



Definitions

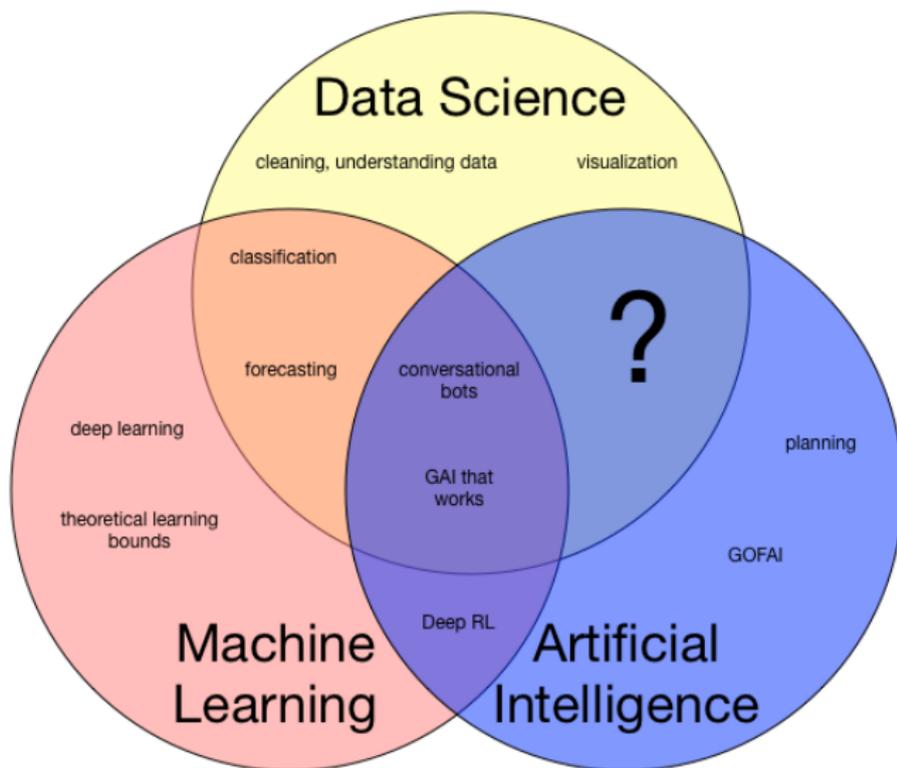
Jordan Boyd-Graber

University of Maryland

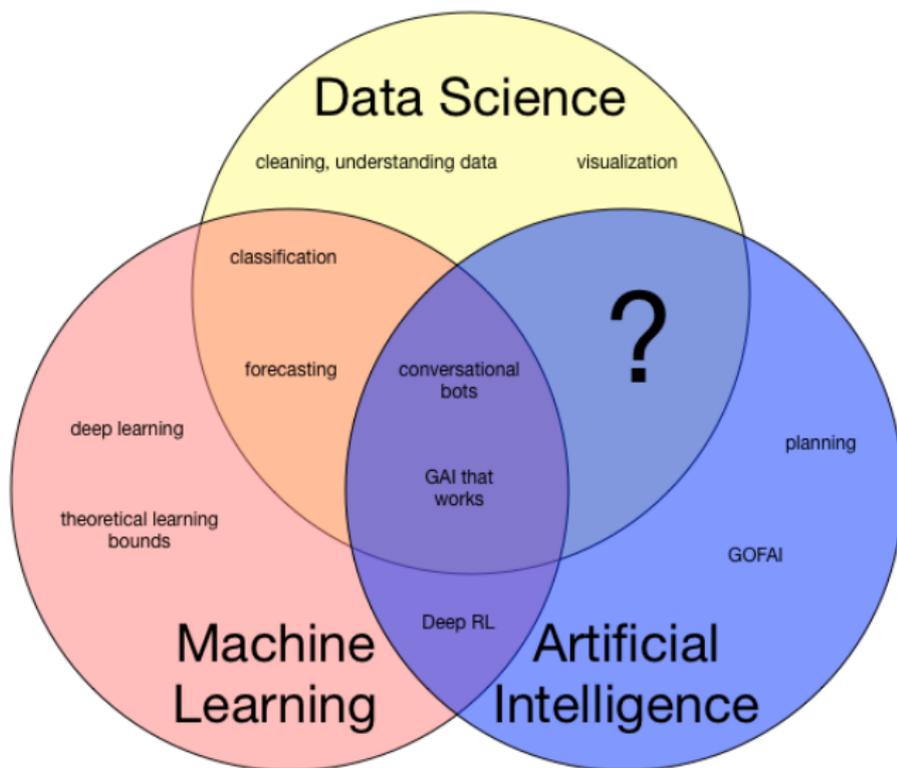
JANUARY 12, 2018

Roadmap

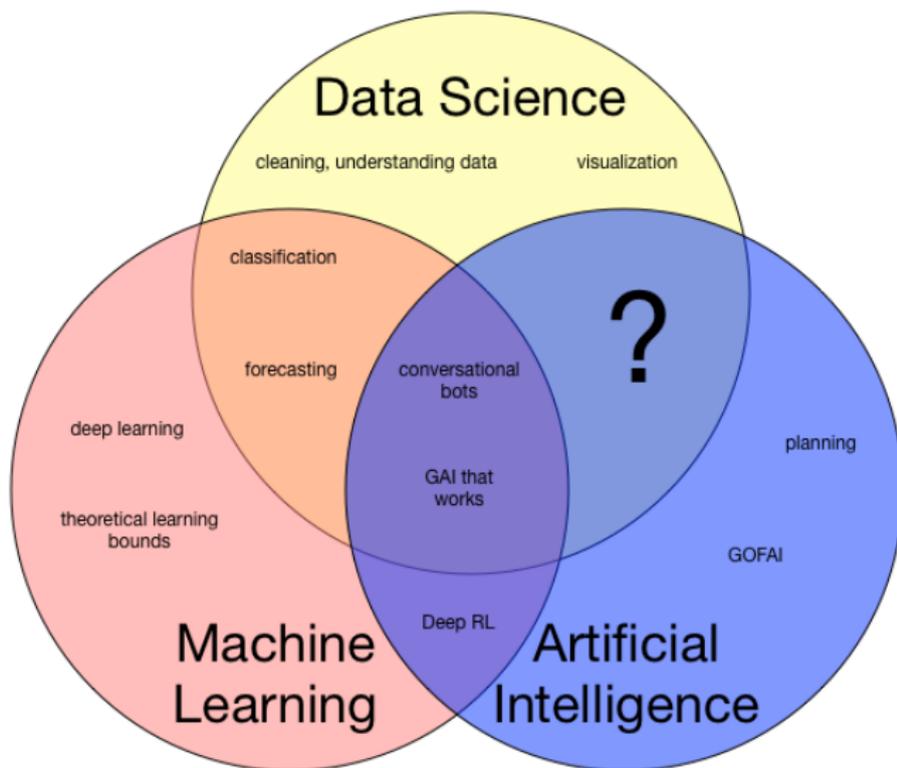
- What is data science / artificial intelligence / machine learning?
- What is possible with these new tools and techniques?
- What are the challenges?



Data provides the foundation



Machine learning provides algorithms



Artificial intelligence defines problems



Nowhere near true AI

Data are everywhere.

