

# Veritabanı Yönetim Sistemleri (335)

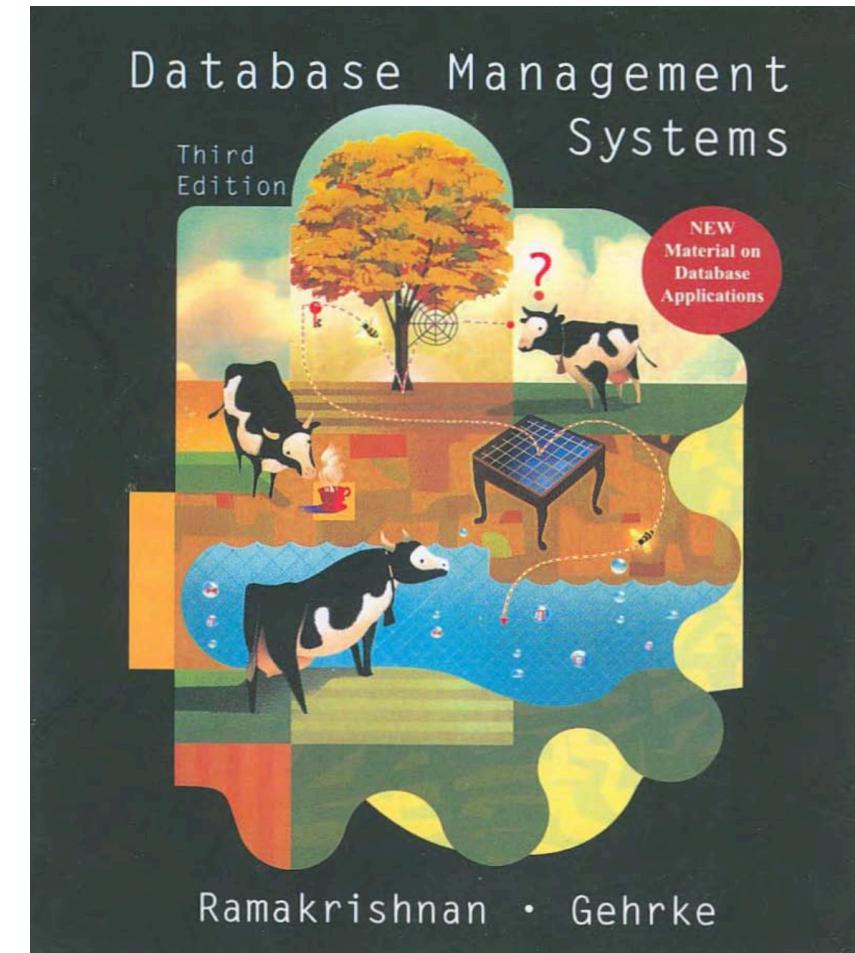
Dr. Ahmet Arif AYDIN

# Hedeflerimiz

- Veritabanı Yönetimi Sistemleri'ni kavramak
- Veritabanına kaydedilecek veriyi modelleyebilmek
- Veritabanı denetim, erişim ve yönetim metotlarını öğrenmek
- Veritabanı uygulamaları geliştirmek

# Ders Materyali ve Ofis Saatleri

- Ders Kitabı
  - Database Management Systems, 3rd Edition
  - Raghu Ramakrishnan , Johannes Gehrke
- Derste yapılan Sunular.
- Labaratuardan Sorumlu Hocalarımız
  - Kenan İNCE
  - Ahmet KARADOĞAN



[https://bayanbox.ir/view/8736593520639826197  
/Ramakrishnan-Database-Management-Systems-  
3rd-Edition-1-1.pdf](https://bayanbox.ir/view/8736593520639826197/Ramakrishnan-Database-Management-Systems-3rd-Edition-1-1.pdf)

# Puanlama

- Vize
  - Labaratuvar Uygulamaları: **50p**
  - Vize : **50p**
- Final
  - Labaratuvar Uygulamaları : **50p**
  - Proje veya Final **50p**



