# Welcome to Sapienza Data Science

Aris Anagnostopoulos
Sapienza Università di Roma

# Who am I?

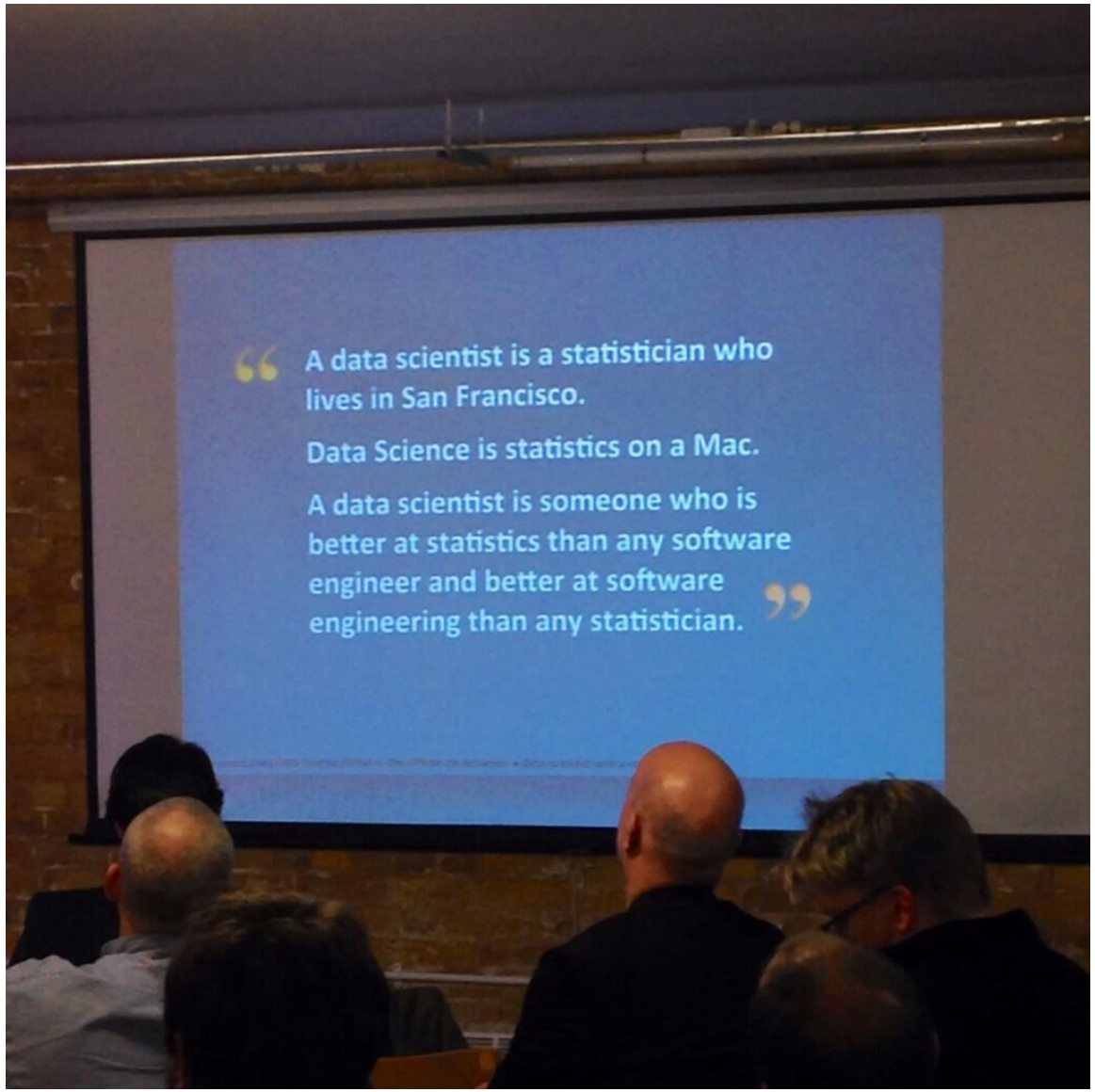
- Professor in computer engineering at Sapienza
- Studied in Greece, USA
- Worked in the US for a bit, then came to Sapienza
- Works on algorithms, data analysis, data science
  - Search engines, social networks, recommender systems, misinformation, biology, finance, medicine, …
- Helped create the Data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

Who are you?

