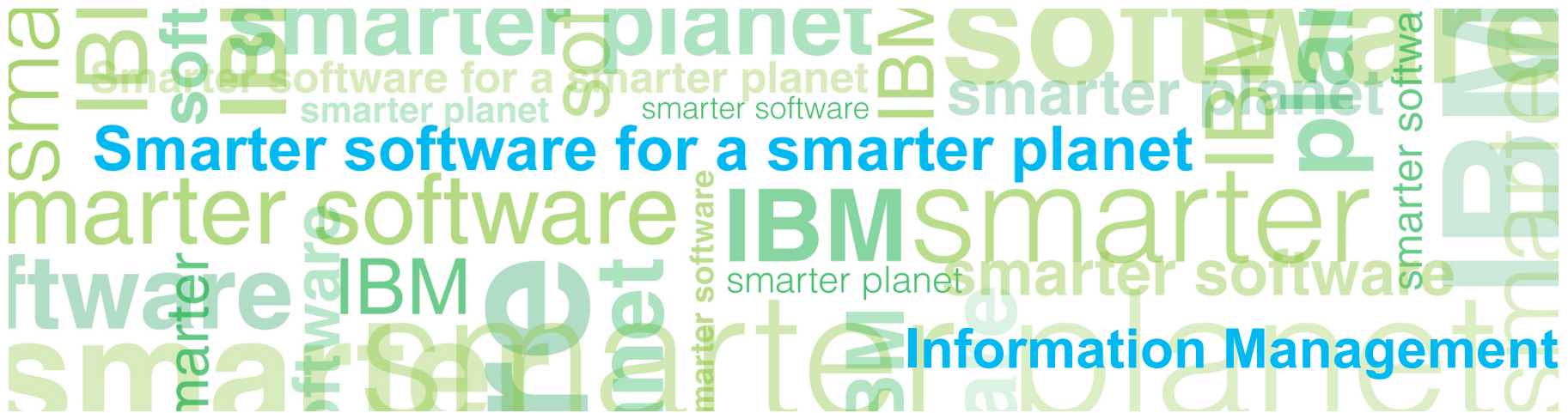
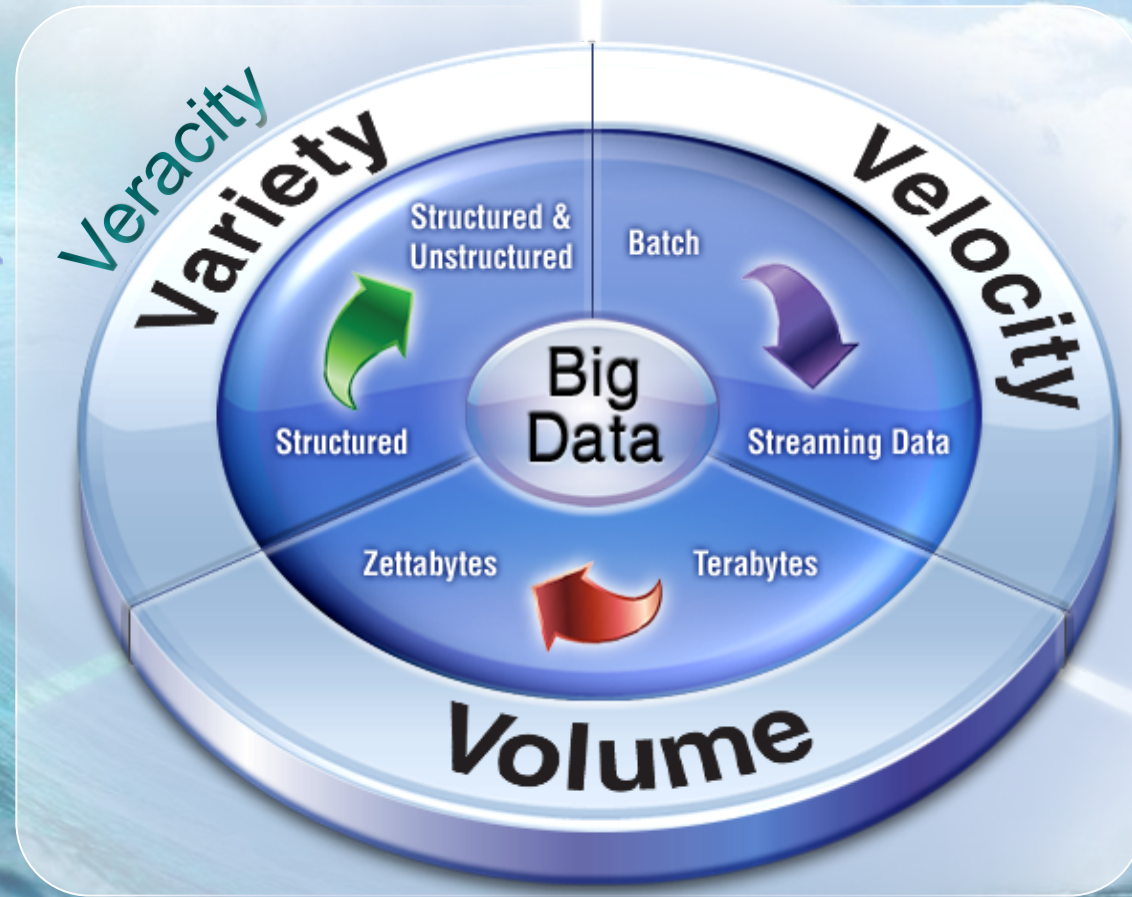


Büyük Veri'nin Gücünden Faydalanın



Big Data

Daha önce analizi mümkün olmayan muazzam büyüklük ve/veya çeşitlilik ve/veya akışkanlıktaki veriler üzerinden bilgiye erişim.



5 Temel Kullanım Alanı



Büyük Veri Arama

Karar alma sürecini geliştirmek için tüm büyük veri kaynaklarını aramak, veriyi anlamak, görselleştirmek



Gelişkin 360° Müşteri Görüntüsü

Yeni İç ve Dış veri kaynaklarıyla mevcut Müşteri Görüntüsünü genişletin



Yeni Güvenlik / İstihbarat Yetkinlikleri

Düşük risk, gerçek zamanlı dolandırıcılık izleme ve siber güvenlik algılama



Operasyon Analizi

Makinelerin ürettiği farklı çeşitlilikteki verilerin analiz edilerek iş sonuçlarının geliştirilmesi

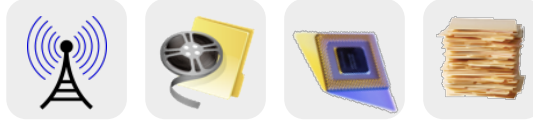


Veri Ambarı Eklentisi

Operasyonel verimliliği artırmak için büyük veri ve veri ambarı özelliklerini entegre etmek

Veri Ambarı Eklentisi : Farklı İhtiyaçlar

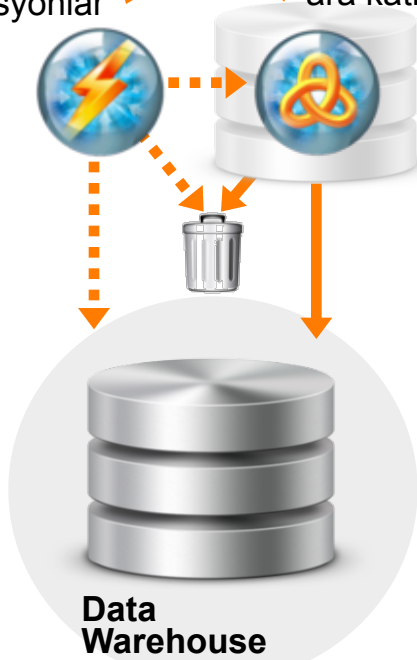
1 Ön-İşleme



Data Explorer

Streams
Gerçek zamanlı operasyonlar

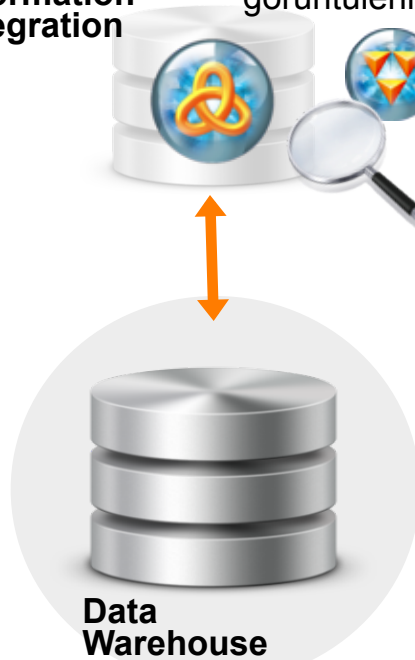
BigInsights
Her tür veri için ara katman



