“Büyük Veri”de Türkiye’den Uygulama Örnekleri

Dr. Güven Fidan

ARGEDOR Bilişim Teknolojileri

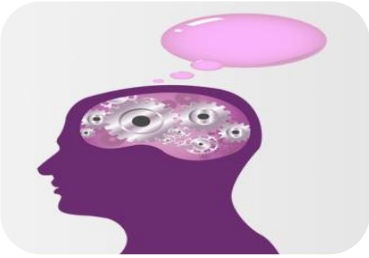
ARGEDOR

- ARGEDOR, şirketlere ve son kullanıcılara yenilikçi bilgiyi işleme çözümleriyle dünya çapında mevcut olan bilgiye erişme, işleme, araştırma ve uyarlama konusunda yardımcı olmaktadır.

- İşgücü verimliliğini arttıran, işletme süreçlerini düzene sokan, müşteri memnuniyetini yükselten ve Web üzerinden satışları arttıran ürünler oluşturmaktadır.

ARGEDOR

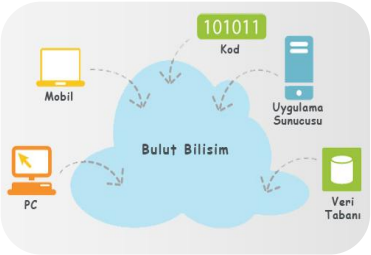
- ARGEDOR aşağıdaki alanlarda yenilikçi ürünler geliştirmektedir.



Veri Madenciliği

Arama Teknolojileri

Öneri Sistemleri



Bulut Bilişim

Duygu Analizi

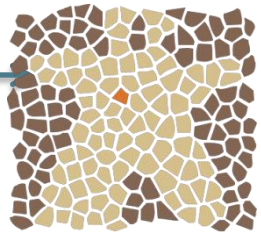
Big Data

SMS
SMS

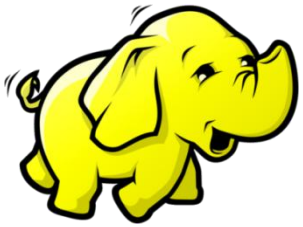
BIG
DATA
SOFTWARE
ANALYTICS
SAN
SIZE
NEEDED
PETABYTES
INTERNET
TECHNOLOGIES
HUNDREDS
DISK
SHARED
TIME
INCLUDE
SYSTEMS
NETWORKS
SEARCH
INFORMATION
RECORDS
MANAGEMENT
SOCIAL
LARGE
RESEARCH
COMPLEX
DATABASES
STORAGE
EVERY
ABILITY
SETS
EXAMPLES
CURRENT

