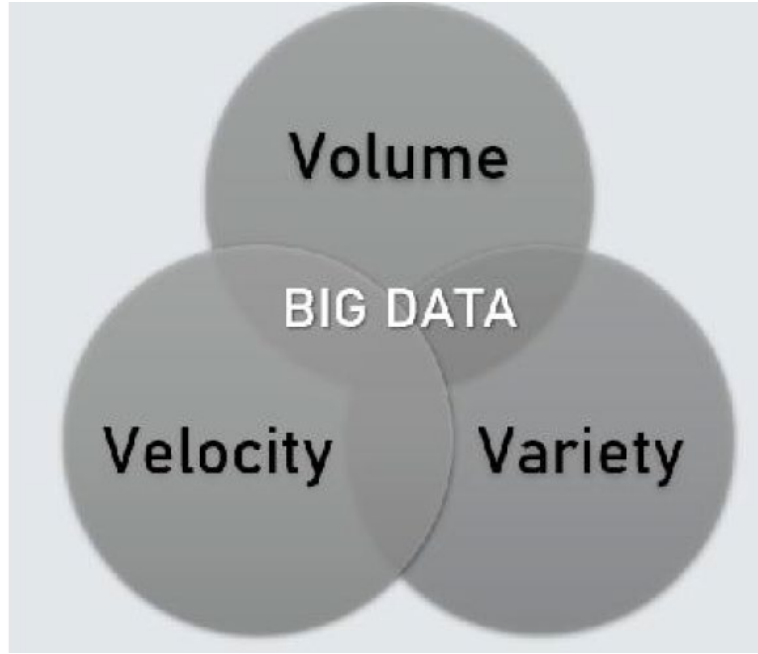


Yazılım Mühendisliğine (YM) Giriş

Büyük Veri / Big Data (devam)
Bilgisayarların Tarihçesi (devam)

2024 GÜZ

- Büyük Data Özellikleri

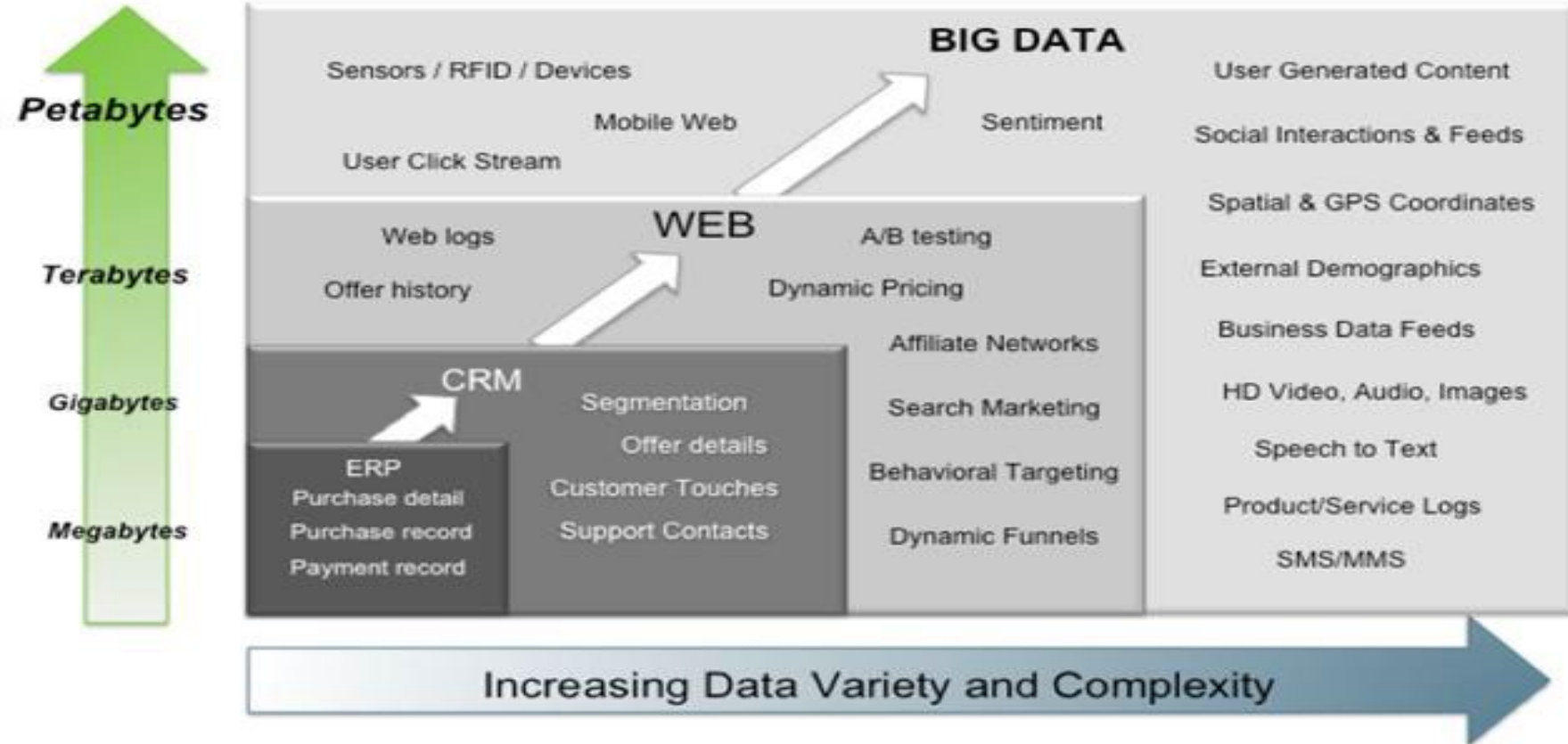


Büyük Verinin 3V Karakteristiklerinden



5V Karakteristiklerine

Büyük Verinin 3V (Volume, Velocity, Variety) Özellikleri



Source: Contents of above graphic created in partnership with Teradata, Inc.