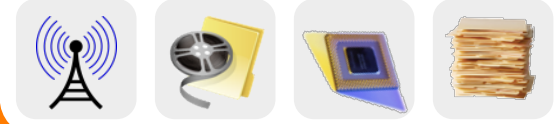
2 Sorgulanabilir Arşiv

BigInsights
Information Integration

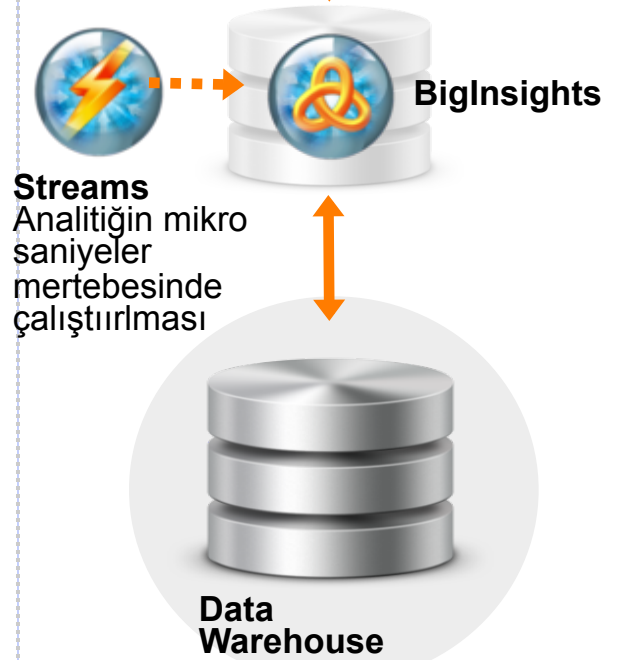
Data Explorer
Verinin aranıp görüntülenmesi



3 Keşifsel Analiz



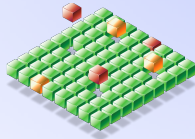
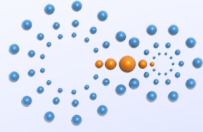
Yapısal olmayan verilerle birlikte analiz



Büyük Veri fırsatları her yerde bulunmakta...

Akan Verinin İncelenmesi

- Akıllı Ölçüm Cihazları
- Çok modlu gözetim
- Gerçek Zamanlı Promosyon / Kampanyalar
- Siber Güvenlik
- Yoğun Bakım Ünitesi Gözetim
- Hisse Senedi İşlemleri
- Click-stream analizi
- CDR işleme
- IT log analizleri
- RFID takip & analiz



Çok Büyük Hacimli Verilerin Analizi

- Yeni servis ve ürünlerin geliştirilmesi
- Sahtecilik modelleme & tespiti
- Risk modelleme & yönetimi
- Sosyal Medya/Duygu Analizi
- Çevre Analizi



Her Tipten Verinin İncelenmesi

- Sosyal Medya/Duygu Analizi
- Coğrafi Analiz
- Marka Stratejisi
- Bilimsel Araştırmalar
- Erken Uyarı Sistemleri
- Pazar Araştırmaları
- Video analizi
- Ses analizi

Keşif ve Deney

- Duygu Analizi
- Marka Stratejisi
- Bilimsel Araştırmalar
- Ad-hoc analizler
- Model geliştirme
- Hipotez testleri
- Yeni servis ve ürünlerin geliştirilmesi



Yönetim ve Planlama

- Operasyonel Analitik – İş Zekası Raporları
- Planlama, Tahminleme, Öngörüsül Analizler

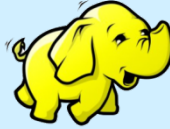
Big Data yeni yetkinlikler gerektiriyor...

Federe veri kaynakları
üzerinde arama ve keşif



Federated Discovery and Navigation

Her çeşit büyük veriyi
saklama ve yönetme



Hadoop File System
MapReduce

Yapısal verilerin hızlı
analizi, kitlesel paralel işlem



Data Warehousing

Akışkan verilerin yönetimi



Stream Computing

Yapısal olmayan verilerin
analizi



Text Analytics Engine

Veri kaynaklarının
entegrasyonu ve sahipliği



Integration, Data Quality, Security,
Lifecycle Management, MDM

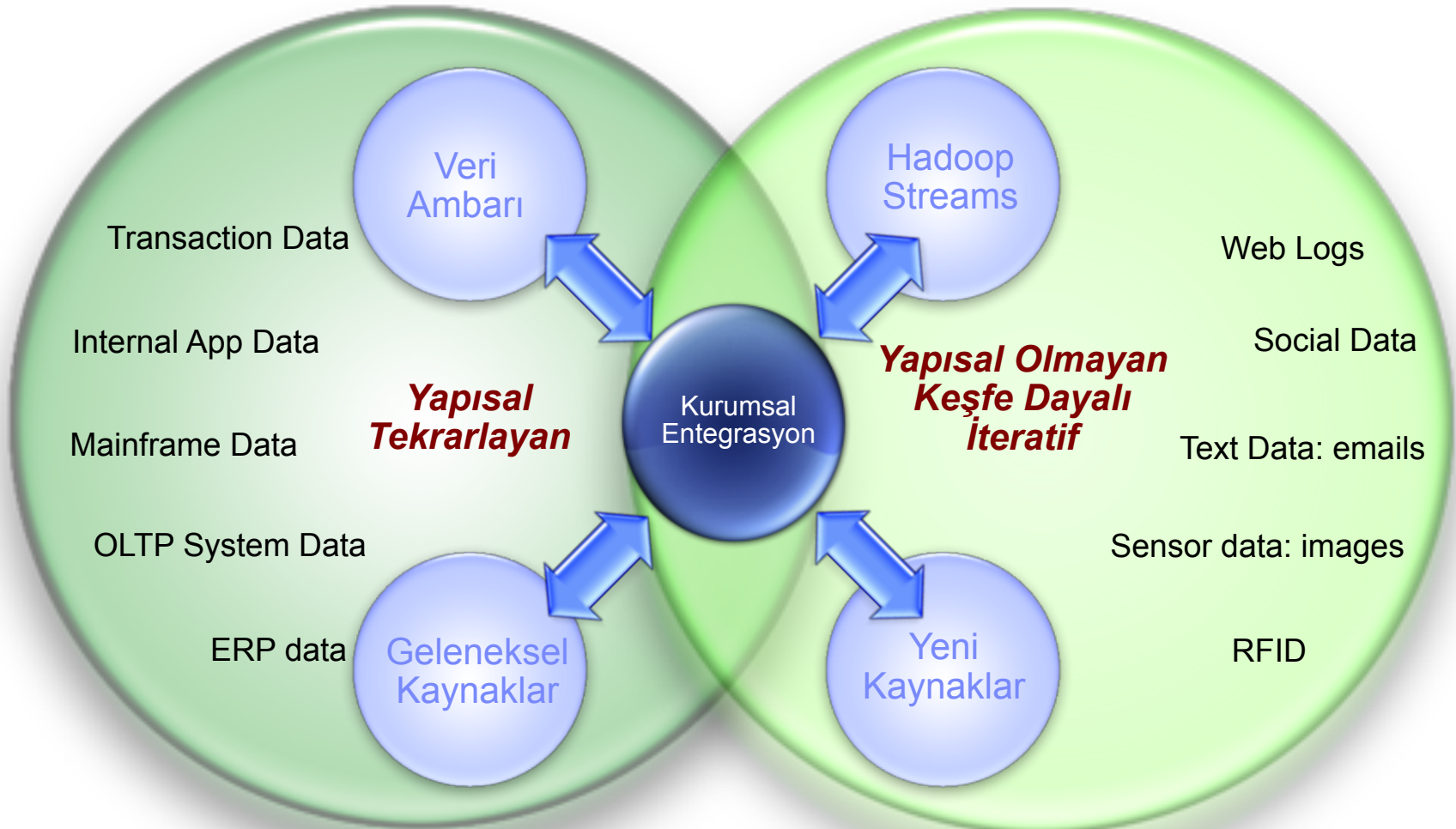
Farklı Uygulama Alanları için Birbirini Bütünleyen Yaklaşımlar

Geleneksel Metodlar

Yapısal, Analitik, Tekrarlayan

Yeni Yaklaşımlar

Yaratıcı, bütünsel görüş, sezgisel



IBM Big Data Stratejisi: Analitiği Veriye yaklaştırın

Analitik Uygulamalar

BI / Raporlama | Analiz/ Görselleştirme | Fonksiyonel Uyg. | Endüstri Uyg. | Veri Madenciliği | İçerik Analitiği

IBM Big Data Platformu

Görselleştirme
ve Keşif

Uygulama
Geliştirme

Sistem
Yönetimi



Hızlandırıcılar

Hadoop
Sistemi



Akışkan
veri işleme



Veri
Ambarı



Veri Entegrasyonu ve Sahipliği

- Tipi, biçimi, boyutu ve akışkanlığı ne olursa olsun tüm verilerinizi entegre edilebilmesi ve yönetimi
- İleri düzey analitik fonksiyonların verinin doğal halinde uygulanması
- Elinizdeki tüm verileri görselleştirebilmesi
- Yeni nesil analitik uygulamalar geliştirebilmek için geliştirme araçları
- İşyükü optimizasyonu ve zamanlama
- Güvenlik ve Veri Sahipliği