Hadoop Ekosistem

Kafka



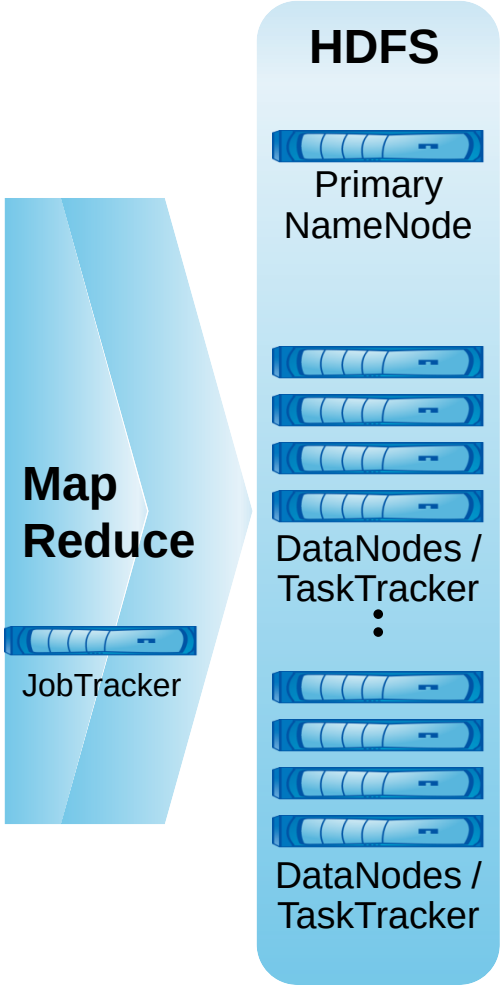
APACHE GIRAPH



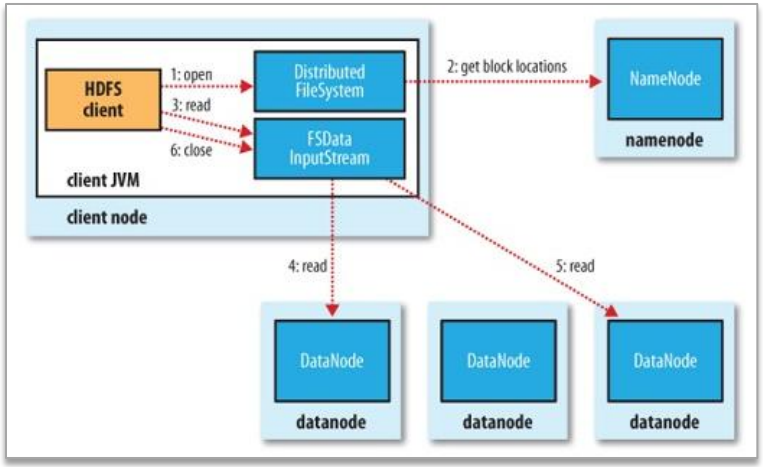
Hadoop Nedir?

Apache Hadoop

- Google FS ilham verdi
- Dağınık Dosya Sistemi, HDFS
- Google MapReduce benzeri paralel programlama modeli
- Google Bigtable benzeri paralel veritabanı
- Binlerce node ve petabytelarca veri adresleme imkanı
- Büyük dosyalar ve yüksek performans ihtiyacı

