



#### Düzenli İfadeler & Kullanım Alanları

sezai.yilmaz@pro-g.com.tr

## Düzenli İfadeler - İçerik



- Düzenli İfadeler (ing. regular expressions)
- Sablonlar (ing. patterns)
- Düzenli ifade şablonları (ing. regular expression patterns)
- Düzenli ifadelerde kullanılan karakterler
- Örnekler
- Kullanım alanları
- Standartlar
- Kaynakça

## Belge Lisansı – GFDL



Bu belgeyi, Free Software Foundation tarafından yayınlanmış bulunan GNU Özgür Belgeleme Lisansının 1.2 ya da daha sonraki sürümünün koşullarına bağlı kalarak kopyalayabilir, dağıtabilir ve/veya değiştirebilirsiniz. Lisans'ın bir kopyasını http://www.gnu.org/copyleft/gfdl.html adresinde bulabilirsiniz.

Bu belgedeki bilgilerin kullanımından doğacak sorumluluklar ve olası zararlardan belge yazarı sorumlu tutulamaz. Bu belgedeki bilgileri uygulama sorumluluğu uygulayana aittir.

Tüm telif hakları aksi özellikle belirtilmediği sürece sahibine aittir. Belge içinde geçen herhangi bir terim, bir ticarî isim ya da kuruma itibar kazandırma olarak algılanmamalıdır. Bir ürün ya da markanın kullanılmış olması ona onay verildiği anlamında görülmemelidir.

### Düzenli İfadeler



- Belirli bir kurala (şablona) uyan dizgeleri eşleyen bir formüldür
  - ilk bakışta çok karmaşık olduğu düşünülür
  - [+-]?(([0-9]{1,3}(\.[0-9]{3})\*)|([0-9]+))(, [0-9]+)?
  - Biraz pratikle çok kolay yazılabildiği görülür
- Kullanım amacı
  - Yoğun veriden ihtiyaç duyulan bilginin çekilmesi
  - Kullanıcı girdisinin denetimi
  - Verinin amaca uygun biçime sokulması

## Şablonlar (ing. Patterns)



- Aranan karakter dizgesidir (ing. search string)
- Karakterlerin yan yana yerleştirilmesi ile meydana gelen dizgelerdir
  - penguen
    - Metin içinde penguen dizgesini eşler
  - → H20
    - Metin içinde H2O dizgesini eşler

## Şablonlar - Örnek



```
# grep system /etc/passwd
rpm:x:13:101:system user for rpm:/var/lib/rpm:/bin/false
rpc:x:70:70:system user for portmap:/:/bin/false
xfs:x:71:71:system user for XFree86:/etc/X11/fs:/bin/false
apache:x:72:72:system user for apache-conf:/var/www:/bin/sh
postfix:x:73:73:system user for postfix:/var/spool/postfix:/bin/false
rpcuser:x:74:74:system user for nfs-utils:/var/lib/nfs:/bin/false
sshd:x:75:75:system user for openssh:/var/empty:/bin/true
ftp:x:76:76:system user for proftpd:/var/ftp:/bin/false
squid:x:77:77:system user for squid:/var/spool/squid:/bin/false
mysql:x:78:78:system user for MySQL:/var/lib/mysql:/bin/bash
qdm:x:79:79:system user for qdm:/var/lib/qdm:/bin/false
ups:x:80:423:system user for nut:/var/state/ups:/bin/false
postgres:x:81:424:system user for postgresgl:/var/lib/pgsgl:/bin/bash
snort:x:82:426:system user for snort:/var/log/snort:/bin/false
# grep root /etc/passwd
root:x:0:0:root:/root:/bin/bash
```

## Düzenli İfade Şablonları



- Aranan, kurallı karakter dizgesidir
- Özel karakterlerin (ing. meta characters) yardımıyla, şablonun belirli bir formüle uydurulması
  - se[zv]gi
    - Sadece sezgi ve/veya sevgi dizgelerini eşler
  - **1234?** 
    - Sadece 123 ve/veya 1234 dizgelerini eşler

### Düzenli İfade Karakterleri



- Alfasayısal karakterler
  - Sıradan harfler, rakamlar
- Meta karakterler
  - Özel anlam taşıyan noktalama işaretleri

## Alfa Sayısal Karakterler



#### Harfler

- a, b, c, d, ..., z

#### Rakamlar

© 0, 1, 2, 3, ..., 9

#### Düzenli ifade içinde her biri kendini ifade eder

- unix
  - u, n, i, x harflerinin yanyana bulunduğu dizgeleri eşler
- **14** 
  - 1, 4 rakamlarının yanyana bulunduğu dizgeleri eşler