IBM Pure Data System for Analytics

Analitik İhtiyaçlarınız için İş Yükü Optimize Sistem

Powered by
Netezza Technology

Analitik Uygulamalar

BI / Raporlama | Analiz/ Görselleştirme | Fonksiyonel Uyg. | Endüstri Uyg. | Veri Madenciliği | İçerik Analitiği

IBM Big Data Platformu

Görselleştirme
ve Keşif

Uygulama
Geliştirme

Sistem
Yönetimi



Hızlandırıcılar

Hadoop
Sistemi



Akışkan
veri işleme



Veri
Ambarı



Veri Entegrasyonu ve Sahipliği

PureData System for Analytics



İş Amacı Optimize Yeni Teknolojilerle
ve Kitlesel Paralel İşlem Mimarisi ile
Veri Ambarınızı hızlandırın



InfoSphere Streams – Akan Veri İşleme Teknolojisi

Analitik Uygulamalar

BI / Raporlama | Analiz/ Görselleştirme | Fonksiyonel Uyg. | Endüstri Uyg. | Veri Madenciliği | İçerik Analitiği

IBM Big Data Platformu

Görselleştirme
ve Keşif

Uygulama
Geliştirme

Sistem
Yönetimi



Hızlandırıcılar

Hadoop
Sistemi



Akışkan
veri işleme



Veri
Ambarı



Veri Entegrasyonu ve Sahipliği

- Hareket halindeki verilerin analizi
 - Eş zamanlı çok kanallı akışkan veri girişi
 - Ölçeklenebilir mimari
- Farklı tipten verilerin işlenebilmesi
 - Yapısal / yapısal olmayan, video, ses
 - İleri Düzey Analitik operatörler
- Adapte olan Gerçek Zamanlı Analitik
 - Veri Ambarı ile birlikte
 - Hadoop ile birlikte

Akan Verilerin İşlenmesi – Yeni bir Paradigma

Geleneksel Veri İşleme



Tarihsel verilerin analizi

Disk üzerindeki verilerin taranması ve analizi

Gecikmeli

Sorgu tetiklemeli, statik görüntünün analizi



Akan Veri İşleme



Anlık gerçeklerin işlenmesi

Hareket halindeki verinin, diske yazılmadan analizi

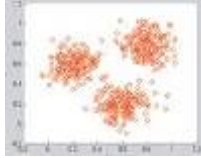
Düşük geçikme, anında aksiyon

Veri tetiklemeli – verinin analitik süreçlerden geçmesi

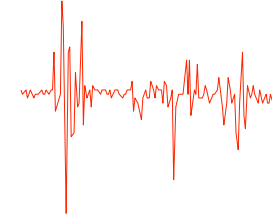




InfoSphere Streams – Akan Veri Analizi



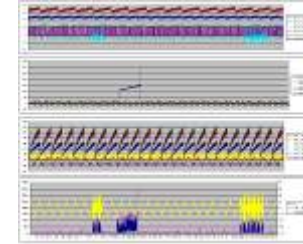
**Mikrosaniyeler Mertebesinde
Veri Madenciliği**



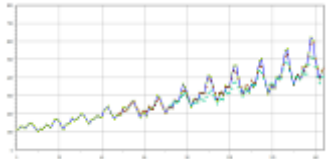
Akustik Analiz

Metin
(listen, verb),
(radio, noun)

**Basit & Gelişkin Metin
Analitiği**



**İleri Düzey
Matematiksel
Modeller**



Tahminleme



**İstatistiksel
Fonksiyonlar**

$$\sum_{\text{population}} R(s_t, a_t)$$


**Coğrafi
Analizler**

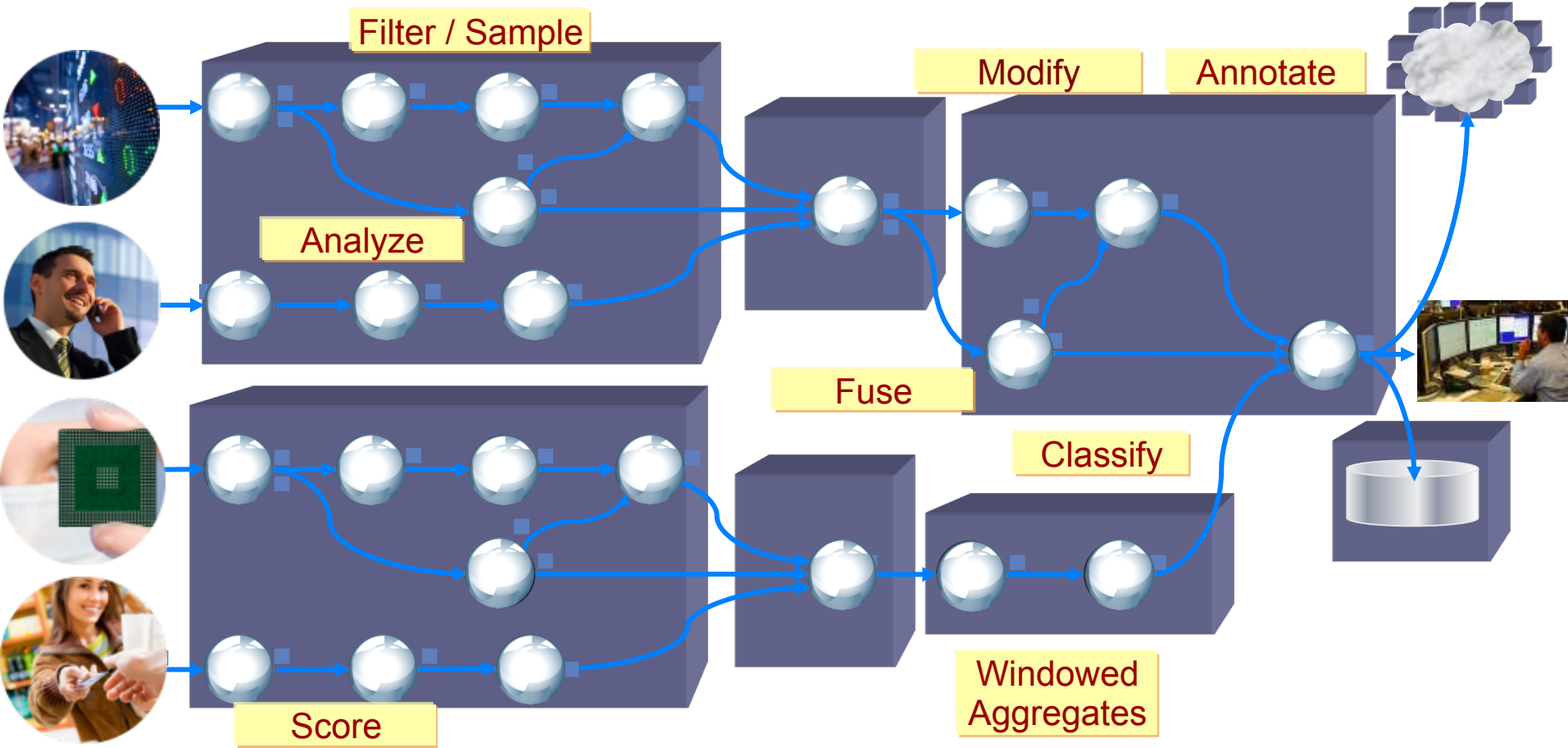


**İmaj & Video
İşleme**





InfoSphere Streams – Gerçek Zamanlı Big Data





BigInsights, Hadoop yetkinliklerini kurumunuza getiriyor

Analitik Uygulamalar

BI / Raporlama	Analiz/ Görselleştirme	Fonksiyonel Uyg.	Endüstri Uyg.	Veri Madenciliği	İçerik Analitiği
----------------	------------------------	------------------	---------------	------------------	------------------

