

# Welcome to Sapienza Data Science

Aris Anagnostopoulos  
Sapienza Università di Roma

# Who am I?

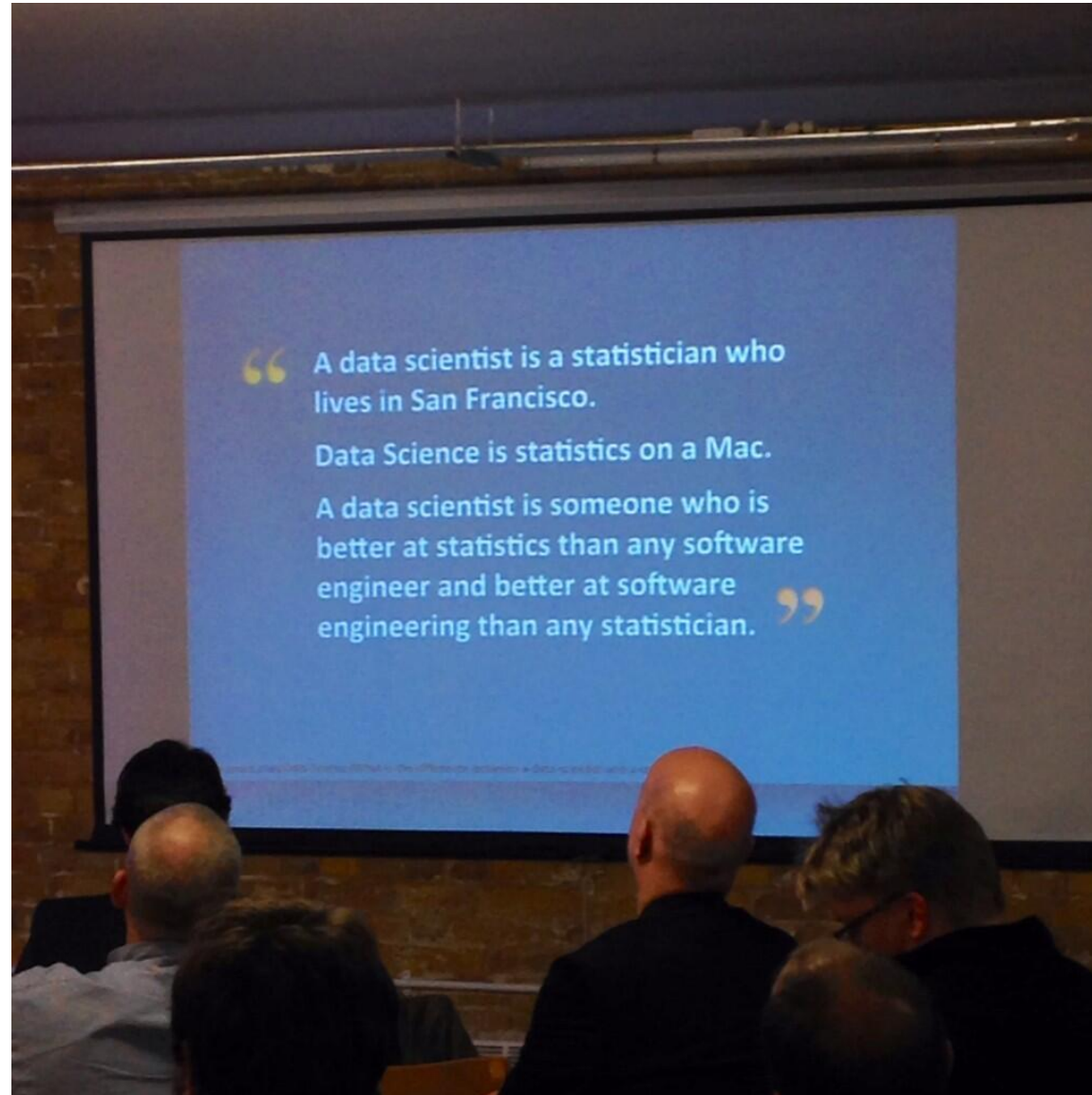
- Professor in computer engineering at Sapienza
- Studied in Greece, USA
- Worked in the US for a bit, then came to Sapienza
- Works on algorithms, data analysis, data science
  - Search engines, social networks, recommender systems, misinformation, biology, finance, medicine, ...
- Helped create the Data Science master in Sapienza

Who are you?

# What is Data Science?

Boh...

