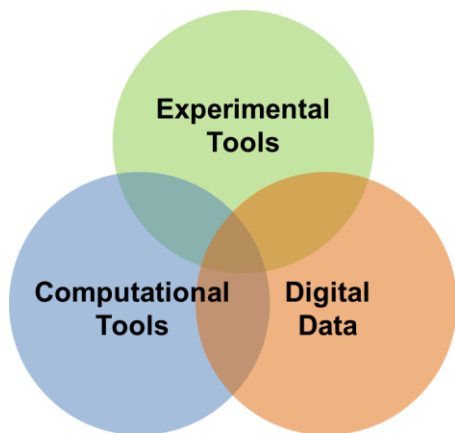
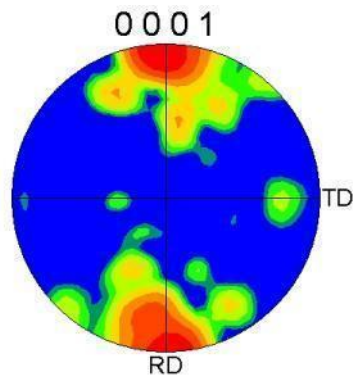


Materials Data Science

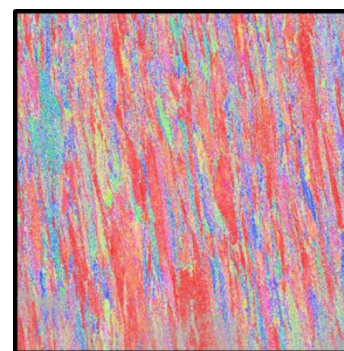
Presenter: Ayman Salem, PhD



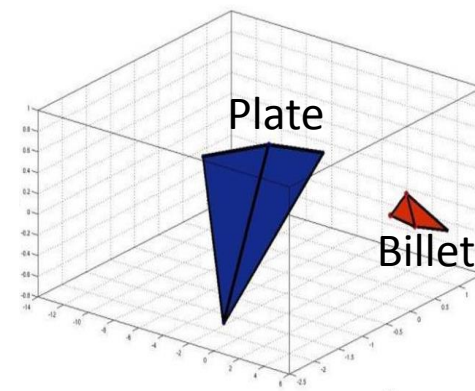
Materials Data Science



Texture Analysis



Large area EBSD



Data Mining

What is Data Science?

Data science [1,2] **incorporates** techniques and theories from many **fields**, including math, statistics, data engineering, pattern recognition and learning, advanced computing, visualization, uncertainty modeling, data warehousing, and high performance computing with the **goal** of extracting meaning from data and **creating data products**.

Data scientists are able to work with various elements of mathematics, statistics, computer science within a **specific domain**. There is probably no living person who is an expert in all of these disciplines - if so they would be extremely rare. This means that data science must be practiced as a **team**.

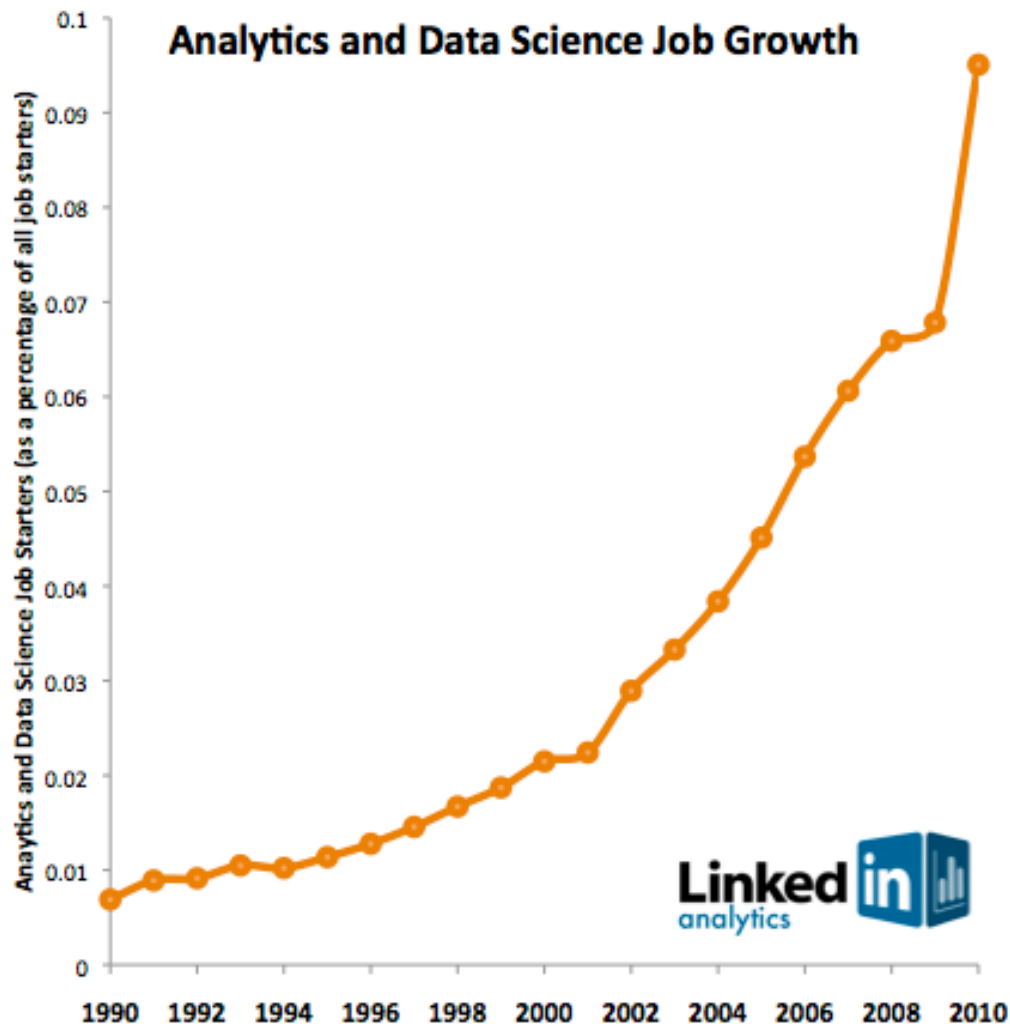
*“A new role is fast gaining prominence in organizations: that of the **Data Scientist**. Data scientists are the people who understand **how to fish out answers** to important business questions from today’s **tsunami** of unstructured **information**.”[3]*

