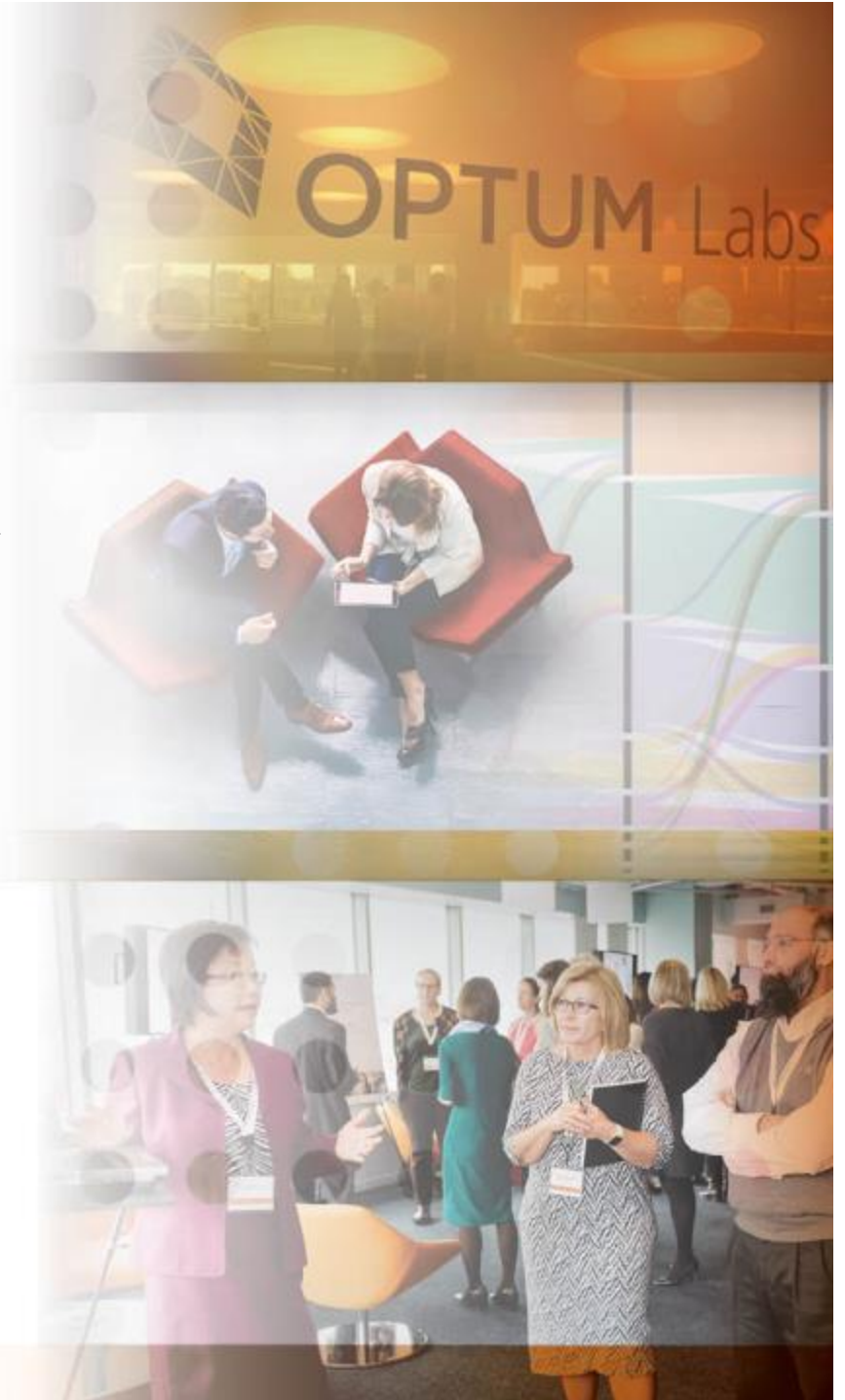


A Sampling of “Big” Data in Health Care

SOMA, January 2019

Paul Bleicher, MD PhD
CEO, OptumLabs
Cambridge, MA



Topics

- What is “big data?”
- A bit of big data at OptumLabs
- Prediction, machine learning, deep learning and artificial intelligence
- Chatbots, Internet of Things, Blockchain, Individual Health Records
- Limitations and risk to big data and AI