Çeşitlilik (variety) : Yapılandırılmamış veriler, yarı yapılandırılmış veriler ve ham veriler dahil olmak üzere farklı veri türlerinin çeşitliliği ve aralığı. Veriler sayılar, metinler, e-postalar, fotoğraflar ve videolar, ses kayıtları gibi çok farklı biçimde gelebilir.

Hacim (volume): Şirketlerin yönettiği ve analiz ettiği büyük verinin boyutu ve miktarı.

Yıllar geçtikçe veri depolama maliyetlerinin düşmesi kuruluşların verileri daha kolaylıkla depolayabilmesini sağlıyor.

Hız (velocity) : Firmaların verileri alma, depolama ve yönetme hızı. Örneğin bir gün, bir ay veya farklı bir zaman biriminde alınan belirli sayıda sosyal medya gönderisi veya arama sorgusu.

Yıllar geçtikçe gelişen teknolojiyle birlikte veriler daha hızlı hareket eder. Daha hızlı olması, verinin gerçek zamana yakın şekilde çözümlenebilmesini sağlar.

Büyük Veriyi 5V'ye Tamamlamak

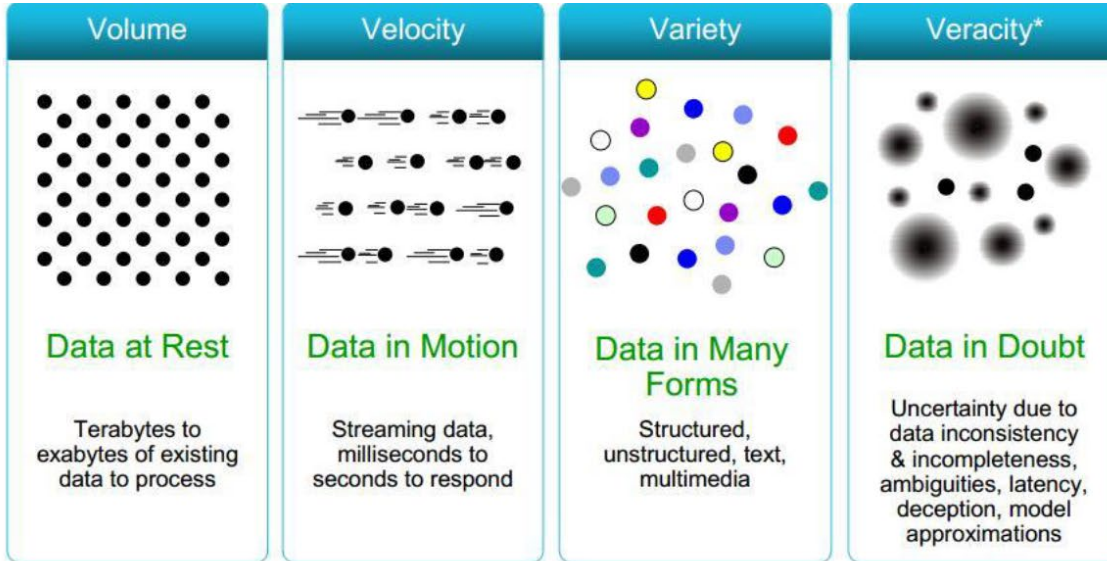
Volume, Velocity, Variety, **Veracity**, and Value

Doğruluk (Veracity) : Genellikle yönetici düzeyinde güveni belirleyecek olan veri ve bilgi kaynaklarının "gerçekliği" veya doğruluğu.

Milyonlarca farklı kaynaktan gelen milyarlarca verinin analiz edilmesi, sonuçların doğruluğunu artırır.

❑ **Değer (value)** : Firma bakış açısından incelendiğinde en önemli "V" olan büyük verinin değeri.

Genellikle daha etkili operasyonlar, daha güçlü müşteri ilişkileri ve iş fırsatlarının net ve ölçülebilir olmasıyla ortaya çıkartılan bir özelliktir.



Value*



İki tane daha Büyük Veri Özelliği : 7V Visualization (Görselleştirme) ve Validity (Geçerlilik)

❑ **Görselleştirme (Visualization):** Büyük veride veri görselleştirmenin önemine atıfta bulunur.

Görselleştirme araçları, analistlerin ve karar vericilerin karmaşık veri kümelerini hızlı bir şekilde anlamalarına yardımcı olmak için kullanılır.

❑ **Geçerlilik (Validity):** Geçerlilik, verilerin gerçekten temsil ettiği iddia edilen şeyi temsil edip etmediğiyle ilgilidir.

Yanlış veya alakasız veriler, büyük veriden çıkarılacak öngörüler üzerinde önemli bir etkiye sahip olabilir ve yanlış sonuçlara ve kararlara yol açabilir.

Büyük Verinin Karakteristikleri



Bir verinin büyük veri olduğu nasıl anlaşılır?

Bilgisayar Bilimci:

Bilgisayarım bu kadar veriyi saklayacak (depolayacak) düzeyde belleğe sahip değil.

Bu büyük veridir

Hayırrrrr.

Çalışmakta olduğum veri 2TB 'ın üstünde. Bu veriyi nerede depolayabilirim?

Bu büyük veridir.