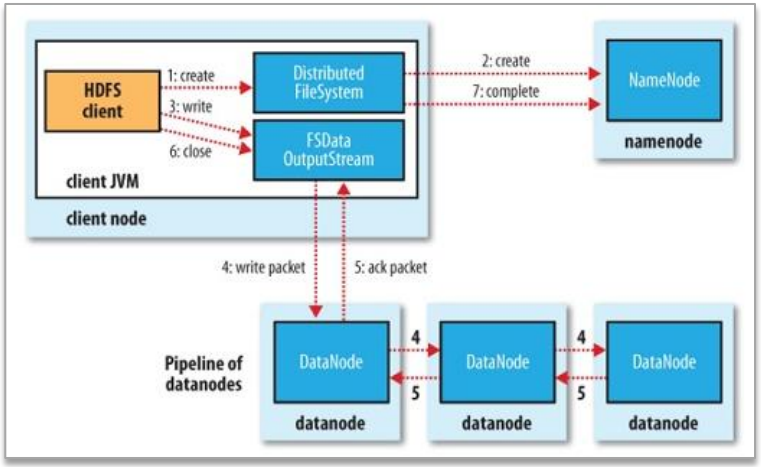


HDFS



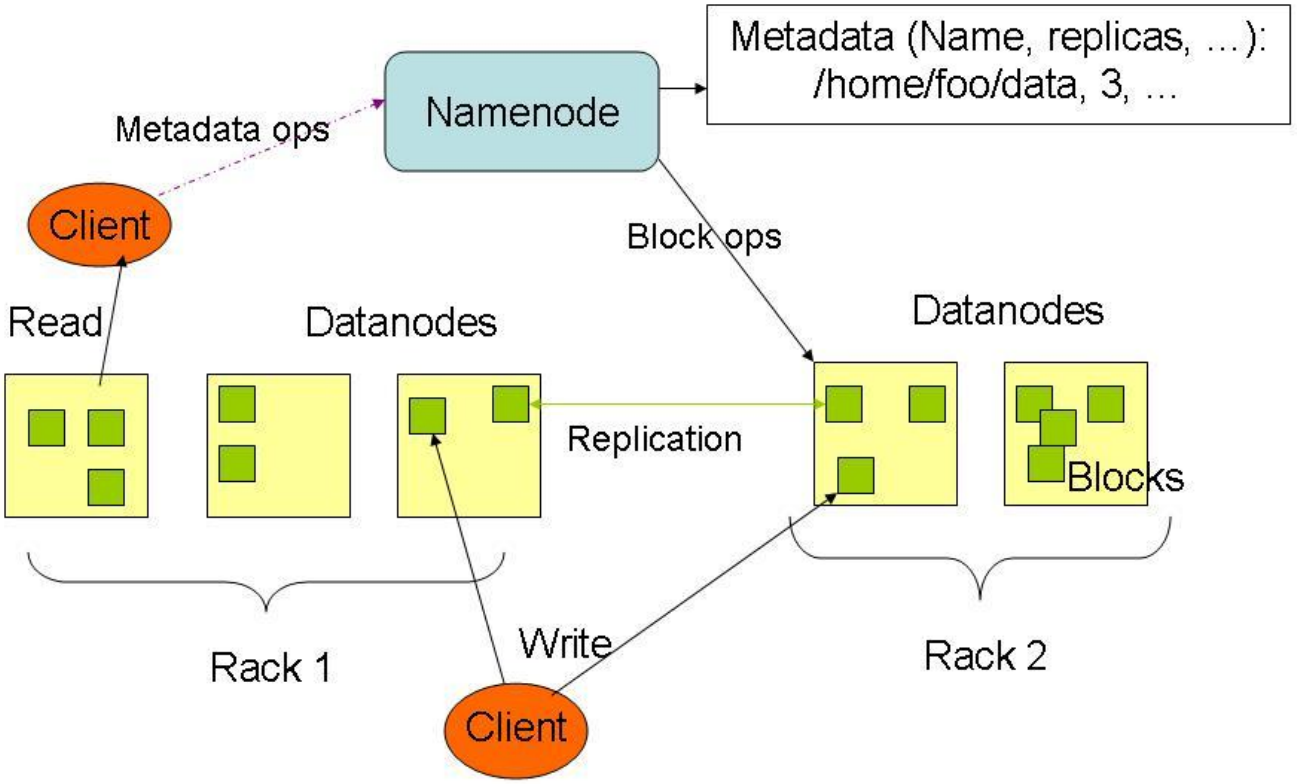
Hadoop Distributed File System

Google File System'den esinlenmiştir

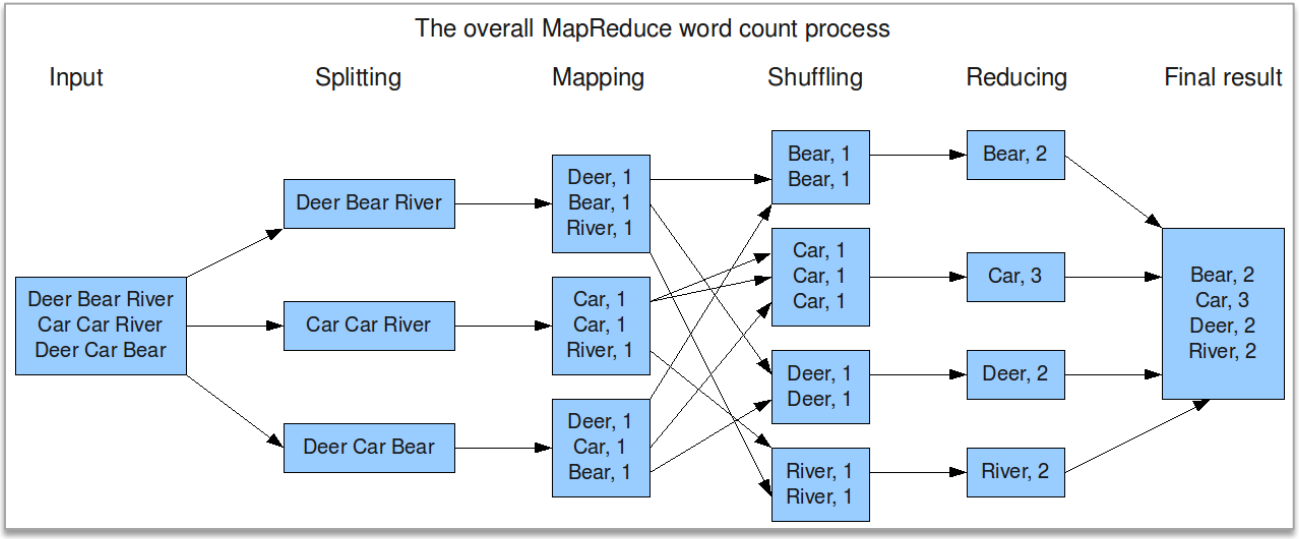


HDFS

HDFS Architecture



MapReduce



MAP

- Üzerinde çalışılacak dosya *split* adı verilen parçalara ayrılır
- Her bir *split* için bir *map* görevi başlatılır
- *Map* görevleri, işlem sonuçlarını lokal diske yazarlar (HDFS)

REDUCE

- *Map* sonuçları *Reduce* için kaynak dosya olur
- Konsolide bir çıktı verir
- Nodelar arasında ağ üzerinden haberleşirler
- İşlem sonucu HDFS üzerinde saklanır

HBase

- HDFS üzerinde çalışan ilişkisel olmayan sütun-tabanlı veritabanı
- Büyük veri üzerinde random ve gerçek zamanlı hızlı veri okuma ve yazma amacıyla kullanılır
- HDFS'in tüm avantajlarından faydalanabilirsiniz ve MapReduce programla modelini kullanabilirsiniz
- Güvenli okuma ve yazma sunar
- Lineer ve modüler ölçeklenme sağlar
- HBase API'ı genişletilebilir ve özelleştirilebilir

Hive

- Hadoop sitemine bir çeşit veri ambarı yapısı kurarak, hızlı veri özetleme, raporlama işlemlerini büyük veri setleri üzerinde gerçekleştirebilir