## Karakter Değişkenleri



- "." Herhangi bir karakteri eşler
  - a.i
    - ..., a3i, a4i, ..., aAi, aBi, ..., aai, abi, aci, adi, ...,
- "[...]" Listelenmişlerden birisini eşler
  - a[lbd]i
    - ali, abi ve/veya adi dizgeleri eşlenir
- "[^...]" Listelenmemişlerden birisini eşler
  - a[^lbd]i
    - ..., a,i, a3i, a4i, ..., aAi, aBi, ..., aai, aci, aei, ...,

### [...] Aralık Tanımı



- "-" karakteri kullanılır
- ☞ [a-z]
  - ASCII tablosundaki a'dan z'ye karakterlerin hepsi
  - Solunda ve sağında karakter olmalı
  - Soldaki karakter sağdakinden küçük olmalı
- ☞ [-Z]
  - "-" veya "z" karakteri, "-" nin aralık anlamı yok
- ☞ [a-]
  - "a" veya "-" karakteri, "-" nin aralık anlamı yok

## [^...] Değilleme



- [^A-Z]
  - ASCII tablosunun A'dan Z'ye kadar olan büyük harfler bölümü haricindeki diğer tüm karakterler
  - "^" karakteri "[" karakterinden sonraki ilk karakter olduğu sürece değilleme anlamı taşır
- [^abc^0-9]
  - ASCII tablosundaki a, b, c, ^, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 karakterleri haricindeki tüm karakterler

# [...] Diğer Özelliği



Özel anlamı olan karakterlerin kendilerini ifade etmek için kullanılır

- - "." karakterinin kendisini ifade eder
- **⑤** 
  - "[" karakterinin kendisini ifade eder
- (\*)
  - "\*" karakterinin kendisini ifade eder

## Basit Niceleyiciler



- ? hiç veya bir kez
  - Kendisinden önceki karakterin hiç veya bir kez yanyana bulunabileceğini belirtir
- \* hiç veya birçok kez
  - Kendisinden önceki karakterin hiç veya birçok kez yanyana bulunabileceğini belirtir
- + en az bir kez
  - Kendisinden önceki karakterin bir veya birçok kez yanyana bulunabileceğini belirtir

# Basit Niceleyiciler – Örnek



- selami?
  - selam ve/veya selami dizgelerini eşler
- selami\*
  - selam, selami, ..., selamiiii, ... dizgelerini eşler
- selami+
  - selami, selamii, ..., selamiiiii, ... dizgelerini eşler
- test[234]+
  - test2, test4, test42, test23, test3342, test222, .... dizgelerini eşler

## Karmaşık Niceleyiciler



- {n} tam olarak n kez
  - Kendisinden önceki karakter tam olarak n kez yanyana bulunabilir
- {n,m} n ile m arasında kez
  - Kendisinden önceki karakter n ile m arasında kez yanyana bulunabilir
- {n,} n'den fazla kez
  - Kendisinden önceki karakter n ve daha fazla kez yanyana bulunabilir

# Karmaşık Niceleyiciler – Örnek



- selami{4}
  - Sadece selamiiii dizgesini eşler
- selami{2,4}
  - Sadece selamii, selamiii ve/veya selamiii dizgelerini eşler
- selami{3,}
  - selamiii, selamiiii, selamiiii, .... dizgelerini eşler
- test.{2}
  - testci, tester, test;a, test%4, ... dizgelerini eşler

