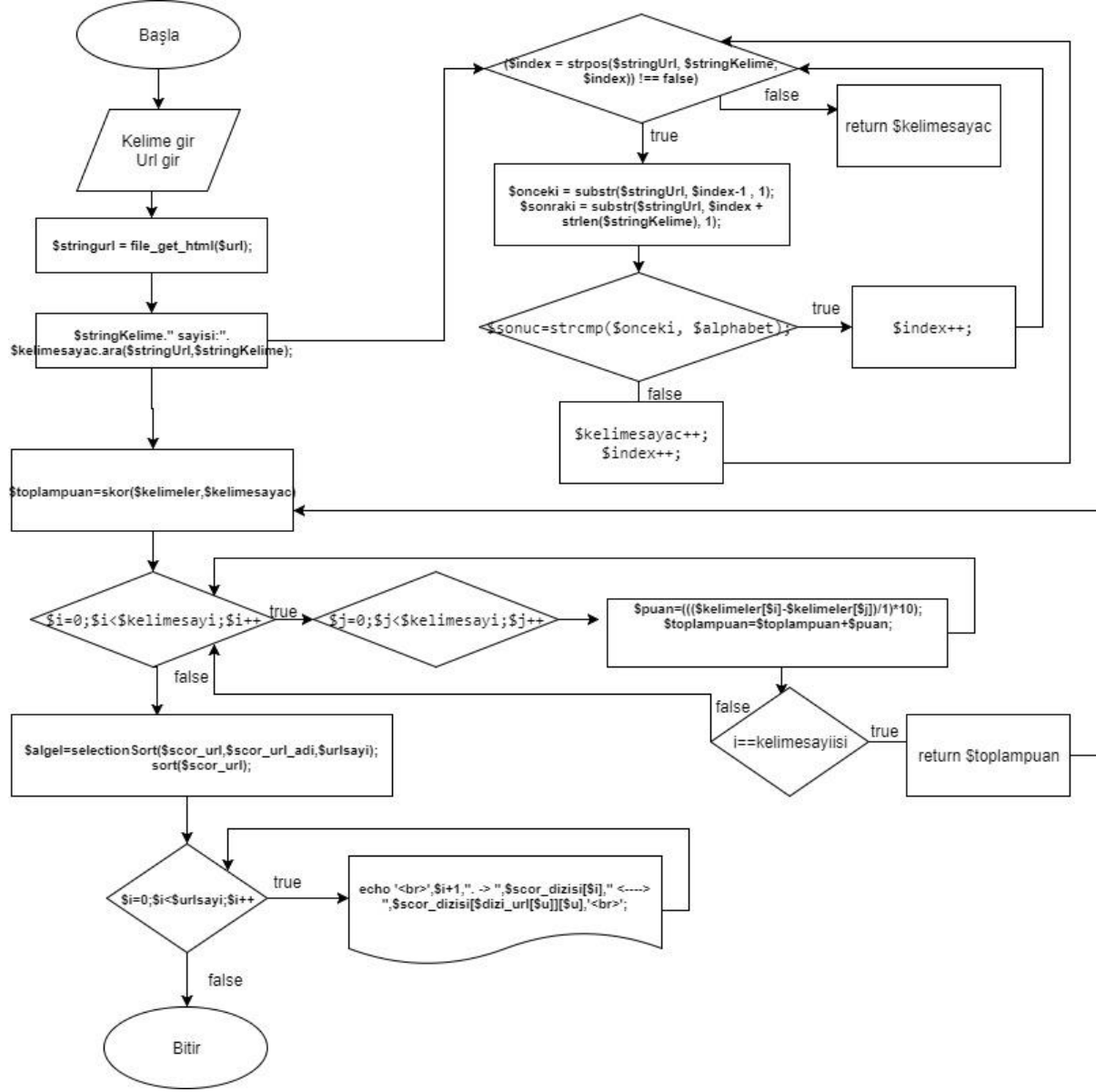
### 3.Akış Şeması

#### Arama-1



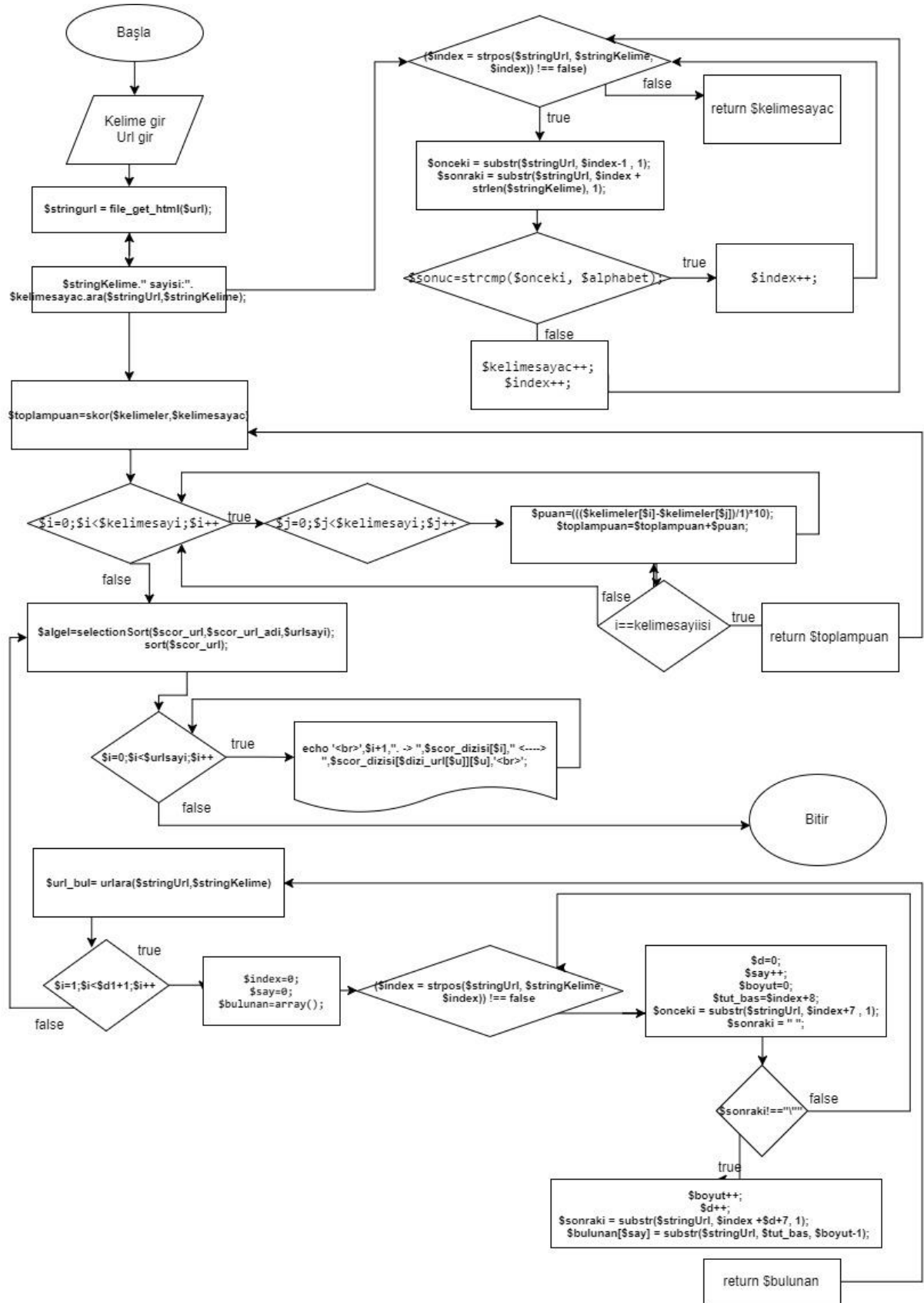
Resim-1

## Arama-2



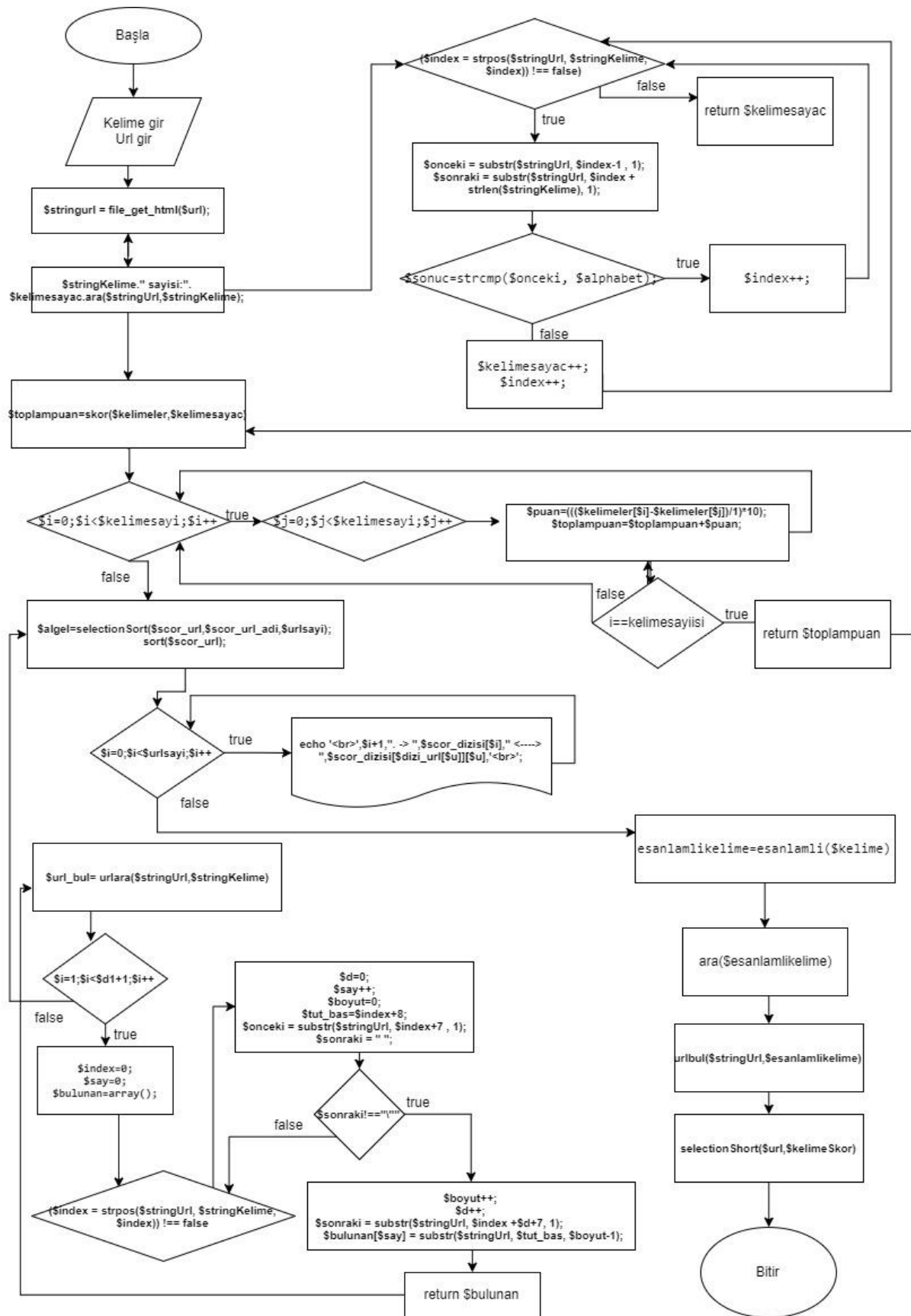
Resim-2

### Arama-3



Resim-3

## Arama-4

*Resim-4*

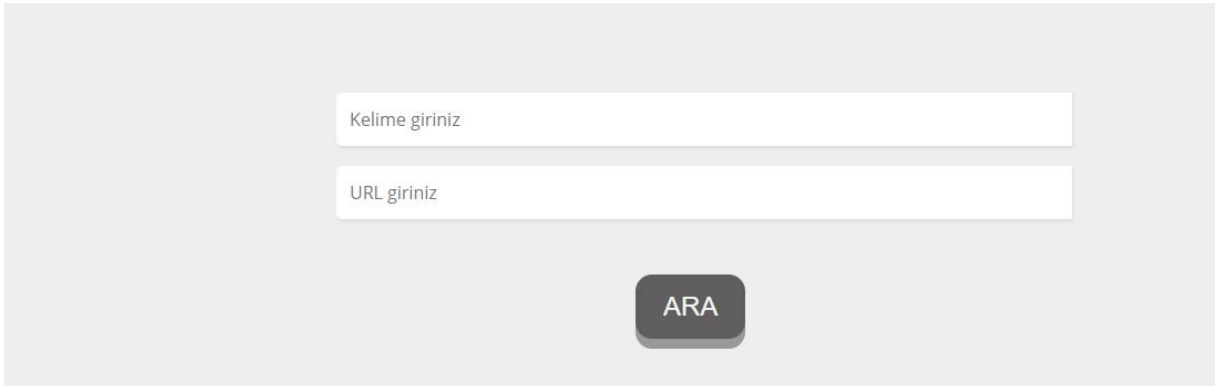
#### 4.Kullanım Kılavuzu

Yerli Arama Motoru



Resim-5

<http://muhammedsara.com/yazlab1-3/anasayfa.php> adresinden projemizi çalıştırdığımızda yukarıdaki resimde görüldüğü gibi bir ekran karşımıza çıkacaktır. Proje dokümanında bizden istenilen arama işlemlerine göre seçim yaptıktan sonra karşınıza resim-2 de verilen ekran gelecektir. Bu ekran da aranacak kelime ve aranacak url