Özete, mevcut veri ile bilgisayarda işlem yapmaktan zorlanılıyorsa bu büyük veridir.

Büyük Verinin Karakteristikleri



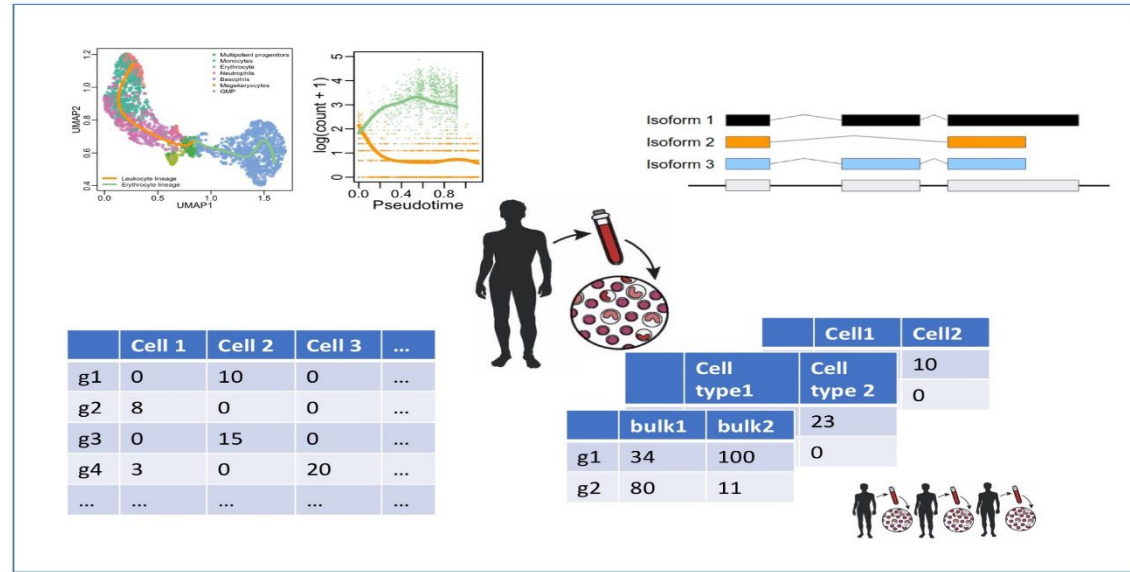
Veri bilimi uzmanı (istatistikçi)

Verinin büyük veri olduğu nasıl anlaşılır?

Soru: Bu işlem ne zaman sonlanır? Çok uzun süredir işlemin bitmesini bekliyorum.

Verinin boyutunun (dimension) çok fazla olduğu durumlarda, bu veriyi kullanarak bilgisayarda istatistiksel bir model oluşturmak mümkün olamaz.

Bu koşuldaki veri , büyük veridir.



How Much Data Is Generated Every Day in 2024?

<https://spacelift.io/blog/how-much-data-is-generated-every-day>

Nesnelerin İnterneti (IOT)

☐ İnternet üzerinden diğer cihaz ve sistemlerle veri bağlantısı ve paylaşımı yapılmasıdır.

☐ Bu amaçla sensörler ve yazılımlar kullanılır.

Böylece, IOT diğer teknolojilerle gömülü olan fiziksel nesnelerin ağını açıklar.

☐ Bu cihazlar, klasik ev nesnelerinden gelişmiş endüstriyel araçlara kadar değişiklik gösterir.

IoT nasıl yaygınlaştı?

- ❑ **Sensör Teknolojisindeki Gelişmeler.** Düşük maliyetli, düşük güçlü sensör teknolojisi ile erişim kolaylaştı
 - ❖ Uygun fiyatlı ve güvenilir sensörler, IoT teknolojisini daha fazla üretici için cazip hale getirdi.
 - ❑ **Bağlantı** İnternet için bir dizi ağ protokolü, sensörlerin buluta ve etkin veri aktarımı için diğer "nesnelere" bağlanmasını kolaylaştırdı.
 - ❑ **Bulut bilişim platformları** Bulut platformlarının kullanılabilirliği arttı.
 - ❖ Hem şirketler hem de müşterilerin herhangi bir yönetme zorunluluğu olmadan ihtiyaçları olan altyapıya erişebiliyor. Ölçeklendirme yapılarak kullandıkları kadarına sahip oluyorlar.
 - ❑ **Makine öğrenimi ve analitik** Bu alanındaki gelişmeler ve bulutta depolanan çeşitli ve çok sayıda veriye erişim sayesinde işletmeler daha hızlı ve daha kolay bir şekilde süreçlerini yönetebiliyor.
- Bu teknolojilerin ortaya çıkması IoT'nin sınırlarını zorlamaya devam etmekte ve IoT tarafından üretilen veriler de bu teknolojileri beslemektedir.
- ❑ **Yapay zeka (AI)** Örüntüleri tespit etmek veya sayıları tahmin etmek için yapay zeka modeli eğitiminde büyük miktarda veri gerektirir. Buradan bulut bilişim hizmetlerine ve veri merkezlerine olan talep sürekli artıyor.

