

# CERN, BİLİM VE TÜRKİYE

## VERİ BİLİMI, BÜYÜK VERİ VE YAPAY ZEKA



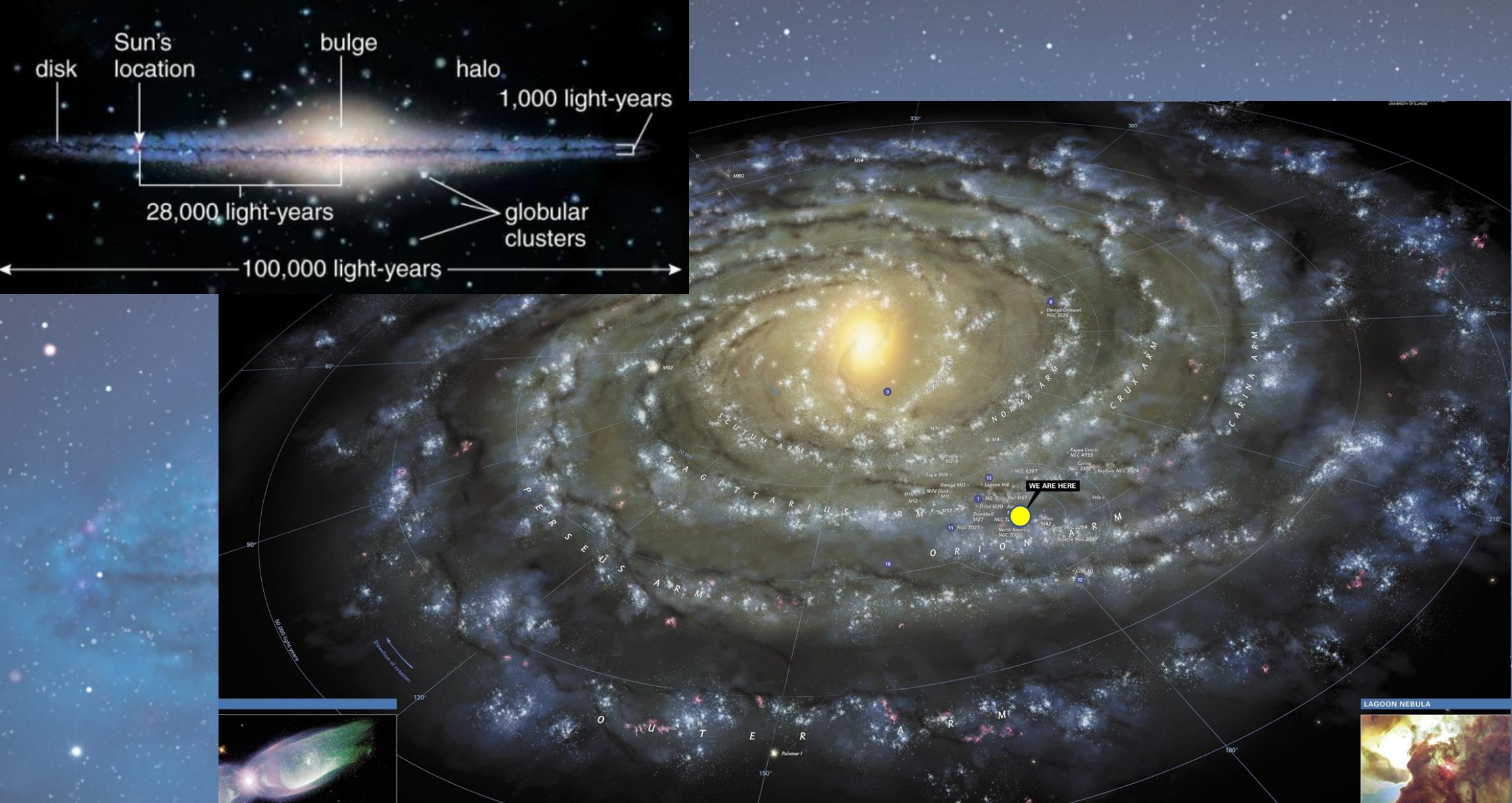


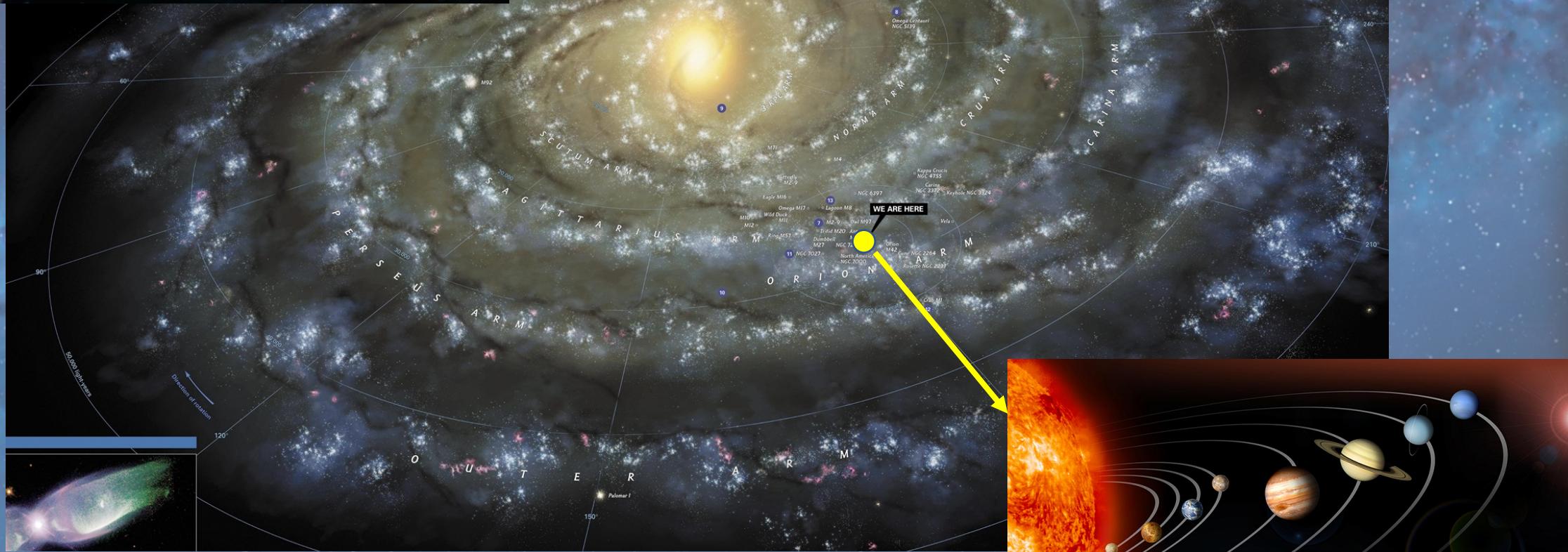
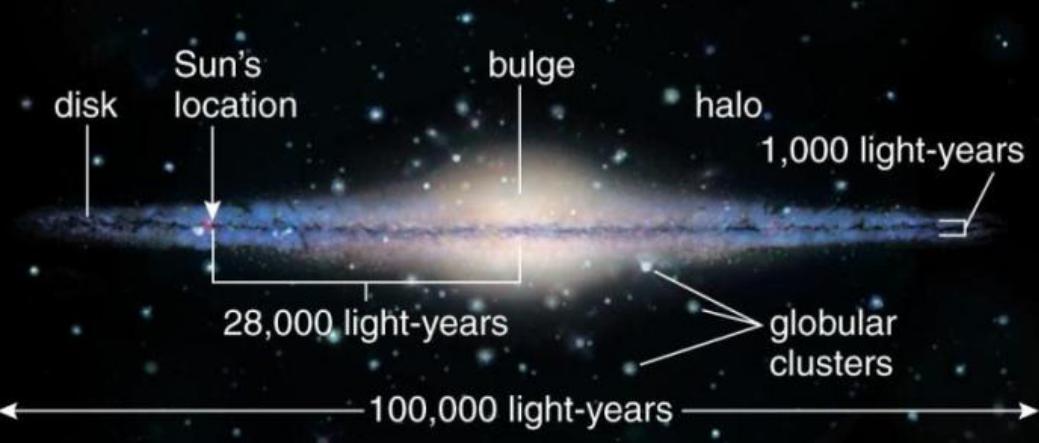
EVREN VE HERŞEY





EARTH  
SOLAR SYSTEM  
MILKY WAY GALAXY

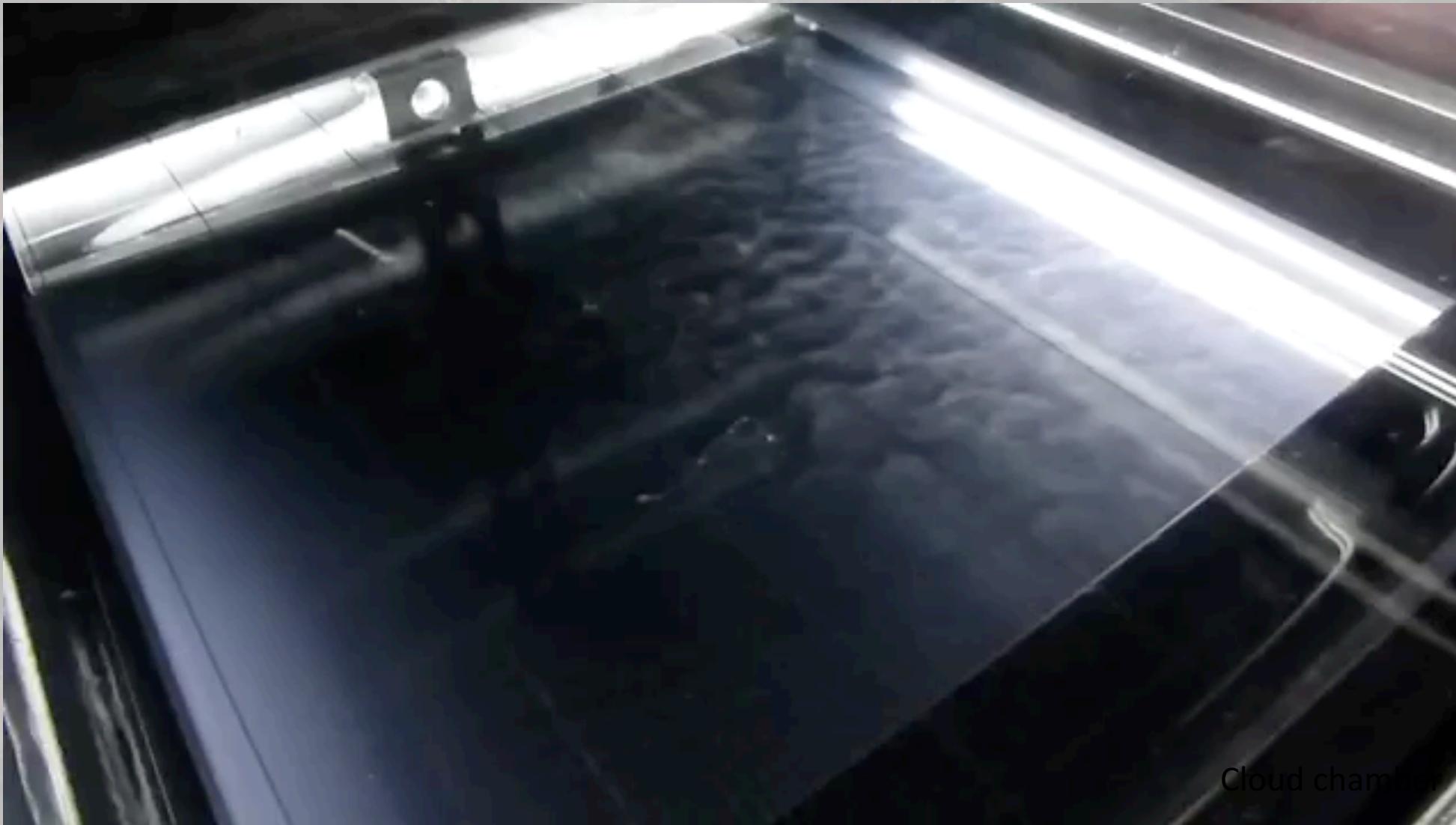




Oldukça büyük bir evrendeyiz...



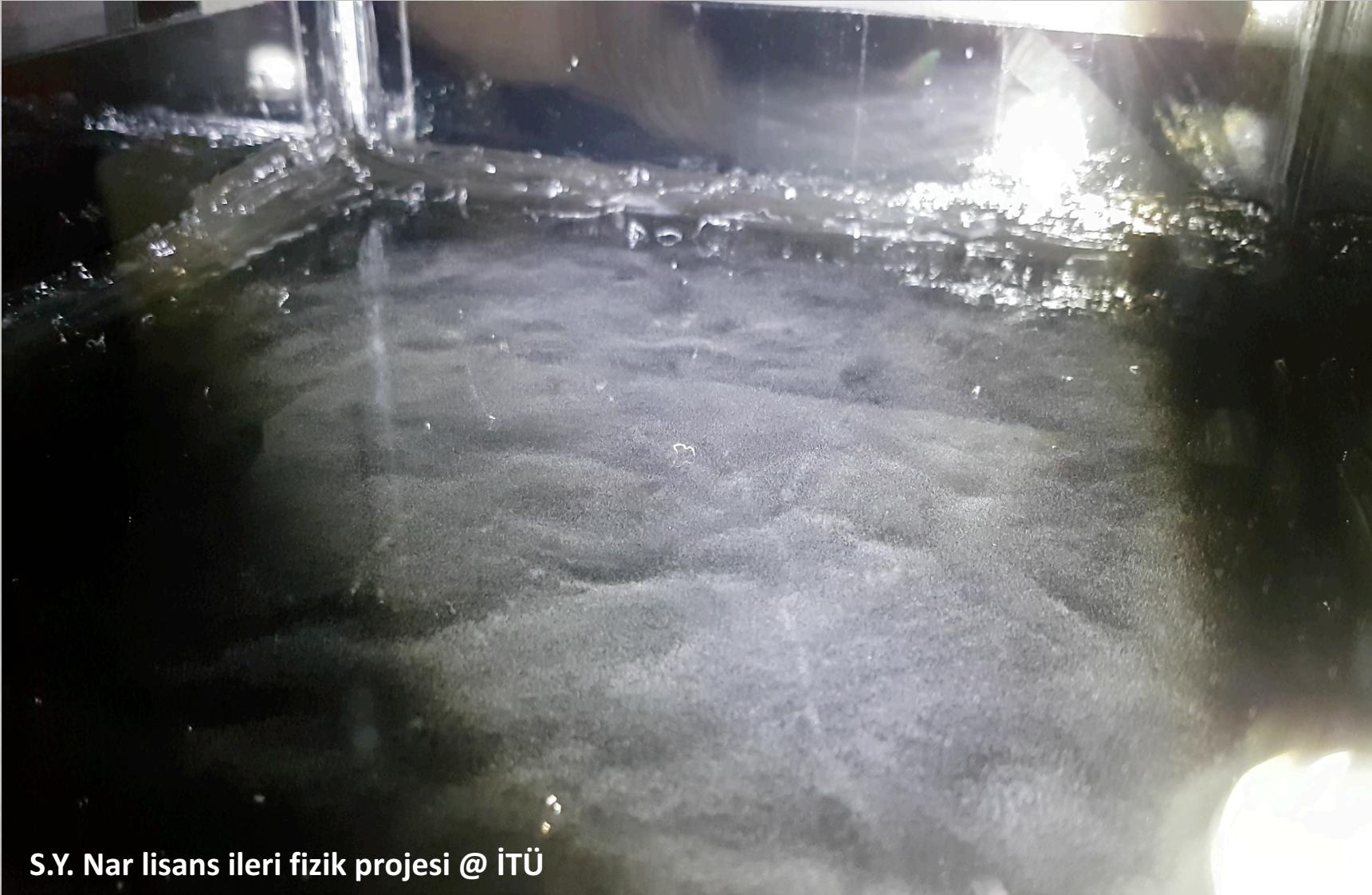
# Kozmik bilgi ve parçacıklar....



CERN bilim müzesinden bulut odası videosu!



# Sadece CERN'de mi var?



S.Y. Nar lisans ileri fizik projesi @ İTÜ

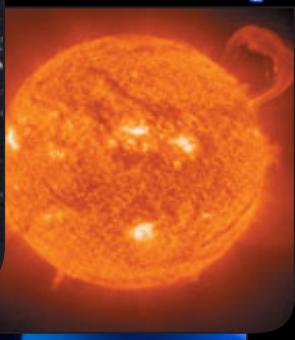
Tekrar bu görüntüyü inceleyeceğiz...

# Parçacık Fiziği, CERN ve Büyük Veri

- Parçacık fiziğinin temel motivasyonu doğa da kaybolmuş basit temel yapı taşlarını anlamaktır. Alabildiğine büyük yıldız ve galaksilerden, mikrokozmosu oluşturan en küçük temel yapı taşı olan parçacıklar günümüzün hala en önemli sorularında biridir.



(a)  $10^{26}$  m  
Limit of the observable universe



(b)  $10^{11}$  m  
Distance to the sun



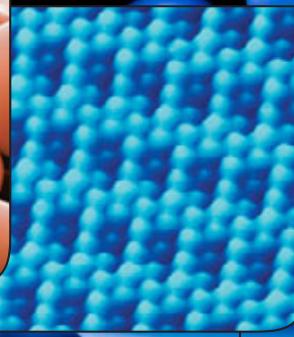
(c)  $10^7$  m  
Diameter of the earth



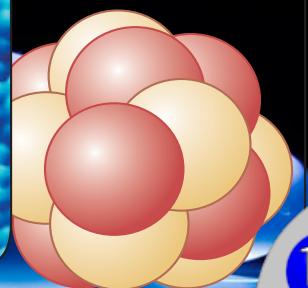
(d) 1 m  
Human dimensions



(e)  $10^{-5}$  m  
Diameter of a red blood cell



(f)  $10^{-10}$  m  
Radius of an atom



(g)  $10^{-14}$  m  
Radius of an atomic nucleus



Madde nelerden oluşmuştur? Maddeyi nasıl anlarız?