# Tavsiyeler

- Dersi ciddiye alın !
- Zamanınızı dikkatli kullanın!
- Sorumluluklarınızı tekrar gözden geçirin!
- İletişim kopukluğu olmasın
  - Ofis saatlerini mutlaka değerlendirin!
    - F Blok 2.Kat No:1
    - Perşembe (13:30-17:00)
    - İÖ: Çarşamba (13:30-17:00)
  - Email'i ciddi olarak kullanın!
    - [arif.aydin@inonu.edu.tr](mailto:arif.aydin@inonu.edu.tr)
  - Duyurular, ödevler ve ders notları için
    - <https://github.com/aaaydin/335DBMS>

# Konular

Hafta	Konular
1	Veritabanı Tasarımı
2	ER-Model
3	UML
4	Relational Model
5	İlişkisel Cebir
6-7	SQL
8	Dosya Yapıları ve İndexleme
9	Sorgu Optimizasyonu
10	Hareket Yönetimi
11	Kilitlenmeler
12	Concurrency
13	Veritabanı Kurtarma Teknikleri

# Big Data Evresi

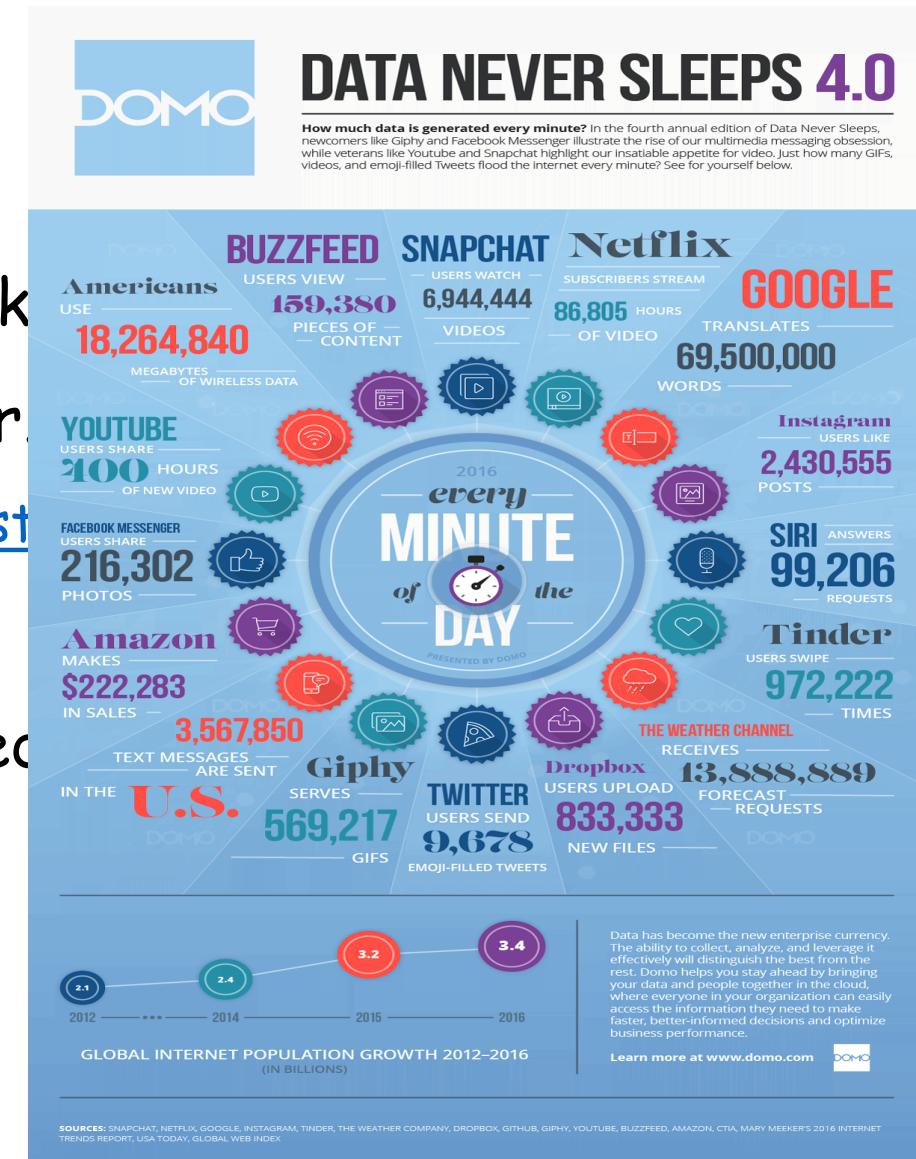
Big data Çağımızdayız !