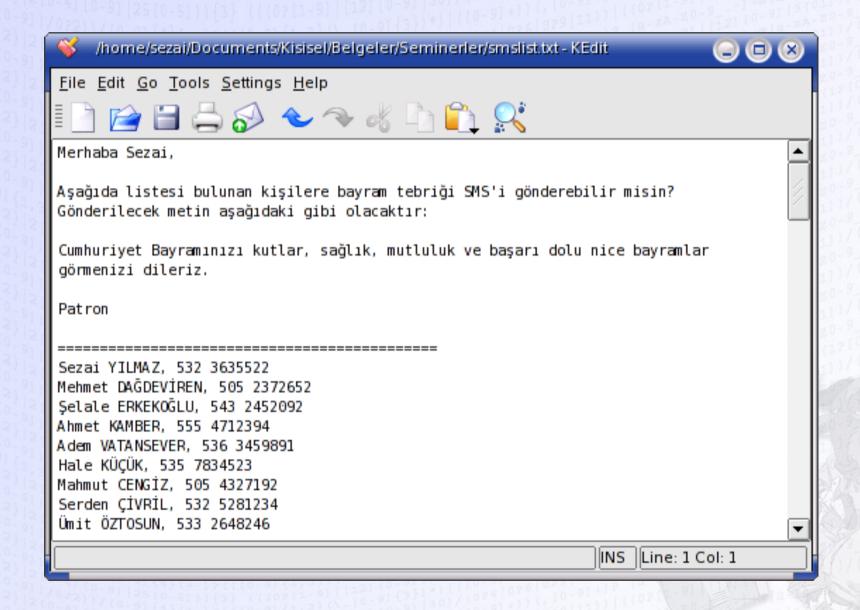
### Örnek – Telefon Numarası



- +90 312 2101494 tarzında numaralar
  - \+90 [123456789][0-9]{2} [1-9][0-9]{6}
- +90 312 2101494 veya +903122101494 tarzında numaralar
  - \+90 ?[1-9][0-9]{2} ?[1-9][0-9]{6}
- © 0(312)2101494 tarzında numaralar
  - 0\([1-9][0-9]{2}\)[1-9][0-9]{6}

### Örnek – SMS Gönderimi - 1





### Örnek – SMS Gönderimi - 2



```
sezai@mtirtil.crusoe.net:/home/sezai/Documents/Kisisel/Belgeler/Seminerler
                                                                             $ grep -P "[0-9]{7}" smslist.txt
Sezai YILMAZ, 532 3635522
Mehmet DAĞDEYİREN. 505 2372652
Selale ERKEKOĞLU. 543 2452092
Ahmet KAMBER, 555 4712394
Adem VATANSEVER, 536 3459891
Hale KÜCÜK. 535 7834523
Mahmut ČENGİZ, 505 4327192
Serden ÇİVRİL, 532 5281234
Ümit ÖZTOSUN, 533 2648246
$ grep -P "[0-9]{7}" smslist.txt | sed "s/[^0-9]//g"
5323635522
5052372652
5432452092
5554712394
5363459891
5357834523
5054327192
5325281234
5332648246
$ grep -P "[0-9]{7}" smslist.txt | sed "s/[^0-9]//g" | sed "s/^/+90/g"
+905323635522
+905052372652
+905432452092
+905554712394
+905363459891
+905357834523
+905054327192
+905325281234
+905332648246
$ grep -P "[0-9]{7}" smslist.txt | sed "s/[^0-9]//g" | sed "s/^/+90/g" > tlf.txt
```

### Örnek – SMS Gönderimi - 3



```
sezai@mtirtil.crusoe.net: /home/sezai/Documents/Kisisel/Belgeler/Seminerler
                                                                                $ cat sendsms.sh
#!/bin/bash
for telefonno in `cat $1`
do
        echo ${telefonno} SMS gönderiliyor.
        gnokii --sendsms ${telefonno} < bayrammesaji.txt</pre>
done
$ chmod 755 sendsms.sh
$ ./sendsms.sh tlf.txt
+905323635522 SMS gönderiliyor.
+905052372652 SMS gönderiliyor.
+905554712394 SMS gönderiliyor.
+905363459891 SMS gönderiliyor.
+905357834523 SMS gönderiliyor.
+905054327192 SMS gönderiliyor.
+905325281234 SMS gönderiliyor.
+905332648246 SMS gönderiliyor.
```

## Özet



- Düzenli ifade, alfasayısal ve meta karakterlerin oluşturduğu şablona uyan dizgeleri eşleyen bir formüldür
- Karakter değişkenleri "." ve "[...]"
  - "." ASCII karakterlerinden birini eşler
  - "[...]" Kümedeki karakterlerden birini eşler
  - "[a-z]" Aralıktaki karakterlerden birini eşler
  - "[^...]"Küme dışındaki karakterlerden birini eşler
- ➢ Niceleyiciler "?", "+", "\*", "{n}", "{n,m}", "{n,}"

#### Meta Karakterler



- (F)
  - Kaçış karakteri (özel anlam kaldırma)
- ♠
  - Satır başı
- **\$** 
  - Satır sonu
- (F)
  - Alternatif düzenli ifade şablonu tanımlama