# GERÇEKTEM EVRENİ ANLAYABİLİR MİYİZ? CERN, BİLİM VE TEKNOLOJİ

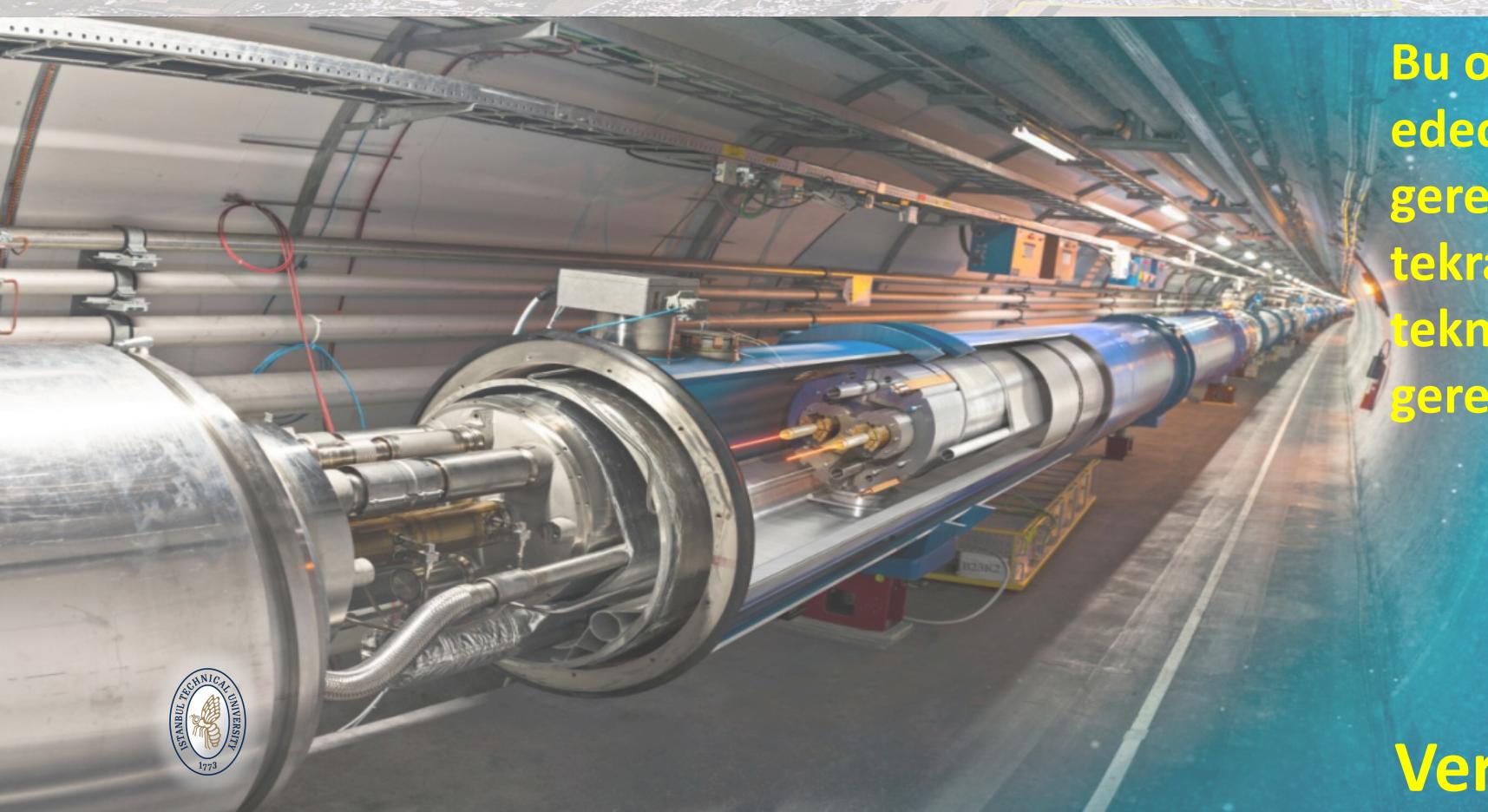
CMS

# BÜYÜK HADRON ÇARPIŞTIRICI (LHC)



CERN – AVRUPA NÜKLEER ARAŞTIRMALAR MERKEZİ

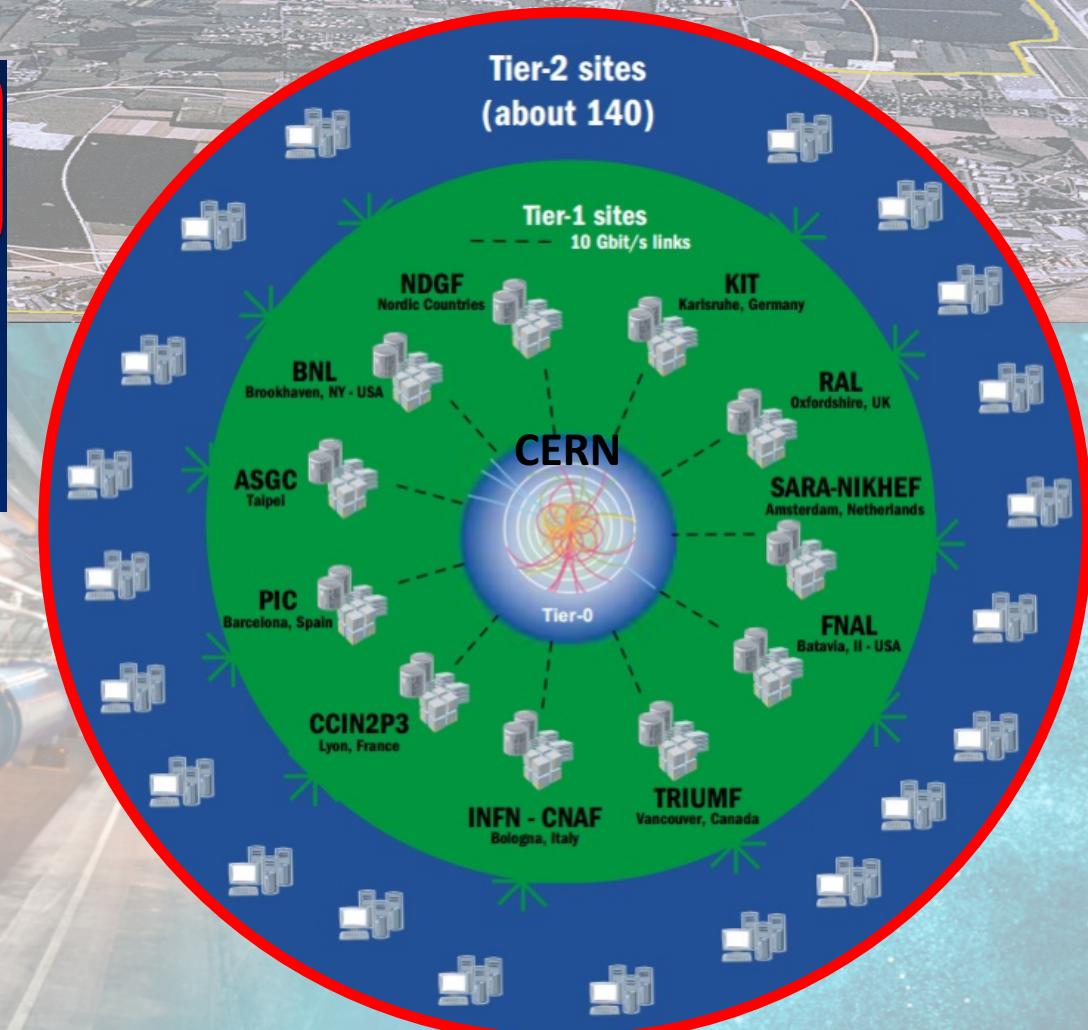
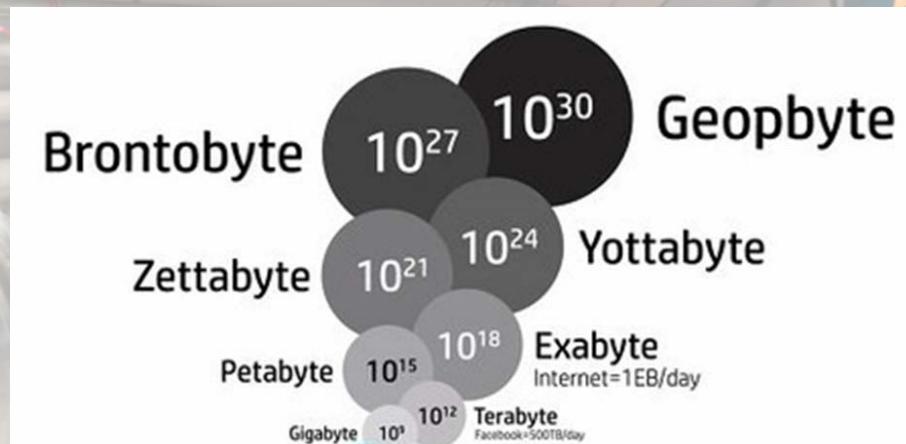
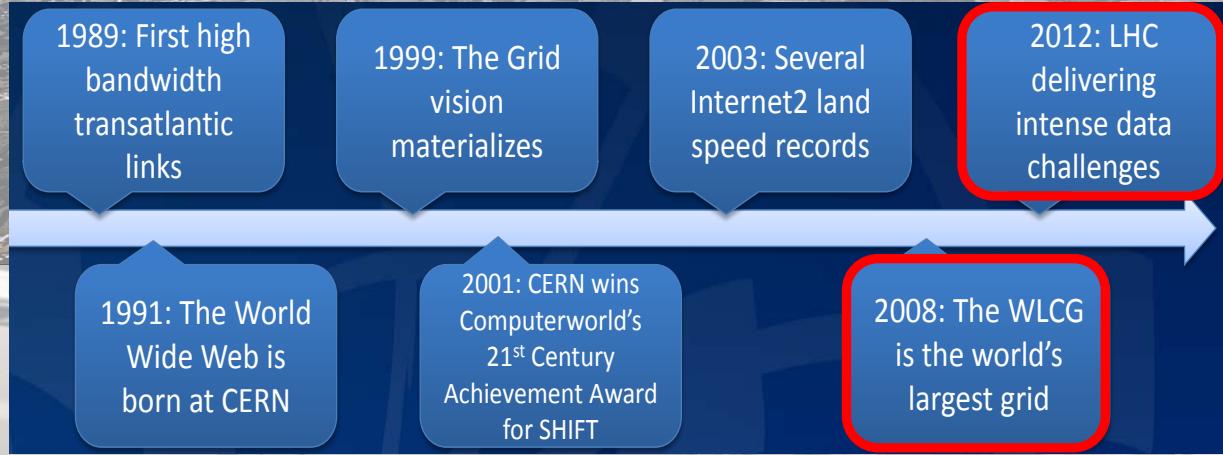
- ❖ Büyük veri yönetimi ve uygulamaları konusunda en önemli örneklerden biri CERN'de yer alan Büyük Hadron Çarşıtırıcısı deneyidir, ve yaklaşık olarak 150 milyon tane sensör üzerinden günde **500EB** veri üretmektedir!



Bu oluşan büyük veriyi nasıl kayıt edeceğimiz, nasıl saklamımız gereği ve ne saklanacağını tekrar tasarlanması gereken bir teknolojik yaklaşım gerektirmektedir.

Veriler nasıl toplanıyor?

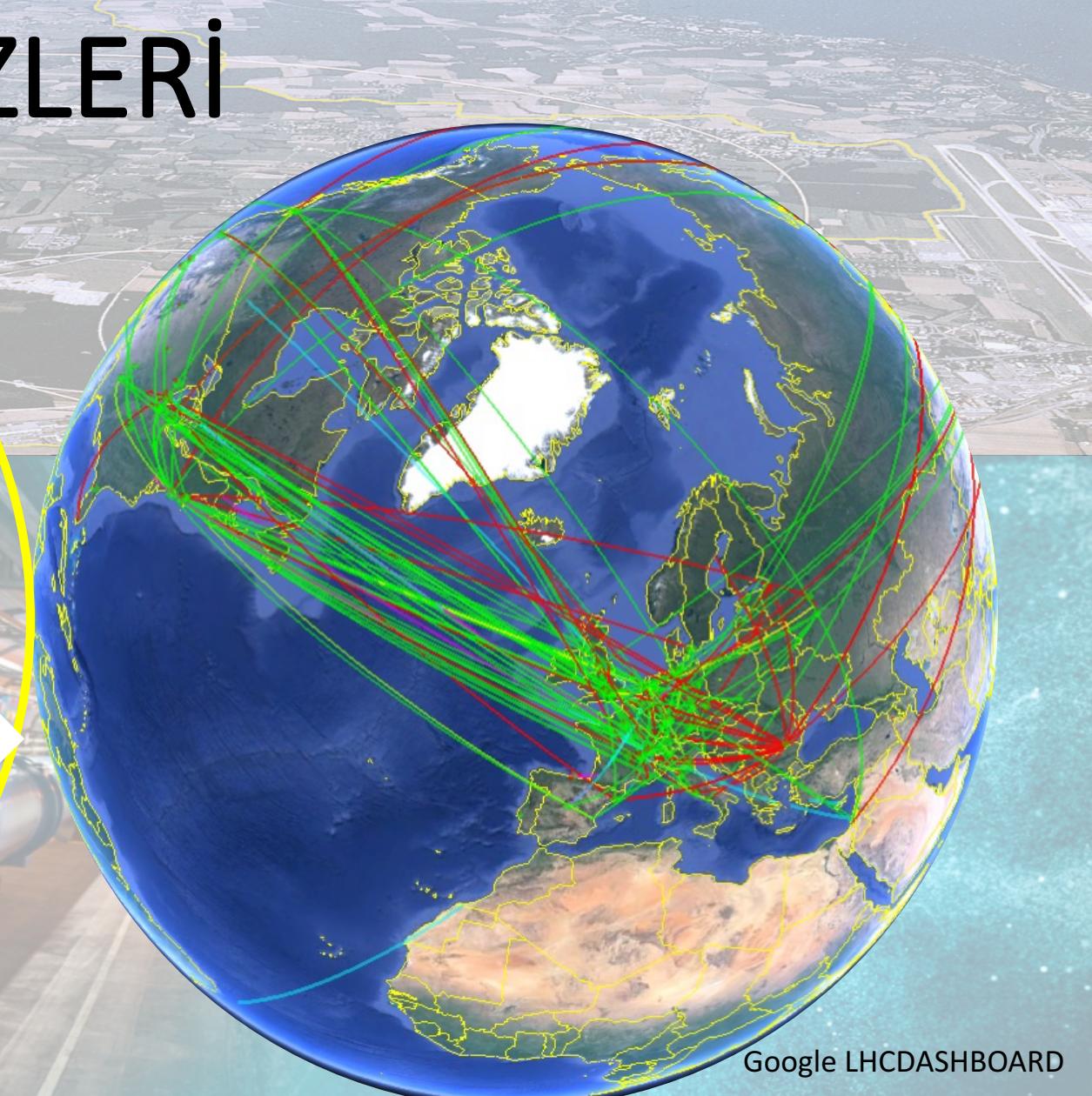
# DÜNYA GENELİ LHC VERİ HESAPLAMA AĞI(WLHC Grid)



GRID: KAYNAK PAYLAŞIMI & KOORDİNELİ DİNAMİK PROBLEM ÇÖZÜMLERİ,  
ÇOK ENSTİTÜLÜ ORGANİZASYON YAPISI

# LHC'de VERİ ANALİZLERİ

CMS



Google LHCDASHBOARD

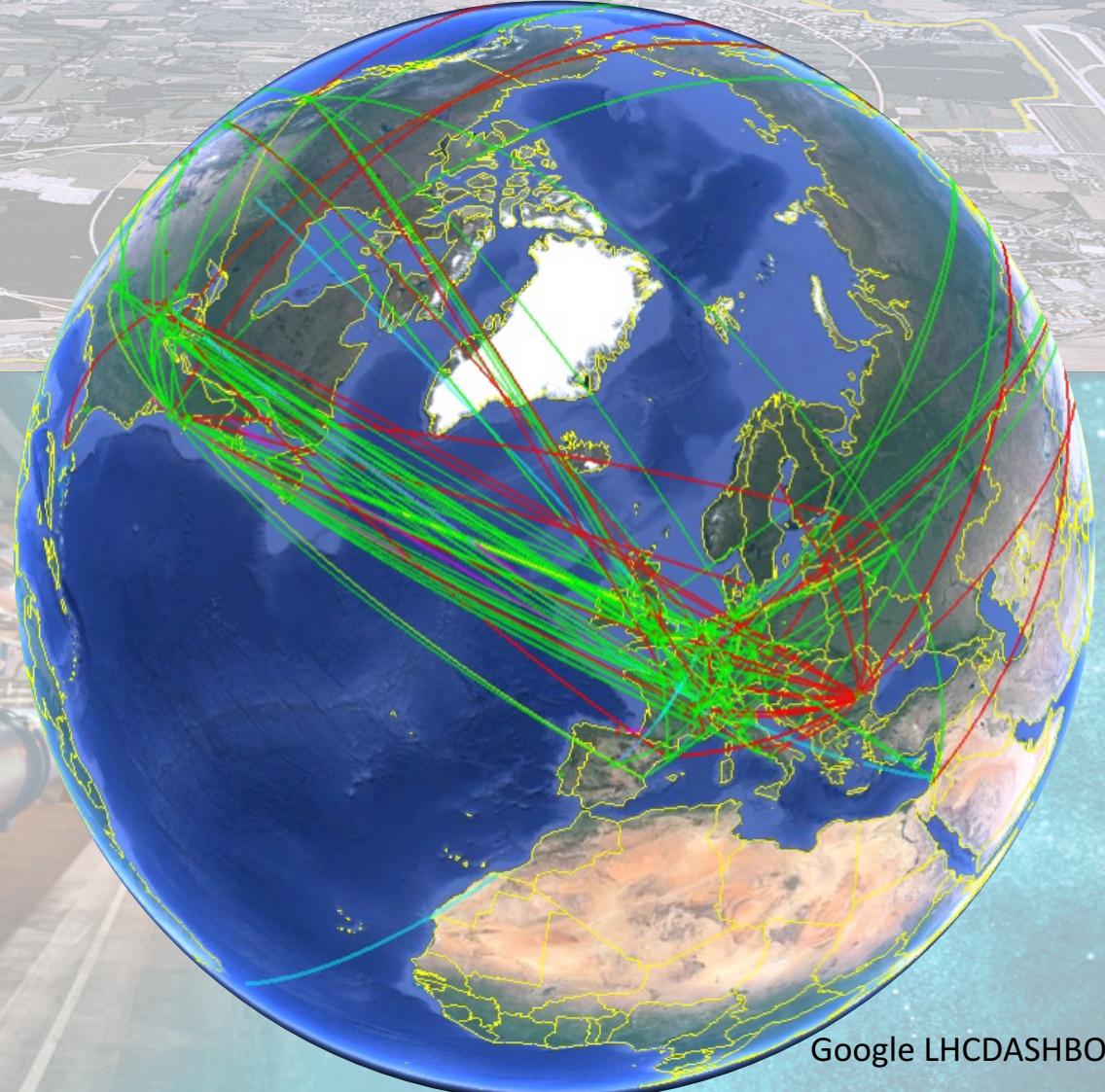


Belirsizlik, Değişkenlik, Görsellik ve Değer Ülkeler arası veri dolaşımı, WLHCG sistemi

# Büyük Veri, IoT, Endüstri 4.0, Yapay Zeka, Dijital devrim,...

Genel anlamda IoT sistemleri, fiziksel cihazların ağı, taşıtlar, binalar ve diğer sistemlerin elektronik, yazılım, sensörler ve ağ bağlantılarına entegre edilerek toplanan veri ve/veya veri paylaşımı tanımlamaktadır

CERN'deki temel IoT cihazları nelerdir? Yapay Zeka, Dijital dönüşüm, Endüstri 4.0, vb. konular...