bilgileri kullanıcı tarafından doldurulması gereklidir. Daha sonra arama işlemi için ara butonuna basılmalı ve kısa bir süre beklenmelidir.



Resim-6

## 5. Genel Yapı

Proje php dili kullanılarak geliştirilmiştir. Projemizde kullanıcı tarafından girilen kelime veya kelime gruplarını bir veya birden fazla sitede arama işlemi gerçekleştirebilmektedir. Url'eri kelimelerin url'de kaç defa geçtiğini bularak sıralamaktayız. Sıralama algoritması olarak selection short algoritması kullandık. Kullanıcı tarafından girilen kelimenin eş anlamlısını bulmak için linkini kaynakçada belirttiğim siteden faydalandık. Kullanıcı tarafından girilen siteyi php simple html dom parser kütüphanesi kullanarak çekip bir stringe atama işlemi gerçekleştirdik. Tasarım olarak css ve html kullandık. Projede bizden istenilenleri gerçekleştirmiş olduk.

## 6.Referanslar

- [1] <https://umitunal.wordpress.com/2010/05/29/php-simple-html-dom-parser-kullanimi/>
- [2] <http://pr.efactory.de/e-pagerank-algorithm.shtml>
- [3] <http://www.r10.net/php/1220400-simple-html-dom-parser-nedir-nasil-kullanilir.html>
- [4] <https://stackoverflow.com/questions/14962359/how-to-use-file-get-contents-or-file-get-html>
- [5] <https://stackoverflow.com/questions/7118823/heck-if-url-has-certain-string-with-php>
- [6] <http://www.free-css.com/free-css-templates/page1>
- [7] <https://www.w3schools.com/CSSref/>
- [8] <https://www.w3schools.com/php/>
- [9] <http://php.net/manual/tr/function.strpos.php>
- [10] [http://www.phpdefteri.com/icerik/35/string\\_fonksiyonlari.html](http://www.phpdefteri.com/icerik/35/string_fonksiyonlari.html)
- [11] <https://www.dersimiz.com/yakin-ve-es-anlamli-anlamdas-kelimeler-sozlugu.asp?>

[12] [http://www.gencprogramci.net/sayfalar/icerik\\_perma/155/php-bot-yapimi-uzak-siteden-vericekmek-curl](http://www.gencprogramci.net/sayfalar/icerik_perma/155/php-bot-yapimi-uzak-siteden-vericekmek-curl)

[13] <https://stackoverflow.com/questions/5465690/php-simple-html-dom-parser-find-string>