## Özel Karakterler



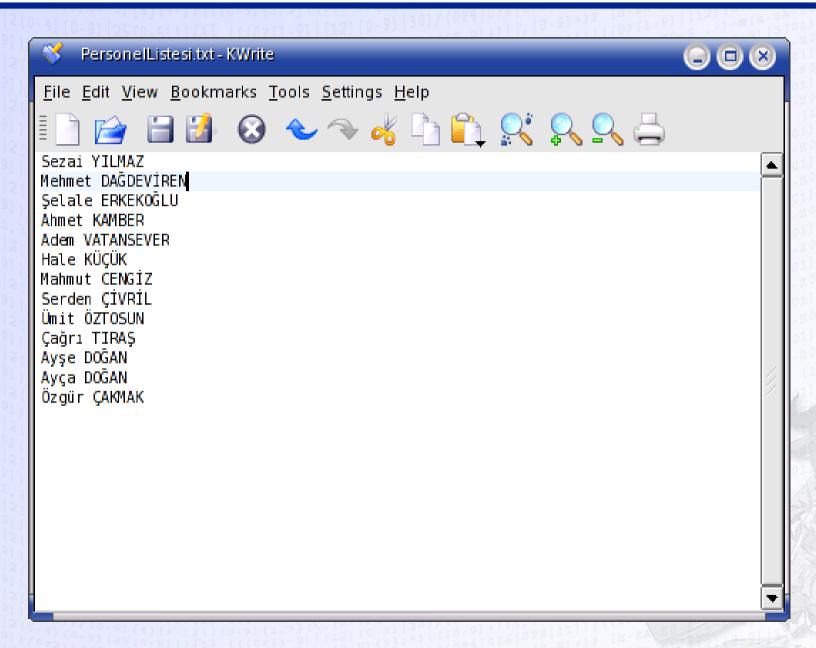
- . herhangi bir karakter
- [ karakter kümesi başlangıcı
- ? hiç veya bir kez niceleyici
- \* hiç veya birçok kez niceleyici
- + en az bir kez niceleyici
- { tam olarak belirli bir sayıda niceleyici
- ( Alt şablon başlangıç ve bitişi

### [...] Meta Karakterleri



- ₩ \
  - Genel kaçış karakteri
- - Eğer ilk karakter ise karakter kümesini değilleme
- **-**
  - Karakter aralığı tanımlama
- - Karakter kümesini sonlandırma







```
sezai@mtirtil.crusoe.net:/home/sezai/Documents/Kisisel/Belgeler/Seminerler/regexp/E-Postamek
                                                                               $ sed -r "y/$ĞÇİÜÖğışçüö/$GCIUOgiscuo/" < PersonelListesi.txt
Sezai YILMAZ
Mehmet DAGDEVIREN
Selale ERKEKOGLU
Ahmet KAMBER
Adem VATANSEVER
Hale KUCUK
Mahmut CENGIZ
Serden CIVRIL
Umit OZTOSUN
Cagri TIRAS
Ayse DOGAN
Ayca DOGAN
Ozgur CAKMAK
$ sed -r "y/$ĞÇİÜÖğışçüö/SGCIUOgiscuo/" < PersonelListesi.txt | tr "A-Z" "a-z"
sezai yilmaz
mehmet dagdeviren
selale erkekoglu
ahmet kamber
adem vatansever
hale kucuk
mahmut cengiz
serden civril
umit oztosun
cagri tiras
ayse dogan
ayca dogan
ozgur cakmak
```



```
sezaj@mtirtil.crusoe.netr/home/sezaj/Documents/Kisisel/Belgeler/Seminerler/regexp/E-Postamek
$ sed -r "y/$ĞÇİÜÖğışçüö/SGCIUOgiscuo/" < PersonelListesi.txt | tr "A-Z" "a-z" |
 sed -r "s/([^ \t]+) +([^ \t]+)/\\1.\\2@pro-g.com.tr/"
sezai.yilmaz@pro-g.com.tr
mehmet.dagdeviren@pro-g.com.tr
selale.erkekoglu@pro-g.com.tr
ahmet.kamber@pro-g.com.tr
adem.vatansever@pro-g.com.tr
hale.kucuk@pro-g.com.tr
mahmut.cengiz@pro-g.com.tr
serden.civril@pro-g.com.tr
umit.oztosun@pro-g.com.tr
cagri.tiras@pro-g.com.tr
ayse.dogan@pro-g.com.tr
ayca.dogan@pro-g.com.tr
ozgur.cakmak@pro-g.com.tr
$ sed -r "y/$ĞÇİÜÖğışçüö/SGCIUOgiscuo/" < PersonelListesi.txt | tr "A-Z" "a-z" |
sed -r "s/([^ \t]+) +([^ \t]+)/\\1.\\2@pro-g.com.tr/" > eposta.txt
$
```