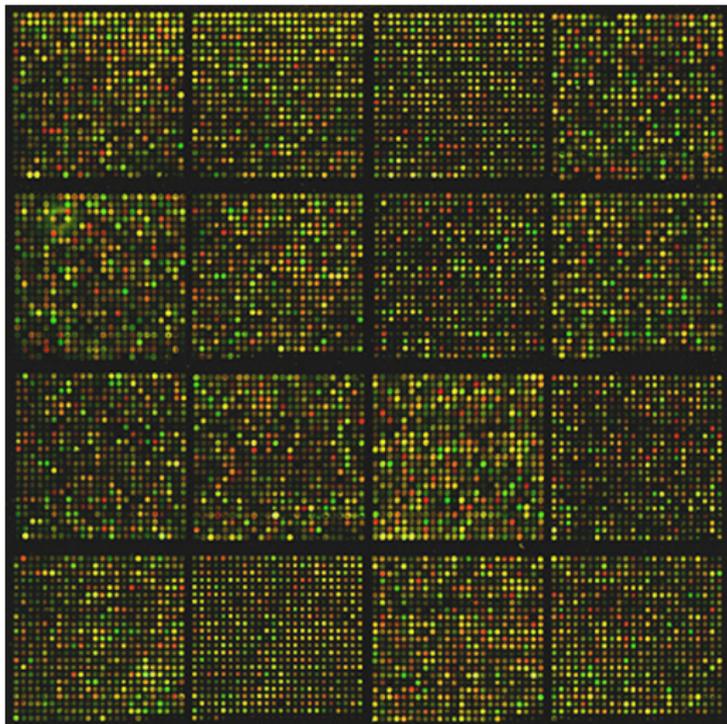
User ratings

Ikiru (1952)	UR	Foreign	
Junebug (2005)	R	Independent	
La Cage aux Folles (1979)	R	Comedy	
The Life Aquatic with Steve Zissou (2004)	R	Comedy	
Lock, Stock and Two Smoking Barrels (1998)	R	Action & Adventure	
Lost in Translation (2003)	R	Drama	
Love and Death (1975)	PG	Comedy	
The Manchurian Candidate (1962)	PG-13	Classics	
Memento (2000)	R	Thriller	
Midnight Cowboy (1969)	R	Classics	