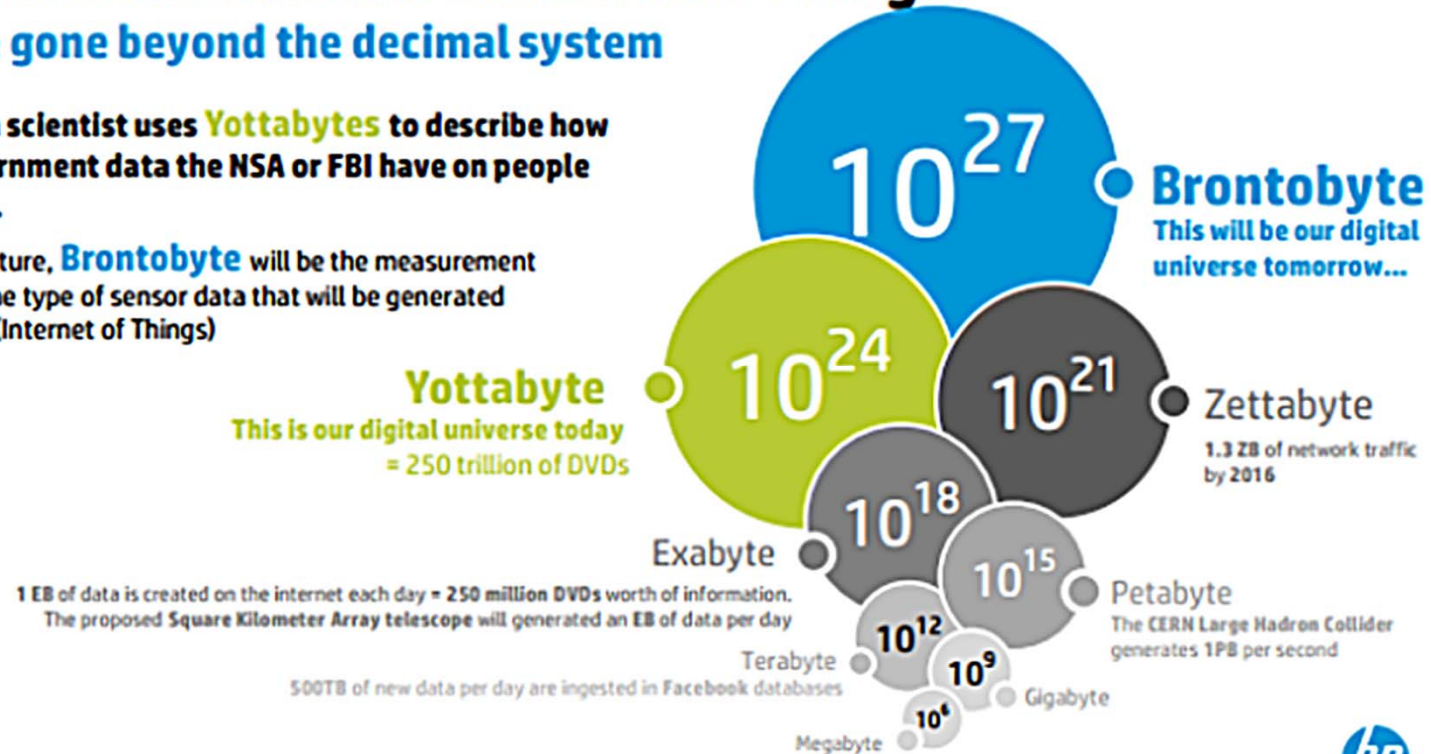
The Scale of Big Data

Information from the Internet of Things:

We have gone beyond the decimal system

Today data scientist uses **Yottabytes** to describe how much government data the NSA or FBI have on people altogether.

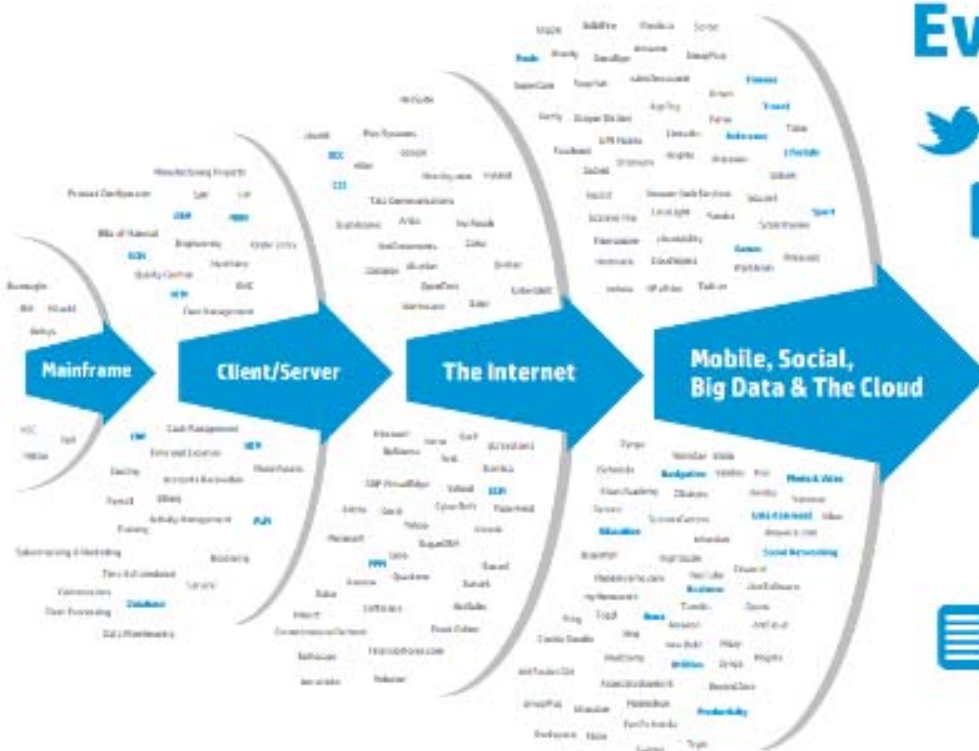
In the near future, **Brontobyte** will be the measurement to describe the type of sensor data that will be generated from the IoT (Internet of Things)










7 © Copyright 2013 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice.



How Do You Get Big Data?