Endüstriyel IOT / IIoT



Enterprise Resource Planning
Kurumsal Kaynak Planlaması

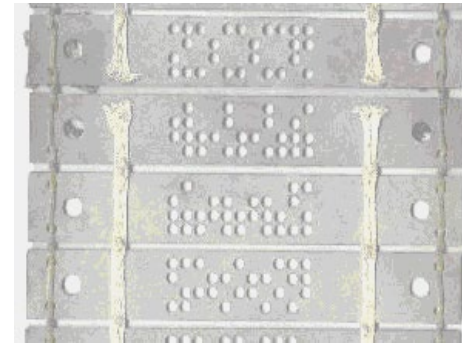
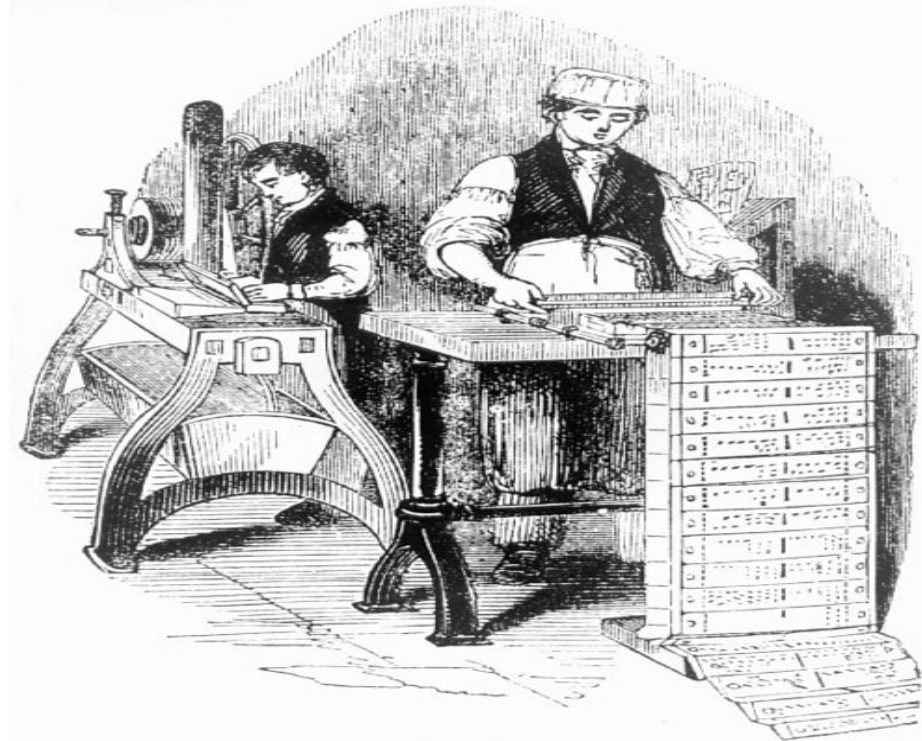
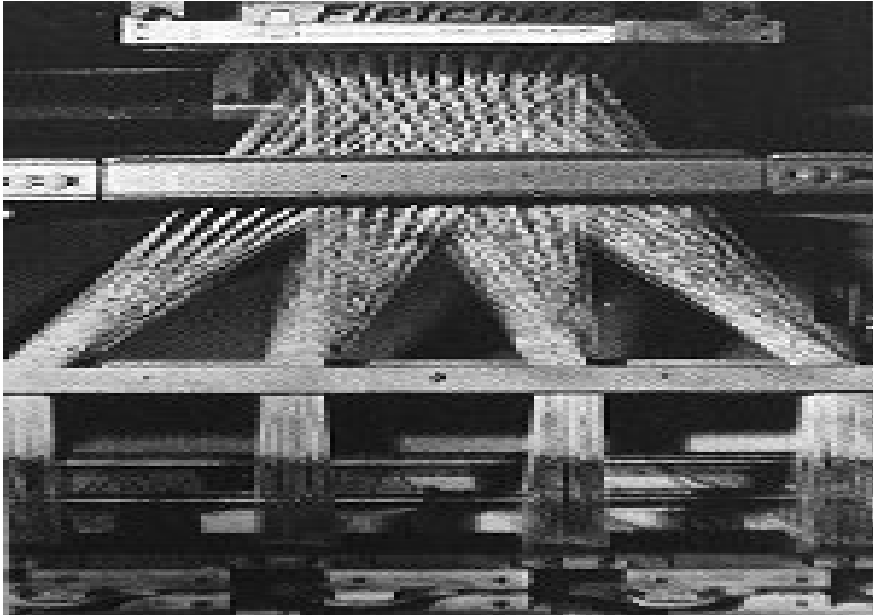
- ❑ IoT, cihazların ve makinelerin internet aracılığıyla veri paylaşmasını sağlar. Böylece , üretim süreçlerinde büyük bir değişim yaratılır. Endüstri 4.0 ise, dijitalleşme ve otomasyonu artırarak üretimi daha akıllı hale getirir. Akıllı fabrikalarda, makineler birbirleriyle konuşur, verileri analiz eder ve insan müdahalesi olmadan optimize edilmiş kararlar alır. Bu dönüşüm, üretim hatlarınının daha hızlı, esnek ve verimli çalışmasını sağlar. Akıllı fabrikalar, aynı zamanda enerji tasarrufu sağlayarak sürdürülebilir üretim yöntemlerini destekler. Bu sayede , üretim maliyetleri düşerken , çevresel etkiler de azaltılır.

BİLGİSAYARLARIN TARİHÇESİ

(devam)

Delgi Kartı – Punched Card Nedir?