1. Cleveland, W. S. International Statistical Review / Revue Internationale de Statistique, 21-26 (2001).

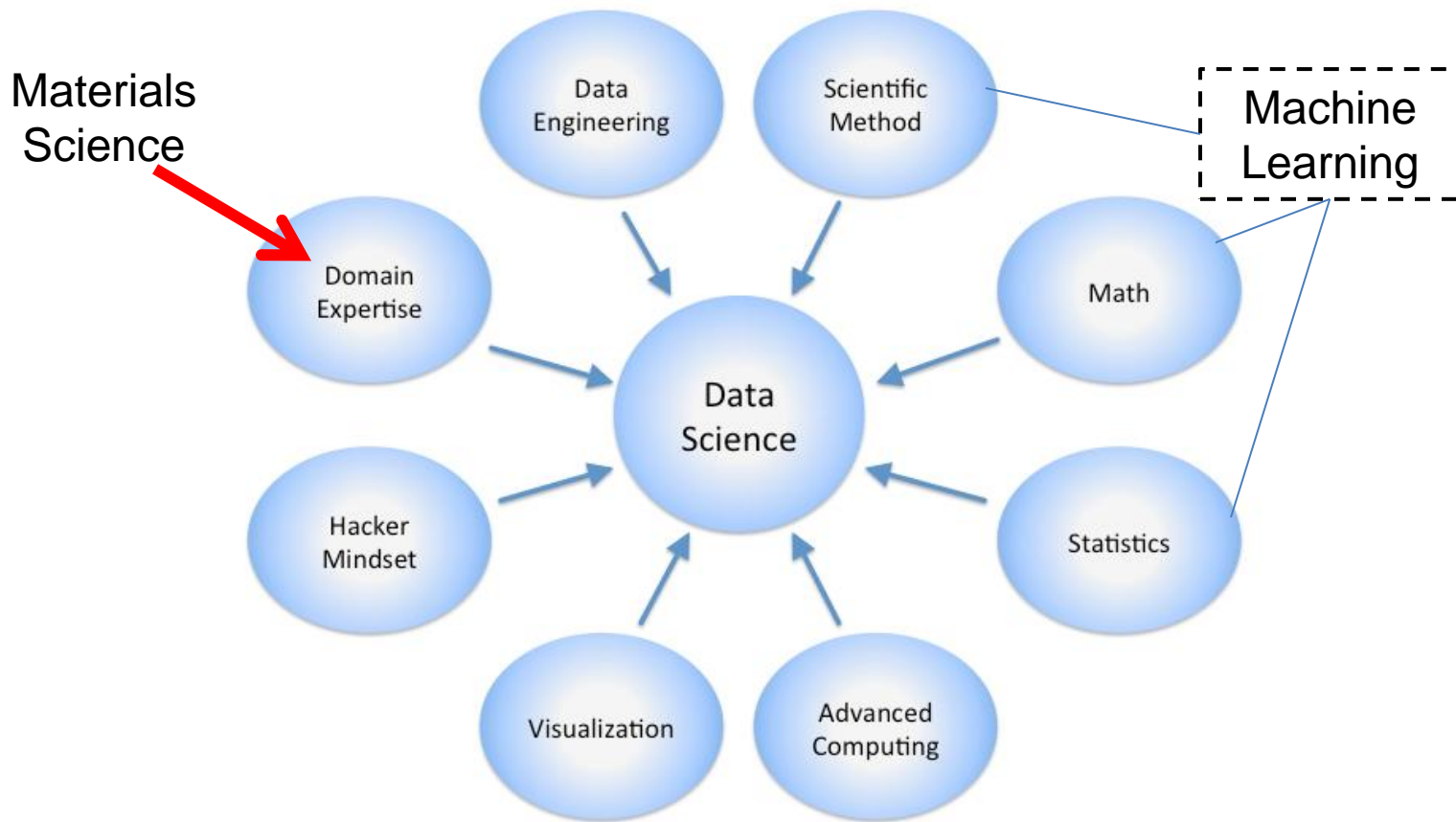
2. http://en.wikipedia.org/wiki/Data_science

The demand for data scientists

McKinsey Global Institute estimates⁶ that the “United States alone faces a shortage of 140,000 to 190,000 people [by 2018] with deep analytical skills”



McKinsey Global Institute, *Big data: The next frontier for innovation, competition, and productivity* (2011).



- Whereas a traditional data analyst may look only at data from a single source – a **Materials data-scientist** will most likely explore and examine data from **multiple** disparate **sources**.
- The Materials data-scientist will sift through all incoming data with the goal of discovering a previously **hidden insight**, which in turn can provide a competitive advantage or address a pressing problem.
- A Materials data-scientist **does not** simply collect and report on data, but also looks at it from many angles, determines what it means, then **recommends** ways to apply the materials data.
- Fashion their own tools and conduct academic-style research
- A hybrid of data hacker, analyst, communicator, and trusted adviser.