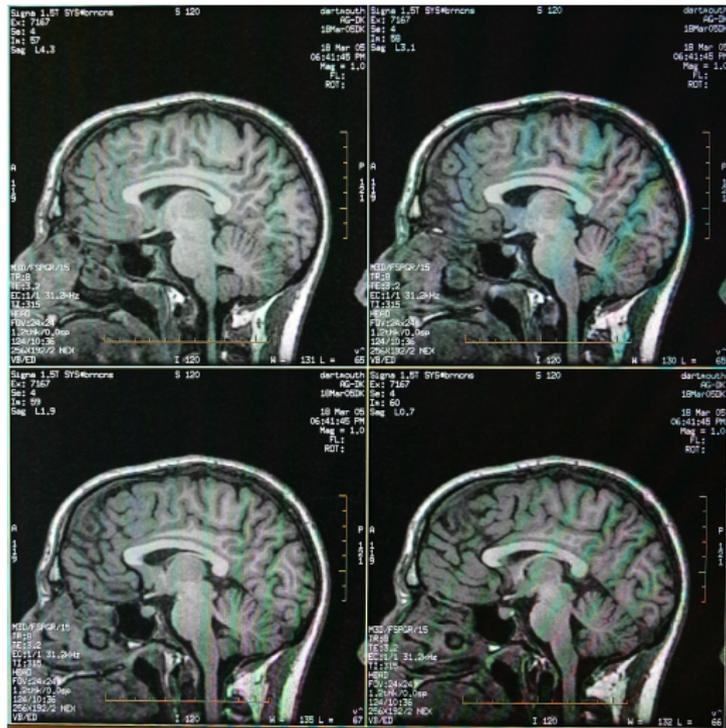
Purchase histories

0.5/0.51 lb	Cheese Cabot Vermont Cheddar	0.51 lb	\$7.99/lb	\$4.07
	Dairy			
1/1	Friendship Lowfat Cottage Cheese (16oz)		\$2.89/ea	\$2.89
1/1	Nature's Yoke Grade A Jumbo Brown Eggs (1 dozen)		\$1.49/ea	\$1.49
1/1	Santa Barbara Hot Salsa, Fresh (16oz)		\$2.69/ea	\$2.69
1/1	Stonyfield Farm Organic Lowfat Plain Yogurt (32oz)		\$3.59/ea	\$3.59
	Fruit			
3/3	Anjou Pears (Farm Fresh, Med)	1.76 lb	\$2.49/lb	\$4.38
2/2	Cantaloupe (Farm Fresh, Med)		\$2.00/ea	\$4.00 S
	Grocery			
1/1	Fantastic World Foods Organic Whole Wheat Couscous (12oz)		\$1.99/ea	\$1.99
1/1	Garden of Eatin' Blue Corn Chips (9oz)		\$2.49/ea	\$2.49
1/1	Goya Low Sodium Chickpeas (15.5oz)		\$0.89/ea	\$0.89
2/2	Marcal 2-Ply Paper Towels, 90ct (1ea)		\$1.09/ea	\$2.18 T
1/1	Muir Glen Organic Tomato Paste (6oz)		\$0.99/ea	\$0.99
1/1	Starkist Solid White Albacore Tuna in Spring Water (6oz)		\$1.89/ea	\$1.89

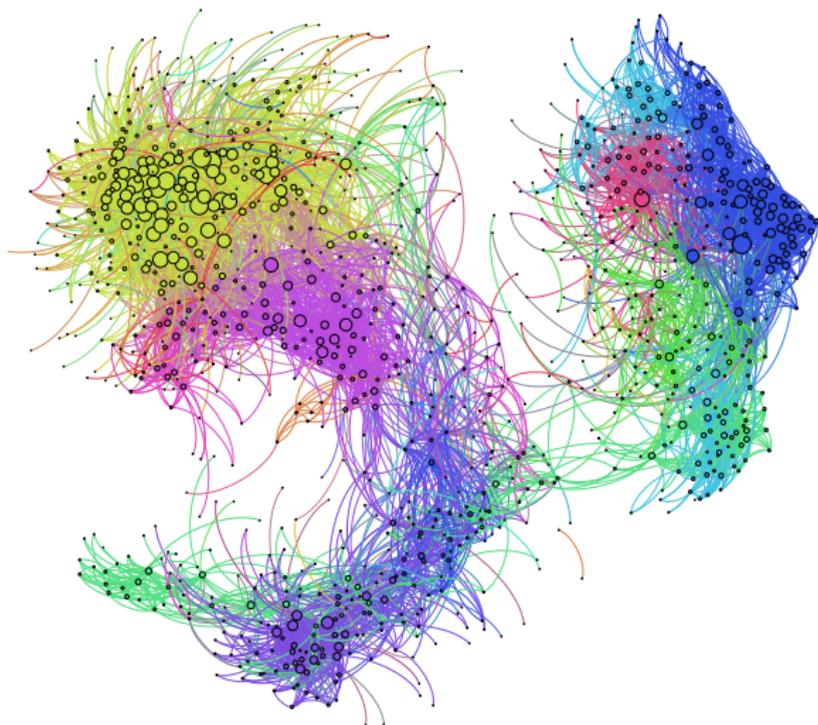
Genomics



Neuroscience



Social networks



Finance

Dow Jones Composite Average (^DJA) - DJI

[+ Add to Portfolio](#)

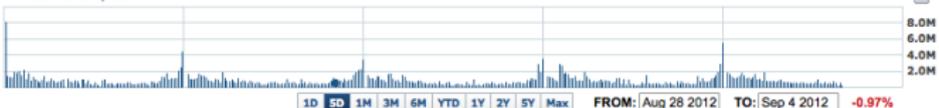
4,372.78 **+20.68 (0.47%)** 1:48PM EDT

Enter name(s) or symbol(s) [GET CHART](#) [COMPARE](#) [EVENTS](#) [TECHNICAL INDICATORS](#) [CHART SETTINGS](#) [RESET](#)