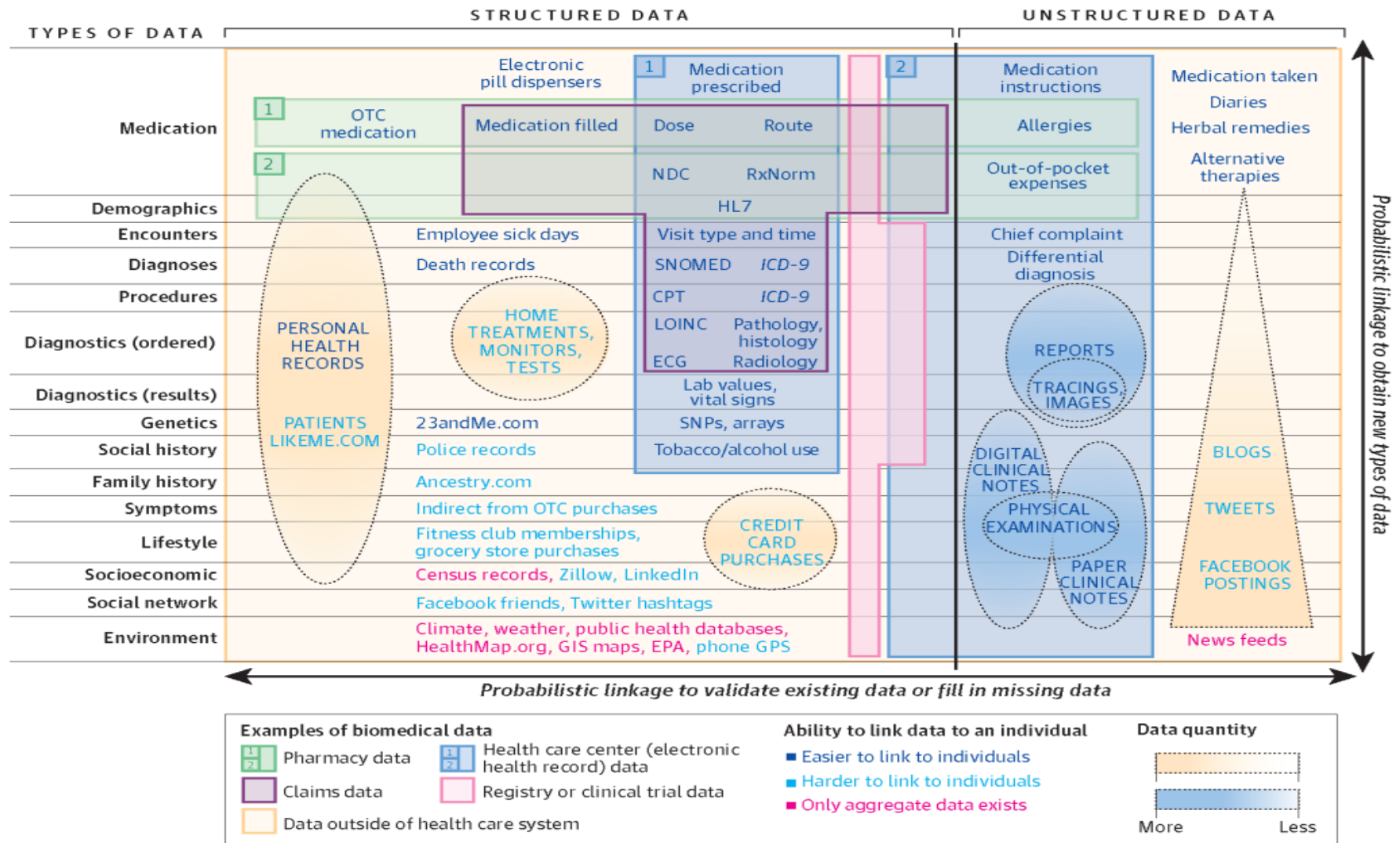
Every 60 seconds

-  **98,000+** tweets
-  **695,000** status updates
-  **11 million** instant messages
-  **698,445** Google searches
-  **168 million+** emails sent
-  **1,820TB** of data created
-  **217** new mobile web users

© Copyright 2011 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice.



How Does Health Care Data Get to Be Big Data?



Probabilistic linkage to obtain new types of data

OptumLabs: A Health Care Data and Innovation Collaborative

We accelerate research, innovation and translation by giving our partners access to the largest U.S. linked patient database, world-class thought leaders and the power of multi-partner collaboration.



DATA & EXPERTISE

Deep knowledge and experience working with industry-leading linked data asset



CONVENING HEALTH CARE & TECHNOLOGY LEADERS

Diverse relationships with health care and technology industry leaders across all sectors



DATA SCIENCE

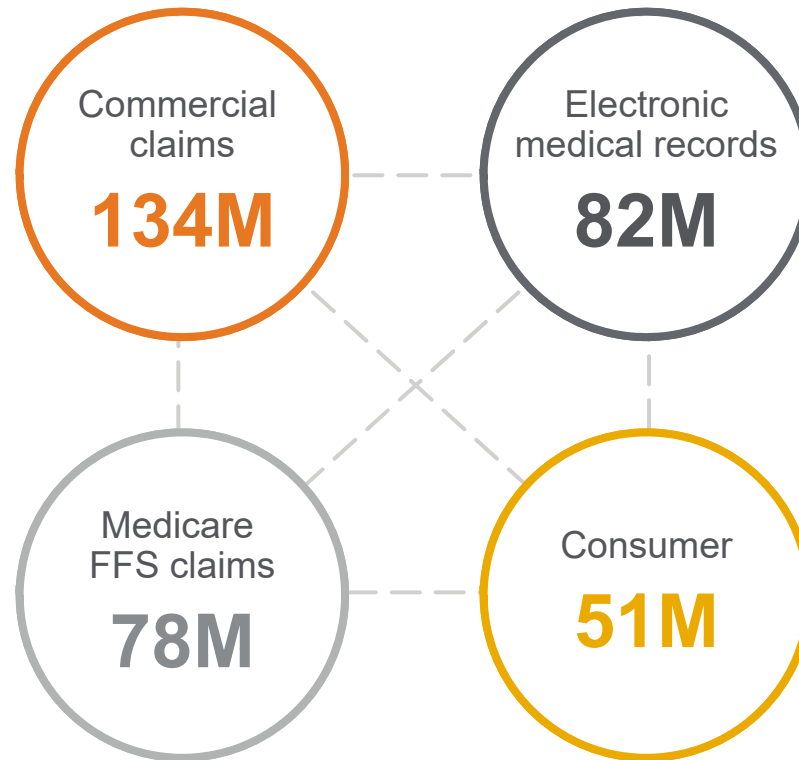
Best-in-class data analytics know-how and data visualization



HEALTH CARE KNOWLEDGE

A collaborative with broad and deep health care knowledge to solve health care's greatest challenges.

Health Care “Big” Data: The OptumLabs Data Warehouse (OLDW)



More than 260 million de-identified lives over 25 years

Health Care Big Data in 48TB

Amazon's Choice



WD 8TB Elements Desktop Hard Drive

by Western Digital

\$149⁹⁹ ~~\$179.99~~ ✓prime

FREE Delivery by Sat, Jan 19

More Buying Choices

\$139.44 (23 used & new offers)

$\times 6 \approx \$900$

What Does OptumLabs Do With “Big” Data – Research...

Discovery by the numbers

Since the inception of OptumLabs, there have been:

190+ RESEARCH PROJECTS INITIATED

220+ CONFERENCE PRESENTATIONS

120+ PUBLICATIONS

20+ RESEARCH GRANTS AWARDED
42 additional grants in process

Annals
of Internal Medicine®

HealthAffairs

JAMA
The Journal of the American Medical Association

the**bmj**

 **JACC**
JOURNAL OF THE AMERICAN COLLEGE OF CARDIOLOGY

Journal of Bone and Mineral Research
JBMR

 **CHEST**
AMERICAN COLLEGE
of CHEST PHYSICIANS

Value
IN HEALTH
The Journal of The International Society for
Pharmacoeconomics and Outcomes Research

JAHA

Journal of the
American Heart
Association

...Translation: Opioid Key Performance Indicators (KPIs)

Claims data-driven, comprehensive metrics

OptumLabs convened experts to develop **29 KPI metrics** that look at the many facets of the opioid crisis.

KPIs were developed in **four domains** that address the opioid epidemic:



Prevention



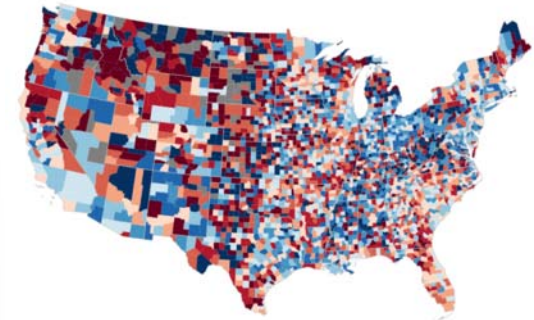
Pain management



Opioid use disorder (OUD) treatment



Maternal and child health



- KPIs enable coordinated UHG benchmarking, target setting, and performance evaluation.
- OptumLabs county-level data and visualizations reveal striking geographic variation in the U.S.
- KPIs support the Optum/UHG Opioid Taskforce to drive an enterprise-wide response to the solve the opioid crisis.
- Metrics have been shared via *Health Affairs* in support of public health.