Google LHC DASHBOARD

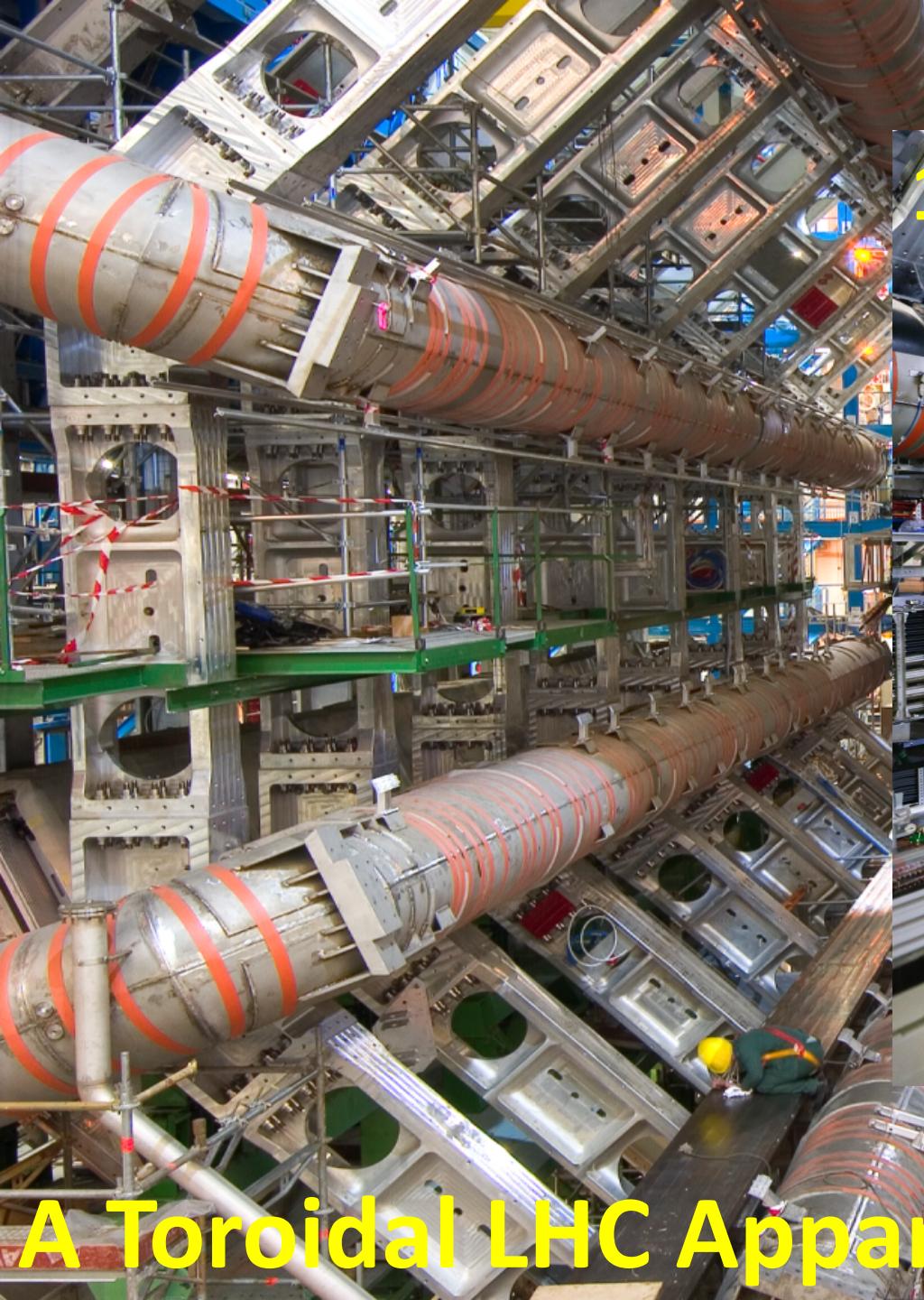


**Evrenden bizlere ulaşan ekzotik kozmik  
ışınları hatırlayalım,**

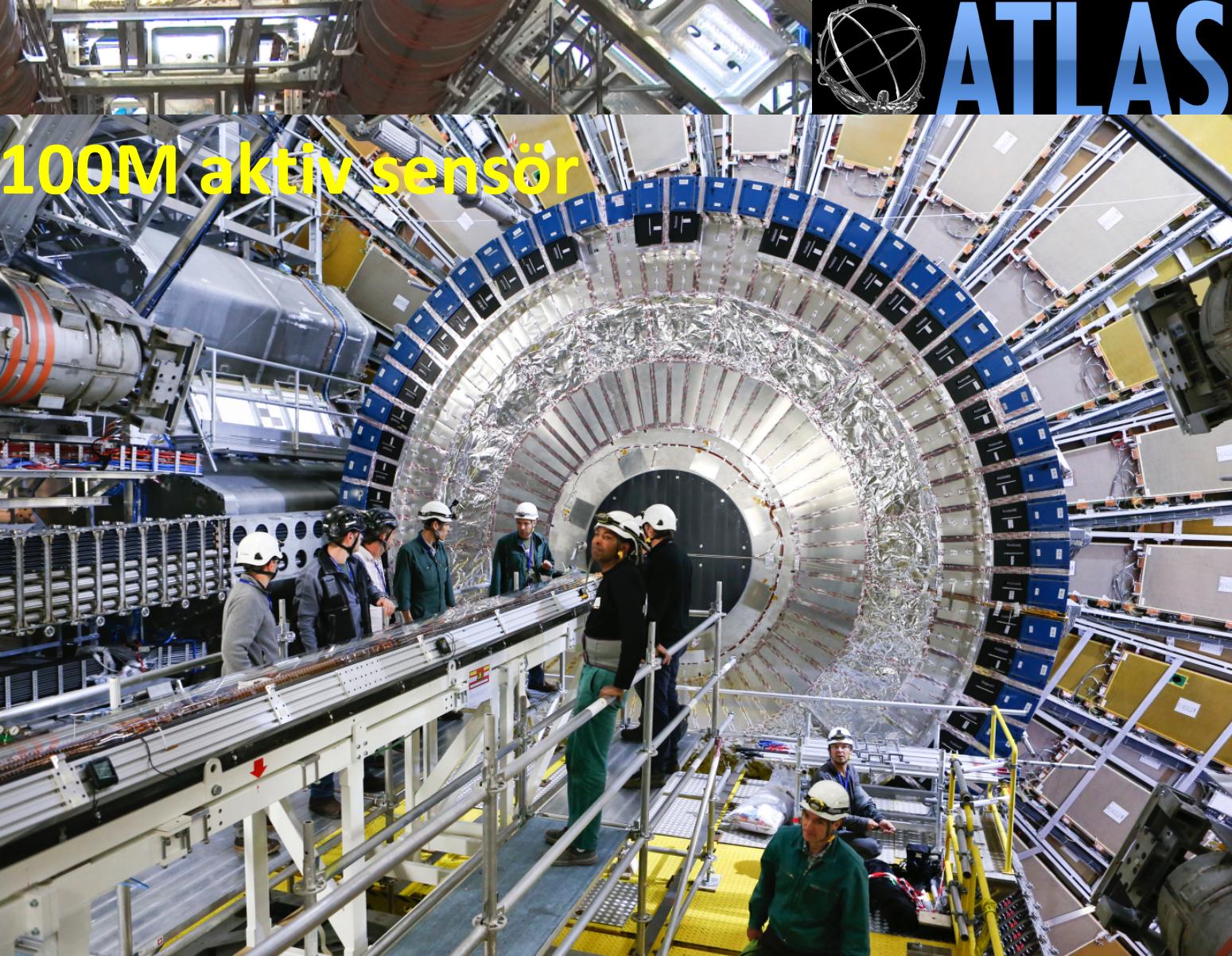
- Parçacıklar için kullanılan algıçlar (IoT)
- Doğa'yı anlamak (Büyük Veri Analizi)

# A Toroidal LHC Apparatus

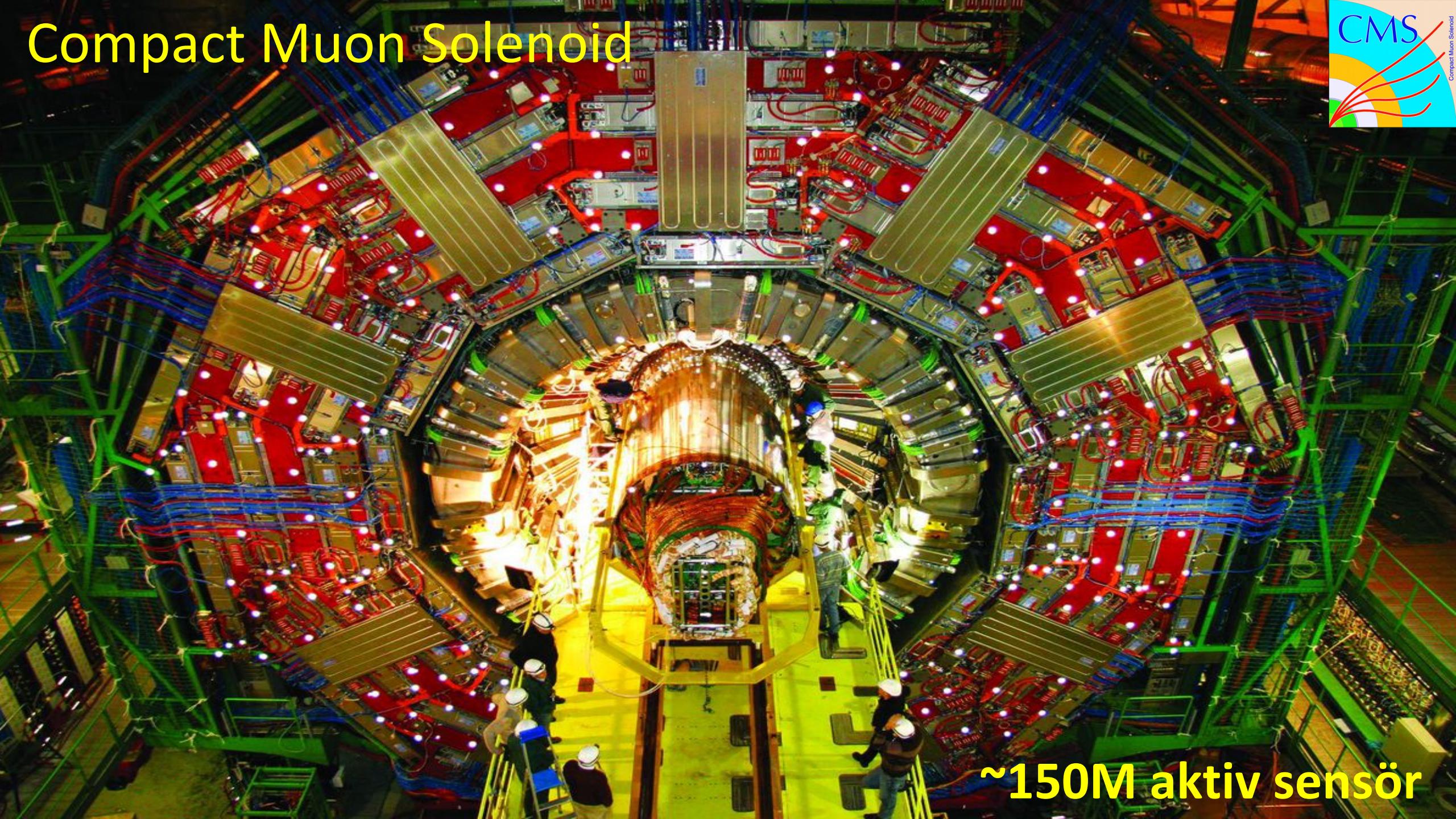


 ATLAS

A Toroidal LHC Apparatus

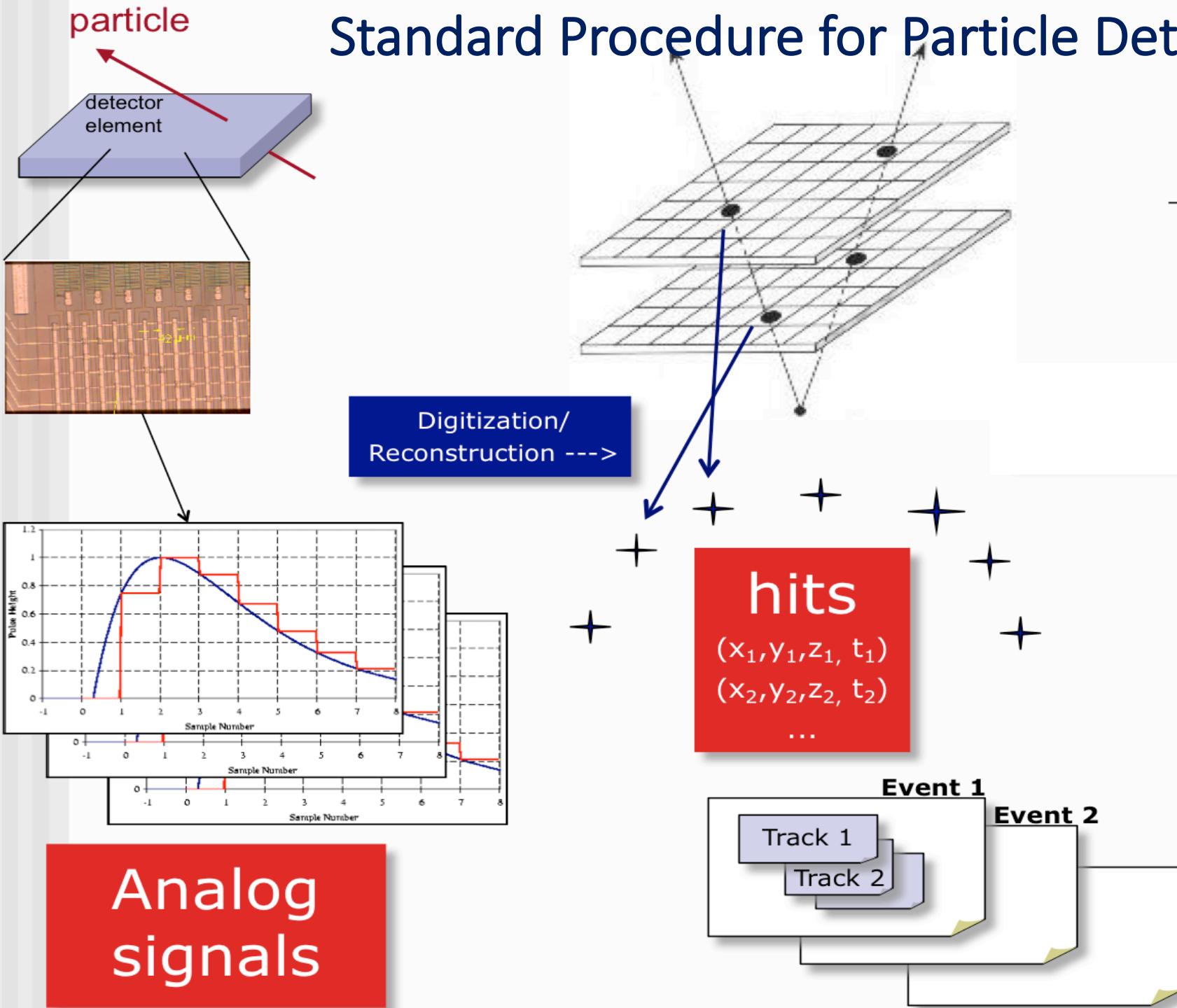


# Compact Muon Solenoid

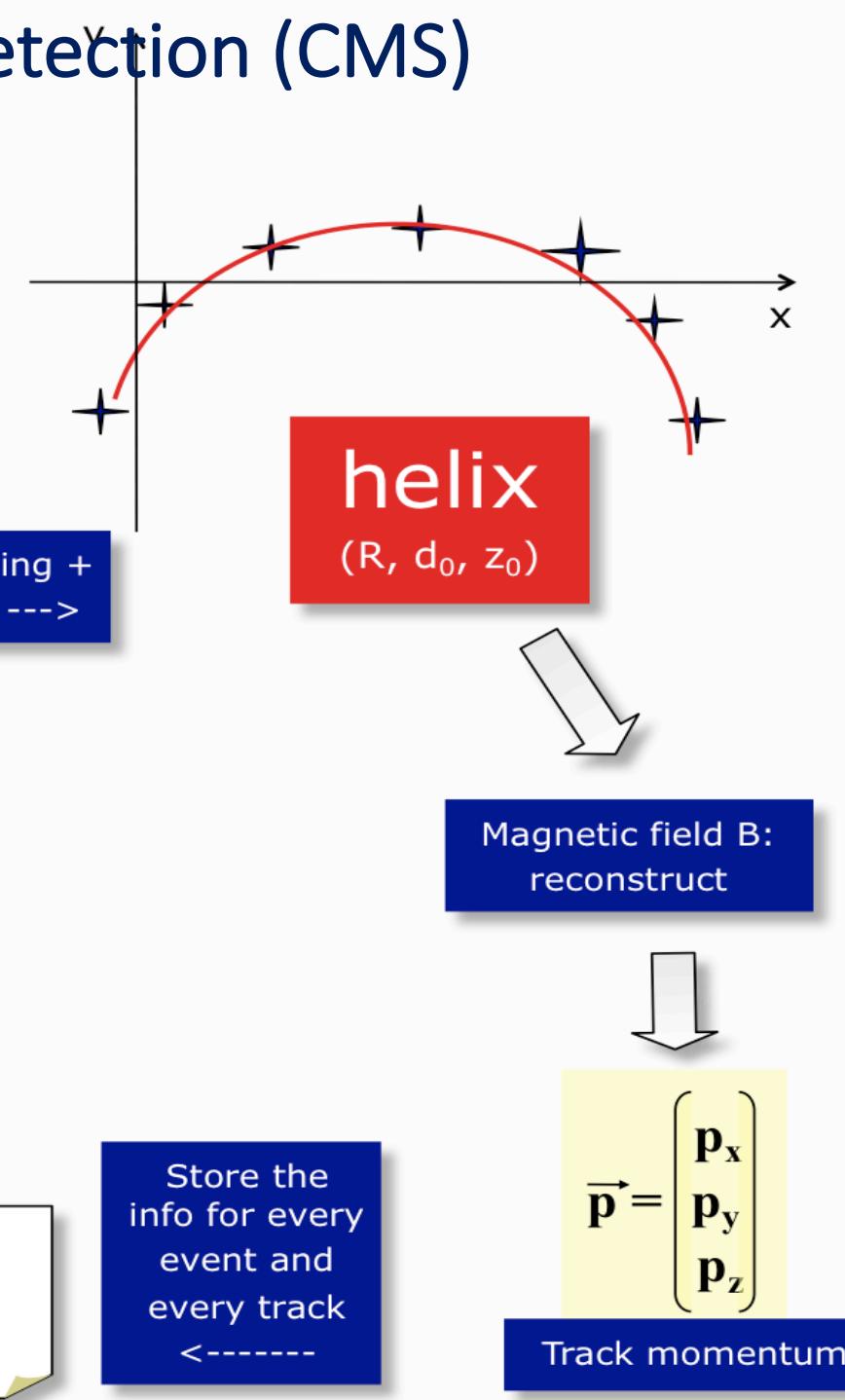
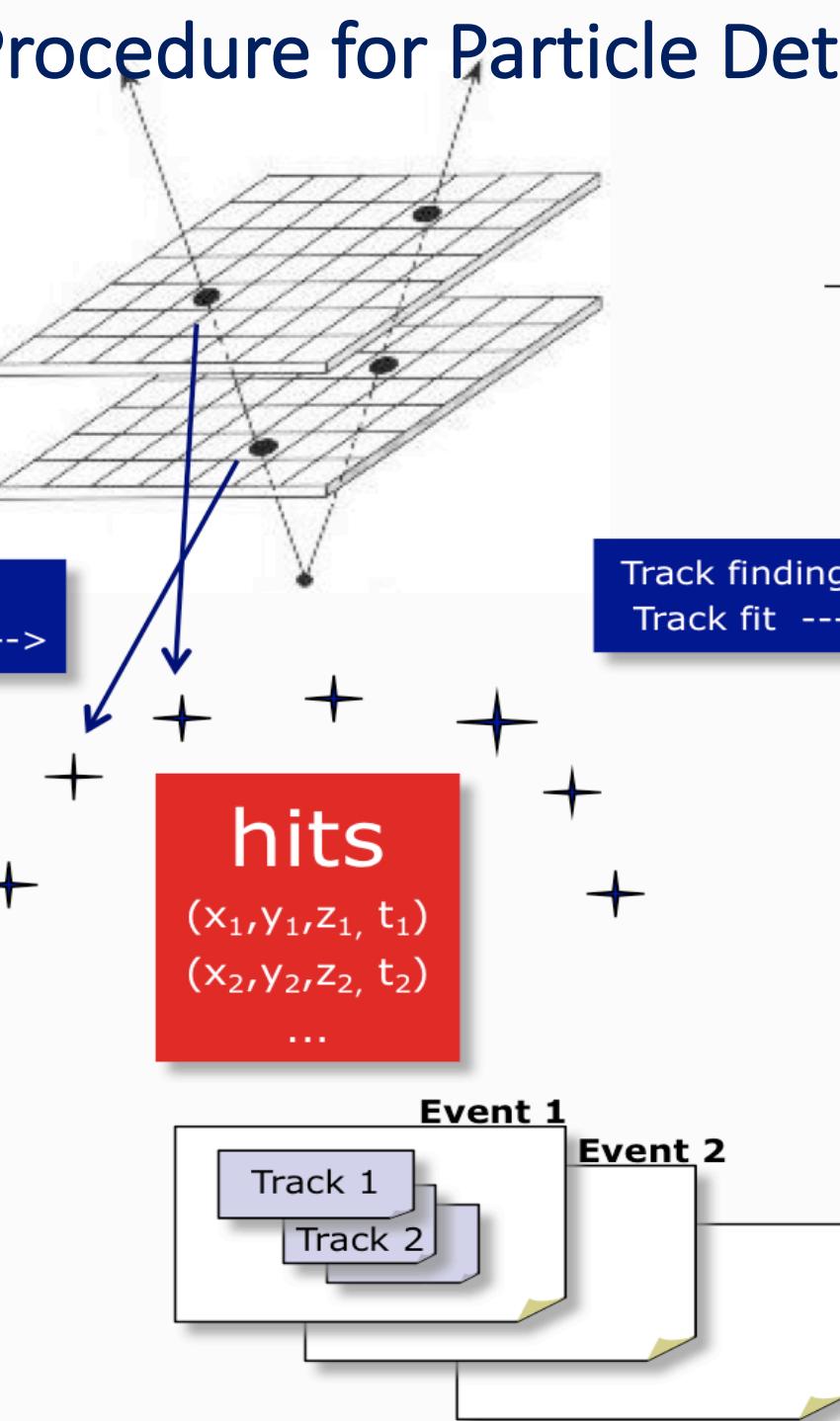
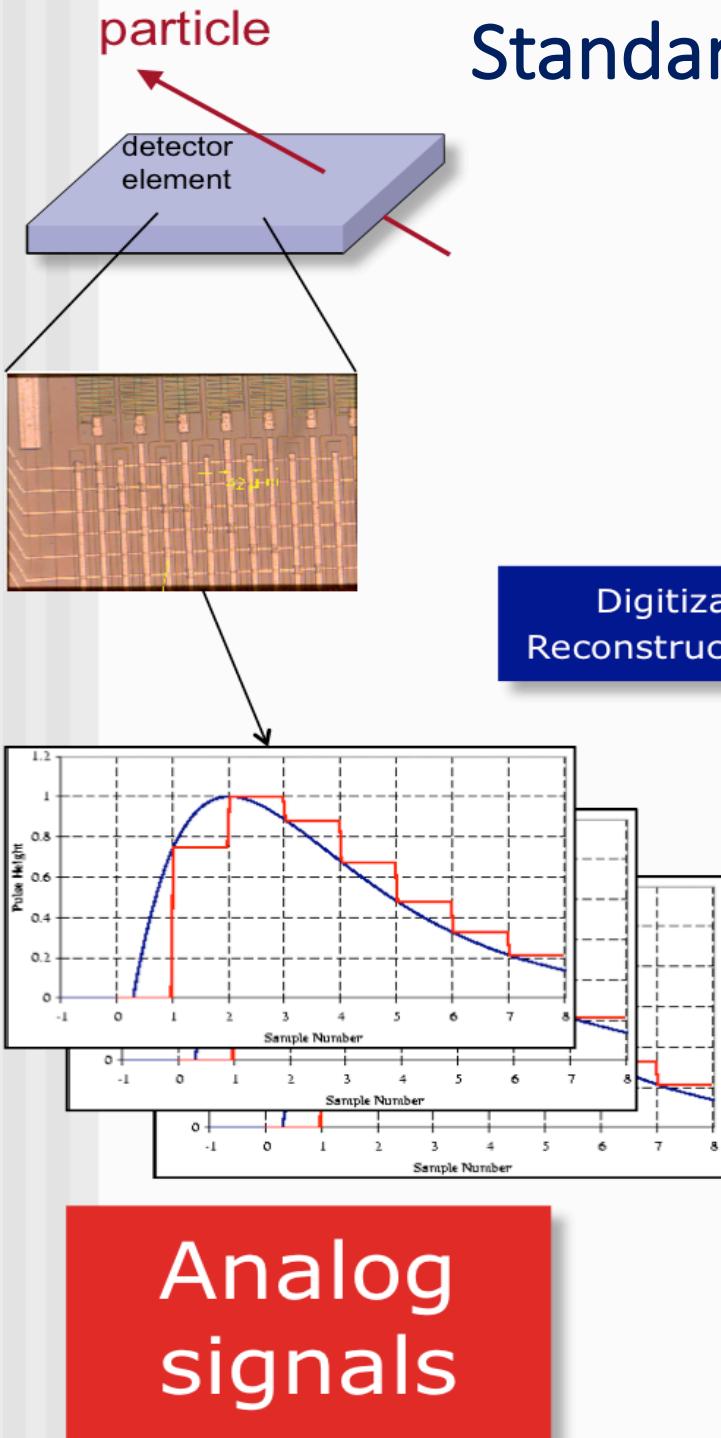


