Overview

Slava Mikhaylov
Professor of Public Policy and Data Science

Institute for Analytics and Data Science, Department of Computer Science
Department of Government
University of Essex
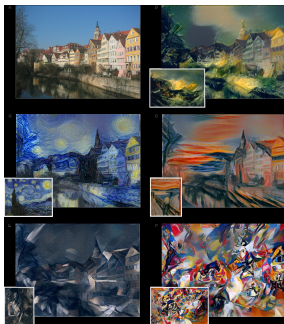
Outline

Data Science

AlphaGo

Prisma

From theory to practice

Leon A. Gatys, Alexander S. Ecker, Matthias Bethge. "A Neural Algorithm of Artistic Style." arXiv:1508.06576.

Data Scientist: The Sexiest Job of the 21st Century

Meet the people who can coax treasure out of messy, unstructured data.
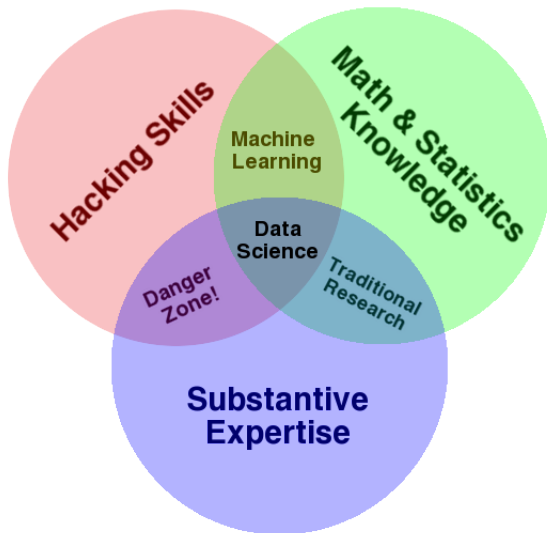by Thomas H. Davenport and D.J. Patil

"I keep saying the sexy job in the next ten years will be statisticians. People think I'm joking, but who would've guessed that computer engineers would've been the sexy job of the 1990s?" Hal Varian (Chief Economist at Google, 2009).

What is data science?

Data Science Problems

- Identify the risk factors for various diseases (e.g. cancer).
- Predict whether someone will have a heart attack on the basis of demographic, diet and clinical measurements.
- Identify the numbers in a handwritten post code.
- Use mobile phone activity for disaster management during floods.
- Identifying risk characters for Airbnb transactions.
- Predict fares at Uber.
- Asking Siri to make you a cup of coffee.

Practical Perspective

Booz Allen Hamilton "The Field Guide to Data Science", 2nd ed.

Data Science is the art of turning data into actions
It's all about the tradecraft. Tradecraft is the process, tools and
technologies for humans and computers to work together to
transform data into insights.

Data Science tradecraft creates data products
Data products provide actionable information without exposing
decision makers to the underlying data or analytics.

Practical Perspective

Booz Allen Hamilton "The Field Guide to Data Science", 2nd ed.

Data Science supports and encourages shifting between deductive (hypothesis-based) and inductive (pattern-based) reasoning

- ▶ This is a fundamental change from traditional analysis approaches.
- ▶ Inductive reasoning and exploratory data analysis provide a means to form or refine hypotheses and discover new analytic paths.
- ▶ Models of reality no longer need to be static.
- ▶ They are constantly tested, updated and improved until better models are found.

Deductive and inductive reasoning

LOOKING BACKWARD AND FORWARD

| FIRST THERE WAS BUSINESS INTELLIGENCE	NOW WE'VE ADDED DATA SCIENCE
Deductive Reasoning	Inductive and Deductive Reasoning
Backward Looking	Forward Looking
Slice and Dice Data	Interact with Data
Warehoused and Siloed Data	Distributed, Real Time Data
Analyze the Past, Guess the Future	Predict and Advise
Creates Reports	Creates Data Products
Analytic Output	Answer Questions and Create New Ones
	Actionable Answer

Practical Perspective

Booz Allen Hamilton "The Field Guide to Data Science", 2nd ed.