- HDFS üzerinde bulunan veri yapısını SQL benzeri bir dille (HiveQL) işleyebilir
- HiveQL geleneksel MapReduce programlama modelindeki mapper ve reducerları dil içerisinde kullanabilir
- Hive API'ı genişletilebilir ve özelleştirilebilir

Pig

- Hadoop üzerinde çalışırken belirli işleri basitleştirmek için geliştirilen bir çeşit programlama dilidir
- Veri yükleme, veriyi yeni bir veri modeline dönüştürme ve sonuçlarını sistemde saklama işleri için kullanılmaktadır
- Pig, yarı-yapılı veri tipleri üzerindeki işlerde kullanılmak için özelleşmiştir: Kullanıcı log bilgileri, ağ trafik bilgileri vb.
- PIG API'ı genişletilebilir ve özelleştirilebilir

Veri Alma ve Verme: Sqoop ve Flume

- Sqoop, ilişkisel veritabanlarından HDFS veya Hive veri modellerine direkt veri almayı sağlar. MapReduce paralalel işleme modelini de veri yazma ve okuma da kullanır
- Sqoop, verileri Hadoop sisteminden ilişkisel veritabanlarına yazmayı da sağlar
- Flume, farklı kaynaklardan Hadoop sistemine, hızlı, güvenilir, hataya toleranslı ve dağıtık veri aktarmayı sağlar
- Flume, merkezi bir yönetim mekanizması ile dinamik veri çekmeyi sağlar