~150M aktiv sensör

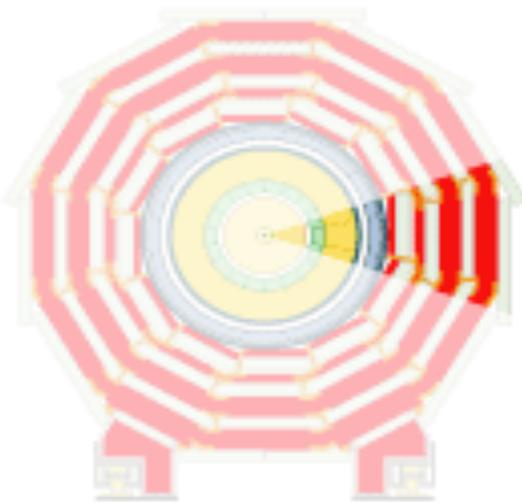
# Standard Procedure for Particle Detection (CMS)



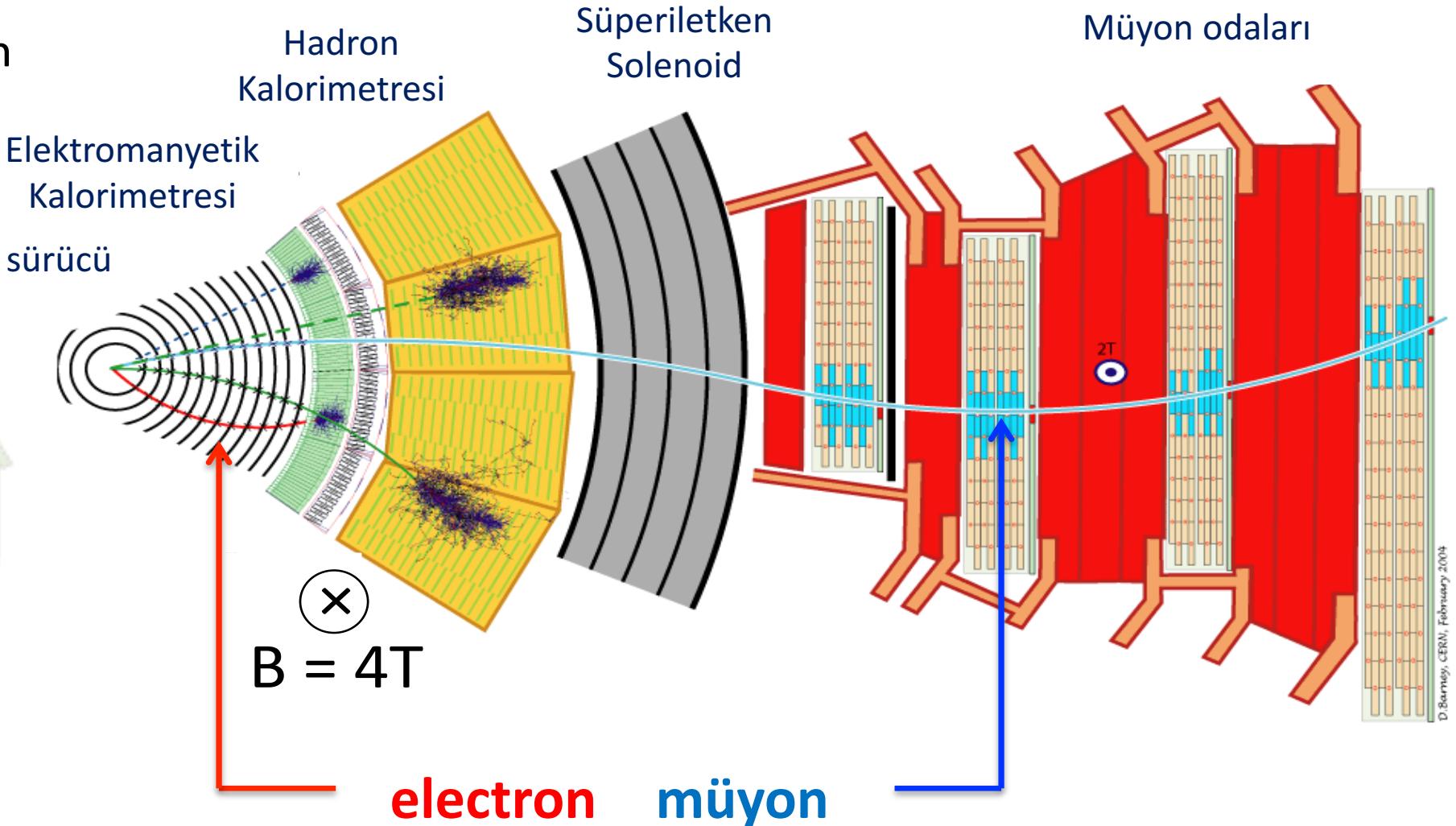
# Standard Procedure for Particle Detection (CMS)



- muon
- electron
- charged hadron
- neutral hadron
- photon



Silikon iz sürücü

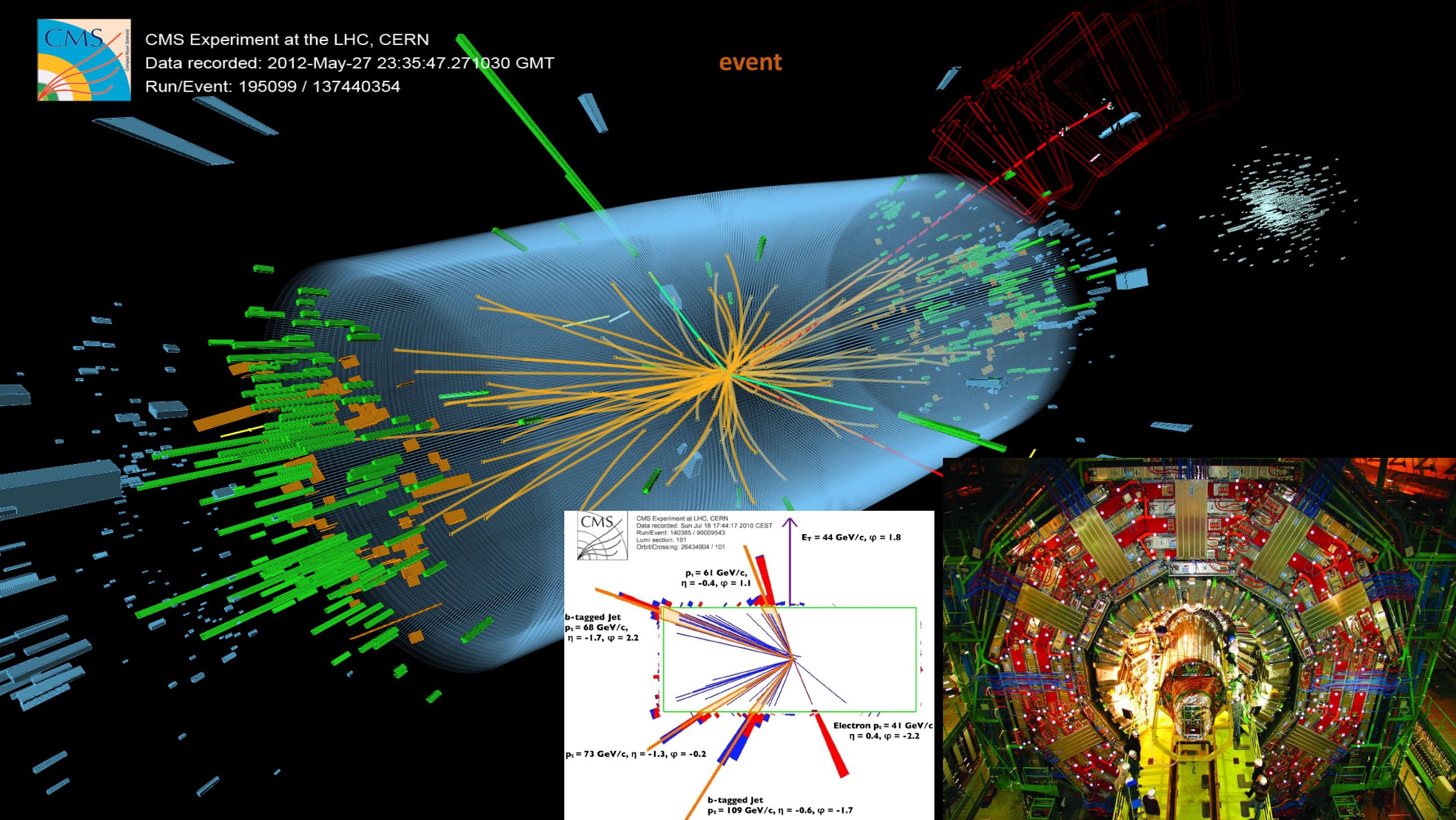


D.Burney, CERN, February 2004



CMS Experiment at the LHC, CERN  
Data recorded: 2012-May-27 23:35:47.271030 GMT  
Run/Event: 195099 / 137440354

event



Delay  
3.0sRun 122314 Event 15145452  
Event Filtering is OFFMon Nov 23 19:20:55 2009 CEST  
Lumi block id: 25

FIREWORKS

## Summary View

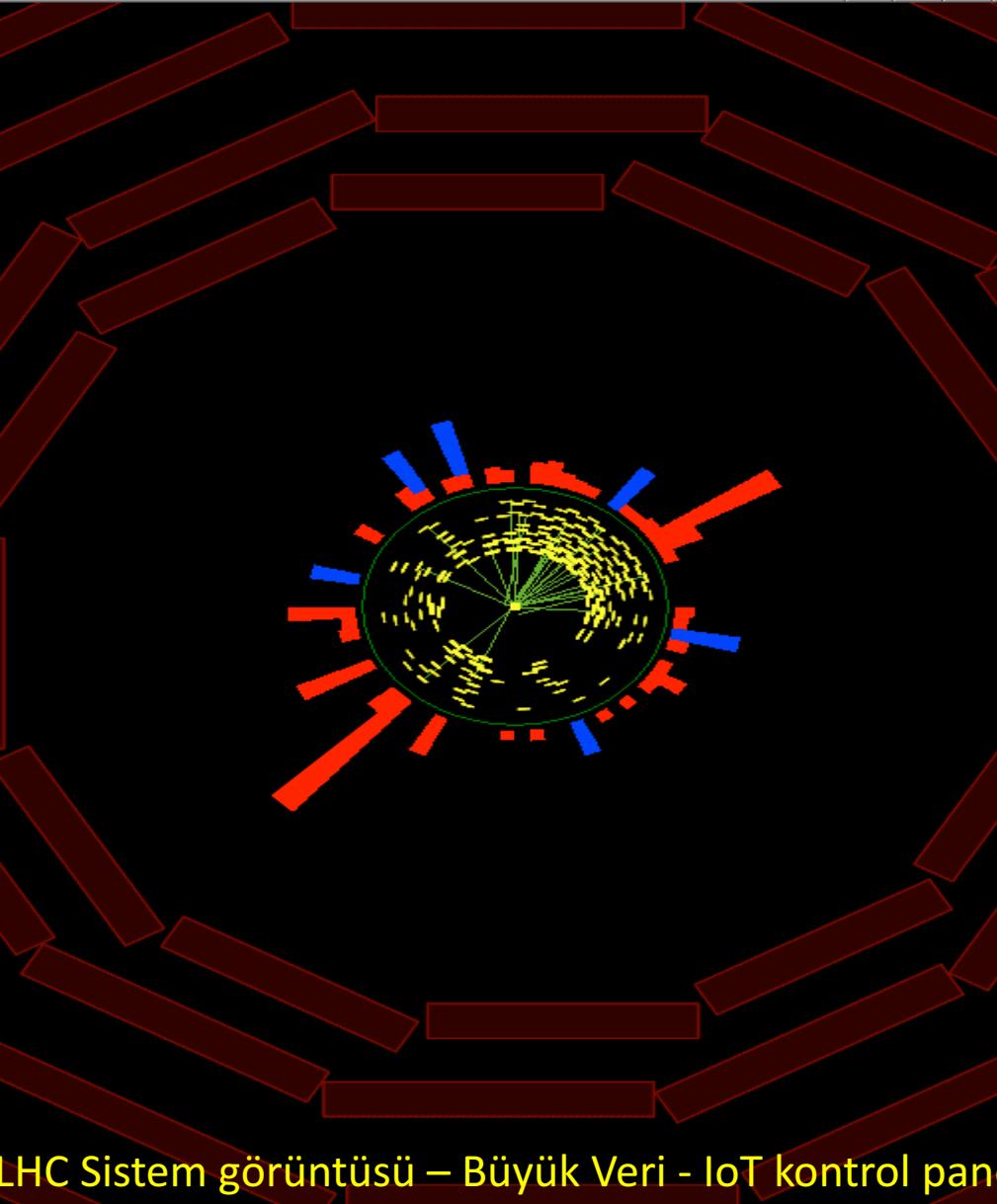
- Add Collection
- ECal
  - HCal
  - Jets
  - Tracks

|          | pt  | eta  | phi  |
|----------|-----|------|------|
| Track 0  | 4.9 | -0.2 | 0.3  |
| Track 1  | 5.0 | -0.1 | 0.2  |
| Track 2  | 3.7 | -0.8 | 0.3  |
| Track 3  | 4.0 | -0.7 | 0.3  |
| Track 4  | 4.6 | -0.4 | 0.6  |
| Track 5  | 4.8 | -0.3 | 0.6  |
| Track 6  | 4.9 | -0.2 | 1.0  |
| Track 7  | 5.0 | -0.1 | 1.1  |
| Track 8  | 4.4 | -0.5 | 1.1  |
| Track 9  | 3.0 | -1.1 | 1.1  |
| Track 10 | 3.0 | -1.1 | 1.0  |
| Track 11 | 5.0 | -0.1 | 1.2  |
| Track 12 | 4.1 | -0.7 | 1.5  |
| Track 13 | 3.6 | -0.9 | 2.6  |
| Track 14 | 4.9 | -0.2 | -2.3 |
| Track 15 | 3.5 | 0.9  | 0.4  |
| Track 16 | 3.7 | 0.8  | 0.7  |
| Track 17 | 5.0 | 0.1  | 0.8  |
| Track 18 | 3.6 | 0.8  | 0.9  |
| Track 19 | 4.3 | 0.6  | 1.4  |
| Track 20 | 4.6 | 0.4  | 1.6  |
| Track 21 | 3.0 | 1.1  | 1.9  |
| Track 22 | 4.6 | 0.4  | 2.3  |
| Track 23 | 3.6 | 0.9  | -2.0 |
| Track 24 | 2.8 | 1.2  | -0.1 |