- Çeşitli kaynaktan üretilen
  - Multimedya
  - Text
  - Ses
  - Video
  - Fotograf
  - Sinyal
  - Konum
  - Uydu



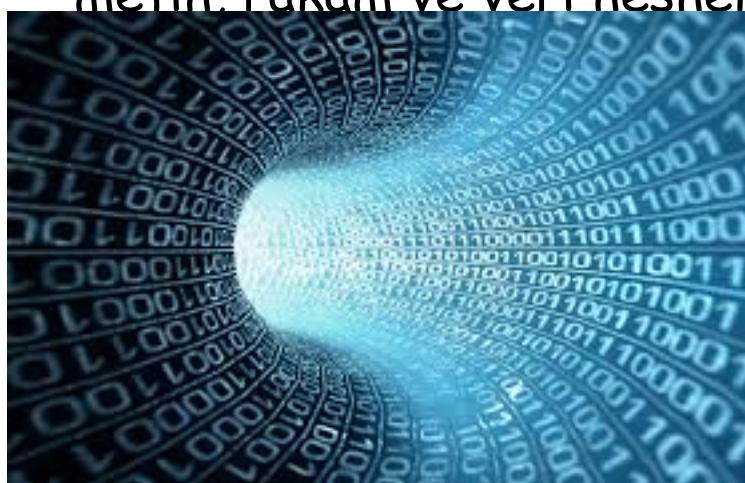
# Veri Her Yerde

- 60 saniye (1 dakika) içerisinde
  - Google da 4 milyon arama gerçekleştirilmektedir.
  - Twitter da 350 bin tweet gönderilmektedir.
    - <http://www.internetlivestats.com/twitter-statistics/>
  - Facebook da 4 milyon ilan işlemi
  - Youtube 400 saatlik yeni video eklenmektedir.



# Veri Nedir?

- **Veri (raw data-ham veri)**
  - Çeşitli kaynaklardan elde edilir
  - **İşlenmemiş**
    - metin, rakam ve veri nesnelerinin oluşturduğu koleksiyondur



# Bilgi Kavramı

- **Bilgi (information)**

- Ham veri kullanılarak
  - çeşitli veri analizi yöntemleri ile faydalı çıkarımların yapılması sonucu elde edilir.
- Data Mining, Machine Learning



# Veri Evresi

- Çok miktarda veri
- 1 TB
  - 1,498 CD
  - Hubble Teleskopu her yıl 10TB yeni data
  - IBM Watson super computer 16 TB RAM
- 1 PB
  - 3.4 yıllık 24/7 Full HD video
  - 1.5 Million CD
- 1 EB
  - 1.5 Milyar CD

<https://www.lifewire.com/terabytes-gigabytes-and-petabytes-how-big-are-they-4125169>



Byte (B)	1
Kilobyte (KB)	1,024 B
Megabyte (MB)	1,024 KB
Gigabyte (GB)	1,024 MB
Terabyte (TB)	1,024 GB
Petabyte (PB)	1,024 TB
Exabyte (EB)	1,024 PB
Zettabyte (ZB)	1,024 EB
Yottabyte (YB)	1,024 YB

# Veri ve Analizi Niçin Önemli?

- Veriden faydalı çıkarımlar yapmak
- Veriyi kullanarak tahminde (predictions) bulunmak
- Müşterilerin isteklerine cevap verebilmek

*“Without big data analytics, companies are blind and deaf, wandering out onto the web like deer on a freeway.”*

– Geoffrey Moore, author and consultant.



<http://www.bestinfographic.co.uk/2012/10/18/ibm-big-data-infographic/>

# Veritabanı (Database) Nedir?

Belirlenen özel amaclar doğrultusunda  
birbirleriyle ilişkili verilerin  
belirli kurallar çerçevesinde  
bir araya getirilmesiyle oluşturulan  
kolleksiyonuna veritabanı (database) denir.



9/28/17



# Veritabanı (Database) Nedir?

- Başarılı veritabanları uzun yıllar kullanılacaklarından
- veritabanı gelecekte istenilecek olan değişimlere (changeability) cevap verebilecek biçimde tasarılanmalıdır (evolvability).