- /([^\t]+) +([^\t]+)/
- sezai yilmaz
- mehmet dagdeviren
- selale erkekoglu
- ahmet kamber
- adem vatansever
- hale kucuk
- mahmut cengiz



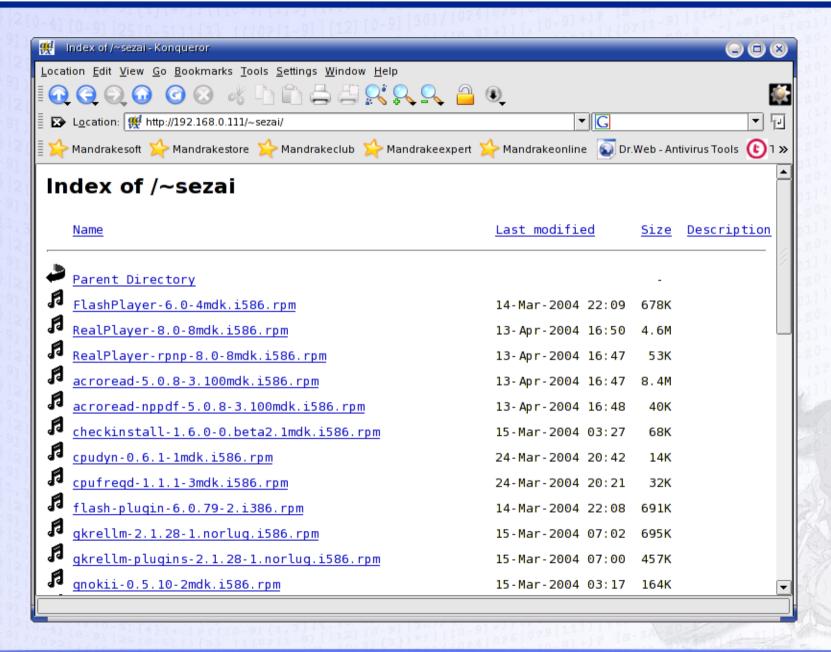
```
sezai@mtirtil.crusoe.net:/home/sezai/Documents/Kisisel/Belgeler/Seminerler/regexo/E-Postamek
$ sed -r "y/ŞĞÇİÜÖğışçüö/SGCIVOgiscuo/" < PersonelListesi.txt | tr "A-Z" "a-z" |
 sed -r "s/([^ \t])[^ \t]+ +([^ \t]+)/\\1.\\2@pro-g.com.tr/"
s.uilmaz@pro-g.com.tr
m.dagdeviren@pro-g.com.tr
s.erkekoglu@pro-g.com.tr
a.kamber@pro-g.com.tr
a.vatansever@pro-g.com.tr
h.kucuk@pro-g.com.tr
m.cengiz@pro-g.com.tr
s.civril@pro-g.com.tr
u.oztosun@pro-g.com.tr
c.tiras@pro-g.com.tr
a.dogan@pro-g.com.tr
a.dogan@pro-g.com.tr
o.cakmak@pro-g.com.tr
$ sed -r "y/ŞĞÇİÜÖğışçüö/SGCIUOgiscuo/" < PersonelListesi.txt | tr "A-Z" "a-z" |
sed -r "s/([^\t])[^\\t]+ +([^\\t]+)/\\1.\\2@pro-g.com.tr/" > eposta.txt
```



- /([^\t])[^\t]++([^\t]+)/
- sezai yilmaz
- mehmet dagdeviren
- selale erkekoglu
- ahmet kamber
- adem vatansever
- hale kucuk
- mahmut cengiz