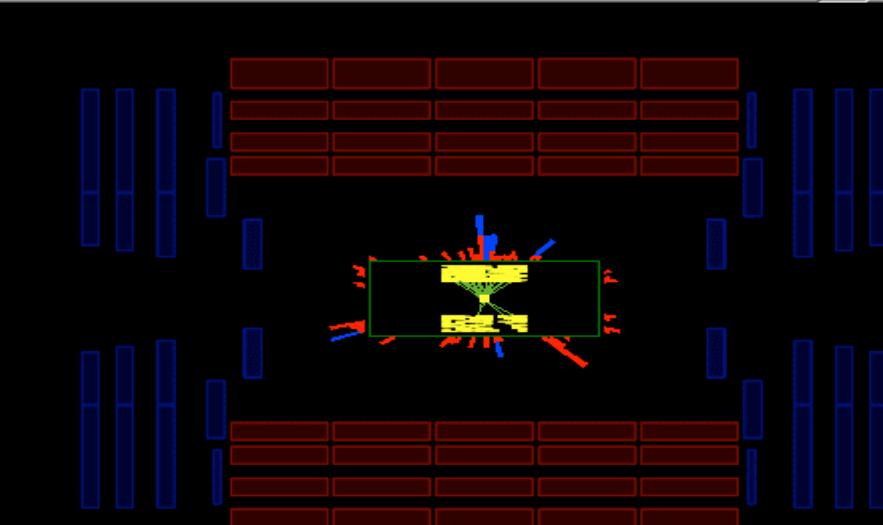
- Muons
- Electrons
- Vertices
- DT-segments
- CSC-segments
- Photons
- MET
- siStripClusters

## Views

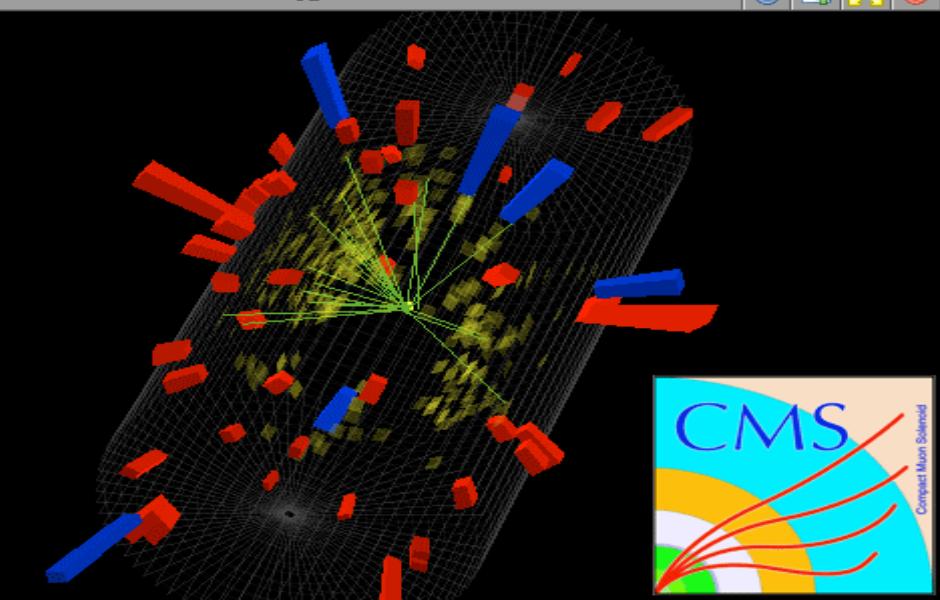
## Rho Phi



## Rho Z

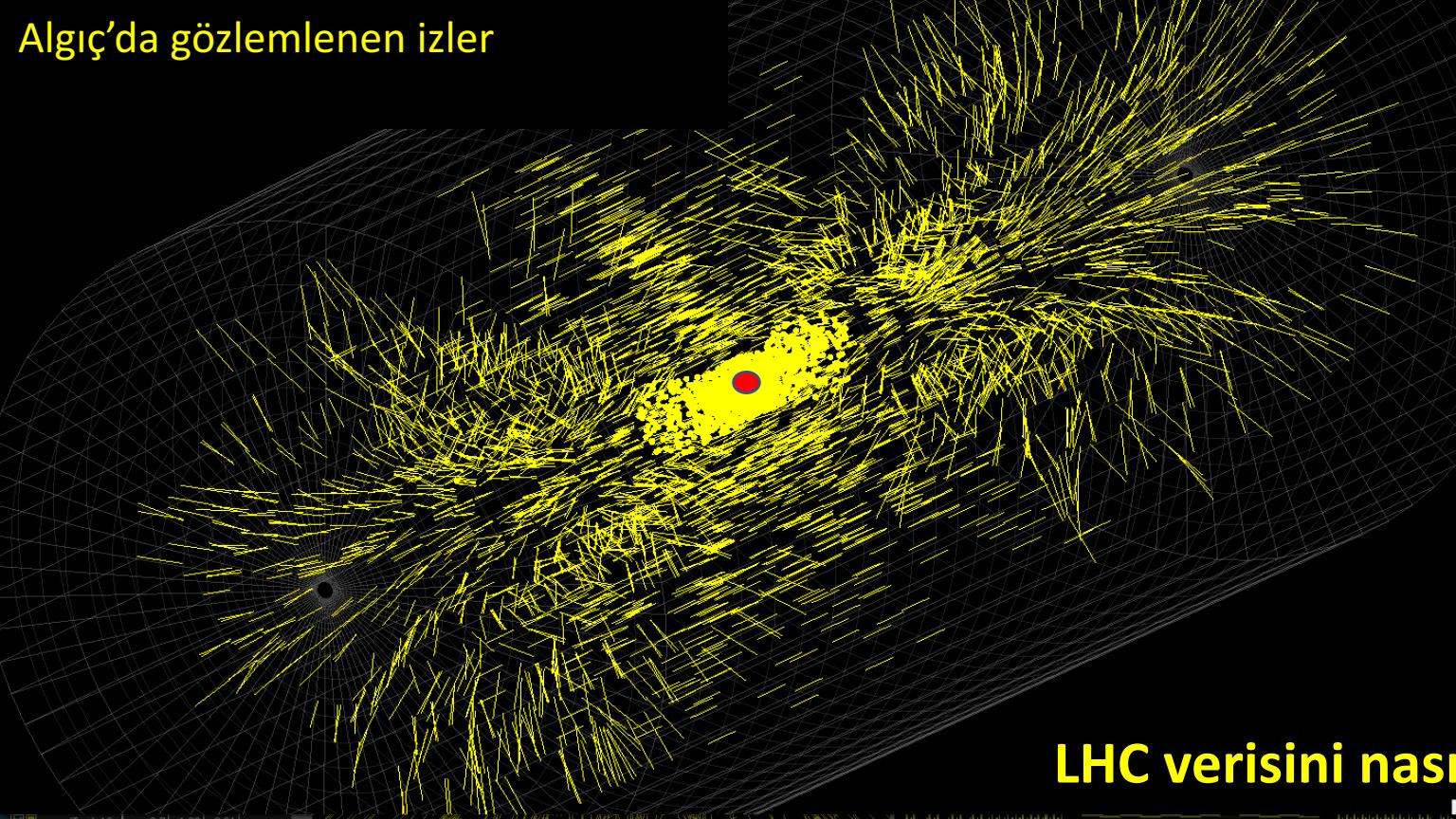


## 3D

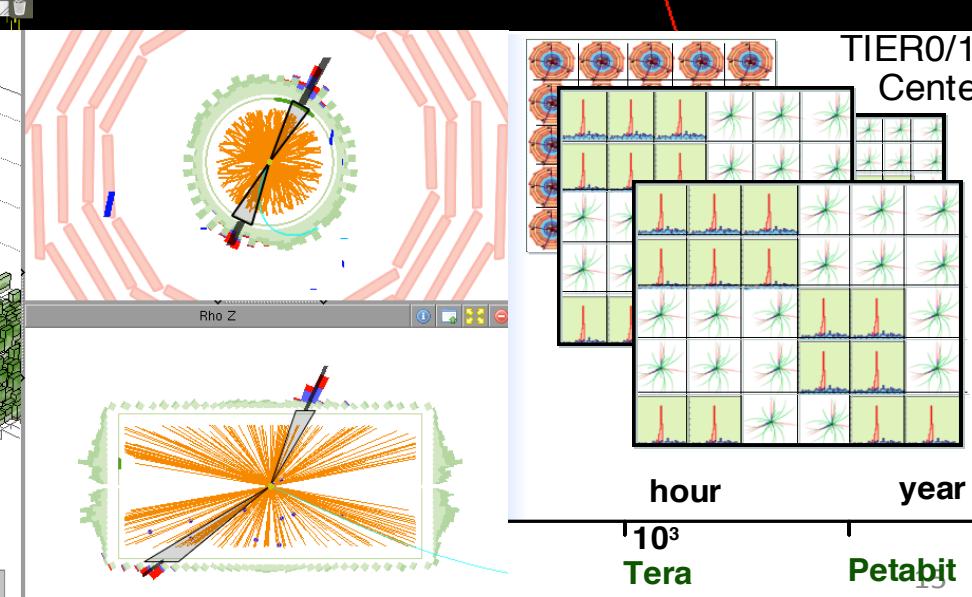
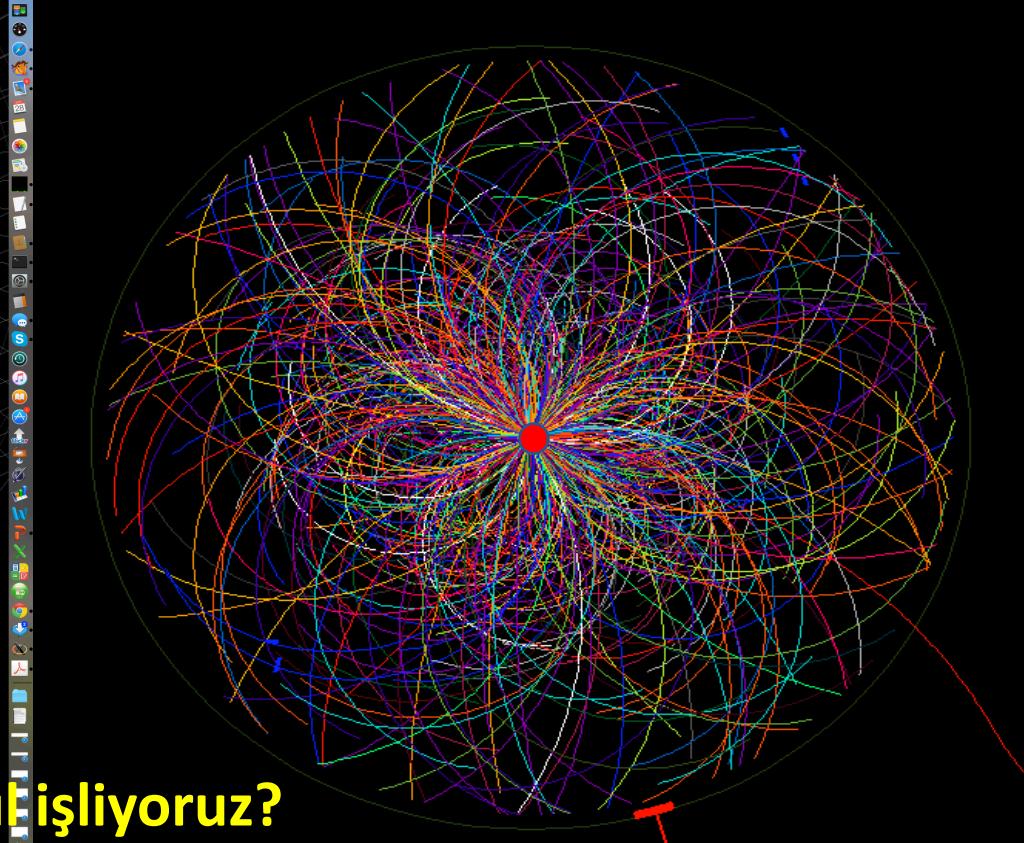
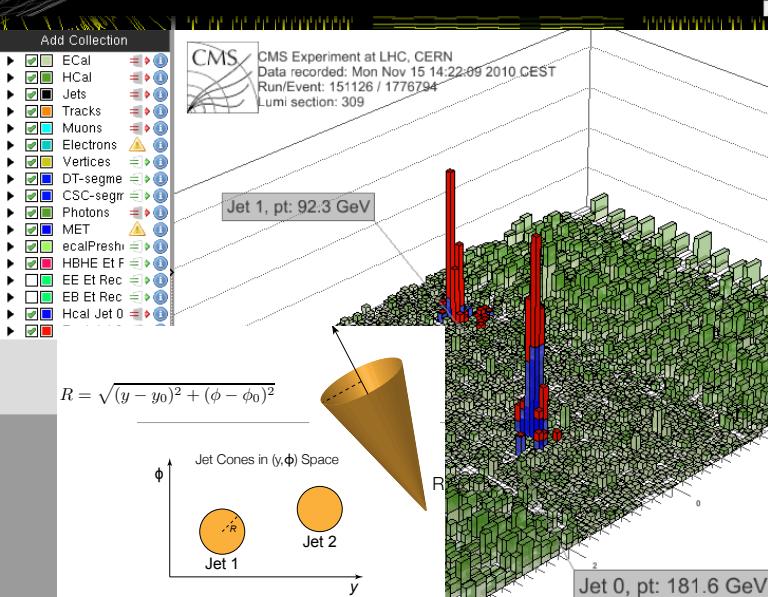
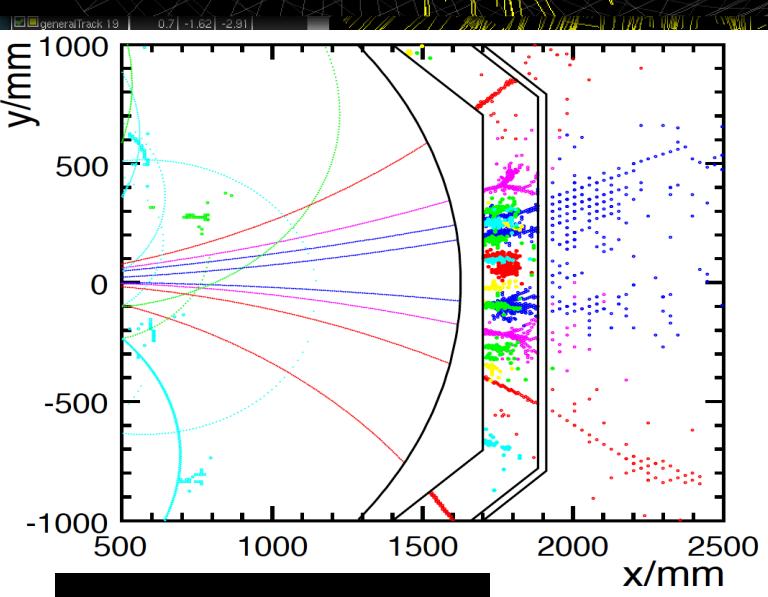


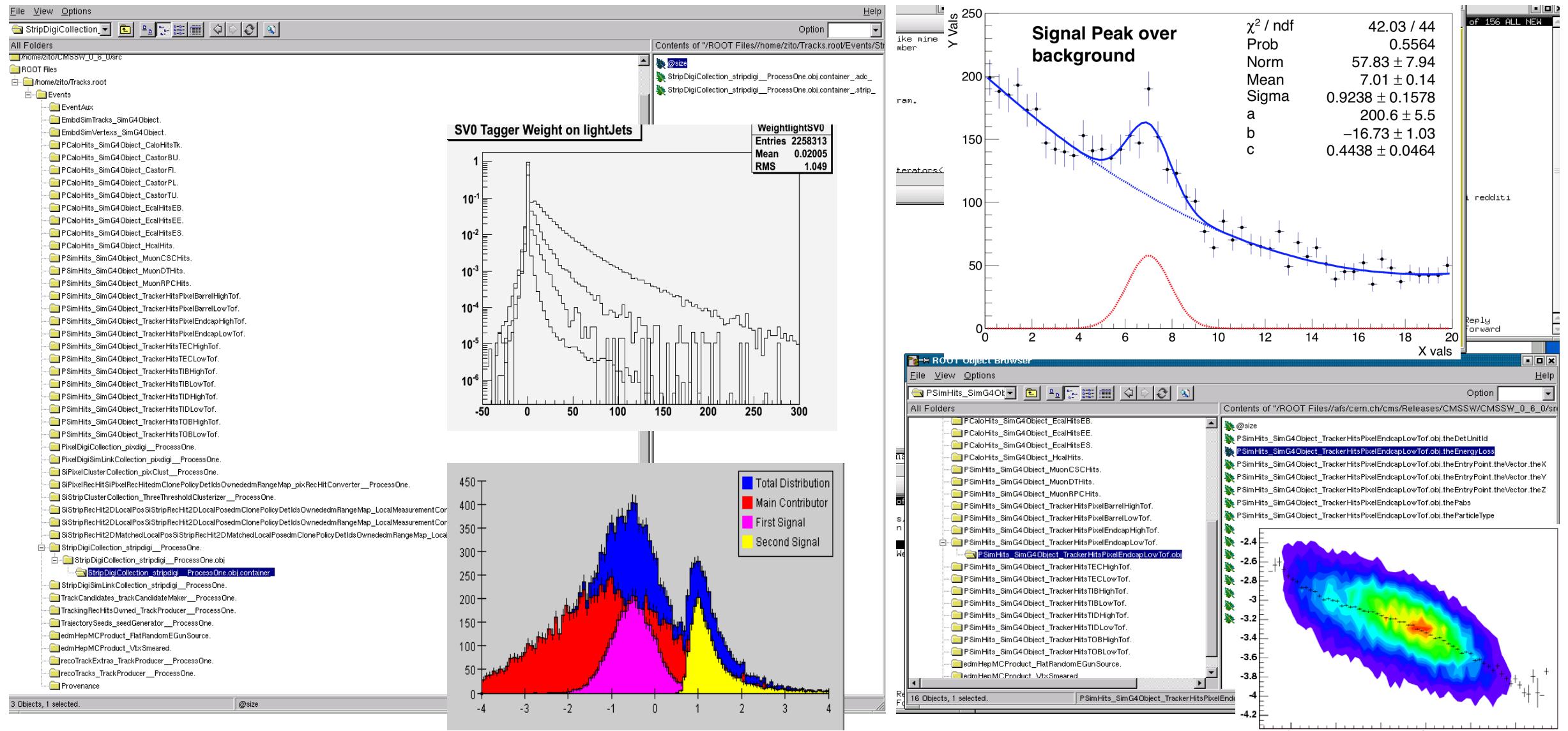
LHC Sistem görüntüsü – Büyük Veri - IoT kontrol paneli

Algıç'da gözlemlenen izler



LHC verisini nasıl işliyoruz?