9/28/17



# Veritabanının Kullanım Alanları?

- Bankacılık sistemi
  - Hesaplar üzerinden gerçekleştirilecek olan her türlü işlem

# Veritabanının Kullanım Alanları?

- E-ticaret (milyonlarca kayıt)
  - Amazon
    - Ürün, müşteri ve satış bilgileri
  - Netflix
    - kullanıcılar, filmler, müşteri bilgileri
  - Youtube
    - Videolar, kullanıcılar
  - Google
    - Aramalar, hesaplar, google drive
  - Twitter
    - Kullanıcılar, hesaplar

# Veritabanının Kullanım Alanları?

- Hava yolu şirketleri
  - online rezervasyonlar, uçuş bilgileri, yolcu bilgileri
- Üretim sektörü
  - Çalışanların, üretim, ihracat ve ithalat bilgileri

# Veritabanının Kullanım Alanları?

- Resmi ve özel kurumlar
  - Üniversiteler: öğrenci bilgileri ve notları, akademik ve idari personel bilgileri
  - E-devlet: milyonlarca kişinin bilgileri
- Akademik yayın sistemleri
  - IEEE, ACM, ELSEVIER
- Diğer organizasyonlar
  - NASA
  - NEIC (National Earthquake Information Center)
  - WHO (World Health Organization)
  - AFAD

# VTYS'nin Tarihsel Gelişimi

- 1960 lı yılların başında IDS (Integrated Data Store)
  - İlk genel amaçlı VTYS sistemi
  - Charles Bachman tarafından General Electric için geliştirilmiştir
  - Bachman bu sistem ile 1973 de ACM Turing odulunu almaya hak kazanmıştır.

# VTYS'nin Tarihsel Gelişimi

- 1960 yılların sonuna doğru IBM IMS (Information Management System)
  - Hiyerarşik veri modelinin temelini oluşturmuştur.

# VTYS'nin Tarihsel Gelişimi

- 1960'lı yıllarda **SABRE** sistemi
  - 1964'de American Airlines ve IBM tarafından kullanılmaya başlanmıştır
  - Ağ üzerinden kullanıma açılan ilk veritabanı olmuştur.
  - SABRE günümüzde halen kullanılmaktadır.

<https://www.sabre.com/files/Sabre-History.pdf>

# VTYS'nin Tarihsel Gelişimi

- **1970'de Edgar Codd**
  - IBM's San Jose Research Lab
  - **İlişkisel veri modeli** (relational data model) kavramını ortaya çıkarmıştır
  - 1981 tarihinde ACM's Turing ödülünü almıştır.

# VTYS'nin Tarihsel Gelişimi

- 1980'de, IBM tarafından RVTYS için SQL dili geliştirildi.
- 1980 - 1990 arasında
  - IBM's DB2, Oracle, Informix
- Internet erişimiyle beraber
  - web-browser yardımıyla VTYS'lerine erişim gerçekleştirildi.

# Örnek

- Bir E-ticaret firmasının
  - Çok sayıda çalışanı, ürün listesi, ve satışları bulunmaktadır
  - Kaydedilen verinin boyutu da yaklaşık olarak 1TB olsun.



# Örnek

- Veri koleksiyonundan istenilen özellikler
  - eş zamanlı olarak birden fazla kullanıcının erişimi sağlanabilecek
  - Yeni veri eklenebilecek
  - Var olan kayıtlar güncellenebilecek veya silinebilecek
  - sorular makul zaman içerisinde cevaplandırılacak
  - Verinin istenilen bazı kısımlarına erişim kısıtlanacak

# Dosya Sistemleri ve Problemler

- 1 TB'lık ana bellek (RAM) olmadığından verinin hepsi aynı anda belleğe alınamayacak. Veri hard disk üzerine kaydedilecek ve her defasında verinin bir parçası hafızaya alınıp işlenmesi gerekmektedir.
- Kullanıcıların sorularını cevaplamak için özel programların geliştirilmesi gerekmektedir.
- Farklı kullanıcılar tarafından eş-zamanlı erişilen verinin tutarlığını korumak gerekmektedir. Ve her bir uygulamanın bunu sağlaması oldukça zordur.
- Sistem çökmelerinde ve sistem üzerinde değişiklikler yapıldığında verinin tutarlı bir biçimde tutulduğu garanti edilmelidir.
- İşletim sistemleri sadece şifre güvenliği sağlamaktadır. Farklı kullanıcılar için izinleri ve şifreleri düzenleme noktasında dosya sistemleri yeterli seviyede esnek değildir.