# What is Data Science?



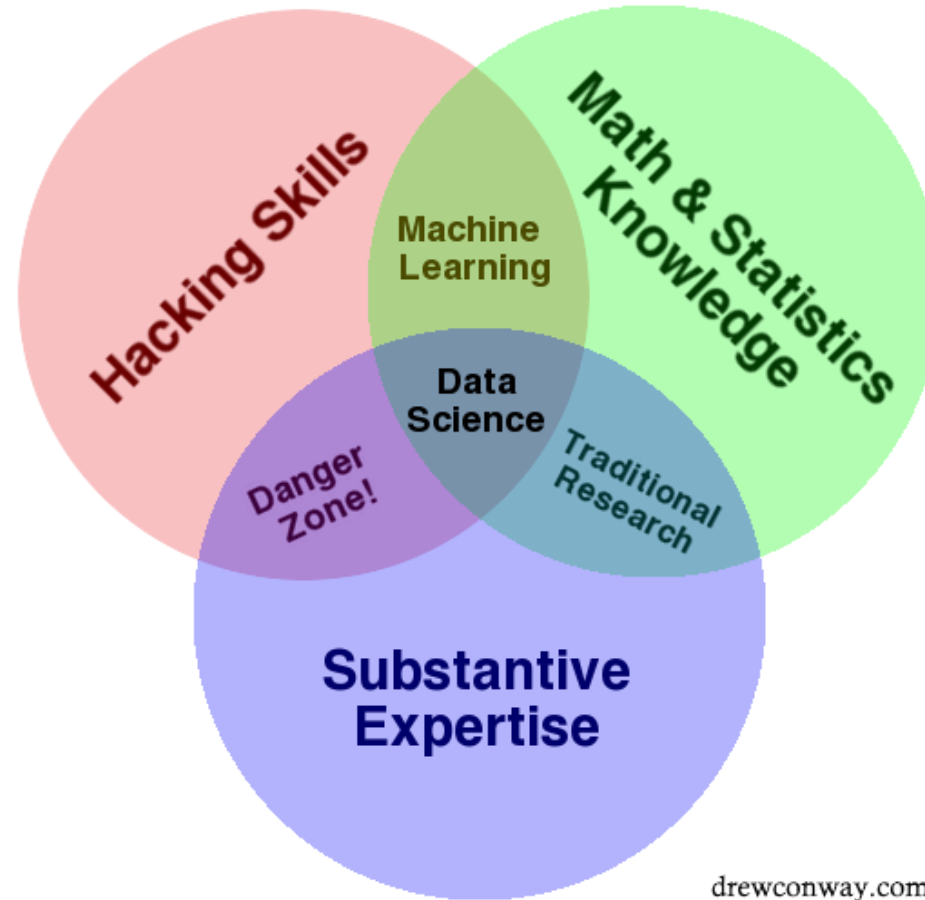
# What is Data Science?

From Wikipedia:

**Data science** is an [interdisciplinary](#) field that uses [scientific methods](#), processes, [algorithms](#) and systems to extract or extrapolate [knowledge](#) and insights from noisy, structured and [unstructured data](#), and apply knowledge from data across a broad range of application domains. Data science is related to [data mining](#), [machine learning](#) and [big data](#).

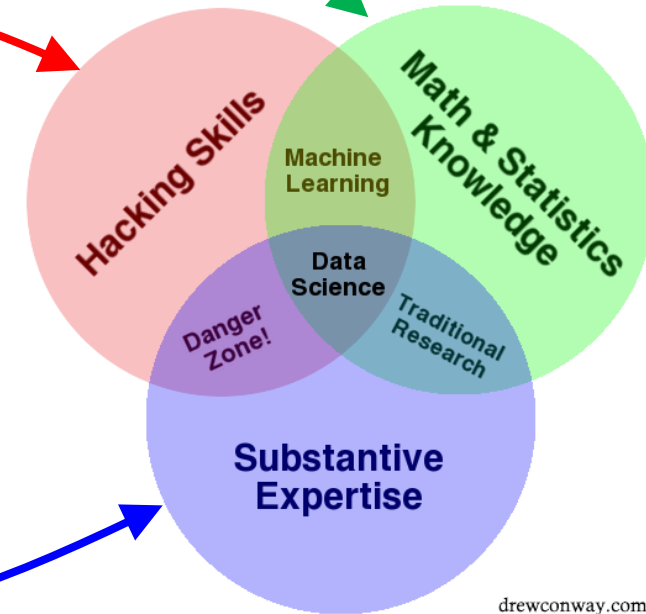
Data science is a "concept to unify [statistics](#), [data analysis](#), [informatics](#), and their related [methods](#)" in order to "understand and analyse actual [phenomena](#)" with [data](#).<sup>[4]</sup> It uses techniques and theories drawn from many fields within the context of [mathematics](#), statistics, [computer science](#), [information science](#), and [domain knowledge](#). However, data science is different from [computer science](#) and information science. [Turing Award](#) winner [Jim Gray](#) imagined data science as a "fourth paradigm" of science ([empirical](#), [theoretical](#), [computational](#), and now data-driven) and asserted that "everything about science is changing because of the impact of [information technology](#)" and the [data deluge](#).