IBM Big Data Platformu

Görselleştirme ve Keşif

Uygulama Geliştirme

Sistem Yönetimi



Hızlandırıcılar

Hadoop Sistemi



Akışkan veri işleme



Veri Ambarı



Veri Entegrasyonu ve Sahipliği

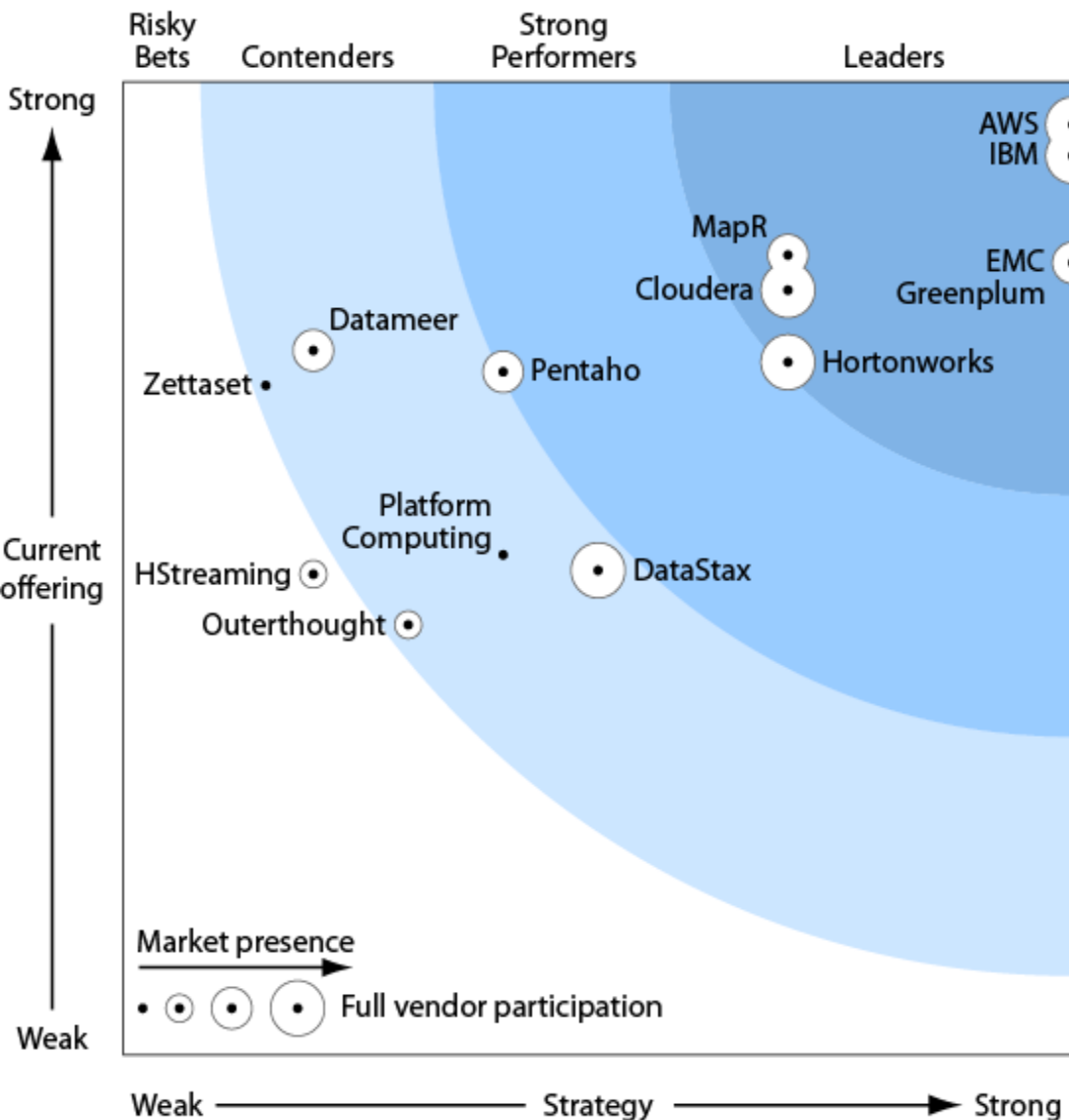
IBM InfoSphere BigInsights

- **Yapısı ne olursa olsun Ham Veriyi İnceleyin**
- **Veri Ambarı maliyetlerinizi düşürün**

Açık kaynak kodlu Hadoop çözümünü tamamlıyor;

- Performans Optimizasyonları
- Geliştirme Araçları
- Kurumsal Entegrasyon Yetkinlikleri
- Analitik Hızlandırıcılar
- Hazır uygulamalar
- Görselleştirme
- Güvenlik
- Gelişkin Analitik Yetkinlikler

Endüstrideki ilk Hadoop Sistemleri Değerlendirmesi

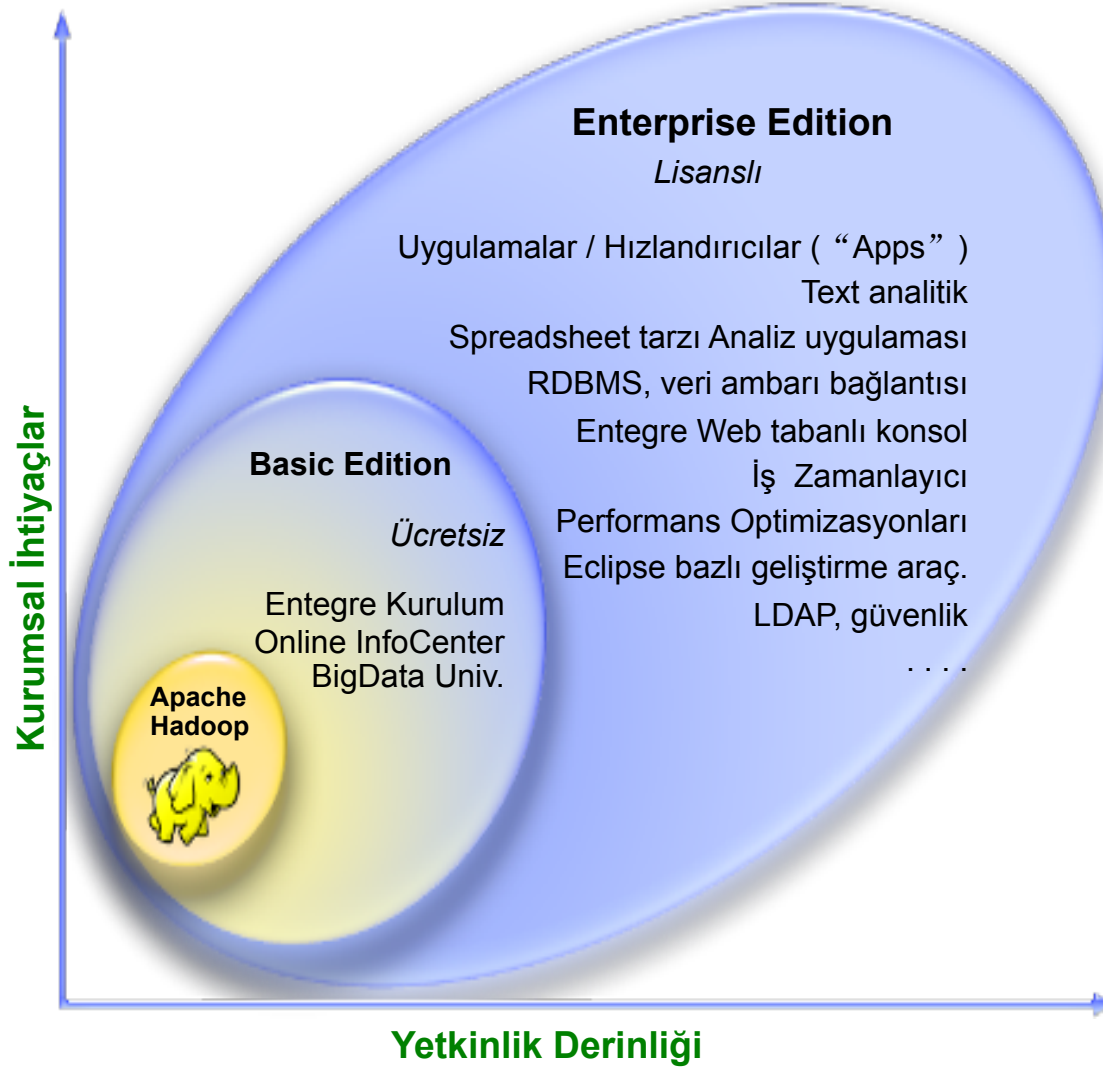


“IBM has the deepest Hadoop platform and application portfolio. IBM, an established EDW vendor, has its own Hadoop distribution; an extensive professional services force working on Hadoop projects; extensive R&D programs developing Hadoop technologies; connections to Hadoop from its EDW.”

–The Forrester Wave™: Enterprise Hadoop Solutions, 1Q12



IBM Infosphere BigInsights





Basic Edition

Open Source

IBM

Infrastructure

Integrated installer

- ZooKeeper
- Jaql
- Pig
- Oozie
- HBase
- Hive
- Lucene
- MapReduce
- HDFS

Connectivity

- JDBC
- Flume



Enterprise Edition

Open Source

IBM

Optional IBM and partner offerings

Analytics and discovery

- Text processing engine and library
- BigSheets

“Apps”

- Web Crawler
- Boardreader
- Distrib file copy
- ...

- DB export
- DB import
- Hive query
- Pig query
- Jaql query

Administrative and development tools

Web console

- Monitor cluster health
- Add / remove nodes
- Start / stop services
- Inspect job status
- Inspect workflow status
- Deploy apps
- Launch apps / jobs
- Work with distrib file system
- Work with spreadsheet interface
- Support REST-based API
- ...