# What is Data Science?

Boh…

# What is Data Science?



# What is Data Science?

From Wikipedia:

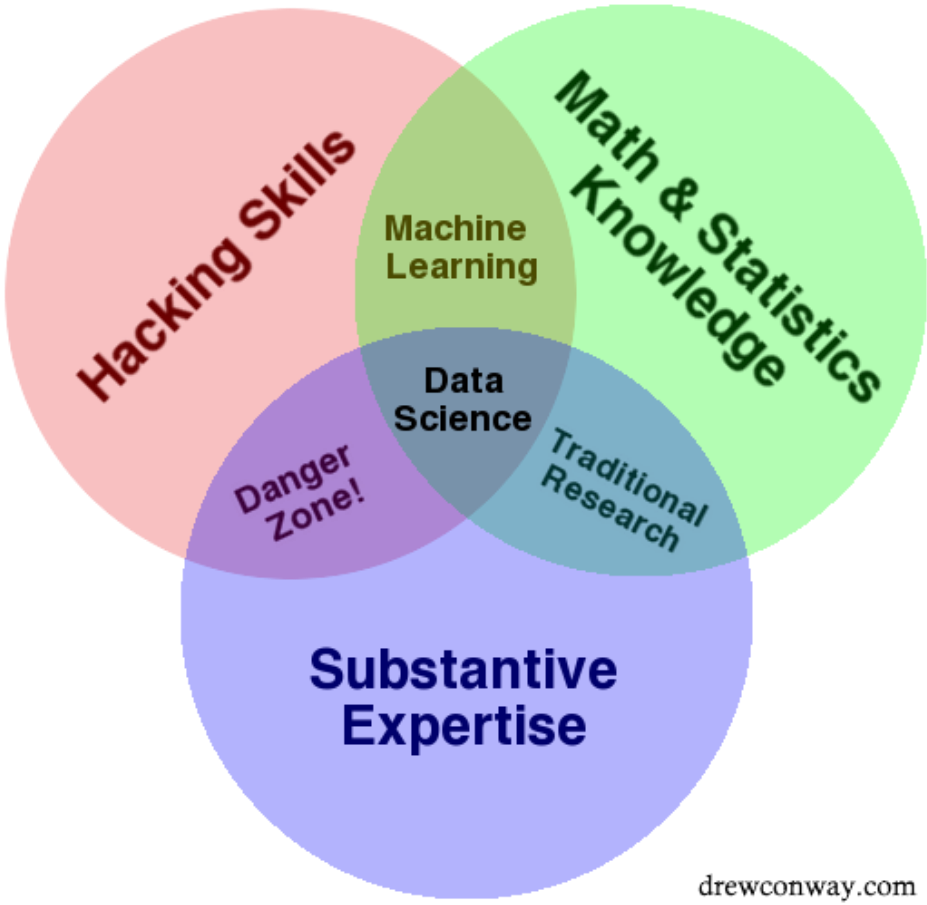
**Data science** is an interdisciplinary academic that uses statistics, scientific computing, scientific methods, processes, algorithms and systems to extract or extrapolate knowledge and insights from noisy, structured, and unstructured data.

Data science also integrates domain knowledge from the underlying application domain (e.g., natural sciences, information technology, and medicine). Data science is multifaceted and can be described as a science, a research paradigm, a research method, a discipline, a workflow, and a profession.

Data science is a "concept to unify statistics, data analysis, informatics, and their related methods" to "understand and analyze actual phenomena" with data. 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.

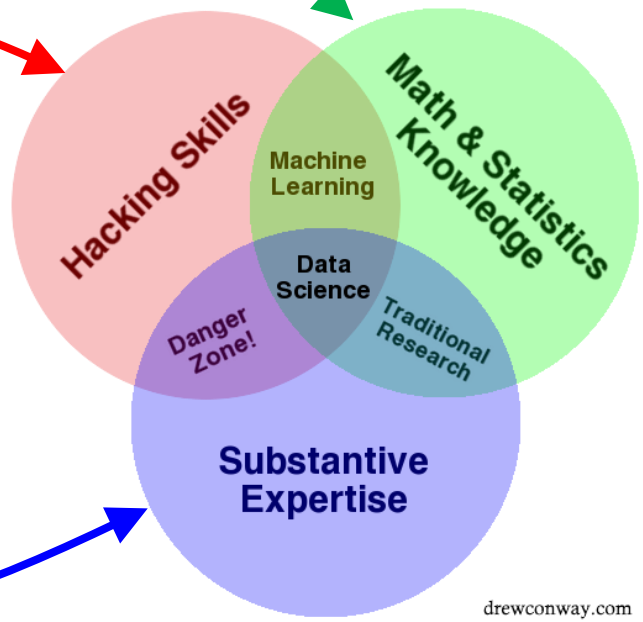
A **data scientist** is a professional who creates programming code and combines it with statistical knowledge to create insights from data.