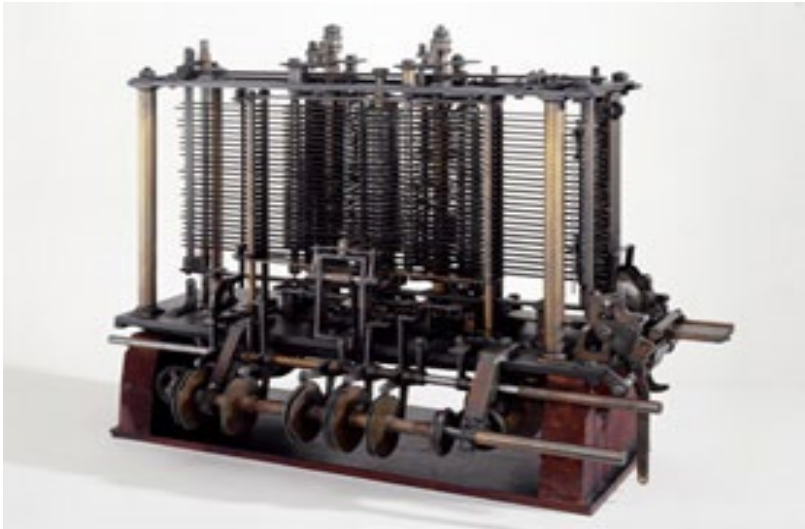
- ❑ Dokuma tezgahlarının kaşifi Jacquard (1801) giysi modelleri dokuyan tekstil tezgahlarının geliştiricisi olarak anılır.
 - ❖ Bu, aynı zamanda 20 yüzyılın ortalarında bilgisayar programlarında kullanılacak delikli kartların (punch-card) kullanıldığı ilk makine olarak da hatırlanabilir.
 - ❖ Dokumalar basılmış delikleri okuyarak ahşaptan yapılmış küçük levhalarda depolanmıştır.



Delgi Kartının veriyi saklama amaçlı kullanımı

❑ İngiliz C. Babbage, ilk mekanik bilgisayarı 1820'lerde tasarladı ama üretemedi.

❖ çizimlerine göre günümüzde yapılan otomatik hesap makinesi çalıştı.



- ❑ Babbage Analitik Makinesi 10-tabanında sabit nokta aritmetiği yapabiliyordu
- ❑ Aynı zamanda sayıları kart üzerine delebiliyordu.
 - ❖ Böylece datanın sonradan yeniden okunabiliyordu.
- ❑ Babbage desen deliklerinin *problem ifadesi* gibi *soyut bir fikir* oluşturabildiğini fark etti.
 - ❖ Bu, aynı zamanda problemin çözümünün ham veri gereksinimi (raw data requirement) idi.

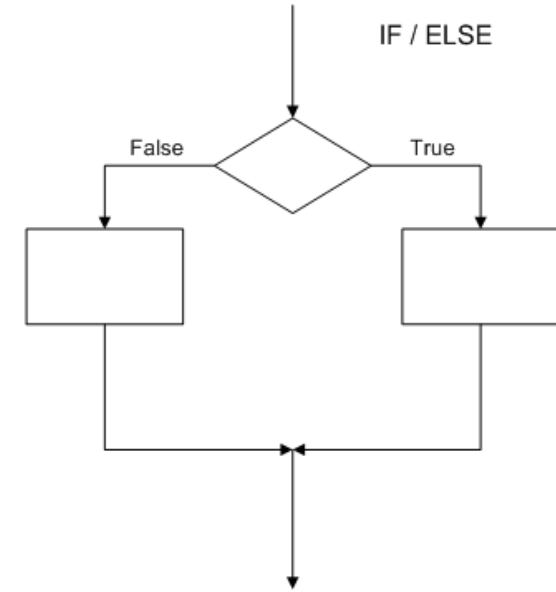
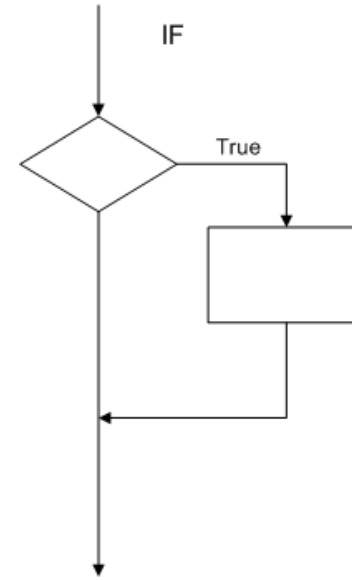
Babbage'in Görüşünün Günümüze Etkileri

❑ Modern bir bilgisayarda bu parçalar *bellek (memory unit)* ve *merkezi işleme birimi central processing unit (CPU)* olarak adlandırılır.

❑ Analitik Makinenin aynı zamanda bilgisayarları hesap makinesinden ayıran bir anahtar fonksiyonu da vardı.

❖ Koşul Deyimi (*conditional statement*).

✓ Bu deyim programın her çalıştırıldığında farklı bir sonucu gerçekleştirmesini sağlar.



Hollerith Makinesi

<https://americanhistory.si.edu/collections/object-groups/tabulating-equipment/from-herman-hollerith-to-ibm>

- ❑ 1880 'li yıllarda verilerin fazlalaşması nedeni ile elle tutulan çizelgeler yeterli olmayınca *elektro -mekanik makine* olan Hollarith makinesi kullanıldı.
- ❑ İşlenen data (yaş, cinsiyet gibi kriterlere göre düzenlendi)
 - ❖ Böylece çok daha kısa sürede tutulmaya başladı. Bu, sayma yapabilecek dişli çark mekanizmasıydı.
- ❑ Veriler delikli kağıda bilgilerin girildiği *kart okuyucu* ile gerçekleştirildi.
 - ❖ Sayma sonuçlarının tutulması sağlanmaktaydı
 - ❖ Kadran göstergesi büyük bir duvarda bulunmaktadır.
 - ❖ Araba hız ölçeri aslında böyle bir hız ölçerdir.