# What is Data Science?



# Programs in Data Science

- In the last 10 years, almost all the large Universities in the world created programs in Data Science
- Provide education that combines
  - Computer science/engineering
  - Math & Statistics
  - Domain knowledge



# Applications

Applications in a lot of areas:

Computer science (All kinds of startups)

Biology

Epidemiology (Prediction)

Medicine

Social sciences (Online experiments)

Politics (Obama)

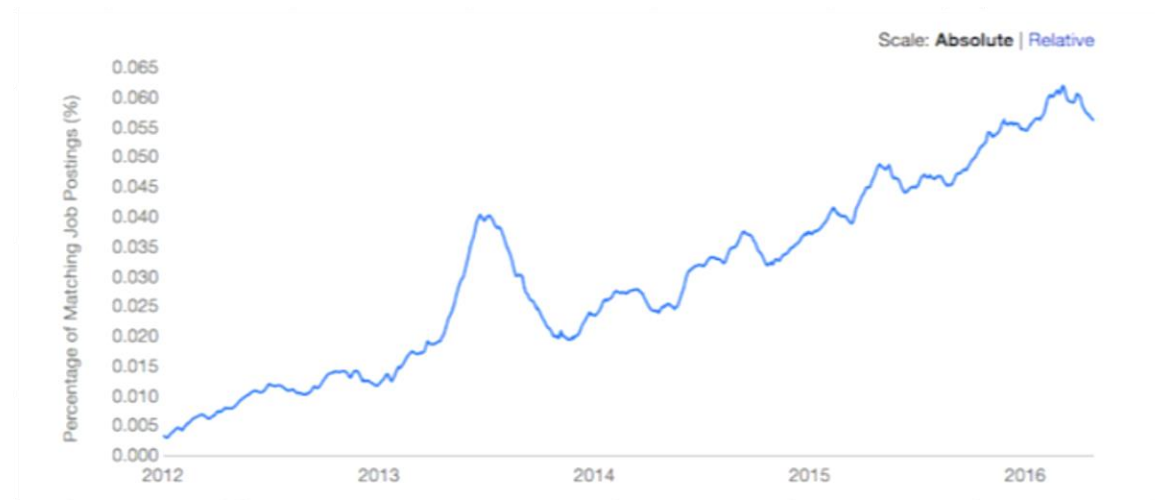
...



# Master Degree in Data Science

The job of big-data analyst is requested for more than 10 years in startups, public agencies and many companies in Fortune 500.

Up to 800 K requests for job in Europe in the area of big data.



# Data Science programs in Italy

**In 2015, 3 Data Science master (laurea magistrale) programs started in Italy:**

- Ca'Foscari
- Sapienza
- U. Torino

**More opened since then:**

- U. Bologna
- U. Firenze
- U. Milano-Bicocca
- U. Padova
- U. Pisa
- U. Trieste + U. Udine (+ SISSA + ICTP)



# Data Science @Sapienza

## Highlights

- Laurea magistrale in Data Science
- 2 years
- After the triennale
- Founded in 2015
- First graduates in October 2017
- Taught in english





# Data Science @Sapienza

## Features

- Multidisciplinary: 4 departments involved:
  - Informatica
  - Ingegneria informatica, automatica e gestionale
  - Ingegneria dell'Informazione, elettronica e telecomunicazioni
  - Scienze statistiche
- Collaboration with the Industry
  - Stages, internships, final project
- Applied from day 1
- Accepts students from various disciplines
- About 1/3 of the students are nonitalians



# Industrial liaison program

ISTAT

Engineering

Accenture

IBM

EY

Poste Italiane

Immobiliare.it

Teradata

SAS

NTT DATA

Oracle

Almawave

Telecom

Microsoft

UNICREDIT

ESA/Esrin

Splunk

Vodafone

CINECA

Sky

KPMG

LAIT Lazio

EMC2

Immobiliare

5M Informatica

Vitrociset



# Program structure

---

4 compulsory courses

1<sup>st</sup> year

Give the basics of DS to everyone

Computer science

Algorithms  
Programming  
Network theory

Probability & Statistics

Hypothesis testing  
Prediction



# Program structure

**7 optional courses**  
**Some courses from each category**  
**Specialize in what interests you**

## Computer science

Databases  
Data mining  
Data privacy & security  
Internet of things

## Probability & Statistics

Advanced statistics  
Machine learning  
Optimization methods

## Specialized domains

Bioinformatics  
Digital epidemiology  
Management  
Data-driven economics

**4 compulsory courses**  
**1<sup>st</sup> year**  
**Give the basics of DS to everyone**

## Computer science

Algorithms  
Programming  
Network theory

## Probability & Statistics

Hypothesis testing  
Prediction



# Program structure

## Practical experience

Training camps (Google, IBM, SAS)  
Stages and internships (Italy, EU, USA)  
Final thesis

**7 optional courses**  
**Some courses from each category**  
**Specialize in what interests you**

### Computer science

Databases  
Data mining  
Data privacy & security  
Internet of things

### Probability & Statistics

Advanced statistics  
Machine learning  
Optimization methods

### Specialized domains

Bioinformatics  
Digital epidemiology  
Management  
Data-driven economics

**4 compulsory courses**  
**1<sup>st</sup> year**  
**Give the basics of DS to everyone**

### Computer science

Algorithms  
Programming  
Network theory

### Probability & Statistics

Hypothesis testing  
Prediction



# 2017 students