### Örnek – Link'ten URL Üretimi





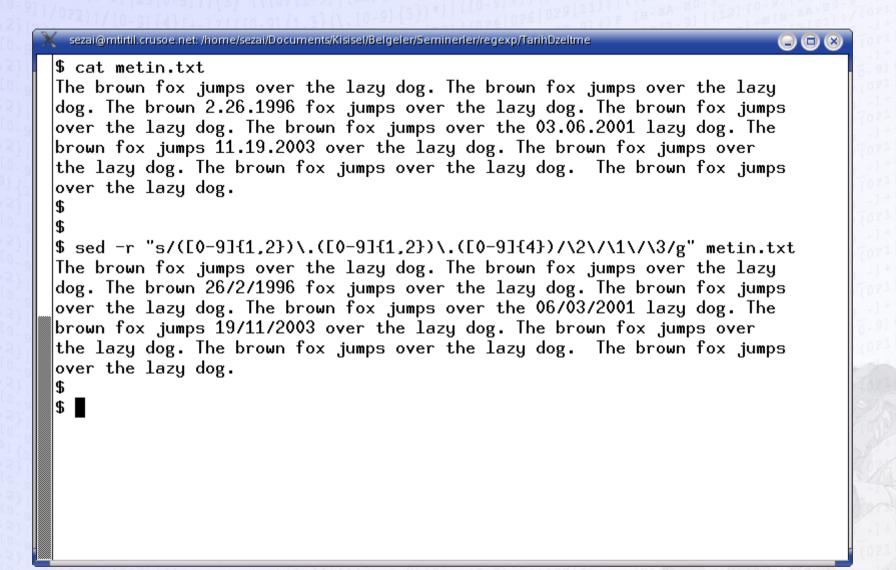
## Örnek – Link'ten URL Üretimi



```
sezai@mtirtil.crusoe.net:/home/sezai/Documents/Kisisel/Belgeler/Seminerler/regexp/HTMLURI
$ grep '<a href' index.html | sed -r 's@.+<a href="([^"]+)">.+@http://192.168.0.1/~sezai/\1@'
http://192.168.0.1/~sezai//
http://192.168.0.1/~sezai/FlashPlayer-6.0-4mdk.i586.rpm
http://192.168.0.1/~sezai/RealPlauer-8.0-8mdk.i586.rpm
http://192.168.0.1/~sezai/RealPlayer-rpnp-8.0-8mdk.i586.rpm
http://192.168.0.1/~sezai/acroread-5.0.8-3.100mdk.i586.rpm
http://192.168.0.1/~sezai/acroread-nppdf-5.0.8-3.100mdk.i586.rpm
http://192.168.0.1/~sezai/checkinstall-1.6.0-0.beta2.1mdk.i586.rpm
http://192.168.0.1/~sezai/cpudyn-0.6.1-1mdk.i586.rpm
http://192.168.0.1/~sezai/cpufreqd-1.1.1-3mdk.i586.rpm
http://192.168.0.1/~sezai/flash-plugin-6.0.79-2.i386.rpm
http://192.168.0.1/~sezai/gkrellm-2.1.28-1.norlug.i586.rpm
http://192.168.0.1/~sezai/gkrellm-plugins-2.1.28-1.norlug.i586.rpm
http://192.168.0.1/~sezai/gnokii-0.5.10-2mdk.i586.rpm
http://192.168.0.1/~sezai/gphoto2-2.1.4-1.norlug.i586.rpm
http://192.168.0.1/~sezai/gtksozluk2-0-2.i686.rpm
http://192.168.0.1/~sezai/hpoj-0.91-4mdk.i586.rpm
http://192.168.0.1/~sezai/j2re-1_4_2_04-linux-i586.rpm
http://192.168.0.1/~sezai/j2sdk-1.4.2_01-1mdk.i586.rpm
http://192.168.0.1/~sezai/java-1.5.0-sun-1.5.0-0.beta1.3jpp.nosrc.rpm
http://192.168.0.1/~sezai/kernel-module-slmodem-2.9.6-2.6.3_4mdk.i586.rpm
http://192.168.0.1/~sezai/libcpufregd-1.1.1-3mdk.i586.rpm
http://192.168.0.1/~sezai/libfftw2-2.1.5-3mdk.i586.rpm
http://192.168.0.1/~sezai/libgnokii1-0.5.10-2mdk.i586.rpm
http://192.168.0.1/~sezai/libmultisync0-0.81-2mdk.i586.rpm
http://192.168.0.1/~sezai/libopenobex1-1.0.1-1mdk.i586.rpm
http://192.168.0.1/~sezai/libstdc++2.10-2.96-0.83mdk.i586.rpm
```