Jacquard ve Hollerith Delgi Kartlarının Farkları

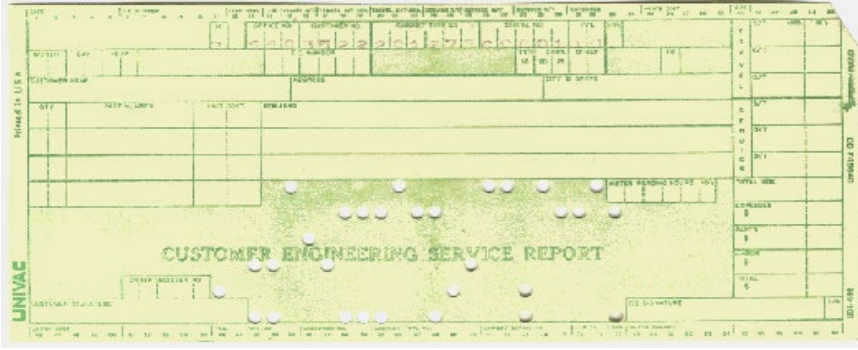
- ❑ Jacquard kartlarındaki desenler belirlendiğinde, tasarlanmış olan resim değiştirilememektedir.
 - ❖Günümüzde bu gösterim bilgi depolamada sadece okunabilir tür (*read-only form*) olarak adlandırılır.
- ❑ Hollerith ise delgi kartlarını günümüzde okuma/yazma (*read/write*) yapılabilir olarak adlandırılan yapıya dönüştürmüştür.



«Herman Hollerith's first tabulating machines opened the world's eyes to the very idea of data processing. Along the way, the machines also laid the foundation for IBM»

IBM'in Doğuşu

- ❑ 1924 yılında Herman Hollerith' in şirketi Computing-Tabulating- Recording Company ismini **International Business Machines** (IBM) olarak değiştirdi.

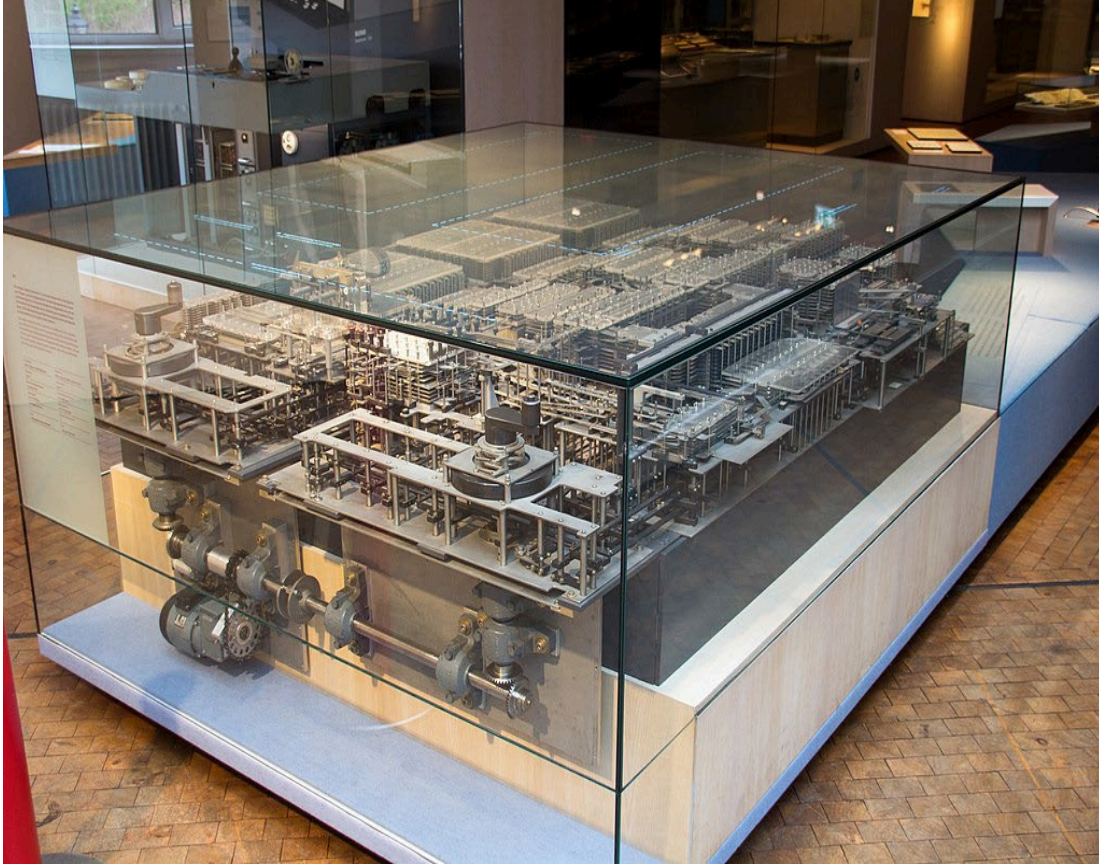


Delgi Kartı

İlk Uygulamaları

- ❑ Gaz faturaları **delgi kartları** ile her ay gelmekteydi. Ödemesi yapılıncaya kullanıcı kartı geri vermekte idi. Delgi kartları kullanıcının hesabında isim, adres, gaz kullanımı vs. şeklinde kaydedilmekteydi.
- ❑ Sürücü otoyola girdiğinde kendisine verilmiş olan delgi kartı yola nereden girdiğini ve nerede çıktığını göstermekteydi. Ödeme, gidilen mesafeye göre hesaplanmaktaydı.
- ❑ Seçimde oy kullanan kişiyi kartı bir delgi kartı idi.