Eclipse plug-ins

- Text analytics
- MapReduce programming
- Jaql development
- Hive query development

Infrastructure

- Integrated installer
- Enhanced security
- ZooKeeper
- Oozie
- Jaql
- HBase
- Pig
- Hive
- Text compression
- BigIndex
- Lucene
- Adaptive MapReduce
- MapReduce
- BigSQL (TP)
- Flexible scheduler
- GPFS
- HDFS

Connectivity

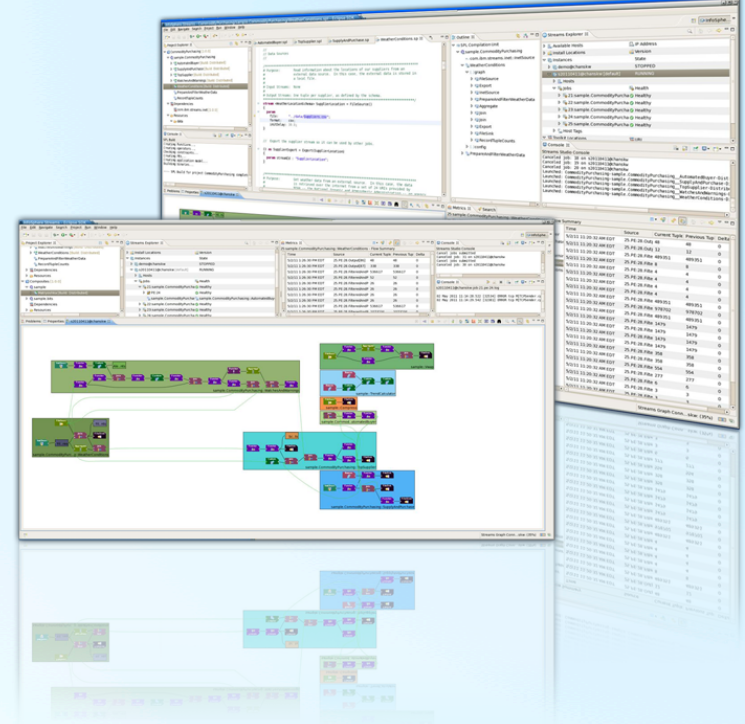
- JDBC
- Flume
- Netezza
- Streams*
- DB2
- Streams
- R

Kullanıcı ve Geliştiriciler için Basit Arayüzler



Son Kullanıcı Görseleştirme

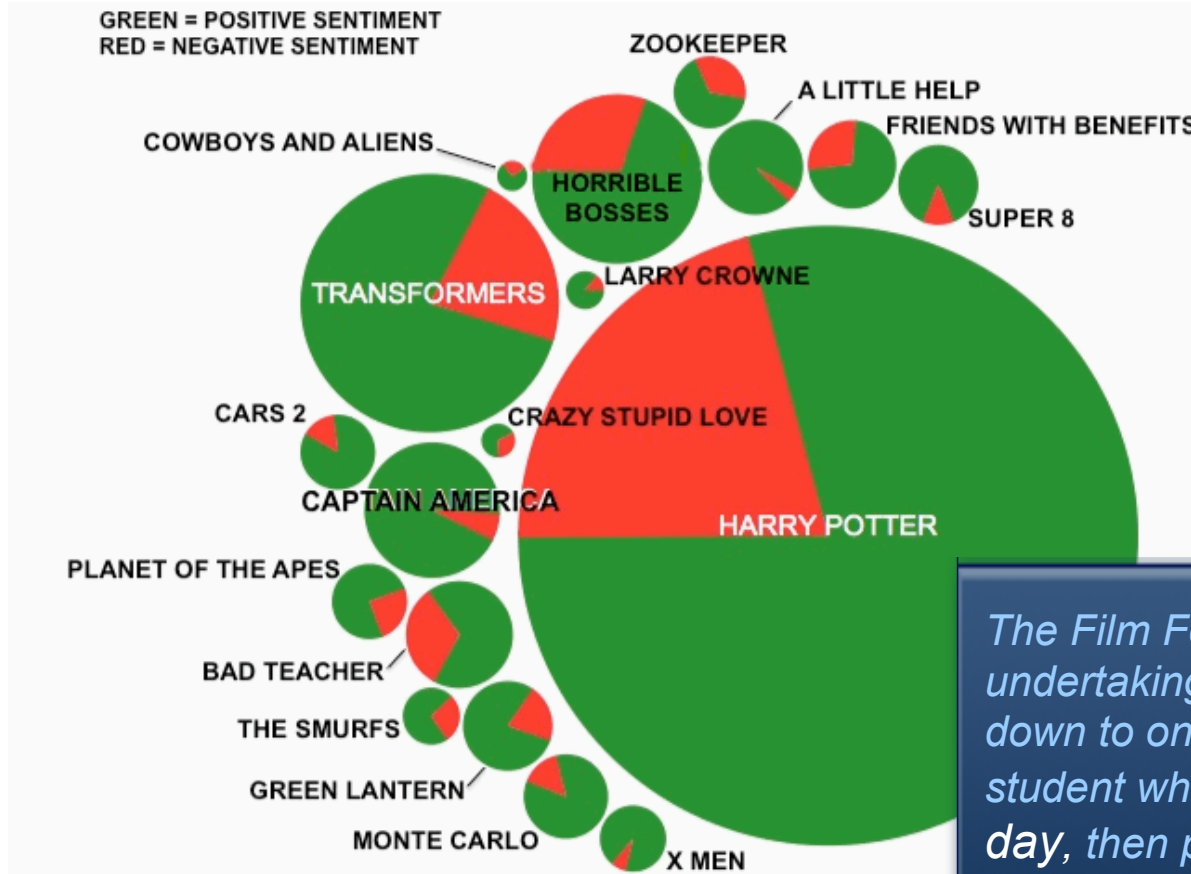
Veri keşfi, veri toplama (crawling), ve analitik



Geliştirme ve Yönetim Araçları

Aşına olduğunuz geliştirme araçları ve ortamı, test ve optimizasyon, sistem yönetimi

Herkes için kolay kullanım ve analiz



The Film Forecaster sounds like a big undertaking for USC, but it really came down to one communications masters student who learned Big Sheets in a day, then pulled in the tweets and analyzed them - Ryan Kim

Analitik Uygulamalar

BI / Raporlama | Analiz/ Görselleştirme | Fonksiyonel Uyg. | Endüstri Uyg. | Veri Madenciliği | İçerik Analitiği

IBM Big Data Platformu

Görselleştirme
ve Keşif

Uygulama
Geliştirme

Sistem
Yönetimi



Hızlandırıcılar

Hadoop
Sistemi



Akışkan
veri işleme



Veri
Ambarı



Veri Entegrasyonu ve Sahipliği

5 Temel Kullanım Alanı



Büyük Veri Arama

Karar alma sürecini geliştirmek için tüm büyük veri kaynaklarını aramak, veriyi anlamak, görselleştirmek



Gelişkin 360° Müşteri Görüntüsü

Yeni İç ve Dış veri kaynaklarıyla mevcut Müşteri Görüntüsünü genişletin



Yeni Güvenlik / İstihbarat Yetkinlikleri

Düşük risk, gerçek zamanlı dolandırıcılık izleme ve siber güvenlik algılama



Operasyon Analizi

Makinelerin ürettiği farklı çeşitlilikteki verilerin analiz edilerek iş sonuçlarının geliştirilmesi



Veri Ambarı Eklentisi

Operasyonel verimliliği artırmak için büyük veri ve veri ambarı özelliklerini entegre etmek

Operasyon Analizi: İhtiyaç



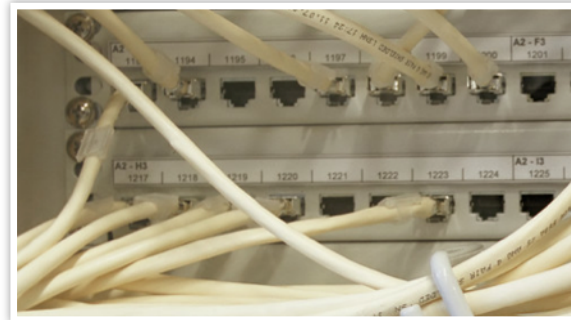
Gelişkin iş sonuçları için, çeşitli cihazların ürettiği verilerin analiz edilmesi

Machine Data Analytics Accelerator

Ne yapar ?