- ► Data Science is necessary for companies to stay with the pack and compete in the future.
- ► Data Science capabilities can be built over time.
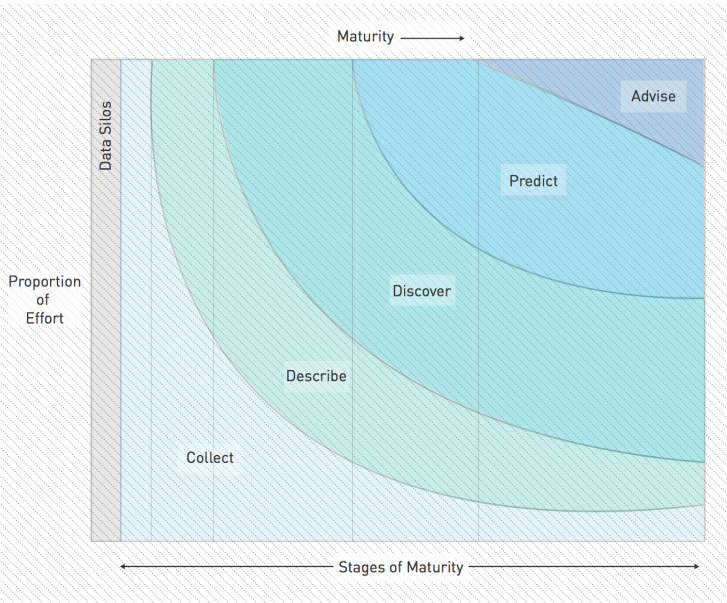- ► Data Science is a different kind of team sport.
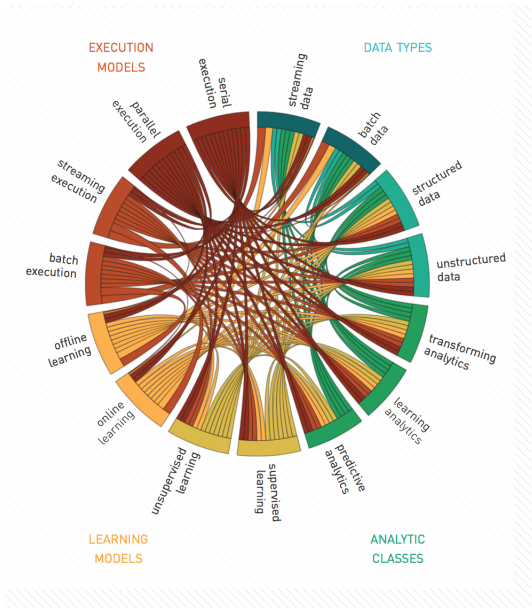
Data science work
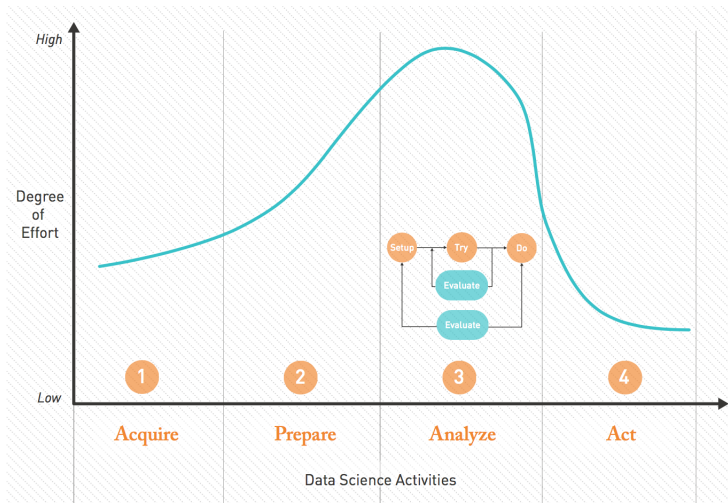
DIKUW Pyramid
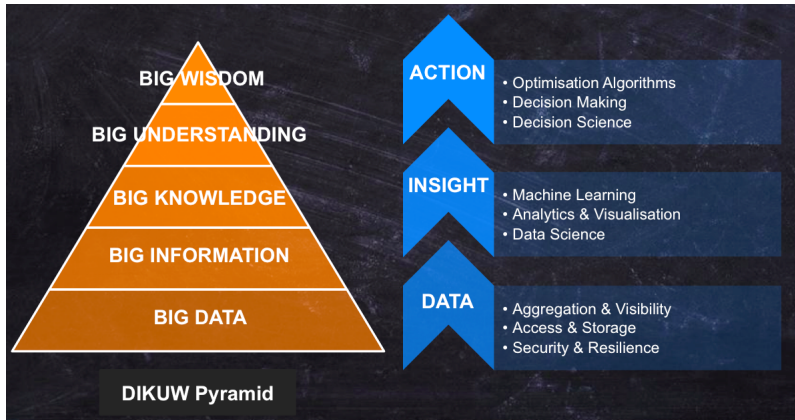
Data Science maturity

The Data Science Maturity Model

Data Science principles

- Be willing to fail.
- Fail often and learn quickly.
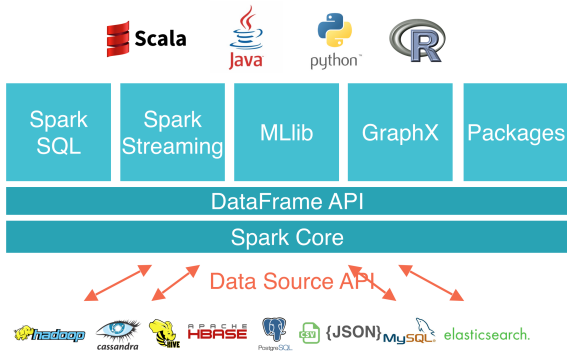- Keep the goal in mind.
- Dedication and focus lead to success.

Data Science components

Interconnection Among the Component Parts of Data Science

Source: Booz Allen Hamilton

Typical data analytic stack

Kaggle Competitions

- Higgs Boson Machine Learning Challenge.
- National Data Science Bowl – Assessing Ocean Health.
- Heritage Health Prize Competition.
- West Nile Virus Prediction.
- Predict Closed Questions on Stack Overflow.
- Flu Forecasting competition.
- Global Energy Forecasting Competition.

Moving forward

deep learning
predictive analytics — machine learning

translation
classification & clustering — natural language processing (NLP)
information extraction

speech to text
text to speech — speech

expert systems

planning, scheduling & optimization

robotics

image recognition
machine vision — vision

Artificial Intelligence (AI)