Aug 29, 2012 2:59 PM - 3:04 PM EDT: ^DJA 4416.03



Volume: 571,700

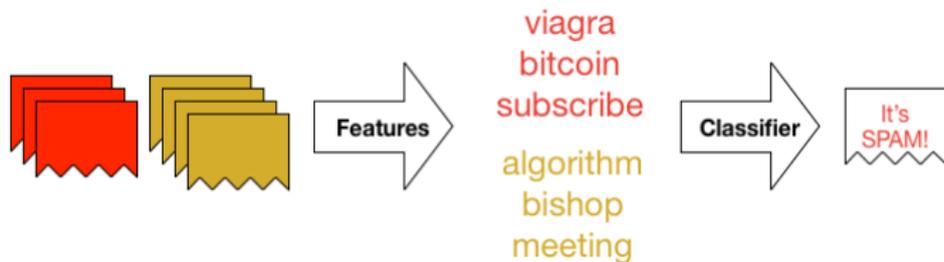


1D SD 1M 3M 6M YTD 1Y 2Y 5Y Max FROM: Aug 26 2012 TO: Sep 4 2012 -0.97%

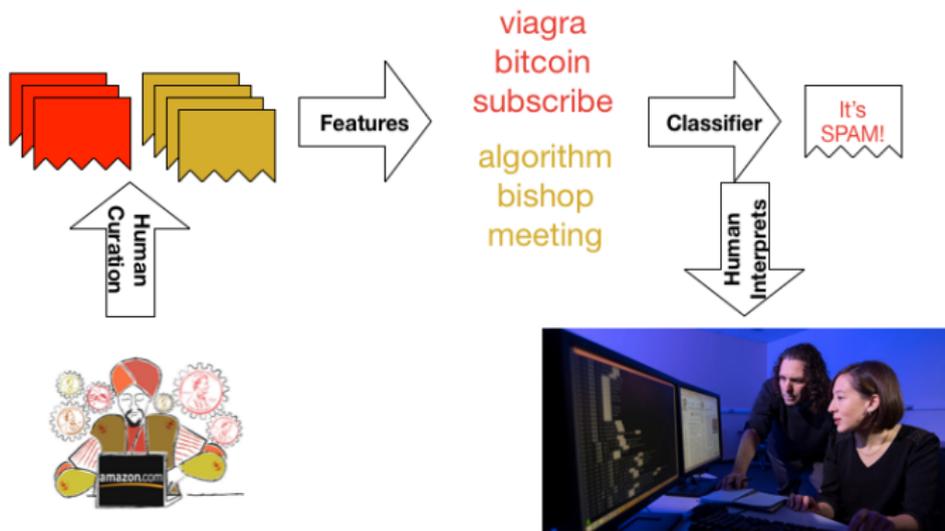


Data are input to algorithms

We can recreate human decisions



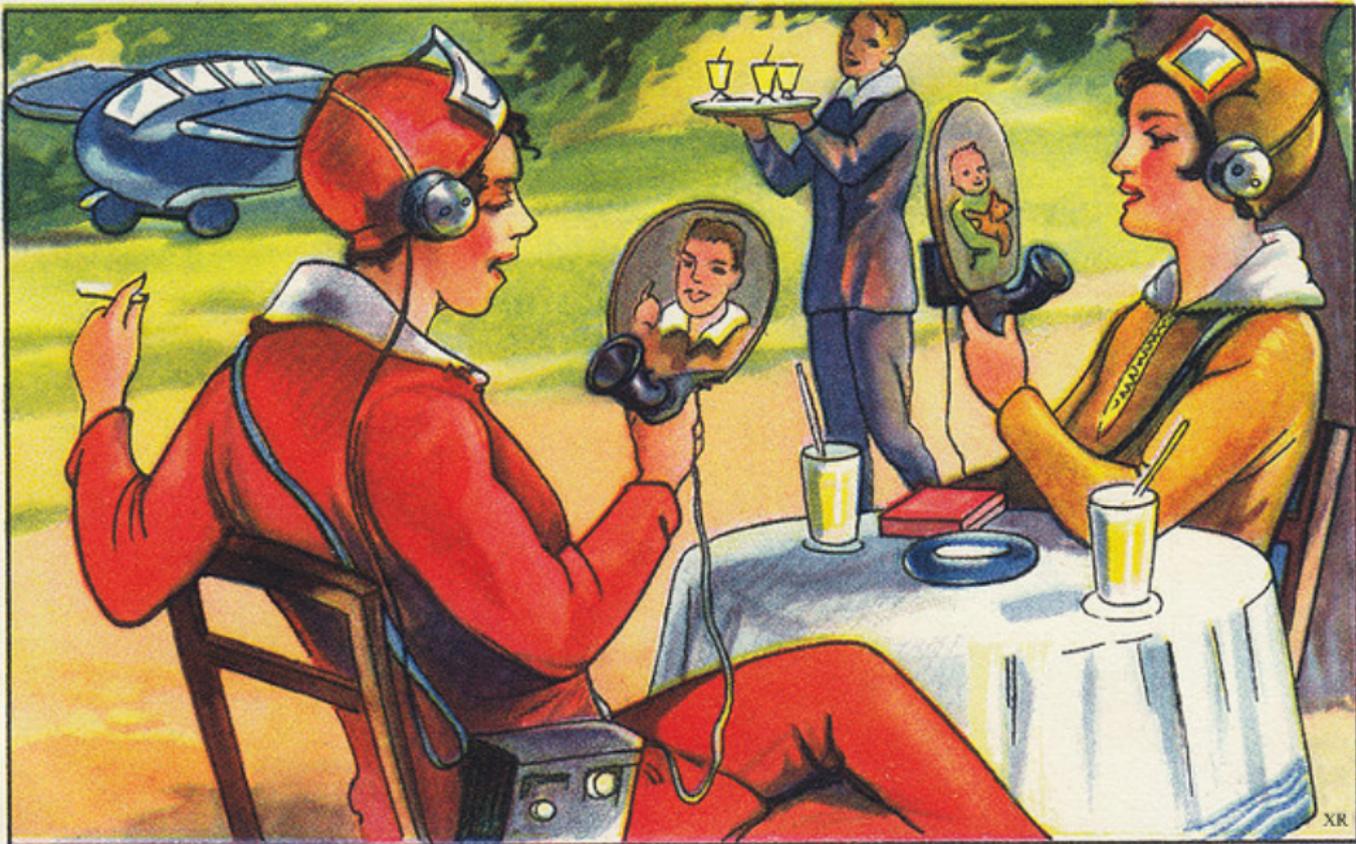
We can recreate human decisions



Algorithms can solve tasks



Waymo driverless car on the road in Silicon Valley



Predictions: Technology changes but society doesn't



Are we prepared for social changes?

Discussions of AI Need Realism

- You can't solve a task without good data
- Good data often requires humans

Discussions of AI Need Realism

- You can't solve a task without good data
- Good data often requires humans
- Unless task is super-straightforward, also need ML expertise

Discussions of AI Need Realism

- You can't solve a task without good data
- Good data often requires humans
- Unless task is super-straightforward, also need ML expertise
- Effective solutions often require infrastructure

Discussions of AI Need Realism

- You can't solve a task without good data
- Good data often requires humans
- Unless task is super-straightforward, also need ML expertise
- Effective solutions often require infrastructure
 - Explanations
 - Retraining
 - Socio-technical interfaces

Discussions of AI Need Realism

- You can't solve a task without good data
- Good data often requires humans
- Unless task is super-straightforward, also need ML expertise
- Effective solutions often require infrastructure
 - Explanations
 - Retraining
 - Socio-technical interfaces
- DS/ML/AI is **hard** (but rewarding!)