# 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

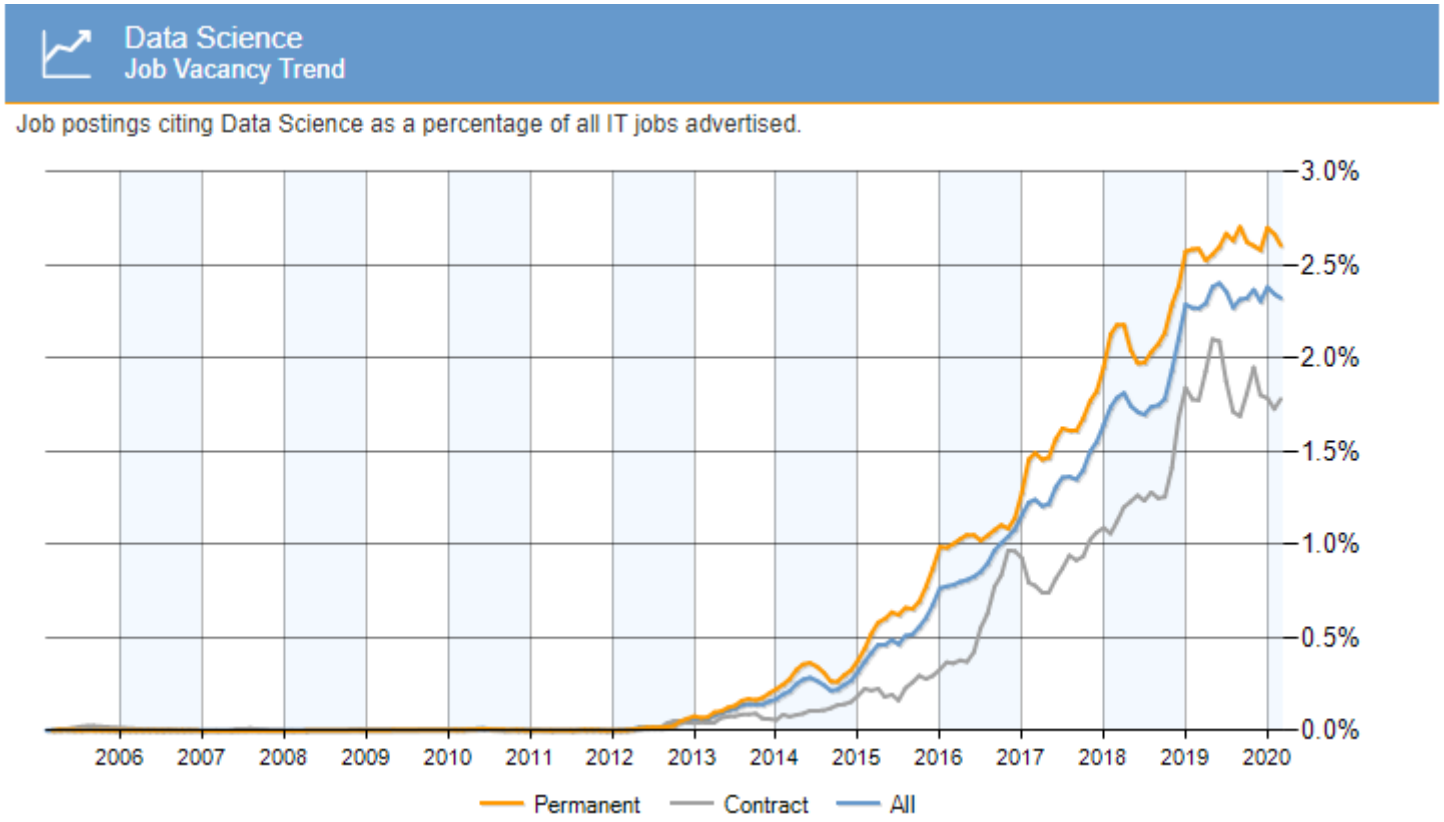
Applications in a lot of areas:

Computer science (All kinds of startups)
Biology
Epidemiology (Prediction)
Medicine
Social sciences (Online experiments)
Politics (Obama)
Sports
…

# Master Degree in Data Science

The job of bid-data analyst is requested for more than 10 years in startups, pubblic agencies and many copanies 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 trienale
- Founded in 2015
- First graduates in October 2017
- Taught in english


LA STAMPA CULTURA

SEGUICI SU ACCEDI

SEZIONI Cerca...

Apple annuncia un evento il 27 marzo: arrivano nuovi iPad e iPhone SE ...

"Noi, diplomati in quattro anni con il cuore in Europa"

Respinti i primi bimbi non vaccinati

Torino, mail e telefonate per spronare i fuoricorso

Sfregando la lampada di Alatin il latino diventa un gioco online

Alla Sapienza arrivano i primi laureati in Data Science d'Italia

# Data Science @Sapienza

SAPIENZA
UNIVERSITÀ DI ROMA

**Features**

- Multidisciplinary: 4 departments involved:
  - Informatica
  - Ingegneria informatica, automatica e gestionale
  - Ingegneria dell'Informazione, elettronica e telecommunicazioni
  - 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
IBM
Immmobiliare.it
NTT DATA
Telecom
ESA/Esrin
CINECA
LAIT Lazio
5M Informatica
Engineering
EY
Teradata
Oracle
Microsoft
Splunk
Sky
EMC2
Vitrociset
Accenture
Poste Italiane
SAS
Almawave
UNICREDIT
Vodafone
KPMG
Immobiliare

# Program structure

SAPIENZA
UNIVERSITÀ DI ROMA

**4 compulsory courses**
**1st 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**
**1st 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**
**1st year**
**Give the basics of DS to everyone**

**Computer science**
Algorithms
Programming
Network theory

**Probability & Statistics**
Hypothesis testing
Prediction

# 2022 students