### Örnek – Tarih Düzeltme





### Kullanım Alanları



#### Örnek Araçlar

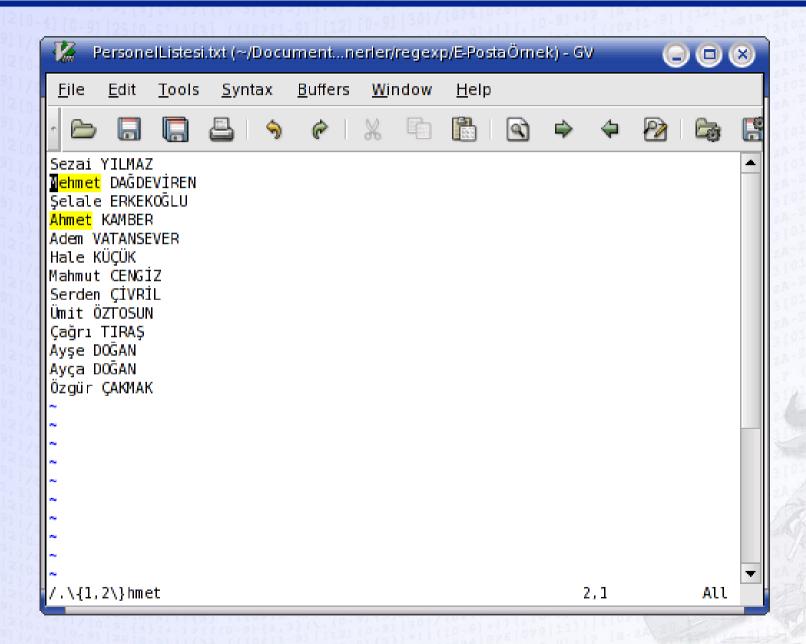
- grep
- ☞ sed
- awk
- ⟨SF Vi
- emacs
- PostgreSQL
- procmail

### Örnek Diller

- php
- perl
- javascript
- ☞ lex
- python
- tcl

# Örnek – gvim





# Örnek – PostgreSQL



X sezai@mtirtil.c	rusoe.net: /home/sezai/Documents/Kisisel/Belgeler/Seminerler/regexp/Say?lar	
deneme=# : haberid	select * from haber_ozetleri;   ozet	
2 3 4 5	Basbakan 12/08/2004 tarihinde Avrupa yolcusu    OpenOffice.org acik kaynak dunyasinin en gozde ofis uygulamasi    Fedora Core 2 01/06/2004'de gorucuye cikiyor    MS-Internet Explorer Mozilla Firefox karsisinda kan kaybediyor    Linus Torvalds kernel development listesinde cekirdek 3.0'i duyurdu.    31/10/2004 tarihinde Linuxant modem suruculeri tekrar ozgur	
deneme=# : haberid	select * from haber_ozetleri where ozet ~ '[0123]?[0-9]/[01]?[0-9]/[0-   ozet	-9]{4}';
3	Basbakan 12/08/2004 tarihinde Avrupa yolcusu   Fedora Core 2 01/06/2004'de gorucuye cikiyor   31/10/2004 tarihinde Linuxant modem suruculeri tekrar ozgur	
deneme=#		

### Örnek – PHP - 1



```
sezai@mtirtil.crusoe.net:/home/sezai/Documents/Kisisel/Belgeler/Seminerler/regexp/PHPmek
                                                                                  <html><head><title>E-Posta dogrulama</title></head>
<body>
<form name=epostaformu method=POST>
E-Posta Adresiniz
<input type=text name=kisieposta</pre>
       value="<?php echo $_POST["kisieposta"]; ?>" size=25 maxlength=80>
<input type=submit name=submitemail value="Gönder">
</form>
<?php
   /* Gecerli E-posta adresi düzenli ifadesi */
   re_{posta} = "/^[a-zA-Z0-9...]+0[a-zA-Z0-9...]+(\.[a-zA-Z0-9...]+)+\.?$/":
   $kisieposta = $_POST["kisieposta"];
   if (!isset($kisieposta) || !preg_match($re_eposta, $kisieposta) ) {
      echo "Gecerli bir E-posta adresi okunamadı<br>\n";
      exit(1):
   echo $kisieposta . " elektronik posta adresini doğru girdiniz! <br/> ';
</body>
</html>
"eposta.php" 23L, 692C written
                                                                      6.1 - 8
                                                                                     A11
```

## Örnek – PHP - 2



/^[a-zA-Z0-9]+@[a-zA-Z0-9]+(\.[a-zA-Z0-9]+)+\.?\$/			
E-Posta Adresiniz admin Gönder			
Gecerli bir E-posta adresi okunamadı			
E-Posta Adresiniz admin@test Gönder			
Gecerli bir E-posta adresi okunamadı			
E-Posta Adresiniz admin@test.com Gönder			
admin@test.com elektronik posta adresini doğru girdiniz!			
E-Posta Adresiniz admin@test.com.tr Gönder admin@test.com.tr elektronik posta adresini doğru girdiniz!			
E-Posta Adresiniz admin@test.com.tr@test Gönder  Gecerli bir E-posta adresi okunamadı			