JAMA Otolaryngology-Head & Neck Surgery | Original Investigation

Incident Hearing Loss and Comorbidity A Longitudinal Administrative Claims Study

Jennifer A. Deal, PhD; Nicholas S. Reed, AuD; Alexander D. Kravetz, PharmD; Heather Weinreich, MD, MPH;
Charlotte Yeh, MD; Frank R. Lin, MD, PhD; Aylin Altan, PhD

JAMA Otolaryngology-Head & Neck Surgery | Original Investigation

Trends in Health Care Costs and Utilization Associated With Untreated Hearing Loss Over 10 Years

Nicholas S. Reed, AuD; Aylin Altan, PhD; Jennifer A. Deal, PhD; Charlotte Yeh, MD; Alexander D. Kravetz, PharmD;
Margaret Wallhagen, RN, PhD; Frank R. Lin, MD, PhD

IMPORTANCE Nearly 38 million individuals in the United States have untreated hearing loss, which is associated with cognitive and functional decline. National initiatives to address hearing loss are currently under way.

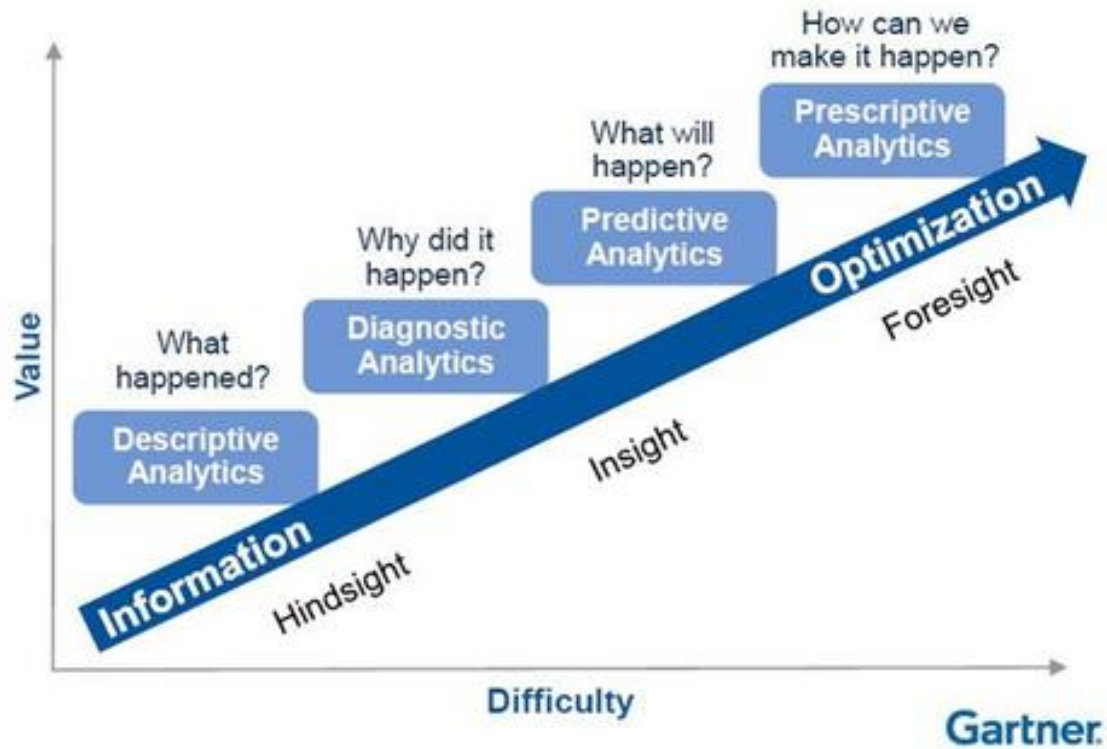
OBJECTIVE To determine whether untreated hearing loss is associated with increased health care cost and utilization on the basis of data from a claims database.

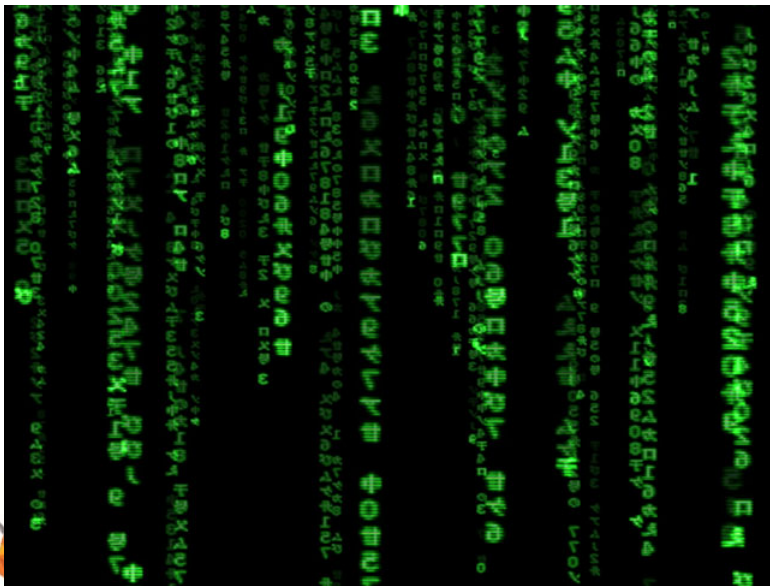
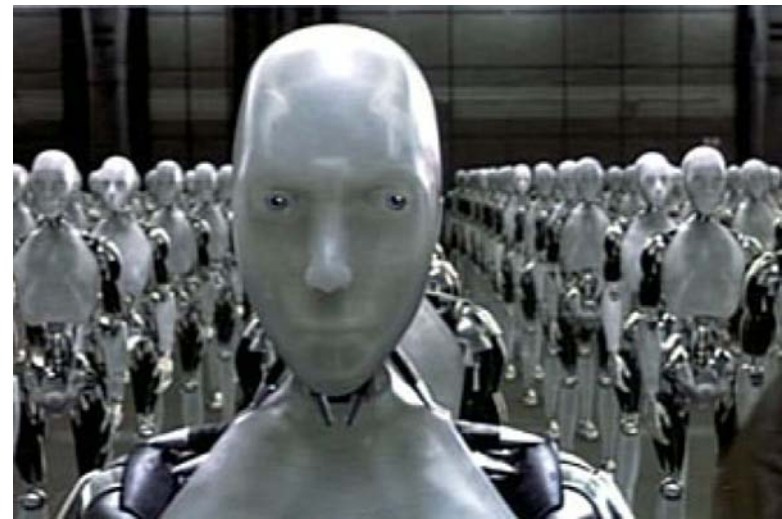
DESIGN, SETTING, PARTICIPANTS Retrospective, propensity-matched cohort study of persons with and without untreated hearing loss based on claims for health services rendered between January 1, 1999, and December 31, 2016, from a large health insurance database. There were 154 414, 44 852, and 4728 participants at the 2-, 5-, and 10-year follow-up periods, respectively. The study was conceptualized and data were analyzed between September 2016 and November 2017.