- Çok çeşitli log ve makine verisinin toplanıp, işlenmesi ve ilişkilendirmesini sağlar

1	11.1.2.4	Oct 21 06:33:45	hex('1043882D078A4D16') syslog	%L2-BDF	Cisco - Num
2	11.1.2.4	Oct 21 06:33:45	hex('1043882D078A4D16') syslog	%L2-BDF	AnyVendor
3	11.1.2.4	Oct 21 06:33:45	hex('1043882D078A4D16') syslog	%L2-BDF	PNOC - Int
4	11.1.2.4	Oct 21 06:33:45	hex('1043882D078A4D16') syslog	%L2-BDF	PNOC - Int
5	11.1.2.4	Oct 21 06:33:45	hex('1043882D078A4D16') syslog	%L2-BDF	AnyVendor
6	11.1.2.4	Oct 21 06:33:45	hex('1043882D078A4D16') syslog	%L2-BDF	Cisco - Num
7	11.1.2.4	Oct 21 06:33:45	hex('1043882D078A4D16') syslog	%L2-BDF	AnyVendor
8	11.1.2.4	Oct 21 06:33:45	hex('1043882D078A4D16') syslog	%L2-BDF	AnyVendor
9	11.1.2.4	Oct 21 06:33:45	hex('1043882D078A4D16') syslog	%L2-BDF	PNOC - Int
10	11.1.2.4	Oct 21 06:33:45	hex('1043882D078A4D16') syslog	%L2-BDF	PNOC - Int
11	11.1.2.4	Oct 21 06:33:45	hex('1043882D078A4D16') syslog	%L2-BDF	AnyVendor
12	11.1.2.4	Oct 21 06:33:45	hex('1043882D078A4D16') syslog	%L2-BDF	AnyVendor
13	11.1.2.4	Oct 21 06:33:45	hex('1043882D078A4D16') syslog	%L2-BDF	Cisco - Int
14	11.1.2.4	Oct 21 06:33:45	hex('1043882D078A4D16') syslog	%L2-BDF	Cisco - Num
15	11.1.2.4	Oct 21 06:33:45	hex('1043882D078A4D16') syslog	%L2-BDF	PNOC - Int
16	11.1.2.4	Oct 21 06:33:45	hex('1043882D078A4D16') syslog	%L2-BDF	PNOC - Int



Makine Verilerinin Analizi Temelde Zorlu Sorunlar Barındırıyor

Veri Kaynakları ve Entegrasyon



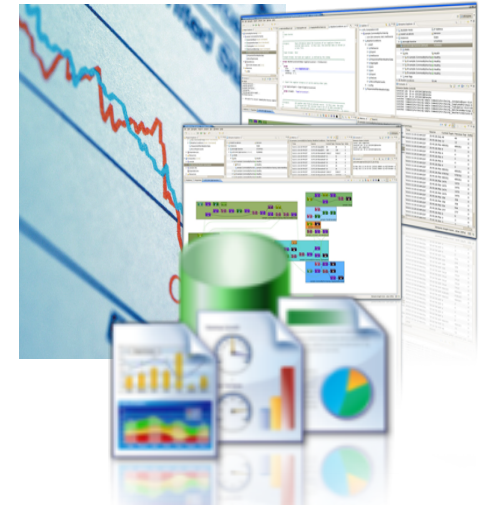
- Standart olmayan kompleks formatlar
- Yüksek hacim
- Kurumsal ve makine verisi karışımı
- Sürekli akan ve biriken veriler
- Birbiri arasında ilişkilendirmede tutarsızlıklar (zaman damgası farklılıkları, IP adres formatları, zaman dilimi vs.)

Analitik

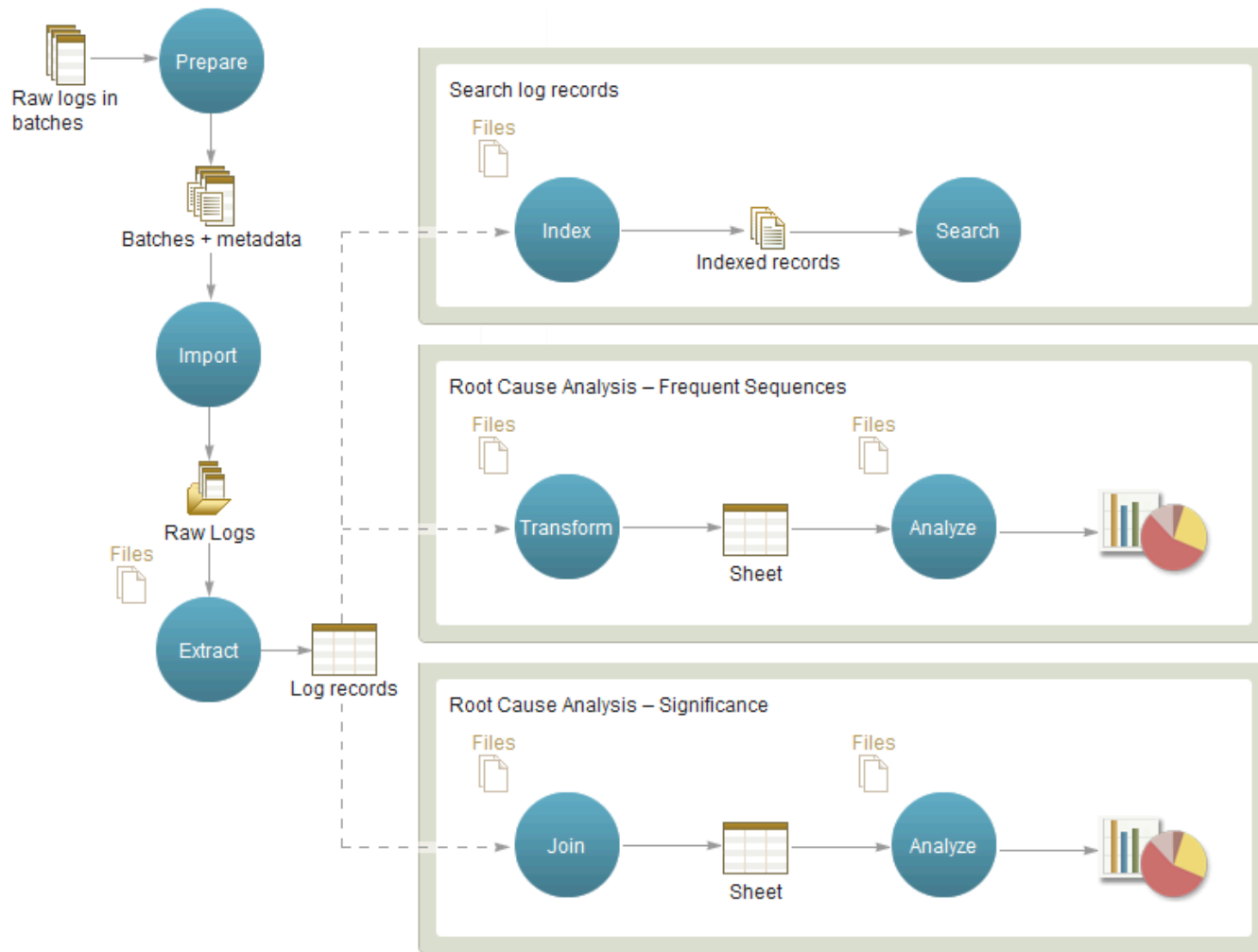


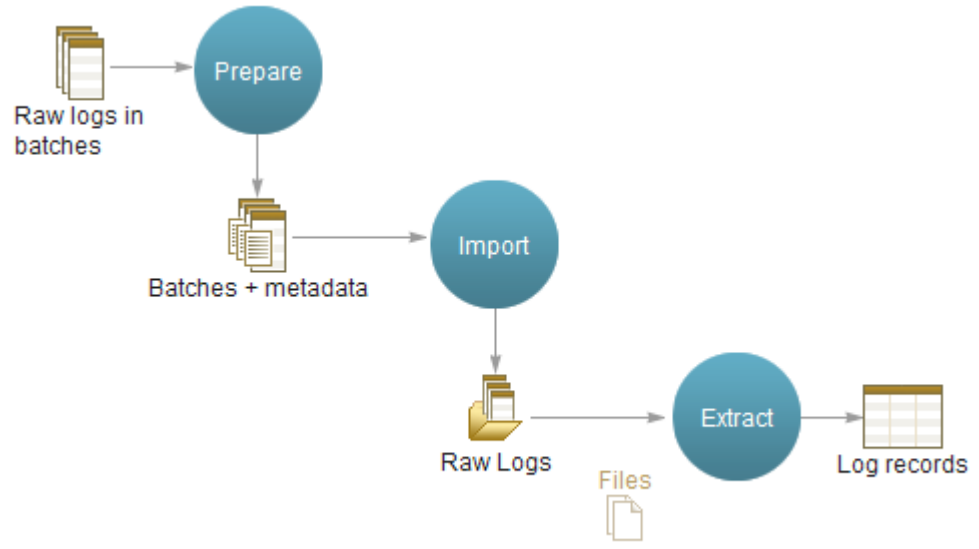
- Yüksek hacimli indeksleme
- Farklı veri setleri arasında korelasyon
- Farklı veri biçimleri için ileri düzey analitik