# VTYS'nin Avantajları

- **Veri Bağımsızlığı (Data Independence)**

- Uygulama programları tarafından verinin hepsine erişim uygun görülmemektedir.
- VTYS'leri verinin genel bir görünüşünü verip diğer detaylarını izole eder.
- Veri bağımsızlığı VTYS'nin kuramsal seviyeleri yardımıyla gerçekleştirilmektedir.

# VTYS'nin Avantajları

- Etkili Veri Bağlantısı (Efficient Data Access)
  - VTYS'leri veriyi kaydetme, sonra tekrar okuma ve elde etme işlemleri için sofistike teknikler sunmaktadır.

# VTYS'nin Avantajları

- Veri Bütünlüğü ve Güvenliği (Data Integrity and Security)
  - VTYS'leri üzerinden erişilen verinin bütünlüğü VTYS tarafından korunur.
  - Birden fazla değişim isteği kısıtlanır.
  - Farklı kullanıcılar için erişim denetimi (access control) sağlanır.

# VTYS'nin Avantajları

- Veri Yönetimi (Data Administration):
  - Ortak kullanılan verinin yönetimi bir çok profesyonel tarafından tasarlanır ve düzenlenir.
  - Verinin sunumunda tekrar ortadan kaldırılarak verinin çağrılmaması optimize edilir.

# VTYS'nin Avantajları

- **Eş zamanlı erişim ve kurtarma (Concurrent Access and Crash Recovery):**
  - VTYS eş-zamanlı erişimlerde her seferinde bir kullanıcının aynı veriye erişimini düzenler
  - Sistem hatalarından ve etkilerinden kullanıcıları ve verileri korur.

# VTYS'nin Avantajları

- Azalan Uygulama geliştirme Zamanı (Reduced Application Development Time):
  - VTYS' ieri önemli fonksiyonları sağladığından tekrar en başından itibaren program yazılması için gereken zaman harcanmasını ortadan kaldırır.

# VTYS'nin Avantajları

- **Veri modelleme**

- Veri kolleksiyonlarının yüksek-seviyeli olarak tanımlanıp ve düşük-seviyeli bilgilerin gizlenmesini sağlayan temsil veri modeli (data model) olarak tanımlanmaktadır.
- VTYS'leri kullanıcılarına veri modelinin oluşturulması ve yönetimini sağlamaktadır.

# Kullanılan VTYS ve Veri Modelleri

- İlişkisel Veri Modeli (Relational Data Model) tabanlı teknolojiler:
  - IBM's DB2,
  - Informix, Oracle, Sybase
  - Microsoft's Access
  - FoxBase
  - Paradox
  - Tandem
  - Teradata
  - PostgreSQL

# Kullanılan VTYS ve Veri Modelleri

- Hiyerarşik Model (Hierarchical model)
  - IBM'in kullandığı IMS sistemi.

# Kullanılan VTYS ve Veri Modelleri

- Ağ modeli (the network model)
  - IDS ve IDMS tarafından kullanılmaktadır.