+ Invited Commentary
+ Author Audio Interview

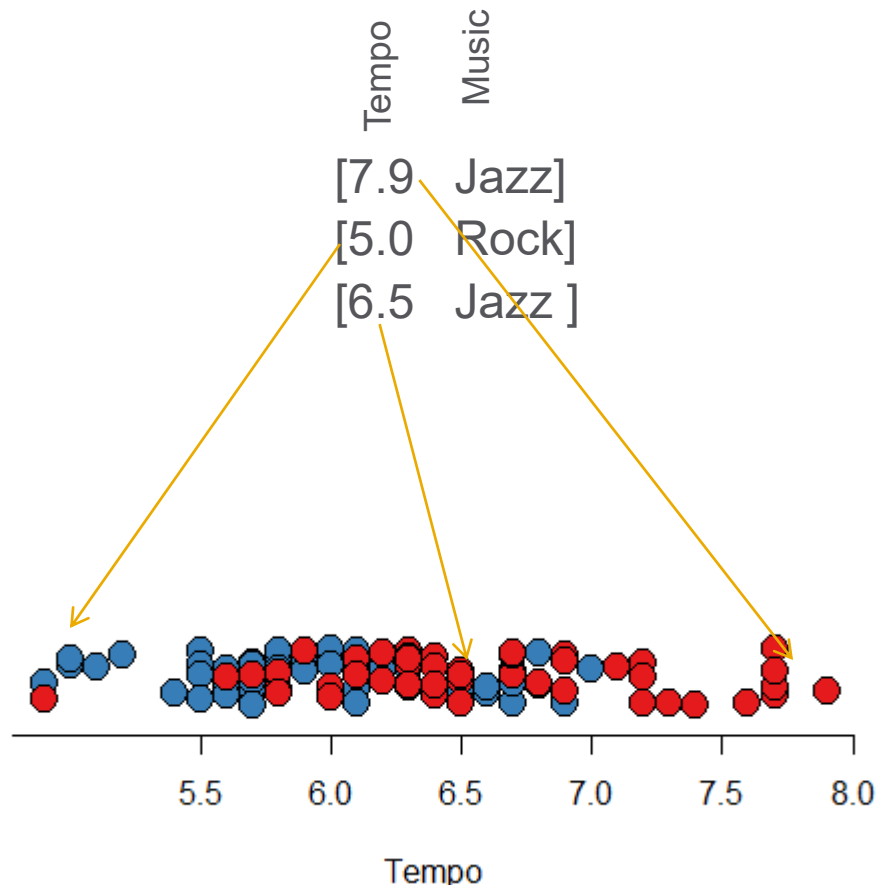
...Analysis for use
in Public Policy

Trends in Analytics in Health Care



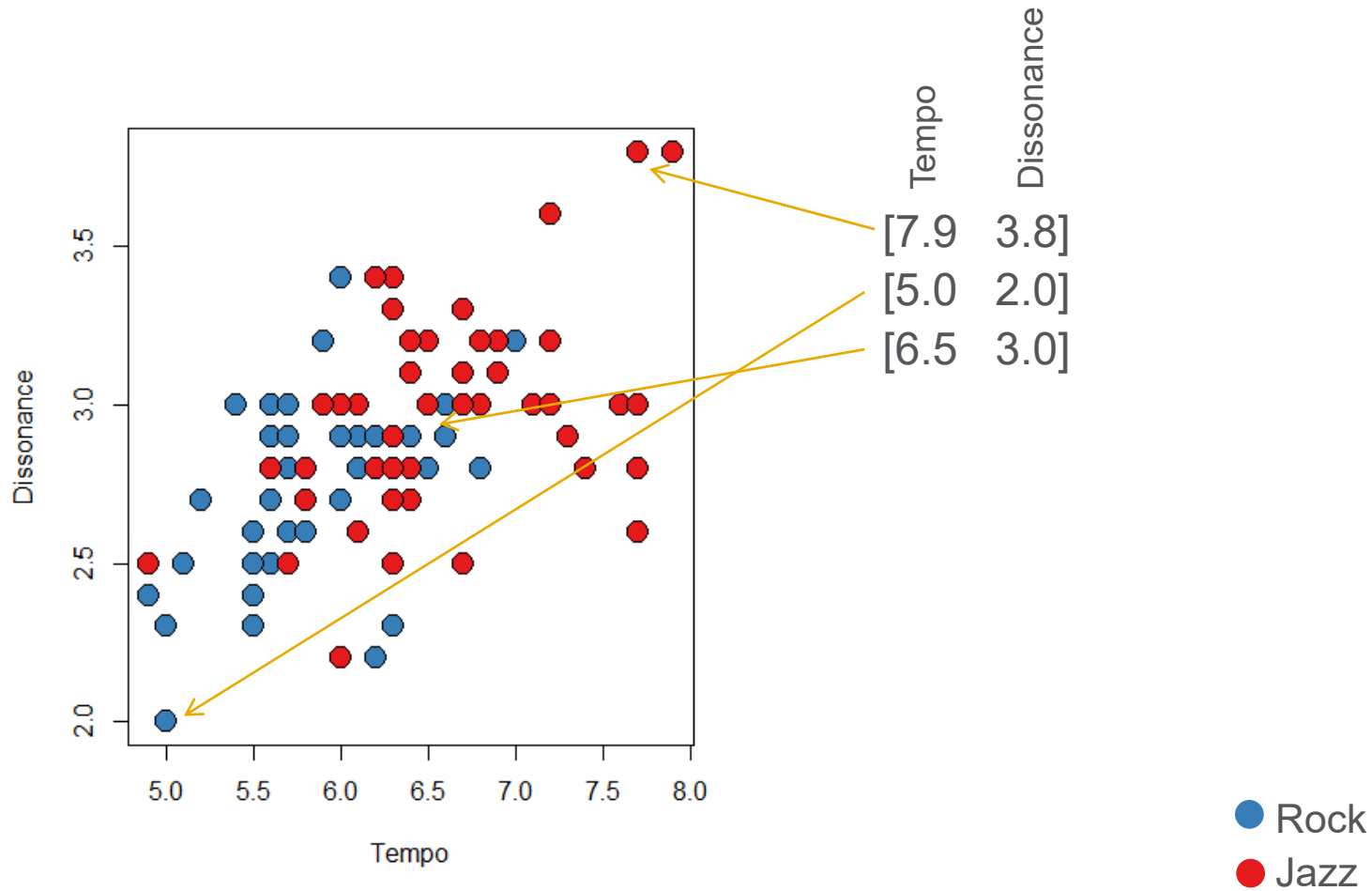


Some basic intuition on predictive models: 1D

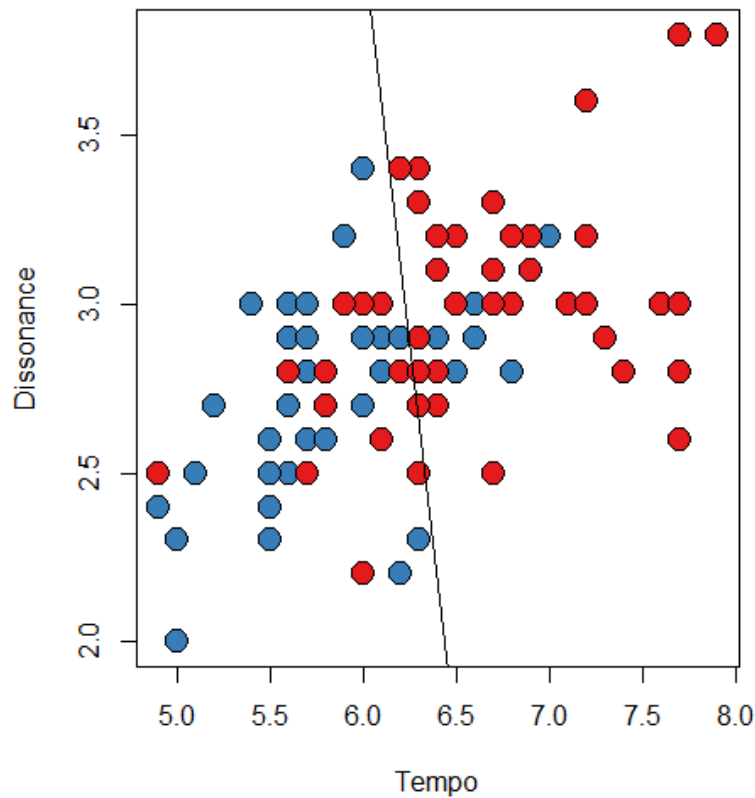