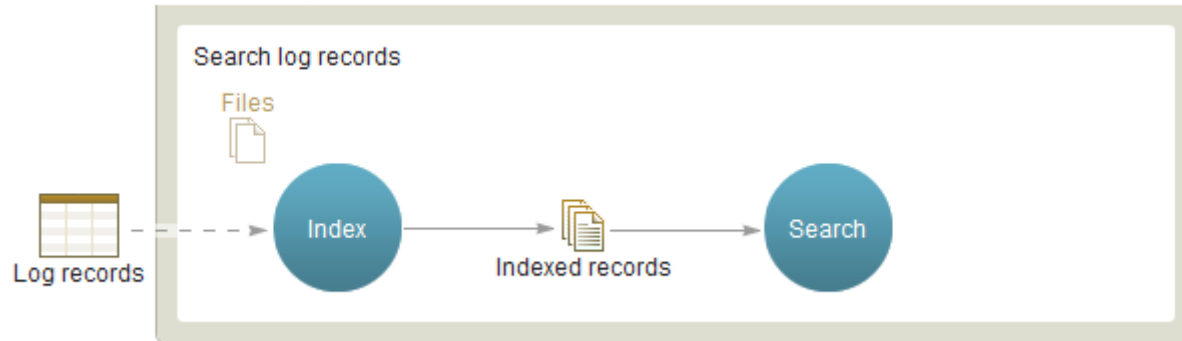
Görselleştirme / Aksiyon ve Çıktı

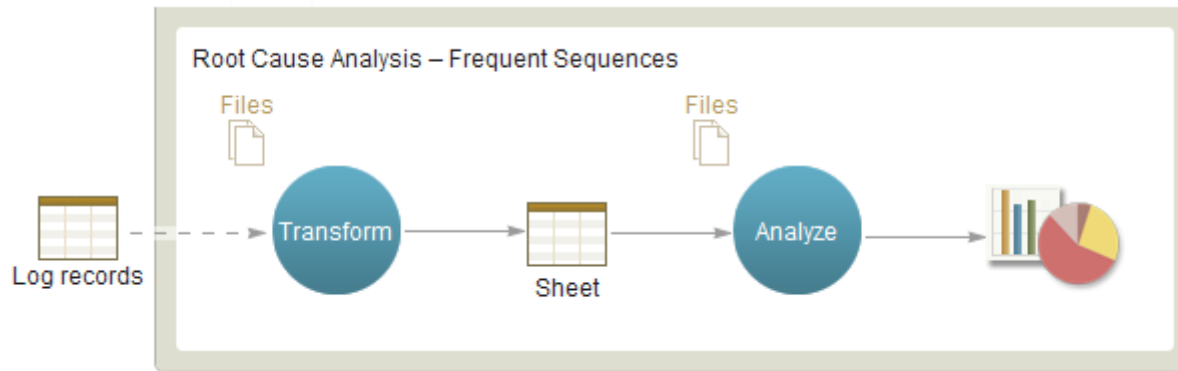


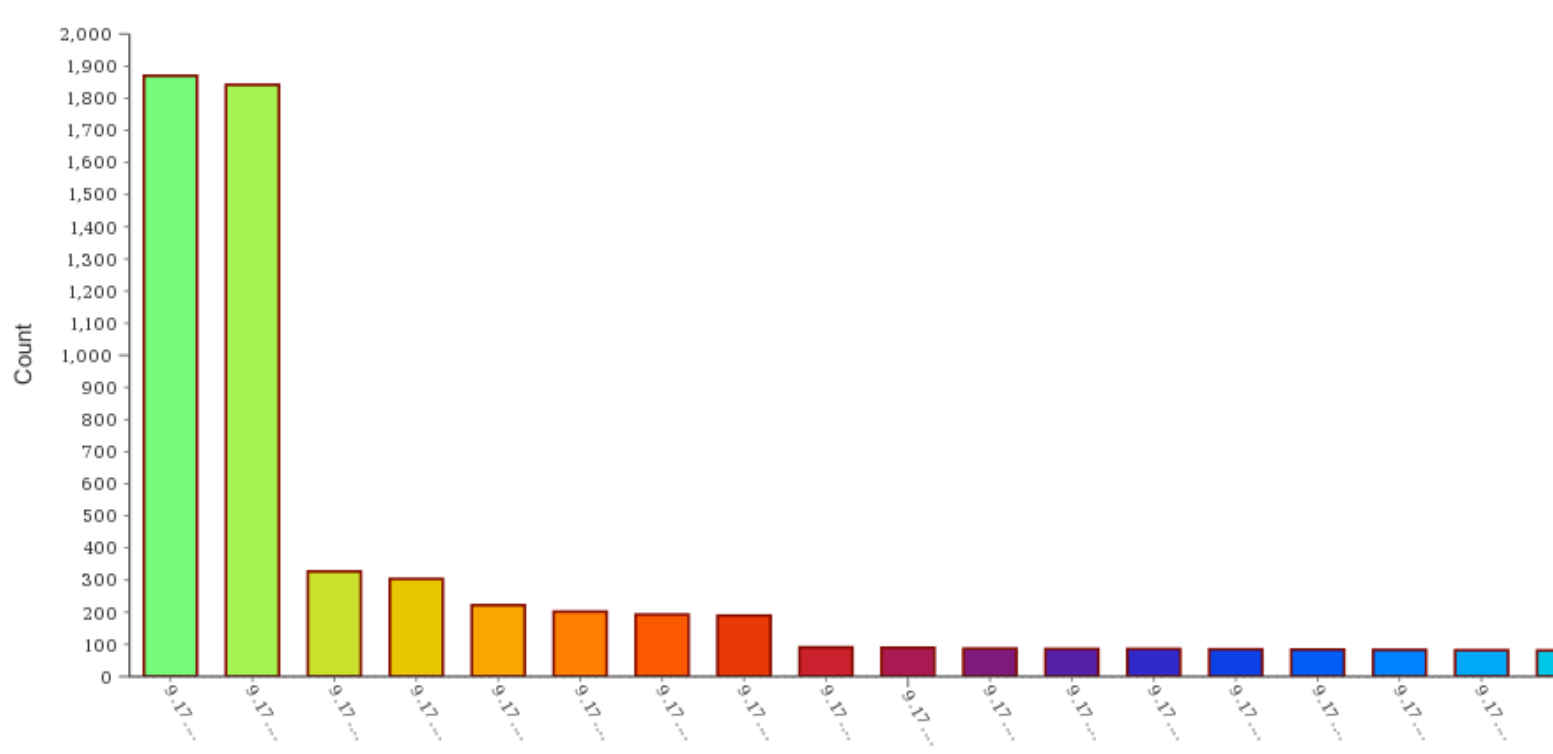
- Akan ve büyük hacimli veriler için yeni görselleştirme yetkinlikleri
- Gerçek zamanlı göstergeler
- Coğrafi zenginleştirme
- Büyük hacimli veriler üzerinde gezinti











Social Data Analytics Accelerator

Ne yapar ?

- Yüksek hacimli farklı tipten sosyal medya verisinin gerçek zamanlı analiz edilebilmesine olanak sağlar

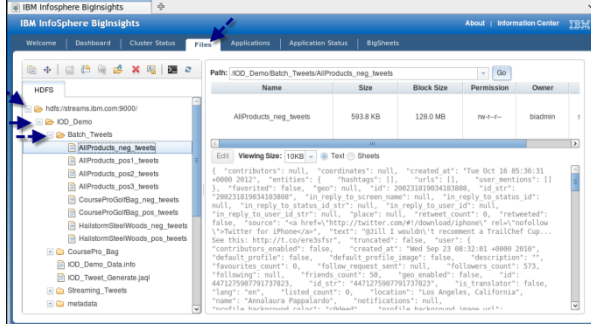
Neden önemli?

- Sizlerin basit bir şekilde iş hedeflerinize ulaşmanızı sağlar:
 - Efektif/Hedef Bazlı Pazarlama Kampanyaları
 - Daha Hızlı Verilebilen Ürün/Pazarlama Kararları
 - Ticari Rekabet Avantajı
 - Müşteri Elde Tutma ve Yeni Müşteri Kazanma

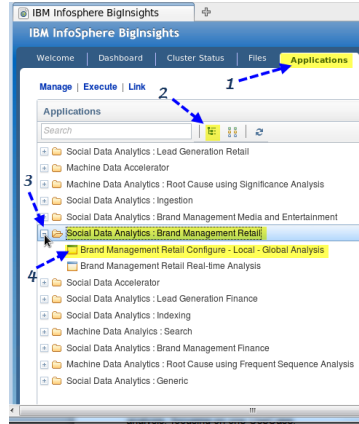


Social Data Analytics

Örnek : “SDA ile Perakende Marka Yönetim Adımları”



Verinin uygun anahtar sözcüklerle içeri aktarılması (Decahose + Powertrack + Boardreader blogs+ forums)

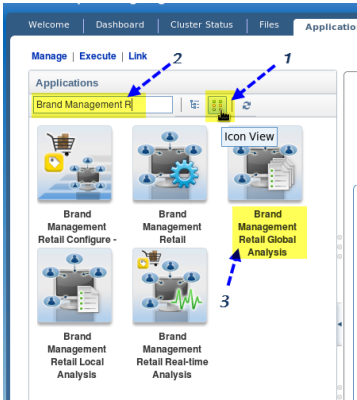


Uygun Uygulamaların Seçimi Ör. Marka Yönetimi

This screenshot shows a table with two columns: 'SearchObject' and 'Category'. The table contains five rows of data related to camping equipment.