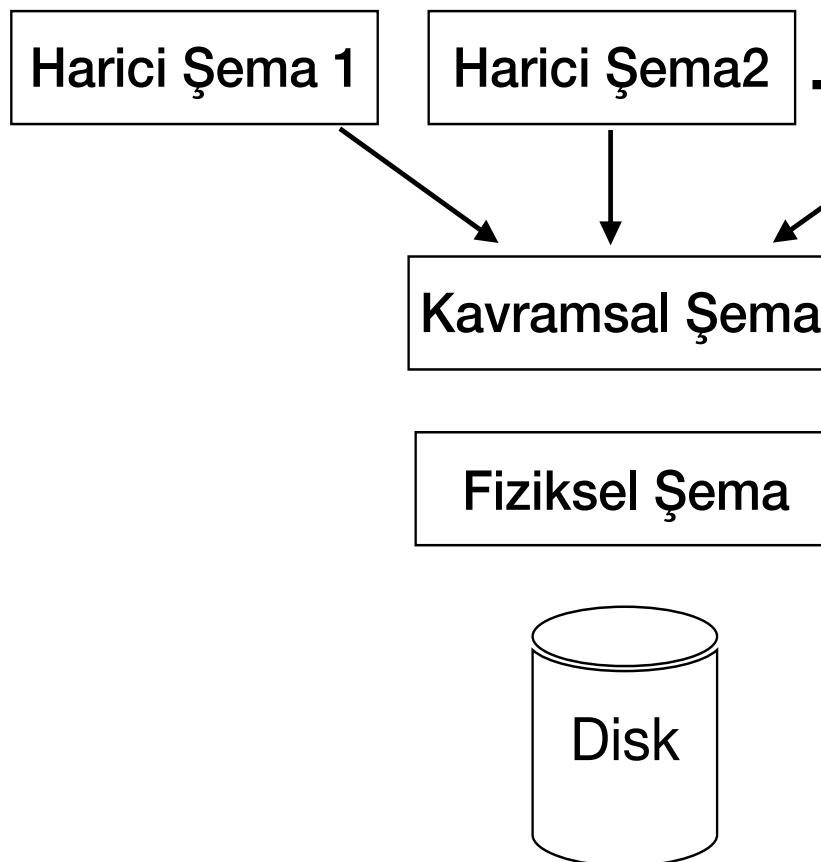
# Kullanılan VTYS ve Veri Modelleri

- Nesne-tabanlı model (object-oriented model)
  - Objectstore ve Versant

# Kullanılan VTYS ve Veri Modelleri

- Nesne-iliskisel model (the object-relational model)
  - IBM, Informix, ObjectStore, Oracle, ve Versant.

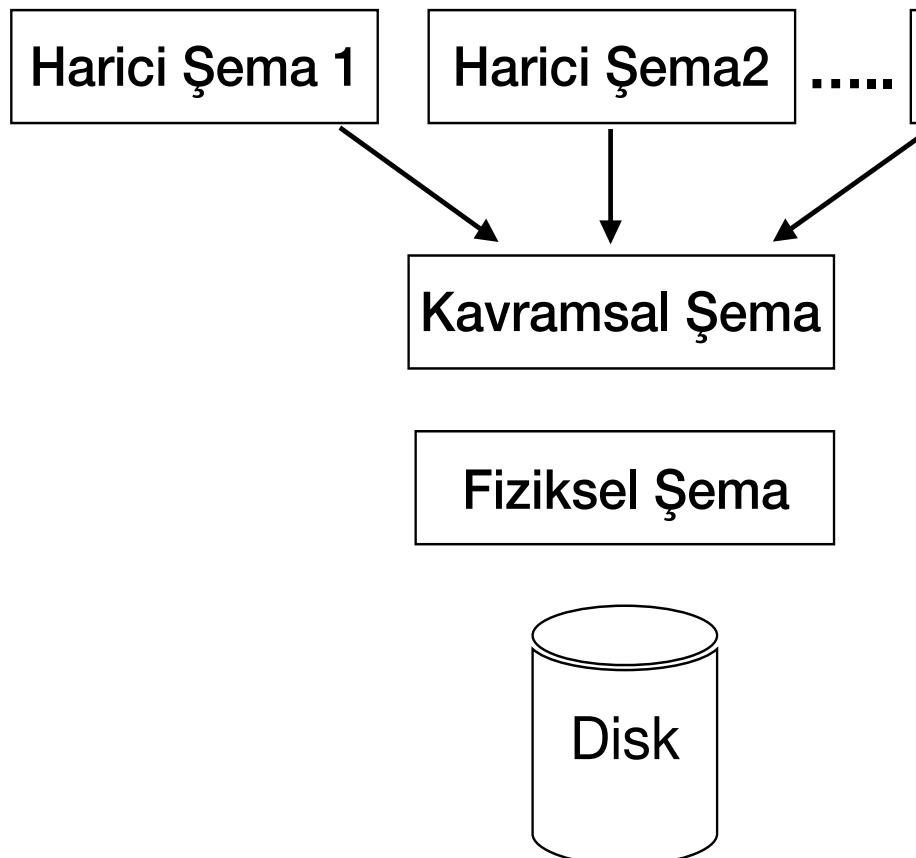
# VTYS'nin Mimarisi ve Kuramsal Seviyeleri



## External Schema

- Birden fazla harici şema olabilir.
- Farklı kullanıcı gruplarının ihtiyaçlarını karşılamak ve çeşitli isteklerine cevap verebilmek için kullanılır
- Harici şema ile verinin farklı alanları ve görünüşleri (view) sunulur.