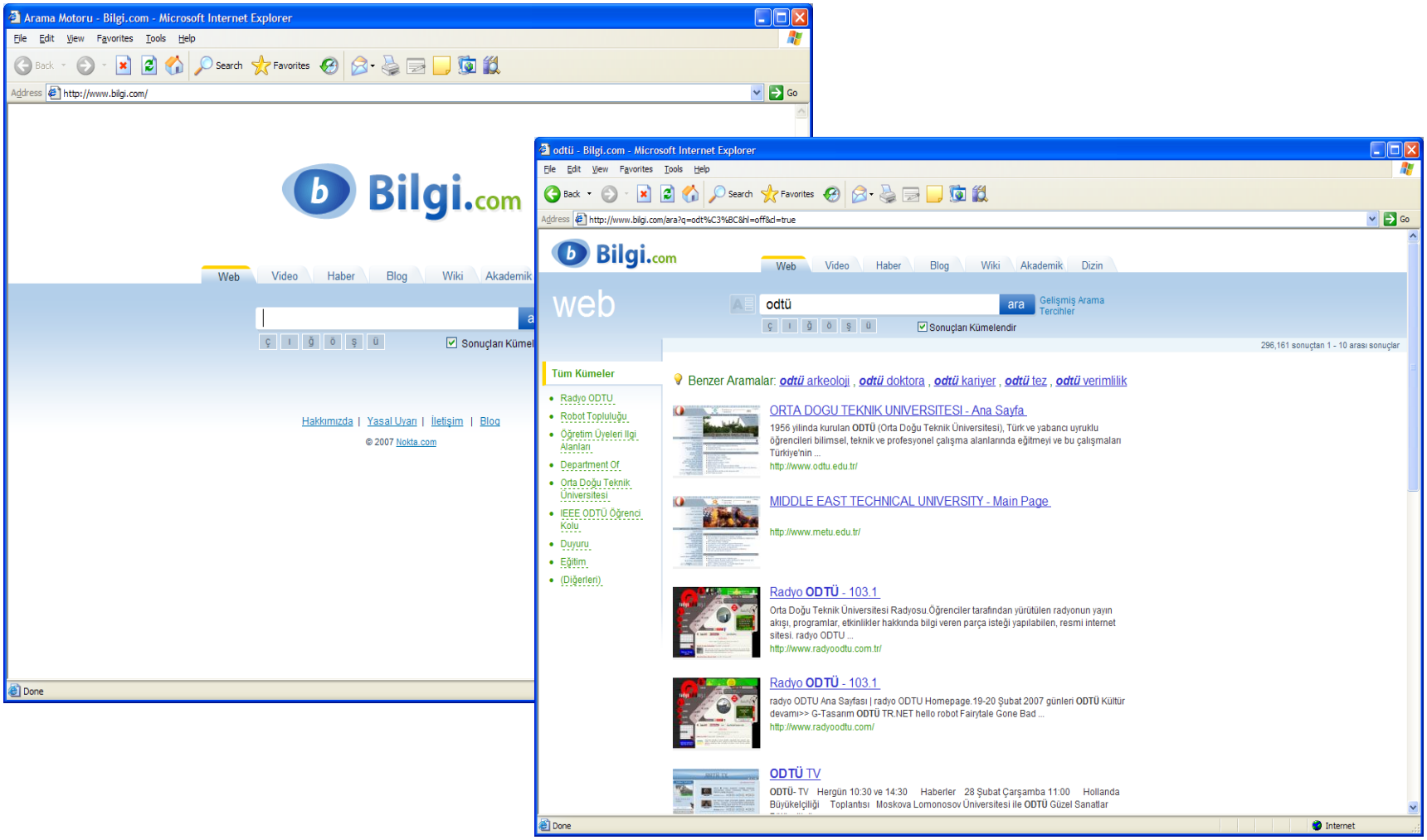
Koordinasyon ve İş Akışı: Zookeeper ve Oozie

- Hadoop kümesi üzerinde yürüyen servisler arttıkça, servislerin koordinasyonu ve adlandırılması önem kazanmaktadır. *Zookeeper* merkezi bir servis yönetim yapısı oluşturarak, servislerin konfigurasyon bilgilerini, isimlendirmelerini, yeni uçlar üzerinde dağıtımını ve servis gruplandırmaların sürekliliğini sağlar.
- *Oozie*, Hadoop üzerinde çalışan servislerin birbirine olan bağımlılığını ve iş akışlarının koordinasyonun sağlar.