SearchObject	Category
TRAILCHEF WATER BAG	CAMPING EQUIPMENT
TRAILCHEF CANTEN	CAMPING EQUIPMENT
TRAILCHEF KITCHEN KIT	CAMPING EQUIPMENT
TRAILCHEF CUP	CAMPING EQUIPMENT
TRAILCHEF COOK SET	CAMPING EQUIPMENT

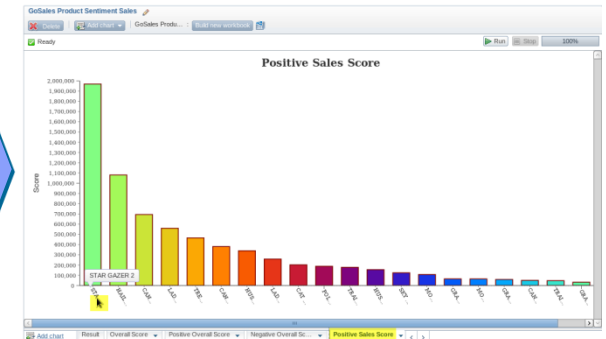
Uygun Perakende şeması kullanılarak SDA için girdilerin sağlanması



Yerel-Toplu Analizlerin Çalıştırılması

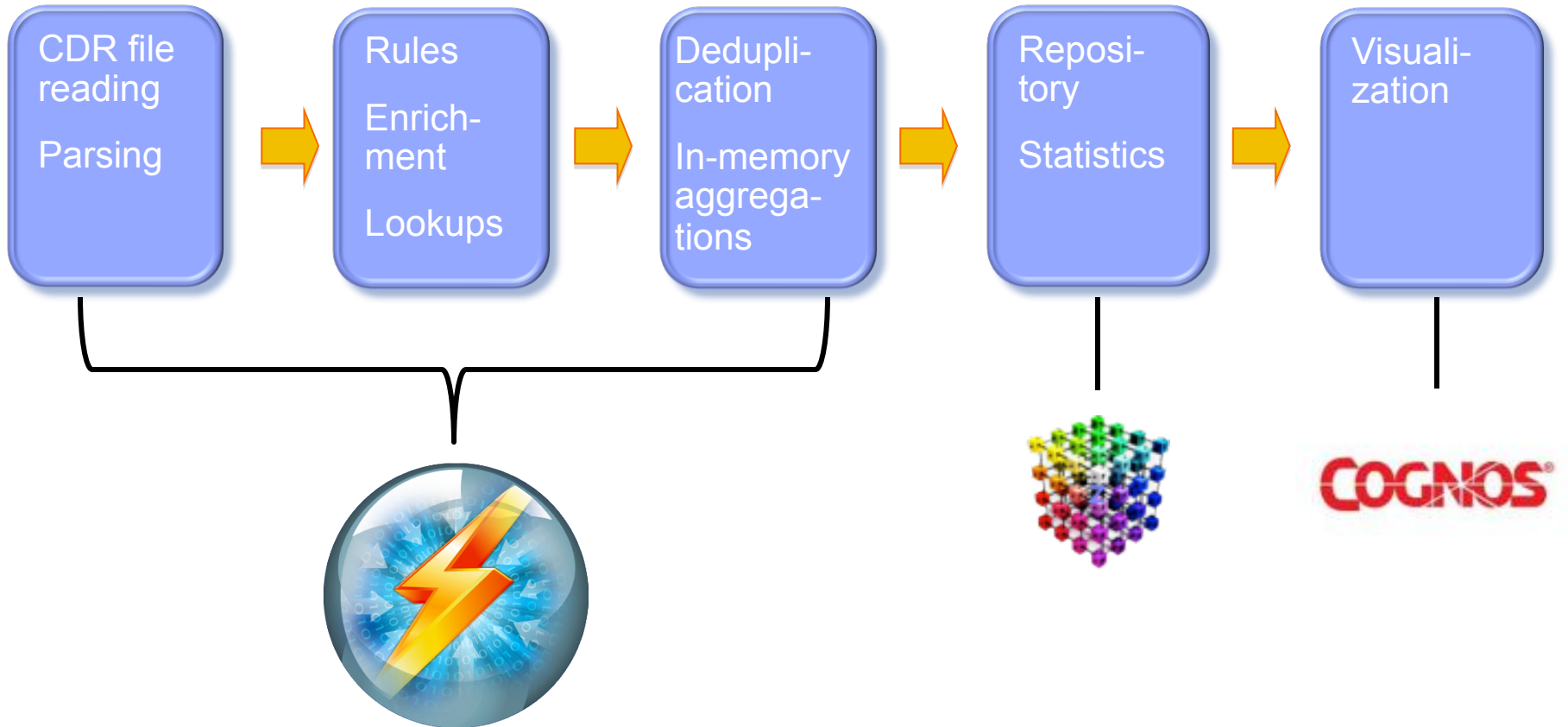
This screenshot shows a 'Workflow Summary' table with columns for Status, Job Status, ID, Name, Type, Start Time, End Time, and Job ID. The table lists several successful jobs related to brand management and retail analysis.

Status	Job Status	ID	Name	Type	Start Time	End Time	Job ID
SUCCEEDED	SUCCEEDED	0000000-121016235110548-ooce-bas-w	sd_a_global_analyt_retail_brand_management	java	Oct 17, 2012 4:20:20 PM	Oct 17, 2012 4:20:39 PM	job_201210162347_0001
SUCCEEDED	SUCCEEDED	0000000-121016235110548-ooce-bas-w	sd_a_global_analyt_retail_brand_management	java	Oct 17, 2012 4:20:39 PM	Oct 17, 2012 4:20:54 PM	job_201210162347_0002
SUCCEEDED	SUCCEEDED	0000000-121016235110548-ooce-bas-w	sd_a_global_analyt_retail_brand_management	java	Oct 17, 2012 4:20:54 PM	Oct 17, 2012 4:21:09 PM	job_201210162347_0003
SUCCEEDED	SUCCEEDED	0000000-121016235110548-ooce-bas-w	sd_a_global_analyt_retail_brand_management	java	Oct 17, 2012 4:21:09 PM	Oct 17, 2012 4:21:21 PM	job_201210162347_0004
SUCCEEDED	SUCCEEDED	0000000-121016235110548-ooce-bas-w	sd_a_global_analyt_retail_brand_management	java	Oct 17, 2012 4:21:21 PM	Oct 17, 2012 4:21:32 PM	job_201210162347_0005
SUCCEEDED	SUCCEEDED	0000000-121016235110548-ooce-bas-w	sd_a_global_analyt_retail_brand_management	java	Oct 17, 2012 4:21:32 PM	Oct 17, 2012 4:21:42 PM	job_201210162347_0006



Gösterge Panelleri kullanılarak görselliğin sağlanması

Telecommunications Event Data Analytics Accelerator



Nasıl başlayabilirsiniz ?

■ Cloud / Bulut Bilişim Sistemleri

- RightScale, ya da Amazon, Rackspace, IBM Smart Enterprise Cloud, ya da özel bulut sistemleri üzerinde
- Sadece kullandığınız kaynak kadar ödersiniz

■ Streams Ürünü Test Sürüşü

- <http://www.ibm.com/developerworks/wikis/display/im/IBM+Virtual+Appliances#IBMVirtualAppliances-ist>
- IBM Big Data : <http://www.ibmbigdatahub.com>

■ Sanal Sınıf

- Ücretsiz eğitim : www.bigdatauniversity.com

■ Kendi Sunucularınızda

- Basic Edition'ı ücretsiz indirebilirsiniz

■ Sınıf içi Eğitimler

BigData University BETA
Learn from the industry's best

English

login sign up

HOME LEARN DOWNLOAD RESOURCES JOBS LEARN Hadoop

search courses go

Hadoop Fundamentals
Start your career on Hadoop with this **FREE** course
Learn with hands-on exercises on a Hadoop cluster
Enroll now!

Why register?

- Easy and Affordable Learning** Hadoop and other Big Data technologies has never been more affordable! Many courses are FREE!
- Latest industry trends** Acquire valuable skills and get updated about industry's latest trends right here. Today!
- Learn from the Experts!** Big Data University offers education about Hadoop and other technologies by the industry's best!
- Learn at your Own Pace!** Find everything right here when you need it and from wherever you are.

sign me up

View sample completion document!

sign up now for a chance to receive: FREE Books
SIGN UP

Study Made Easy!

Student Testimonials go to sign up

Balázs (USA)
The training material has short, easy to digest videos, with separately available transcripts allowing to execute the same commands after watching the videos. Each command line in the transcript can be cut and pasted making the examples easy to reproduce. There is very good support for Windows users both on 32bit and 64bit. Everything works right out of the box as described in the course materials. Online support on the Course Forums is excellent, most questions were answered before even I encountered the issue. The curriculum associates the chapters from the "Getting Started with Hadoop" book.