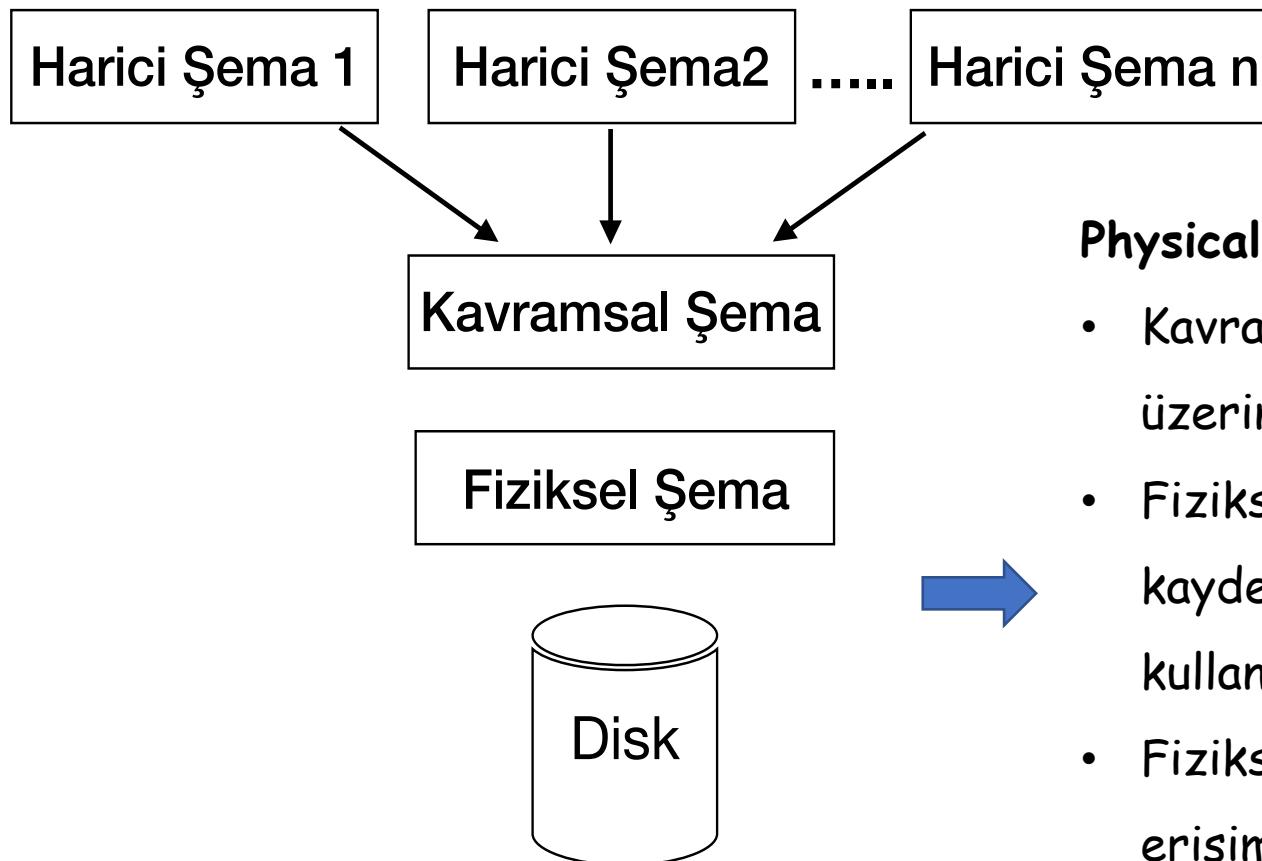
# VTYS'nin Mimarisi ve Kuramsal Seviyeleri



## Conceptual or Logical Schema

- Veritabanında kayıtlı olan verileri veri modelini kullanarak tanımlar ve bütün ilişkileri verir.
- Varlıklar ve varlıklar arasındaki ilişkiler tanımlanır.
- Veri bağımsızlığı ile birlikte kullanıcıların kayıt seviyesindeki değişikliklerden etkilenmelerini minimum seviyeye indirir.