Mahout

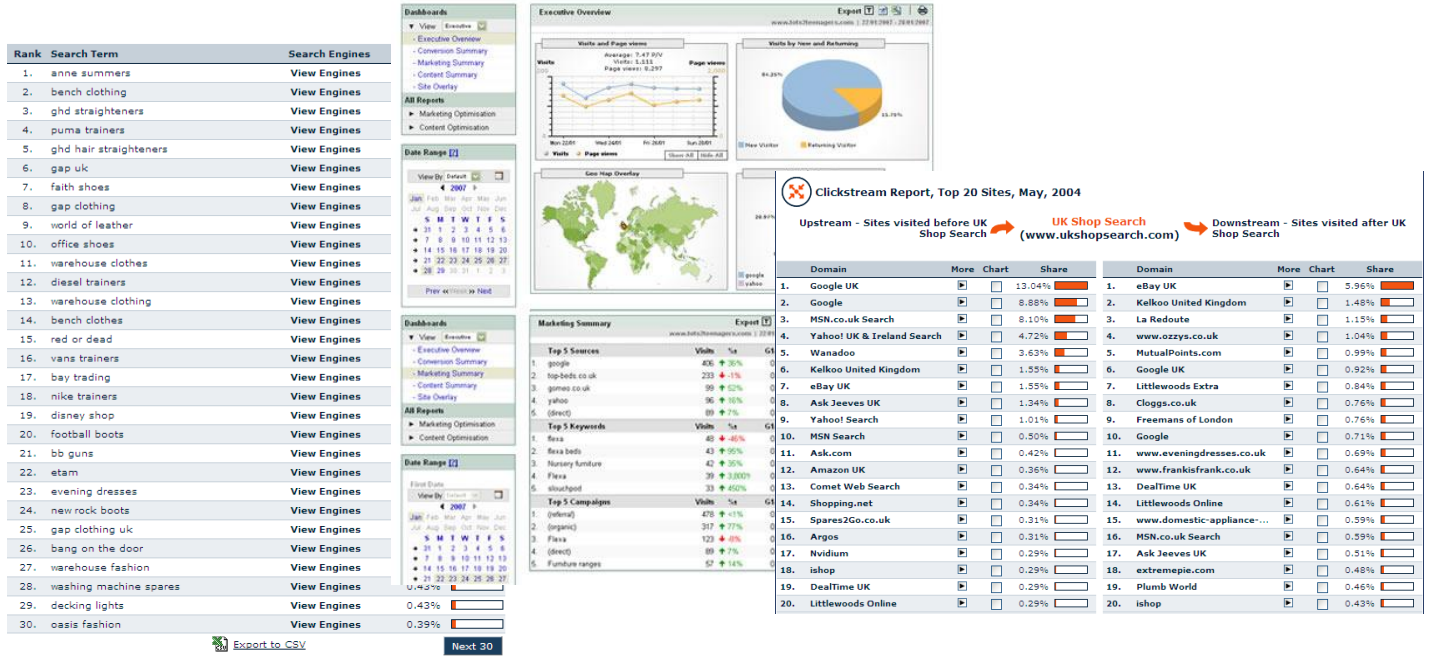
- Mahout, Hadoop üzerinde çalışan makine öğrenme kütüphanesidir
- Büyük veri kümeleri üzerinde veri madenciliği amacıyla kullanılır
- Kümeleme, sınıflandırma, önerme vb. alanlarda uyarlanmış bir çok makine öğrenme metodunu barındırır
- Mahout API'ı genişletilebilir ve özelleştirilebilir