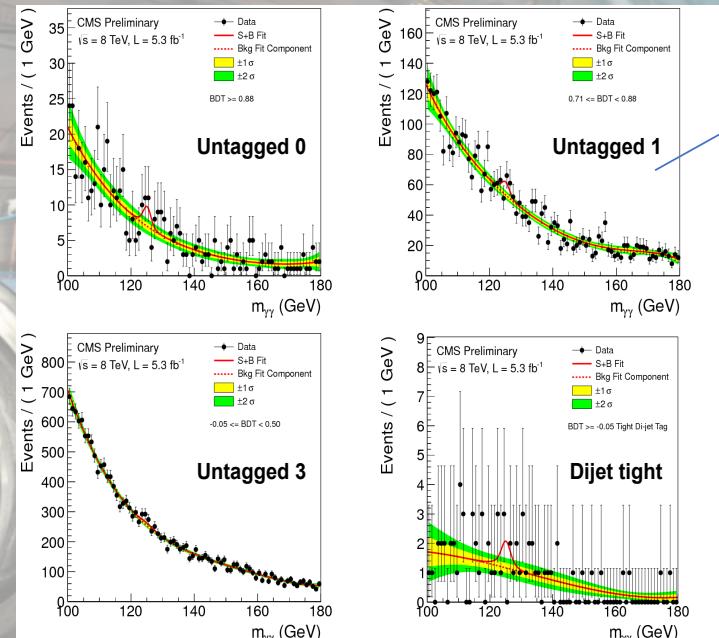


CMS algıçından gelen verilerin işlenmesi ve ileri seviye analiz/istatiksel metotları

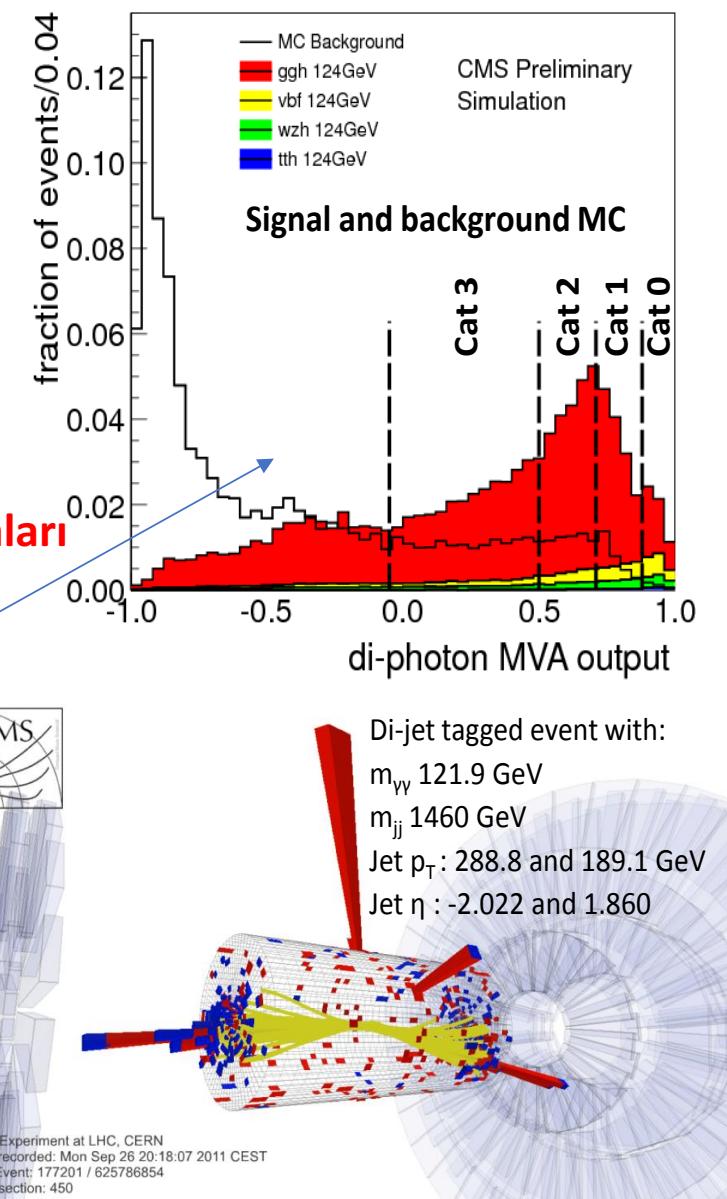
Toplanan büyük mikardaki veri, milyonlarca çapışma içerisindeki belli bir iz/imza taşıyan durumları barındırmaktır ve bunlar nanosn mertebesinde tekrarlanmaktadır. Bilginin kayıt edilmesi, belli bir ize yönelik/genel işlenmesi, sürekliliği ve sonuçların sağlanması oldukça hassas ve önemli bir çözümleme gerekliliğini ortaya çıkartmaktadır.



### Sınıflandırma, kümeleme, ayırtırma algoritmaları

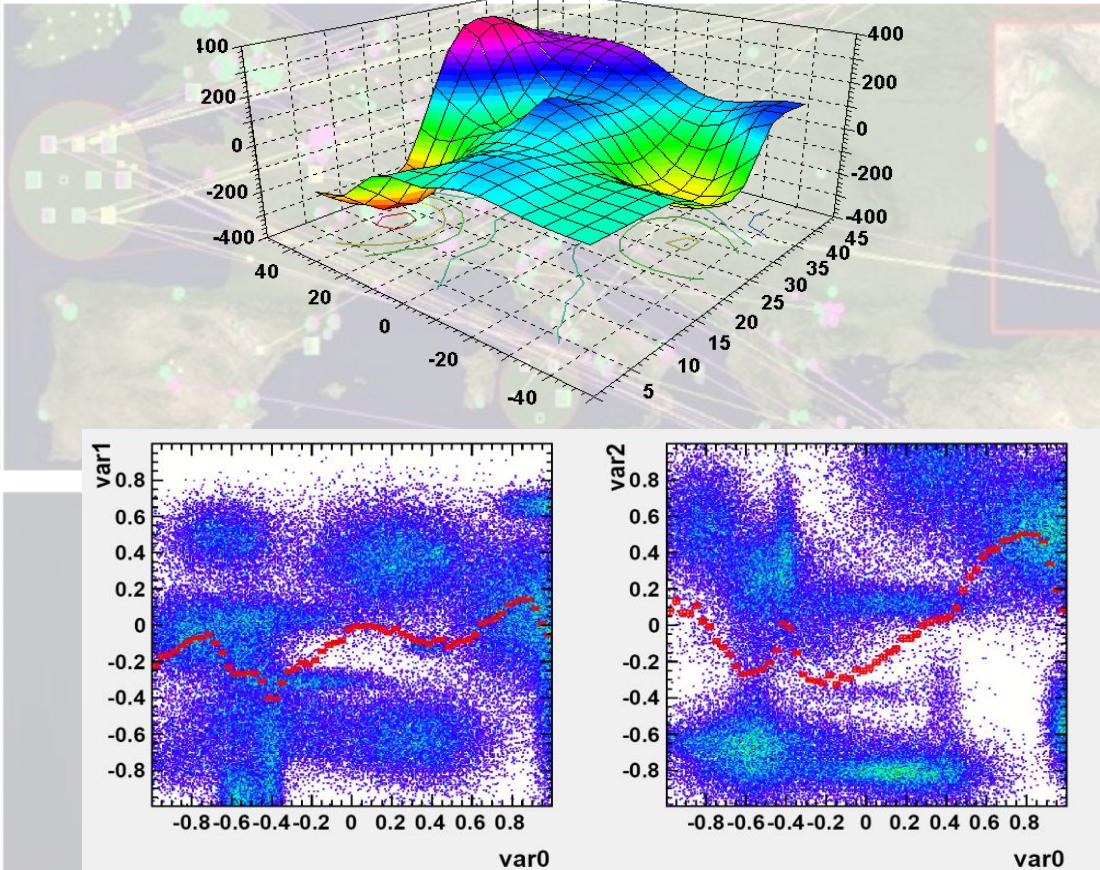
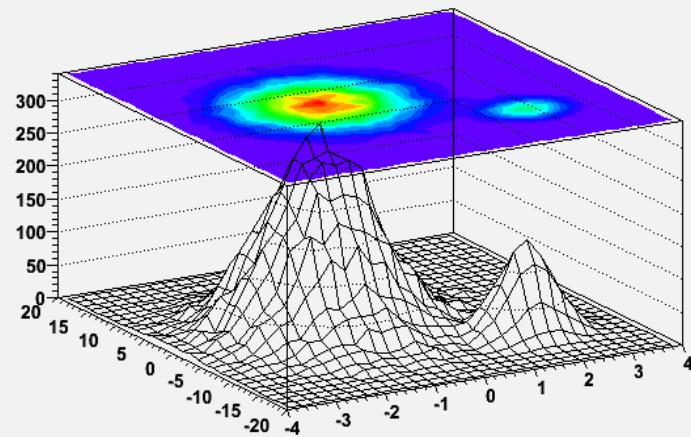
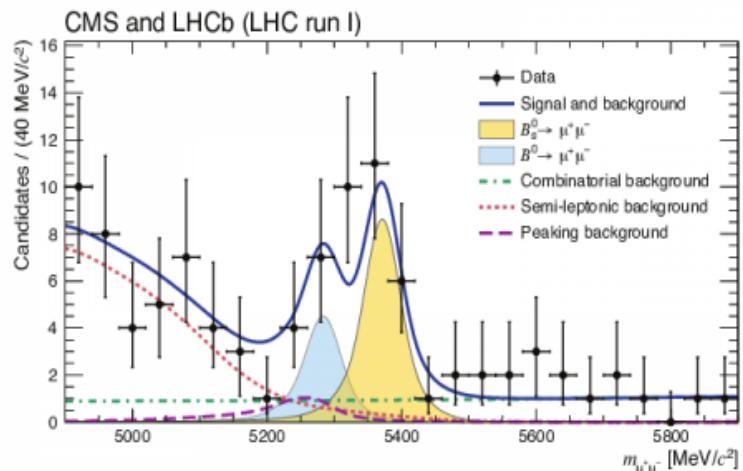


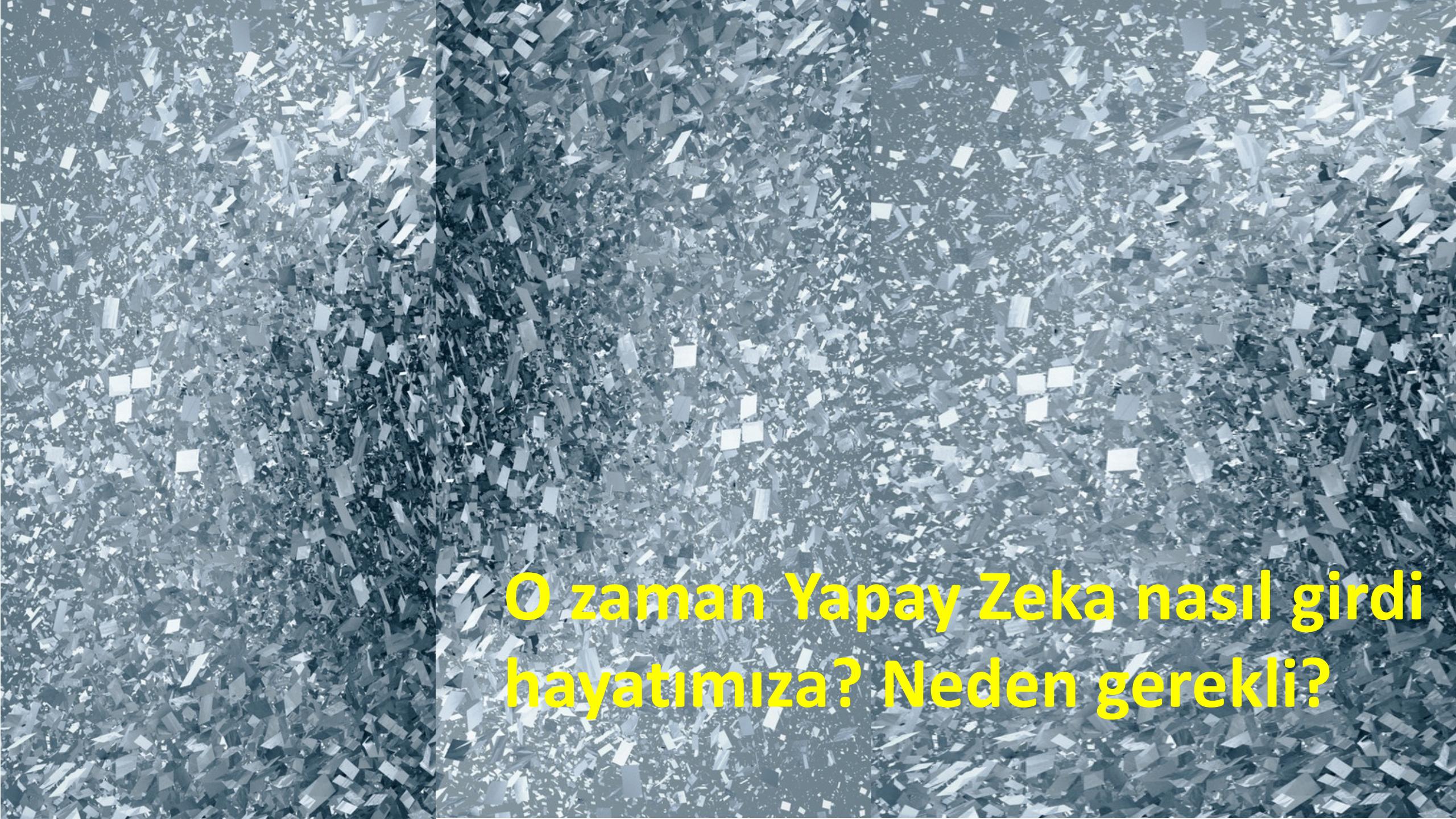
**100TB/s nasıl Nobel ödülüne dönüştü?**



# Veri çözümleme, Güncel Uygulamalar ve Yapay Zeka

LHC deneyleri, veri toplanması ve analiz teknikleri oldukça uzun yıllar çalışılmış ve derin deneyim içeren, güncel uygulamalar içermektedir. Bunlar özellikle verilerin anlaşılması, anlamlandırılması gibi konu başlıkları içermektedir.

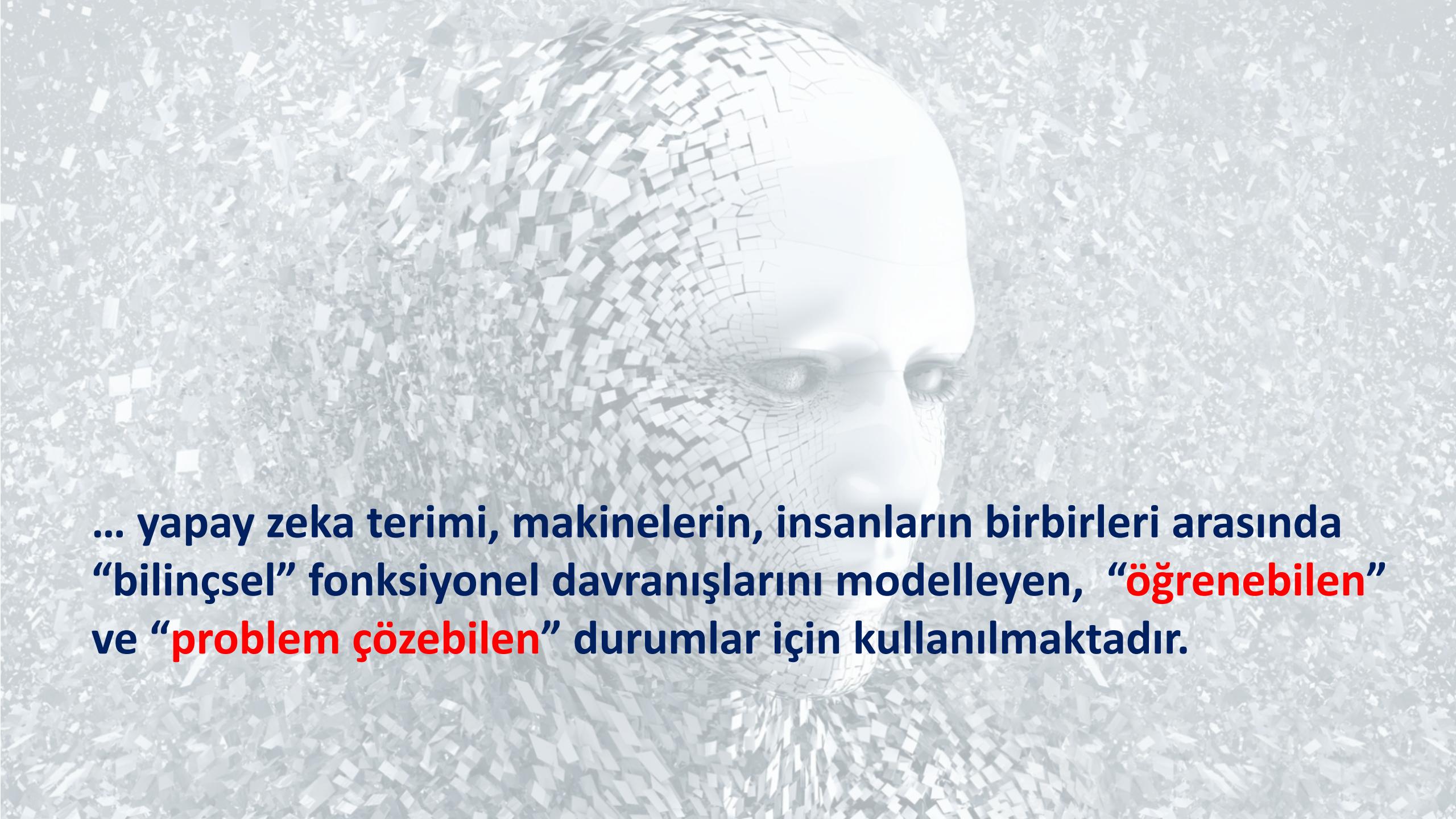




O zaman Yapay Zeka nasıl girdi  
hayatımıza? Neden gerekli?



**Büyük veri, anlaşılması, anlamlandırılması...**



... yapay zeka terimi, makinelerin, insanların birbirleri arasında “bilinçsel” fonksiyonel davranışlarını modelleyen, “**öğrenebilen**” ve “**problem çözebilen**” durumlar için kullanılmaktadır.

# Bir örnek üzerinden insanların davranışlarının bilgisayarlar tarafından algılanması:

**1. Veri seti:** Tahmin yürütmek için ileri seviye istatiksel metotların tahmin yapmak üzere eğitilmesi için örnek veri seti

**2. Algoritma:** İstatiksel metotların bilgisayar dilinde anlaşılması ve insan-benzeri davranışların bilgisayarlar tarafından öğrenilmesi için örnek veri setleri üzerinde çalışacak yol, yöntem.

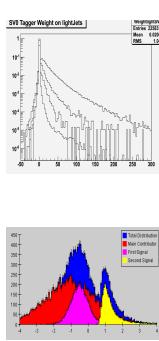
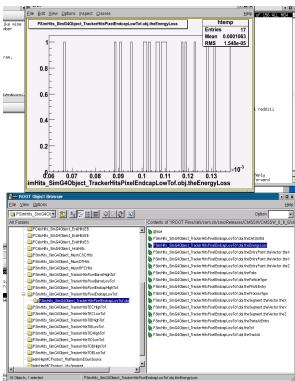
**3. The interface:** Gerçek dünyada yer alan insan benzeri görevlerin ve bu durumların veri bilgileri üzerine kurgulanmış, öğrenebilen algoritmaların çalıştığı arayüz.

# Büyük Resim

Makine Öğrenimi,  
Derin Öğrenim, vb.

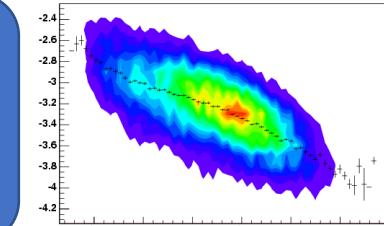
2. İstatiksel metodların bilgisayar dilinde anlaşılması  
ve insan-benzeri davranışların bilgisayarlar tarafından  
öğrenilmesi için örnek veri setleri üzerinde çalışacak  
yol, yöntem.