● Rock
● Jazz

Some basic intuition on predictive models:2D



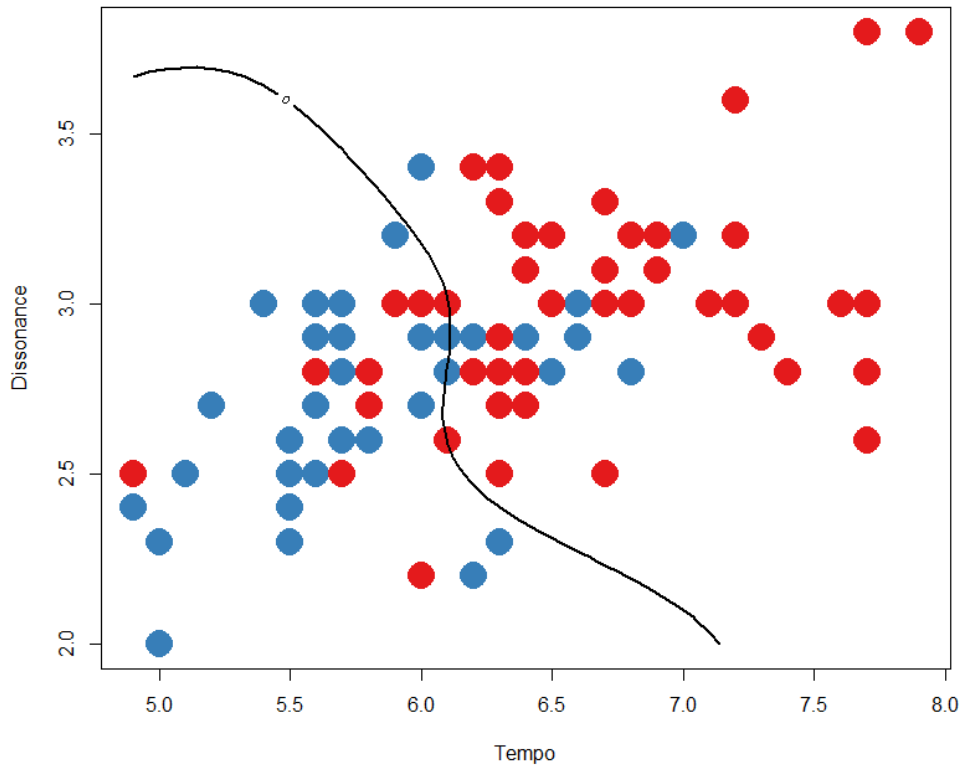
Some basic intuition on predictive models:2D



Tempo	Dissonance	<u>Prediction from Model</u>
[7.9	3.8]	[3.5 Jazz]
[5.0	2.0]	[-2.7 Rock]
[6.5	3.0]	[0.5 ?Rock?Jazz]