İlk Mekanik Bilgisayar Z1



Berlin müzesinde sergilenmektedir.

- ❑ Alman Zuse Z1 isimli ilk mekanik ikili işlemler yapabilen programlanabilir bilgisayarı tasarladı.
- ❑ Z1 hem Boole mantığını kullanan ilk bilgisayardı, hem de ikili kayan noktalı sayılar ilk defa kullanılıyordu.
- ❑ Modern bilgisayarların basit bir mimarisine sahipti
 - ❖ Kontrol birimi, bellek, giriş /çıkış aygıtları içeriyordu.
- ❑ Makinenin mekanik olması nedeni ile güvenilirliği de yüksekti.
- ❑ Z1 makinesinin pek çok kısıtı olmasına rağmen, bilgisayarların gelişiminde önemli rol oynadı.

İlk Mekanik Bilgisayar Z1 ve Kayan Nokta Aritmetiğinin Kullanışı

$$12.345 = \underbrace{12345}_{\text{significand}} \times \underbrace{10^{-3}}_{\text{base}}^{\text{exponent}}$$

- Kayan nokta aritmetiği Z1 mekanik bilgisayar ile kullanılmaya başladı.
- Bu süreç, ilk bilgisayarın keşfi olarak kabul edilir (Alman Zuse -1936)
- Komutlar delinmiş ikili bilgiye sahip «punched typed» üzerinden okunarak mekanik gerçekleştiriliyordu.
- Program kontrollü mekanik bir bilgisayardı ve dünyanın ilk programlanabilir bilgisayarı idi
- Boolean mantığı ve ikili kayan nokta sayıları (floating-point numbers) kullanılıyordu
- Sınırlı işlem yapabilmesi nedeni ile kısıtlı olanakları vardır.

Kayan Noktalı Sayıların Günümüz Bilgisayarlarında 10 'lu Sayı Sisteminde Simgelenişi

Kayan noktalı sayıların simgelenişi tek değildir.

55.66 sayısı 5.566×10^1 olarak, ya da
 0.5566×10^2 olarak, ya da
 0.05566×10^3 olarak ya da

.....

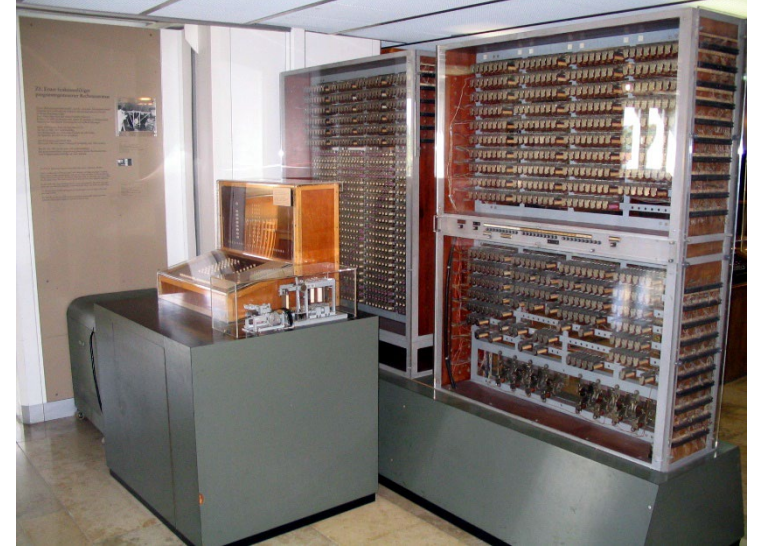
şekilde sonsuz şekilde simgelenebilir.

Bu nedenle:

Kesirli kısmın (fractional part) **normalize** edilmesi gerekir.

Normalize edildiğinde noktadan önce tek bir sıfırdan farklı dijital bulunur.

123.4567 ondalıklı sayı (decimal number) 1.234567×10^2 şeklinde
normalize edilir.



Z3 ise ilk elektro mekanik bilgisayar (Zuse)
Kayan nokta aritmetiği kullanır.

Z1 Makinesinin Modern Bilgisayarlara Etkisi

❑ Z1 bilgisayarı ile kayan noktalı sayı olarak (floating-point number) +12.654 sayısını işlemek mümkün idi .

❖ Burada tamsayı ve kesirli kısım olarak sayıları iki parçadan ifade etmek önemli bir gelişme idi .

❑ Çünkü modern bilgisayarlarda kullanılan bu sayı formatı (internal number format) ilk defa kullanılmakta idi .

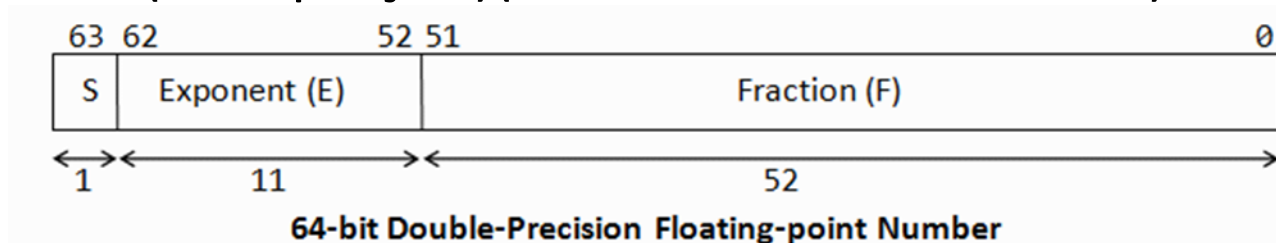
❖ Kayan noktalı sayıların bilgisayarda simgelenişi anlamlı dijitlerin kayıp vermeden simgelenişini gerektirir.