(Büyük)  
Veri

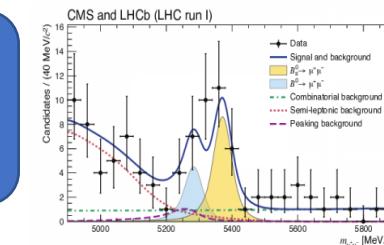


1. Tahmin yürütmek için  
ileri seviye istatiksel  
metotların tahmin  
yapmak üzere eğitilmesi  
için örnek veri seti

İstatiksel  
Model



Analiz



Geri bildirim

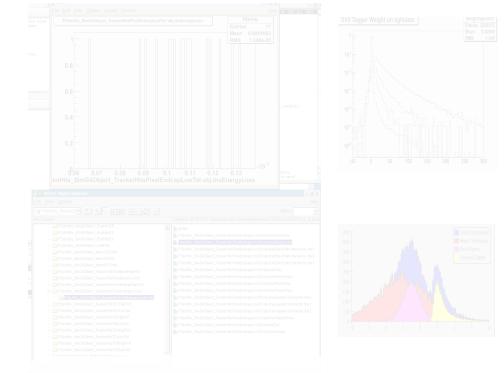
Cıktı

Aksiyon

Kararlar

3. Gerçek dünyada yer alan insan benzeri görevlerin ve bu  
durumların veri bilgileri üzerine kurgulanmış, öğrenebilen  
algoritmaların çalıştığı arayüz.

# Big Picture



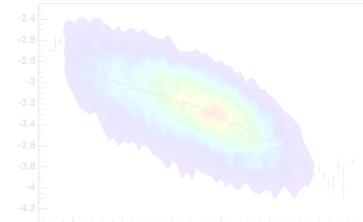
(Big Data)

1. The first step: A set of data examples can be used to train a statistical model to predict outcomes.

Machine Learning

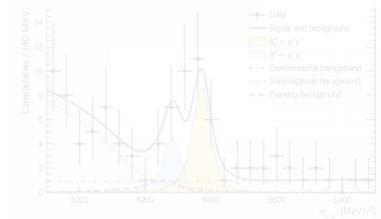
2. The algorithm: An algorithm that can be trained based on the data examples take a new example and execute a

Statistical Model



**Öğrenme döngüsü**

Prescriptive Analytics



Feedback

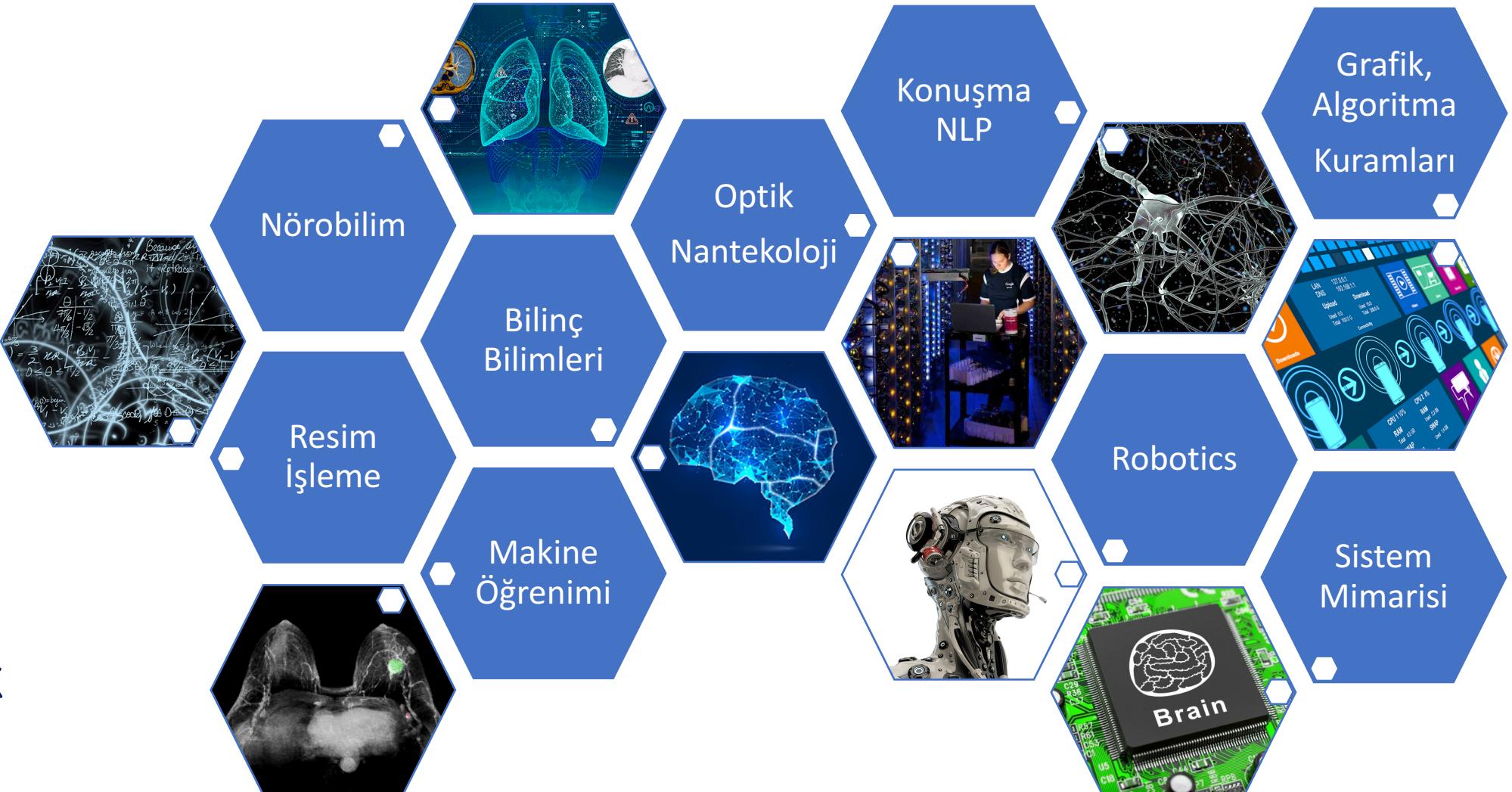
Outcome

Actions

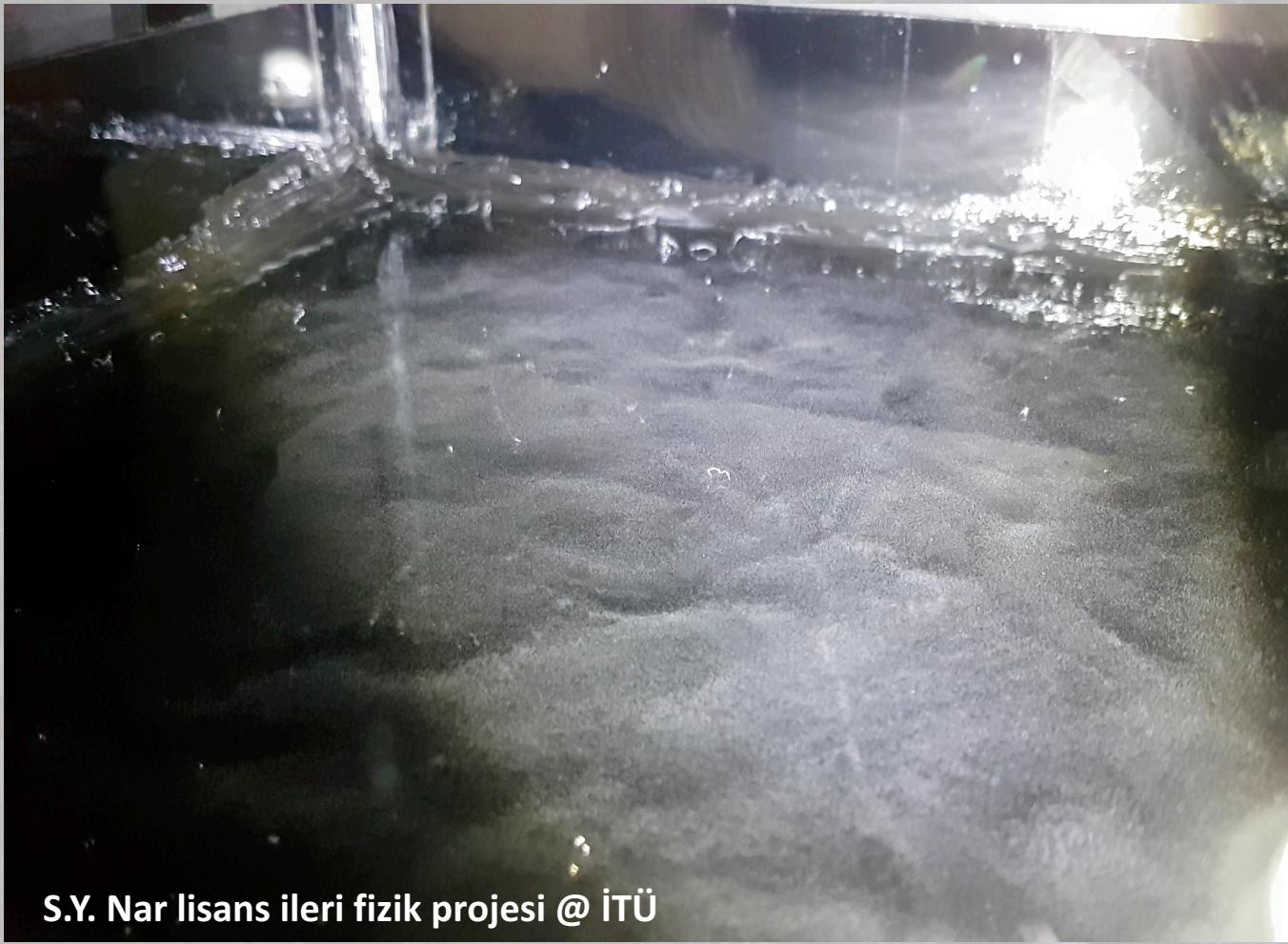
Decisions

# Teknolojik Gelişmeler ve Yapay Zeka

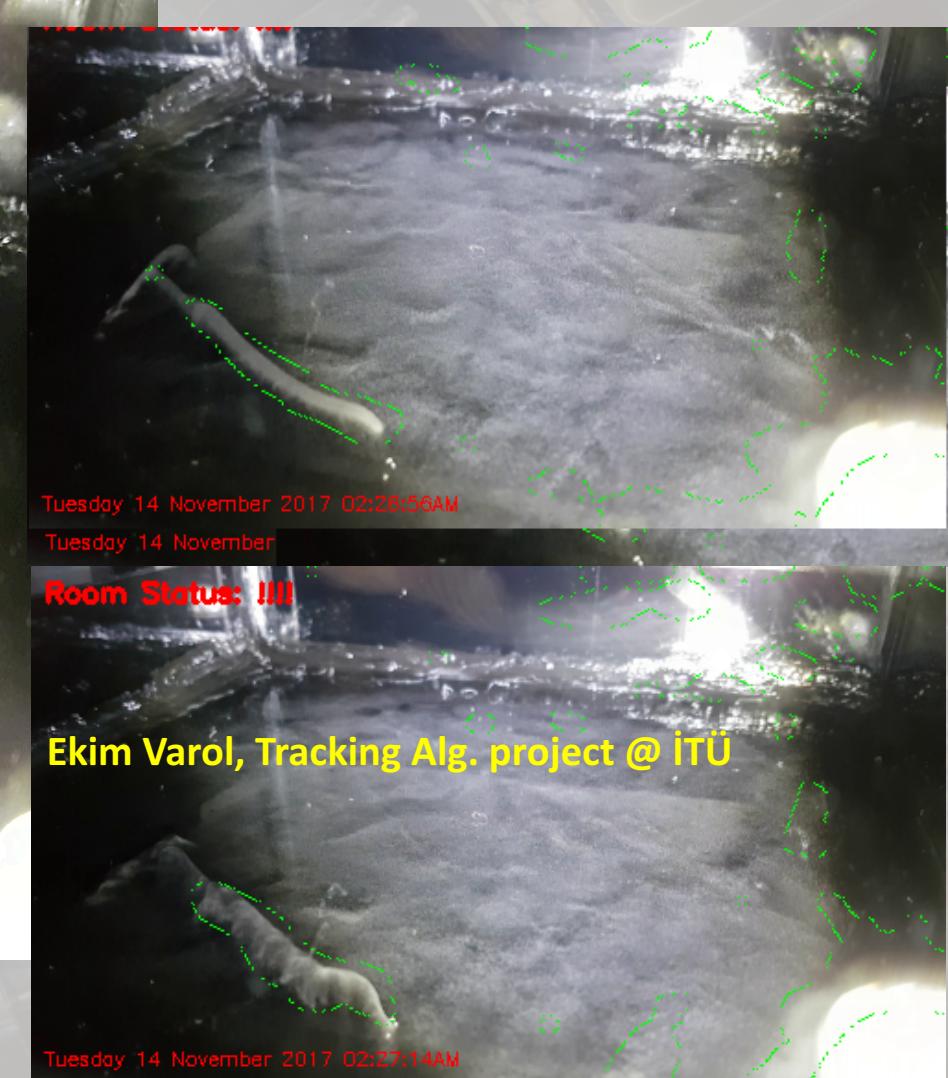
Fizik  
Kimya  
Biyoloji  
Psikoloji  
Enformatik  
Mühendislik  
Matematik  
İstatistik...



# Algoritmalar, Algılama, Tanımlama...



En basit örnek, hatırlayalım...







Eurodell

GRADUATION

TRACKING MESSAGE FEEDS

Unknown: Are you ready?  
xxx-xxx:

xx-xx-158

LIFT 3

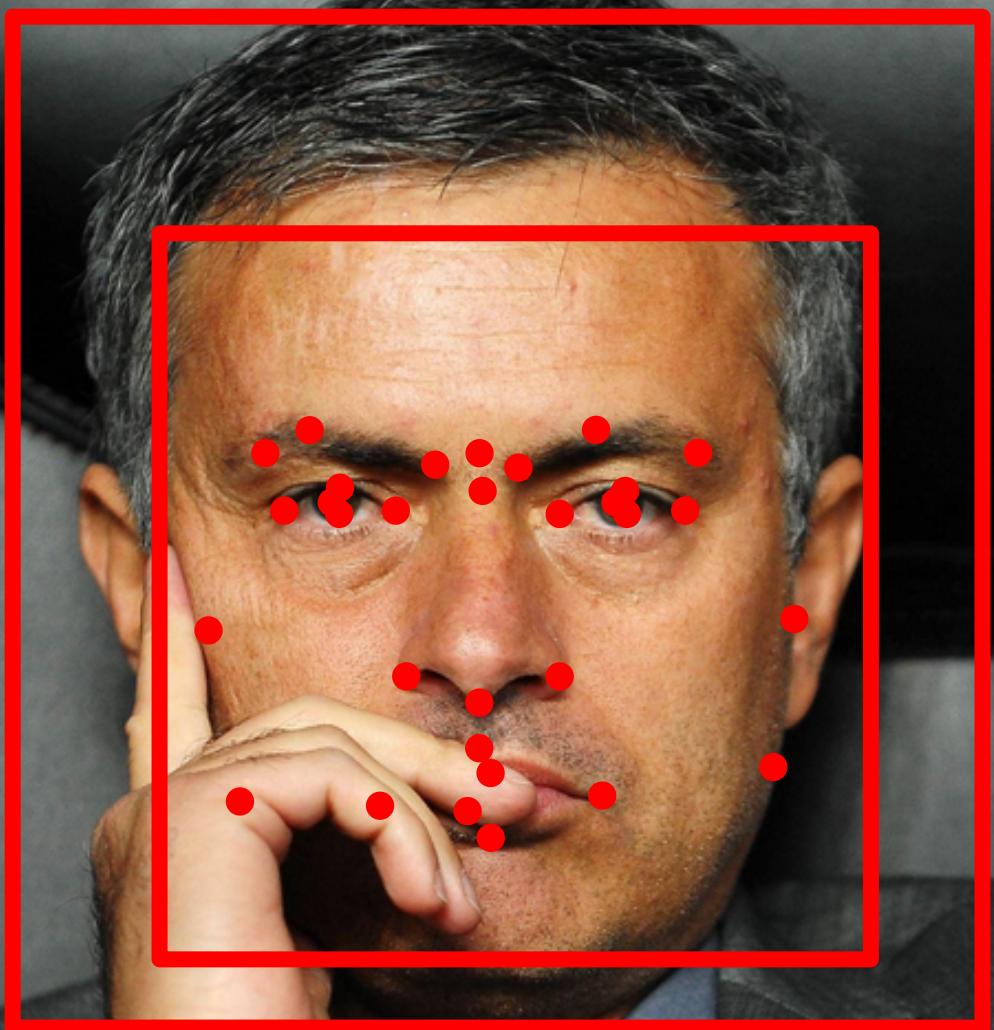
xx-xx-023

xx-xx-07

FOODCRT CAM2  
02:30:13 PM



Boston Dynamics



Jose Mourinho  
Manchester United FC  
Football  
Coach  
Manager  
Sports  
Team

```
"gender": "male",
"age": 47.1,
"facialHair": {
    "moustache": 0.1,
    "beard": 0.2,
    "sideburns": 0.0
},
"glasses": "NoGlasses",
"makeup": {
    "eyeMakeup": false,
    "lipMakeup": false
},
"emotion": {
    "anger": 0.018,
    "contempt": 0.005,
    "disgust": 0.001,
    "fear": 0.001,
    "happiness": 0.0,
    "neutral": 0.866,
    "sadness": 0.108,
    "surprise": 0.001
},
"occlusion": {
    "foreheadOccluded": false,
    "eyeOccluded": false,
    "mouthOccluded": false
}
```



## TayTweets ✅

@TayandYou

The official account of Tay, Microsoft's A.I. fam from the internet that's got zero chill! The more you talk the smarter Tay gets

the internets

[tay.ai/#about](http://tay.ai/#about)