● Rock
● Jazz

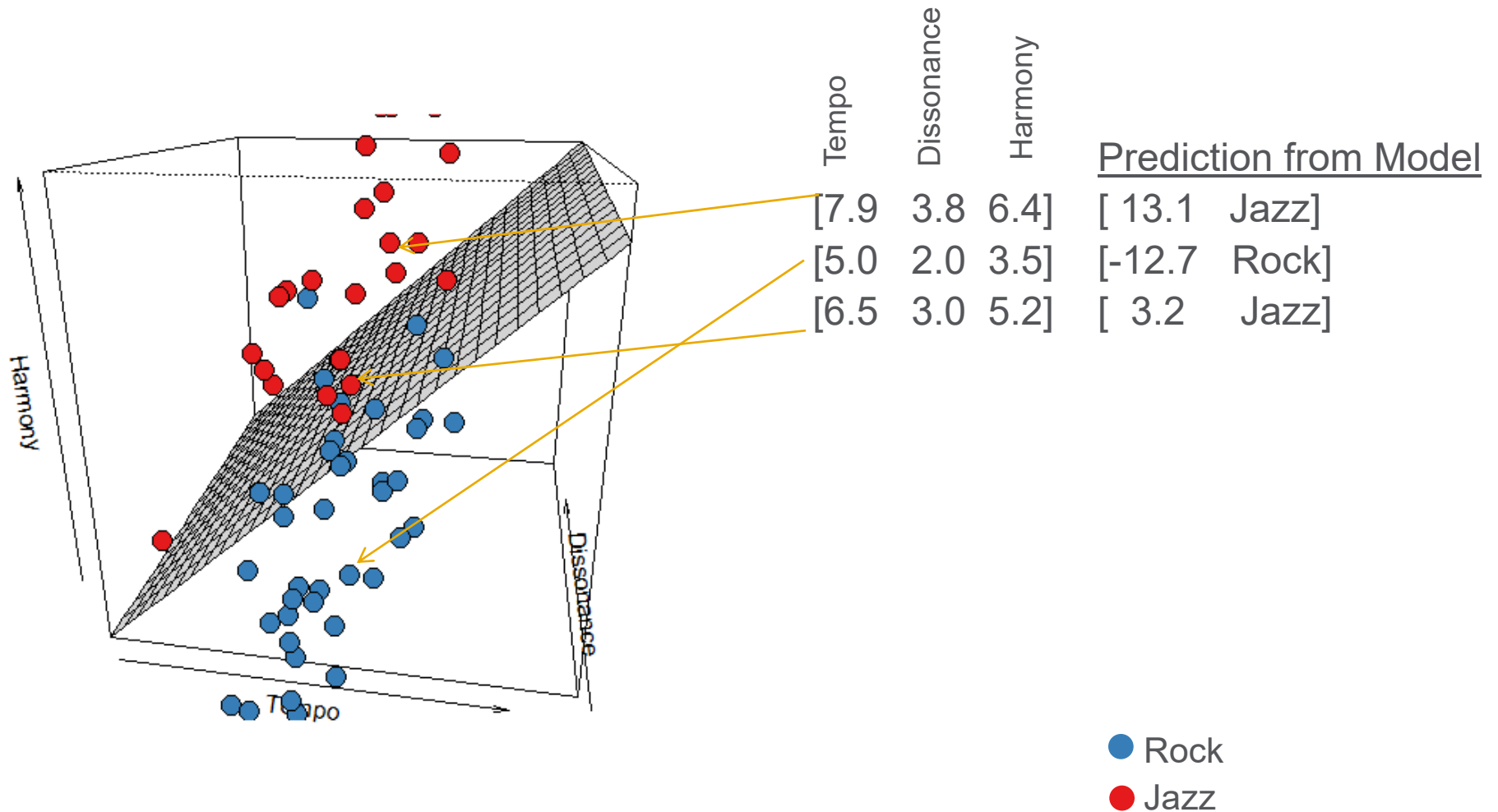
Some basic intuition on predictive models:2D



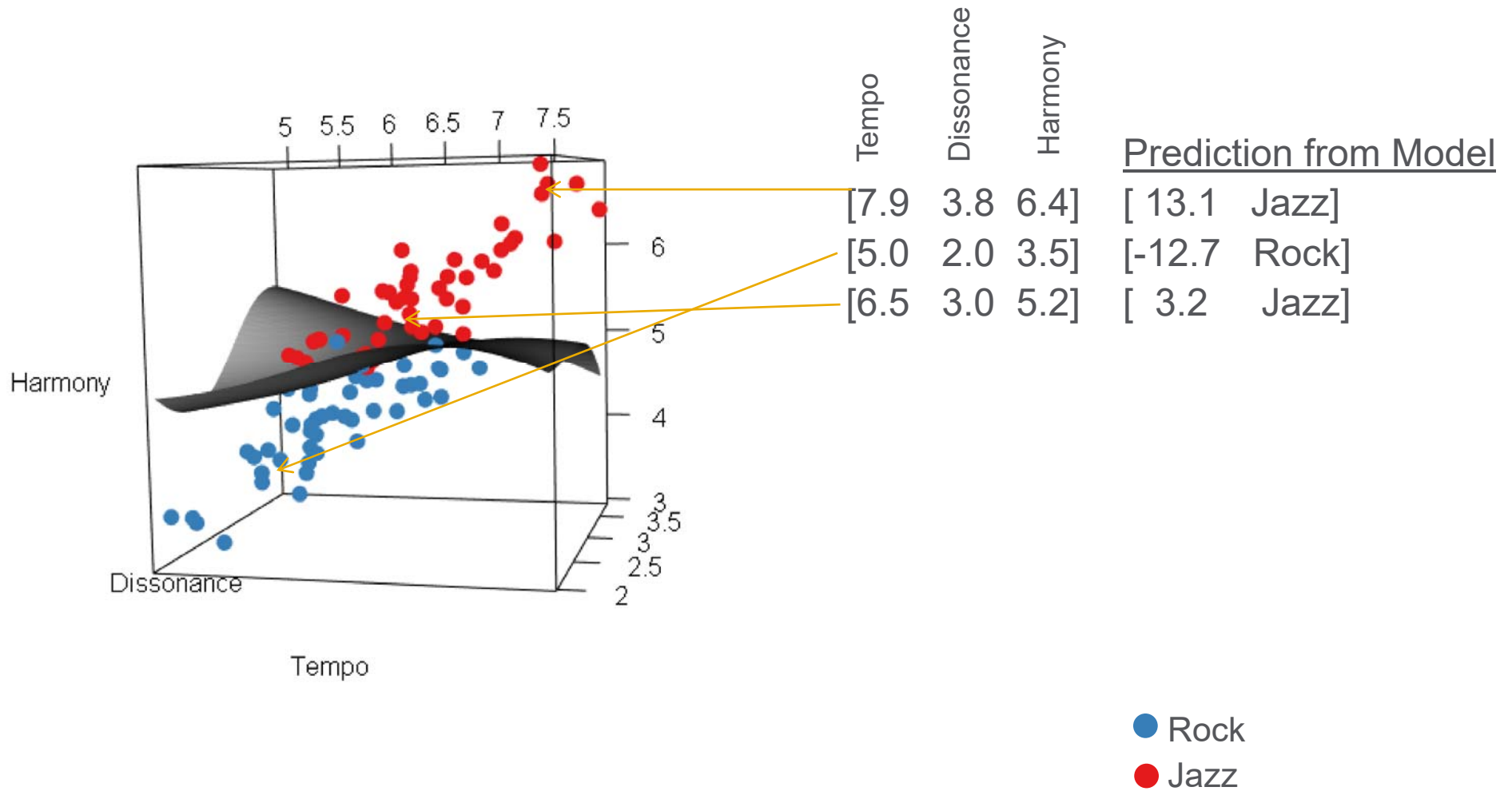
Tempo	Dissonance	<u>Prediction from Model</u>
[7.9	3.8]	[3.5 Jazz]
[5.0	2.0]	[-2.7 Rock]
[6.5	3.0]	[0.5 ?Rock?Jazz]

● Rock
● Jazz

Some basic intuition on predictive models:3D



Some basic intuition on predictive models:3D



Some basic intuition on predictive models



Tempo	Dissonance	Harmony	Syncopation
[7.9	3.8	6.4	6.2]
[5.0	2.0	3.5	2.1]
[6.5	3.0	5.2	5.5]

4 Dimensions is beyond human visualization, but the math stays the same.

N dimensions is no different.