Arama Motoru Teknolojileri

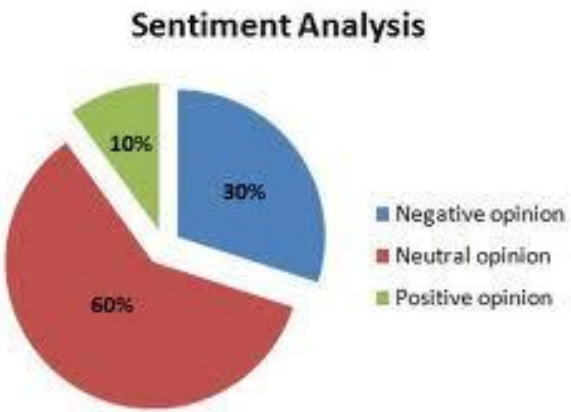


Web Analiz Sistemleri

Bu projede, web sitelerinin kullanıcı kullanım verilerini saklayan, veriler üzerinde çeşitli web kullanım madenciliği yöntemleri uygulayarak, web sitelerini **içerik ve yapısal** olarak iyileştirilmesini sağlayan bir ***web analiz sistemi*** geliştirilmiştir.



Sentiment & Sosyal Medya Analizi



feeling, emotion; attitude, opinion; tender emotion.

Bir dokümanın olumlu ya da olumsuz görüş belirttiğini otomatik olarak saptamaya çalışılmasıdır. Çıkartılmaya çalışılan görüş yazarın konu hakkındaki kararı ya da değerlendirmesi, yazarken hissettiği ruh hali, ya da belirtmek istediği etki olabilir.

Öneri Sistemleri

ALBÜMLER | BİYOGRAFİ | ŞARKILAR | VİDEOLAR | GALERİ

Tarkan | Albümler | Sevdanın Son Vuruşu

ÖNERİLEN ALBÜMLER | ÖNERİLEN SANATÇILAR

Adımı Kalbin...
Tarkan

Metamorfoz
Tarkan

Karma
Tarkan

Aşk
Demet Akalın

Benim Sadık ...
Tarkan

Ben Bugün
Yalın

tüm liste

ÖNERİLEN ALBÜMLER | **ÖNERİLEN SANATÇILAR**

Ajda Pekkan

Sezen Aksu

Sertab Erener

Pinhani

İzel

Ahmet Kaya

tüm liste

Öneri Sistemleri

"aile" için Arama Sonuçları

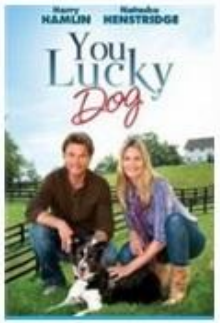
Filmler ve TV Programları

Mirket Ailesi
The Meerkats

2008 - Belgesel

John Carter / Tanıtım
John Carter Trai...

2012 - Aksiyon/Macera



Aile Toplantısı
Family Gathering

2010 - Dram

Karanlıktan Korkma
Don't Be Afraid ...

2010 - Gerilim/Korku

Aşk Parayla Ölçülmez
Love Don't Cost ...

2003 - Romantik Komedi

Tilki ve Avcı Köpeği
The Fox and the ...

1981 - Animasyon

Aile Rehberi - 14.07.2012
Aile Rehberi

2012 - Yaşam & Hobi, Ya...

İbreti Ailem - 11.07.2012
İbreti Ailem

2012 - -

Here I lie
Here I lie

2004 - Kanallar

3D Smart Cities

Konya 3D CBS

Web 2.0/3.0

- Collaborative Content
- Context-Awareness
- Search & Discover
- Recommender Sytems

Partner

ARGEDOR

TÜRK TELEKOM

ITEA 2
INFORMATION TECHNOLOGY FOR EUROPEAN ADVANCEMENT

Celtic-Plus
Smart Connected World
EUREKA

TTNET Müzik

TÜBİTAK

AGMLab
bilişim teknolojileri

SAMSUNG

TTNET

tivibu

BAŞARI®
mobile

Teşekkürler...

Dr. Güven FİDAN

guven.fidan@argedor.com