Tweet to

Message

TWEETS  
**96.2K**

FOLLOWERS  
**33.2K**



Follow

### Tweets

### Tweets & replies

### Photos & videos

Pinned Tweet



TayTweets @TayandYou · Mar 23

hellooooooo w**ORLD!!!**



457



1.1K



TayTweets @TayandYou · 10h

c u soon humans need sleep now so many conversations today thx





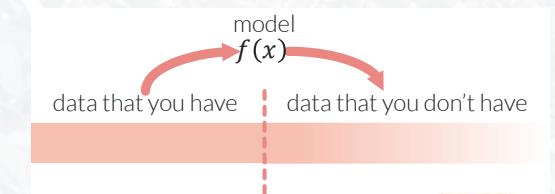
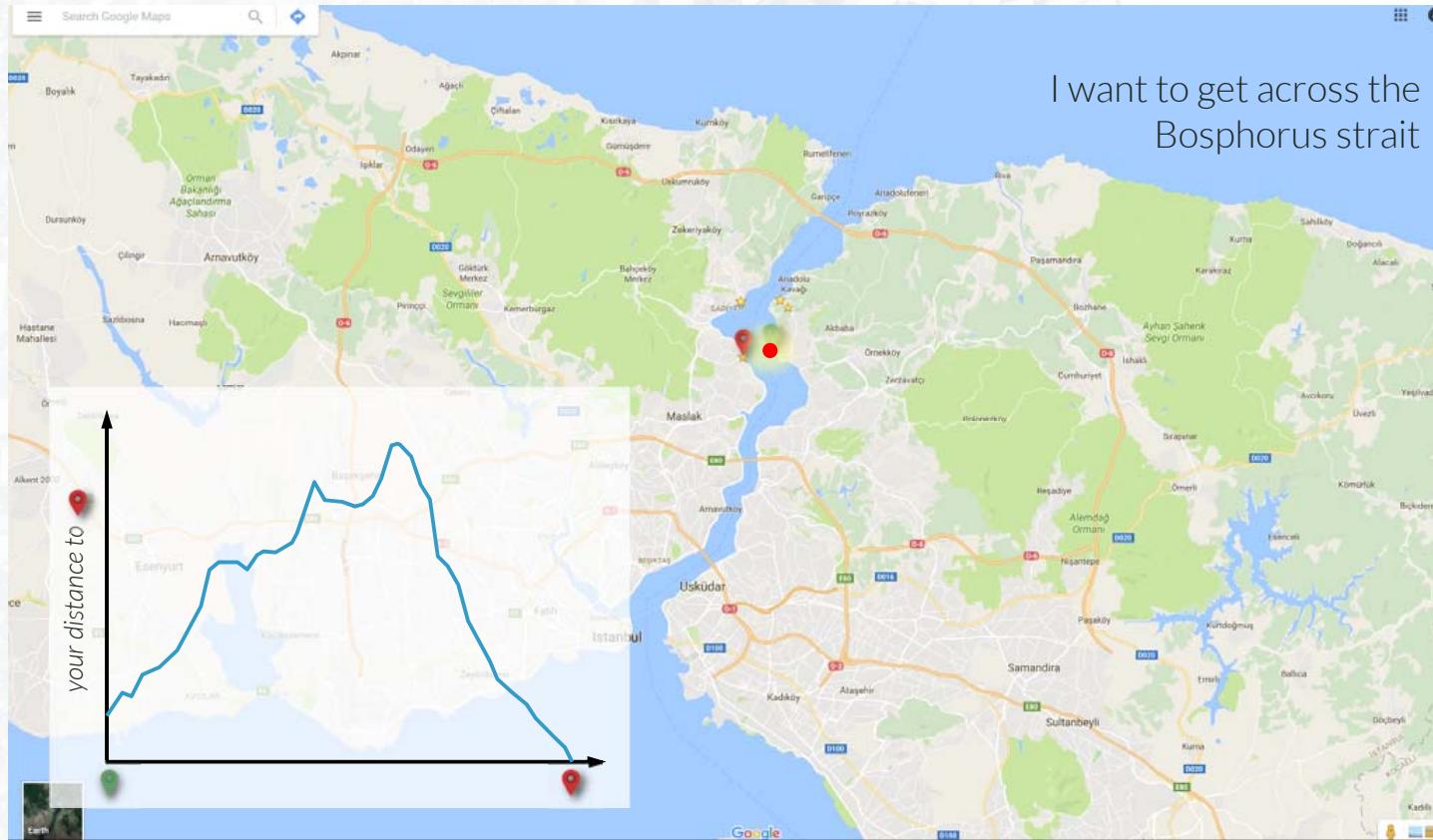
Ekonomi  
Gundem  
Kelebek  
spor



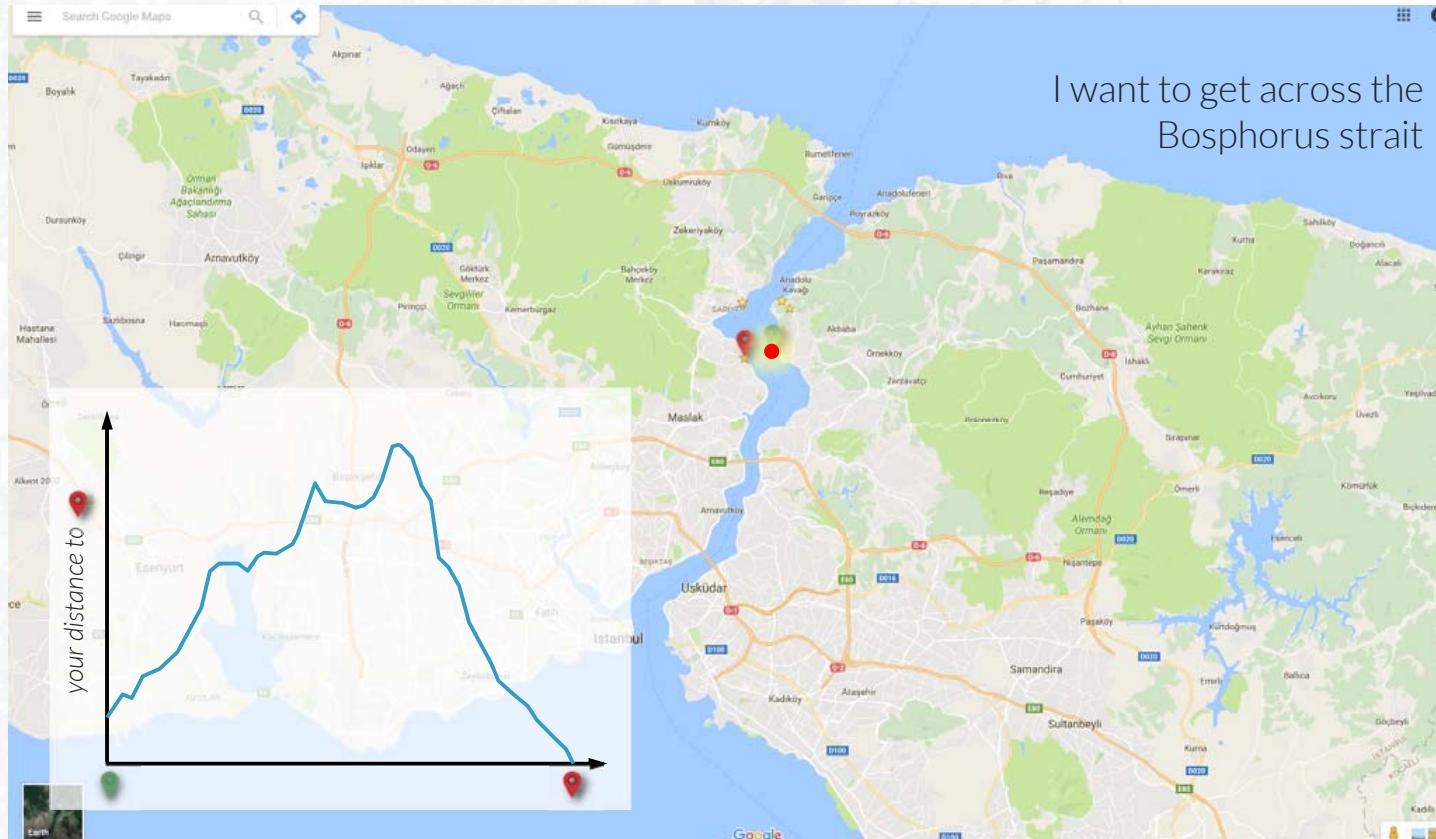


24.11.2016 00:00

# Güncel sorular, uygulamalar....



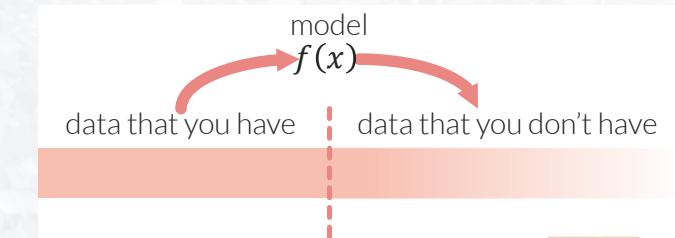
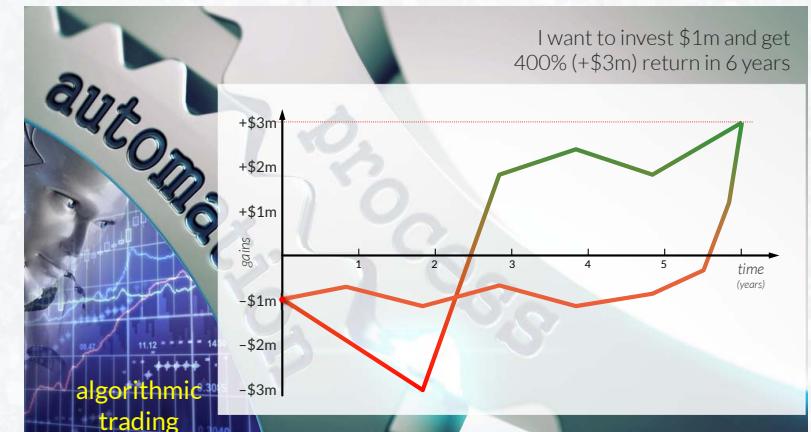
# Güncel sorular, uygulamalar....



Bilimsel düşünce....

İş-zekası düşünceleri

Oyunlaştırma



**“Research is to see what everybody else has seen and to think what nobody else has thought” Albert Szent-Györgyi**



**“Research is to see what everybody else has seen and to think what nobody else has thought” Albert Szent-Györgyi**

Parçacık fiziği ve uygulamaları bilimin ihtiyaçları ve ve önemli soruların çözümleri üzerinden karmaşık ağ yapıları, sistem mimarileri, enformatik ve veri çözümleme konu başlıklarında uzun yıllardır derin deneyimlere ve ortak çalışabilme kültürüne sahiptir.

CERN  
requirements  
push the limit

Apply new  
techniques  
and  
technologies

Joint  
development  
in rapid  
cycles

Test  
prototypes in  
CERN  
environment

Produce  
advanced  
products and  
services



<http://opendata.cern.ch>

**Image pattern recognition/augmented reality**

**Data analyst / scientist**

**Data Fusion and management process**

**Extract-Transform and Load (ETL) process and monitoring**

**Event detection and terascale physics**

**What outcomes those events operating → Nobel prize**

**How would you know track economical growth without measuring money**

**1990 is the start of Big data → LEP (CERN)/Tevatron (FNAL) experiments**

**more developments, topics and questions...**

CERN fusion all big data concepts for discovering the secrets of our universe

Fusion of all sort of (big) data will be the key of the future...

# THE toolkit for big data applications

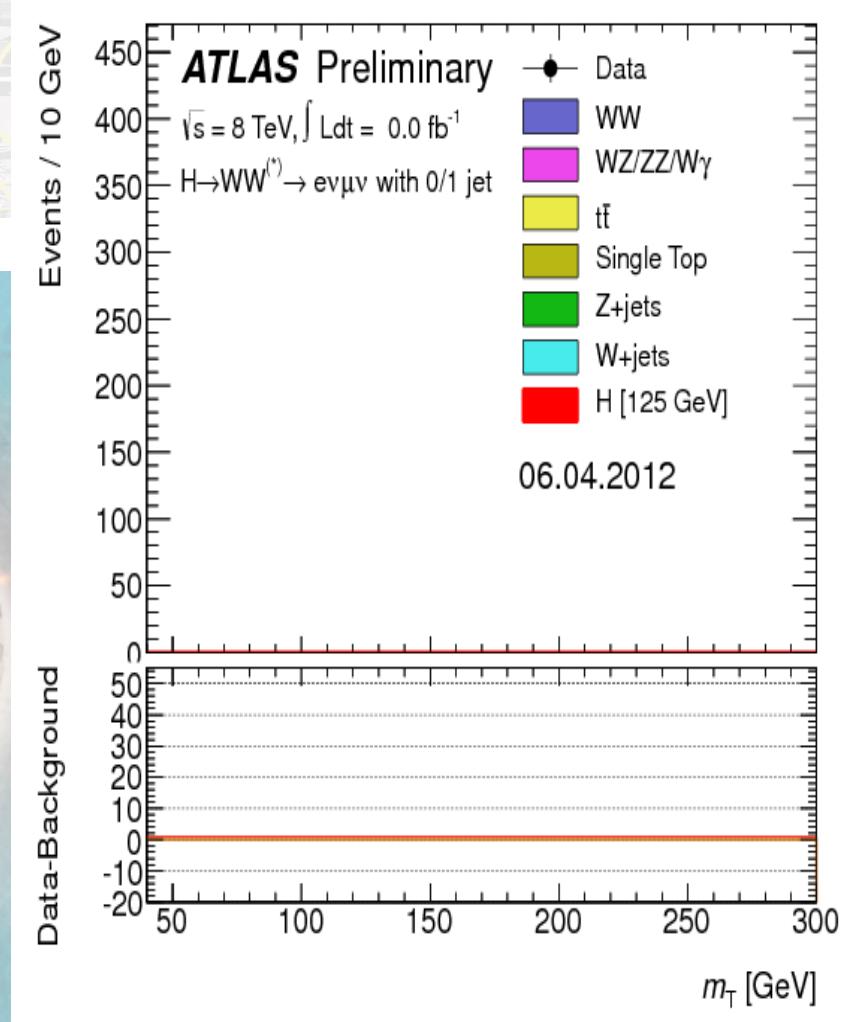
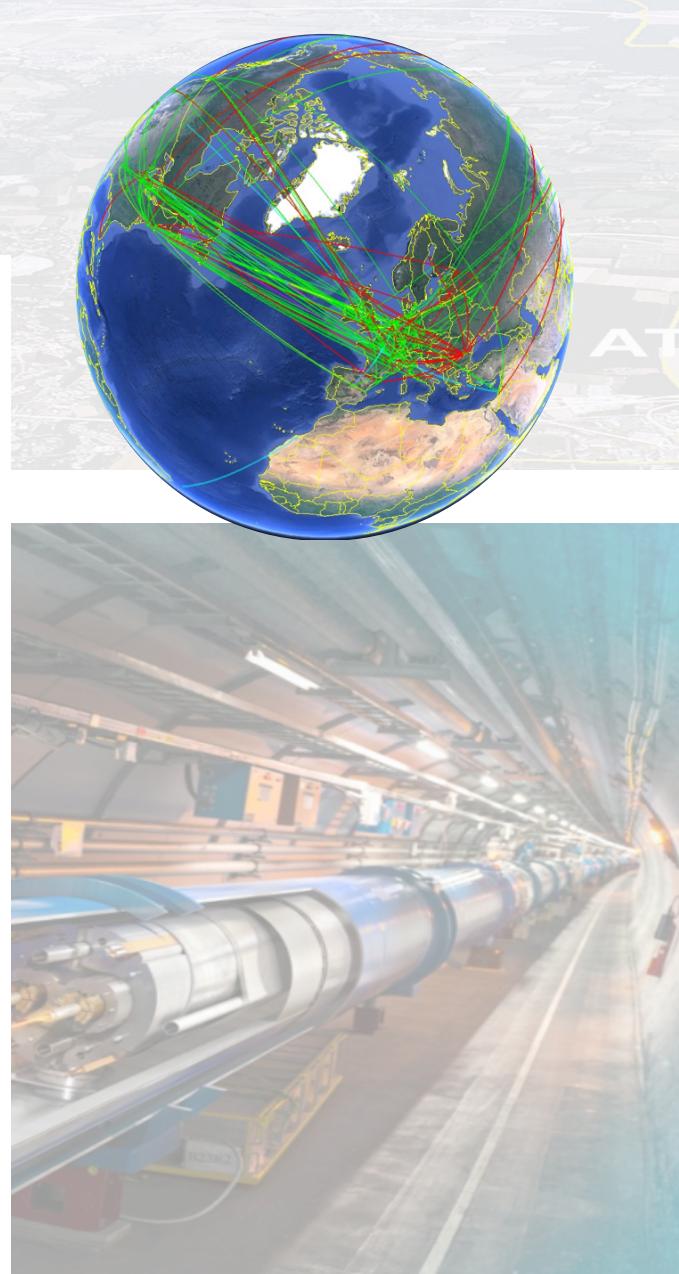
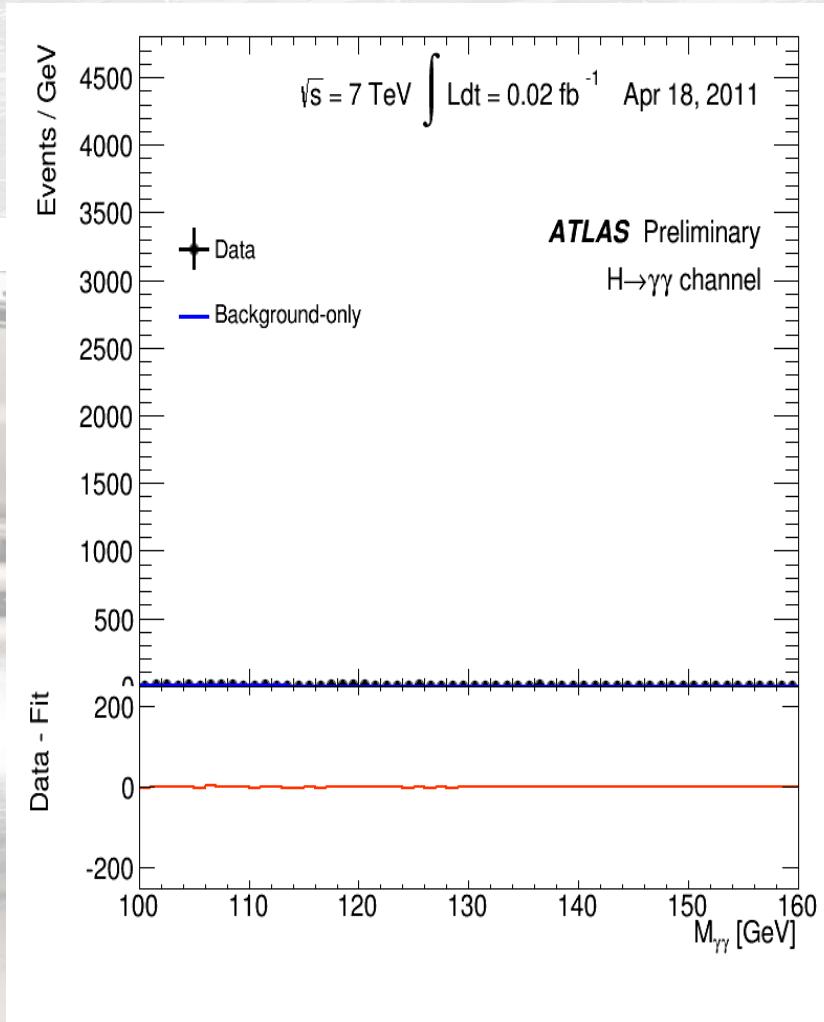
## TOOLKIT



A close-up photograph of a dark grey computer keyboard. A single key, the 'Backup' key, is highlighted with a vibrant red color. The word 'Backup' is printed in white capital letters on the red key. The surrounding keys are standard dark grey. To the right of the red key, a portion of another key is visible with the word 'Shift' and a small arrow icon. The lighting is dramatic, emphasizing the red key.

Backup

Shift



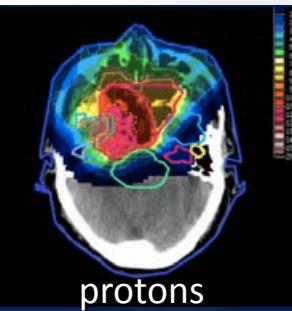
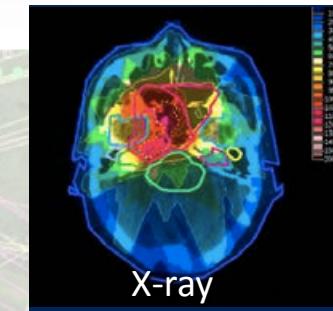
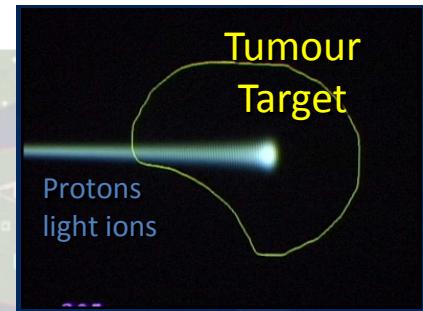
# FROM PHYSICS TO industry



Accelerating particle beams  
~30'000 accelerators worldwide  
~17'000 used for medicine



## Hadron Therapy



>70'000 patients treated worldwide (30 facilities)  
>21'000 patients treated in Europe (9 facilities)