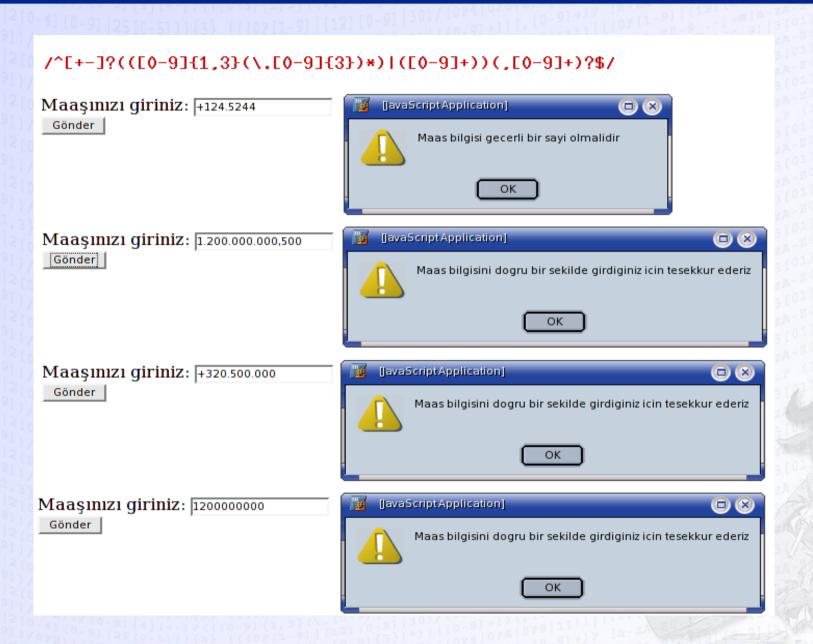
# Örnek – Javascript - 1



```
sezai@mtirtil.crusoe.net:/home/sezai/Documents/Kisisel/Belgeler/Seminerler/regexp/Say?lar
<html><head><title>Javascript ile Düzenli İfade Örneği</title>
<script lang="javascript1.2">
    function alanlaridenetle() {
        var re_sayi;
        re_sayi = /^{+-}?(([0-9]{1,3}(\.[0-9]{3})*)|([0-9]+))(,[0-9]+)?;
        if (document.sayigir.maas.value.match(re_sayi) == null) {
            alert("Maas bilgisi gecerli bir sayi olmalidir");
            return false:
        } else {
            alert("Maas bilgisini dogru bir sekilde "
                 + "girdiginiz icin tesekkur ederiz");
            return true:
</script>
K/head>
<br/>body>
<form name="sayigir" method="post" OnSubmit="return alanlaridenetle();">
Maaşınızı giriniz: <input type="text" name="maas" size="20" maxlength="20"><br>
<input type="submit" name="submitbtn" value="Gönder">
</form>
</bodu>
"test.html" 23L, 718C written
                                                                  15.1-4
                                                                                 A11
```

## Örnek – Javascript - 2





#### Standartlar – 1



#### POSIX

- [:alnum:]
- [:alpha:]
- [:digit:]
- [:lower:]
- [:upper:]
- [:blank:]
- [:whitespace:]
- [:return:]

#### Perl Uyumlu

- \d ondalık rakam
- \D ondalık rak. değil
- \w kelime karakteri
- \W kelime kar. değil
- \s boşluk
- \S boşluk değil
- \b kelime ayıracı

### Standartlar - 2



#### Temel

- **\*** \{\}
- **\(\)**
- ₩ \
- ⟨₩
- **\*** \?

#### Genişletilmiş

- **F** {
- **()**
- **\*** +

## Özet



- Düzenli ifadeler, bir şablona uyan dizgeleri eşleyen formüldür
- Karmaşık görünürler, ama pratikle kullanımı kolaylaşır
- Temel, genişletilmiş, posix ve perl uyumlu türünde düzenli ifadeler vardır
- Yoğun kullanıldığı alanlar
  - Metin düzenleyici, veritabanı, elektronik posta yardımcı programları, akış düzenleyiciler
  - PHP, Java, Javascript, C, C++, Perl, Python, ...

## Kaynakça



- A Tao of Regular Expressions
  - Steve Mansour sman@scruznet.com
  - http://sitescooper.org/tao\_regexps.html
- Using Regular Expressions
  - Stephen Ramsay
  - http://etext.lib.virginia.edu/helpsheets/regex.html
- Regular Expression HOWTO
  - A.M. Kuchling amk@amk.ca
  - http://www.amk.ca/python/howto/regex/