❑ Her sayının simgelenişi 3 kısımdan oluşmakta idi:

❖ Sayının işareti (sign of the number),

❖ İkiye tamamlama gösterimindeki sayının üssü (the exponent of the number in twos complement notation,

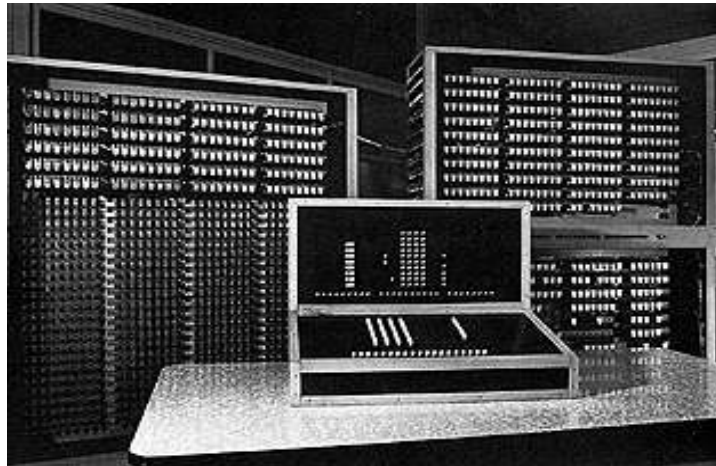
❖ Sayının mantisi (onlu parçası)(mantissa of the number).



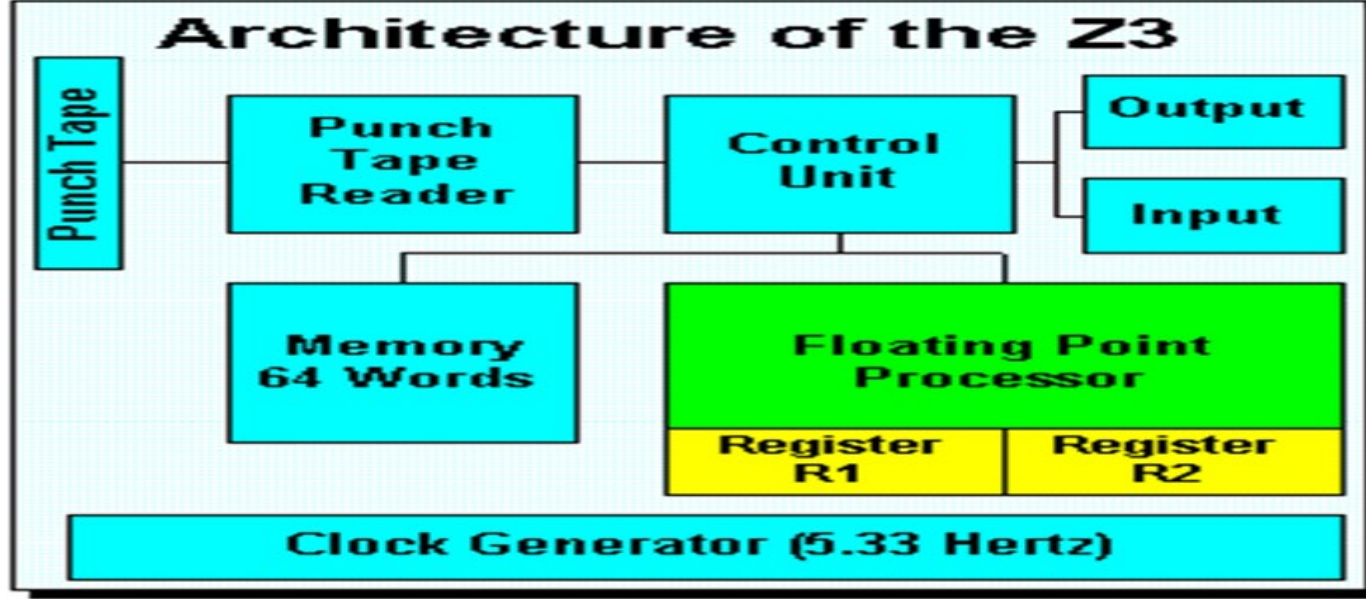
Elektromekanik- (Binary Switch)

Z3 bilgisayarı -1941

- ❑ Hollerith'in makinesi çizelge yapma (cetvel düzenleme) ile sınırlı idi.
 - ❖ Delgi kartları daha karmaşık hesaplamaları gerçekleştiremiyordu.
- ❑ Konrad Zuse (Alman) **karmaşık mühendislik problemlerini** çözebilmek için ilk programlanabilir bilgisayarları tasarladı.
- ❑ **Z3 makinesi** aynı zamanda **ikili sistemde** çalışmaktaydı ve Alman Zuse Babbage'in çalışmasından haberdar değildi.



Z1 ve Z3 Makinesinin Mimarisi



- ❑ Sayıları **depolayabiliyordu** (sayıları depolayabilen bir hafızaya sahipti, fakat depolayabileceği sayıların sayısı sınırlı idi (memory for storing numbers)
- ❑ Hesaplama için **işlemcisi** vardı (processor for computing)
- ❑ Program komutları tekrar kullanılmak, yani depolamak üzere (for storing) bir **«punched tape»** üzerine **yazılıyordu**
- ❑ Aygıtın **giriş –çıkış konsolu** vardı (input-output console)