# VTYS'nin Mimarisi ve Kuramsal Seviyeleri



## Physical Schema

- Kavramsal şemada tanımlanan verinin gerçekte disk üzerine nasıl kaydedileceğinin bilgilerini içerir.
- Fiziksel şema bütün ilişkileri sıralı olmayan dosyalara kaydeder ve her bir şemanın (tablo) ilk kolonunu kullanarak index oluşturur.
- Fiziksel şemanın tasarımı ile ilgili kararlar ilerde veriye erişimi birinci dereceden etkilemektedir.

# VTYS' Kullanıcıları

- **Geliştiriciler (Developers)**
  - IBM, Oracle, DataStax gibi büyük ölçekli firmalar adına çalışmaktadır.
  - **Geliştirilen VTYS'leri**
    - çok çeşitli disiplinlerden
    - farklı tecrübesi olan organizasyonlar kullanmaktadır.

# VTYS' Kullanıcıları

- **Kullanıcılar (Users)**
  - **Teknik bilgisi olmayan kullanıcılar**
    - veritabanı uygulama programlarının hazırladığı form web arayüzlerini kullanarak sorgularını veritabanına yönlendirirler.
  - **Sofistike kullanıcılar**
    - VTYS'ler kendi sorgularını SQL ile yazmalarına imkan sağlamaktadır

# VTYS' Kullanıcıları

- Uygulama Programcıları (Application Developers)
  - veritabanının sorgulama dilini kullanarak istenilen sorguları içeren uygulamaları geliştirirler.
    - raporlar, istatistikler, tablolar
  - Bu uygulamalar harici şema kullanılarak yazılmaktadır.

# VTYS' Kullanıcıları

- VTYS Admin (Database Administrator (DBA))
  - Büyük ölçekli firmaların veritabanlarının tasarıımı, yönetimini ve mevcut durumunu koruması için profesyonel destek (DBA) gerekmektedir.

# VTYS' Kullanıcıları

- Veritabanı yöneticisinin görevleri
  - Kavramsal ve fiziksel Şema tasarımı (Conceptual and Physical Schema Design)
  - Güvenlik ve Yetkilendirme (Security and Authorization)
  - Verinin kullanılabilirliği ve hatalardan kurtarılması (Data availability and Recovery from failures)
  - Veritabanının Düzenlemesi (Database Tuning)
    - Kullanıcı ihtiyaç ve teknik değişikliklere cevap verebilme

# VTYS Sıralaması

334 systems in ranking, September 2017

Rank			DBMS	Database Model	Score		
Sep 2017	Aug 2017	Sep 2016			Sep 2017	Aug 2017	Sep 2016
1.	1.	1.	Oracle	Relational DBMS	1359.09	-8.78	-66.47
2.	2.	2.	MySQL	Relational DBMS	1312.61	-27.69	-41.41
3.	3.	3.	Microsoft SQL Server	Relational DBMS	1212.54	-12.93	+0.99
4.	4.	4.	PostgreSQL	Relational DBMS	372.36	+2.60	+56.01
5.	5.	5.	MongoDB	Document store	332.73	+2.24	+16.74
6.	6.	6.	DB2	Relational DBMS	198.34	+0.87	+17.15
7.	7.	↑ 8.	Microsoft Access	Relational DBMS	128.81	+1.78	+5.50
8.	8.	↓ 7.	Cassandra	Wide column store	126.20	-0.52	-4.29
9.	9.	↑ 10.	Redis	Key-value store	120.41	-1.49	+12.61
10.	10.	↑ 11.	Elasticsearch	Search engine	120.00	+2.35	+23.52
11.	11.	↓ 9.	SQLite	Relational DBMS	112.04	+1.19	+3.41
12.	12.	12.	Teradata	Relational DBMS	80.91	+1.67	+7.84
13.	13.	↑ 14.	Solr	Search engine	69.91	+2.95	+2.95

# Bölüm Soruları

- Veri ile bilgi arasındaki fark nedir?
- Veritabanı nedir?
- VTYS'nin kullanım alanları nelerdir?
- VTYS kullanmanın avantajları nelerdir?
- Dosya sistemleri hangi neden VTYS yerine kullanılmazlar?
- VTYS'lerini kimlerin kullandığını yazınız.
- Günümüzde kullanılan VTYS'nin veri modelleri nelerdir?
- VTYS'lerinin kuramsal seviyeleri nelerdir?
- Veritabanı yöneticisinin görevleri nelerdir?