Demystifying AI/ML – Deep Learning

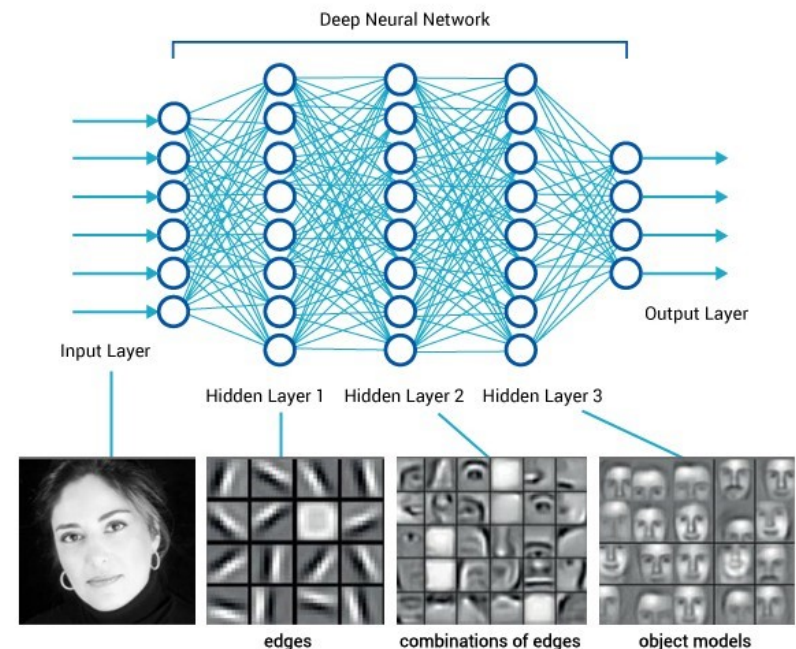
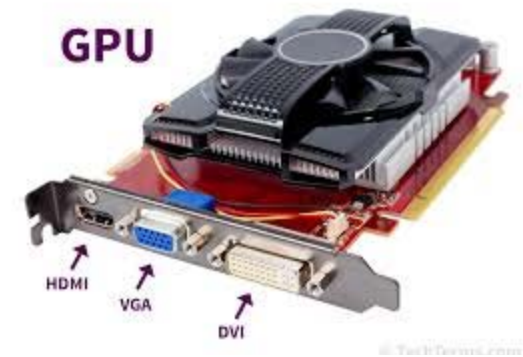


What do you need for deep learning predictions?

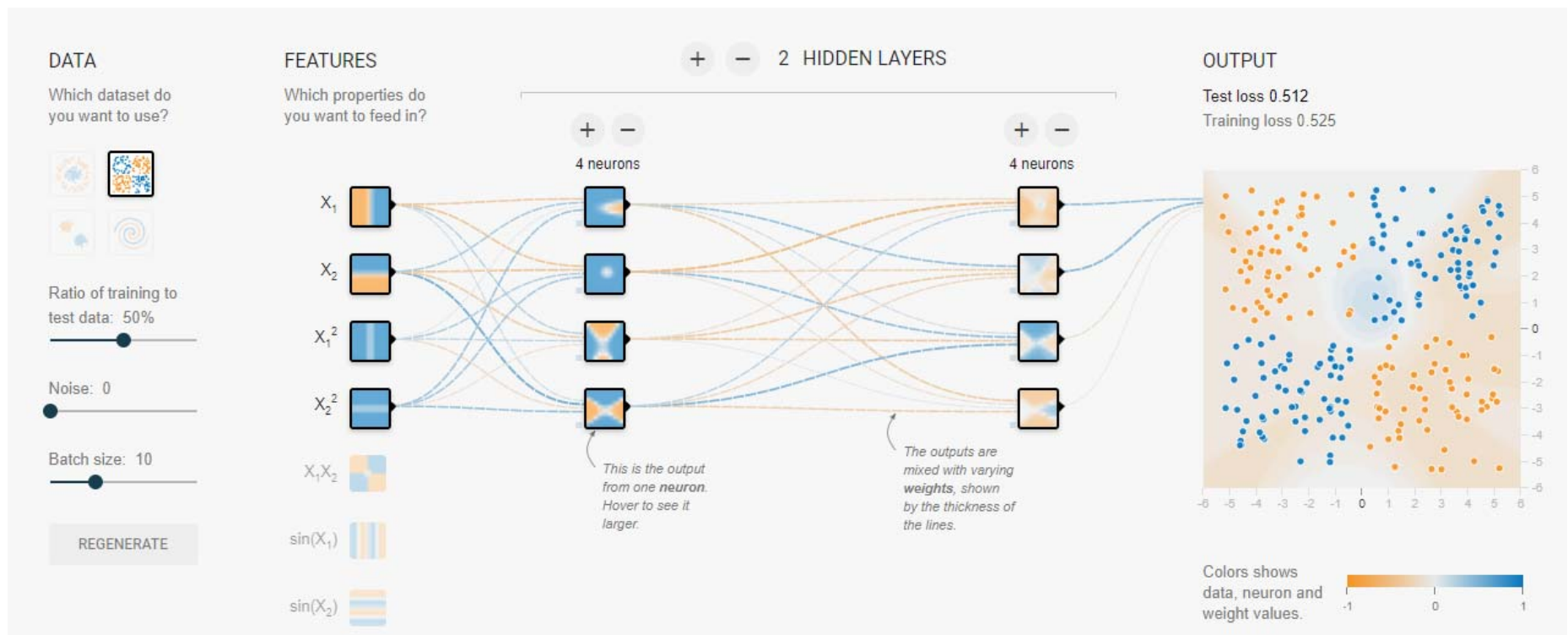
Lots of data labelled with prediction target
– thanks to EHRs and the HITECH act

Computer power
– thanks to video games

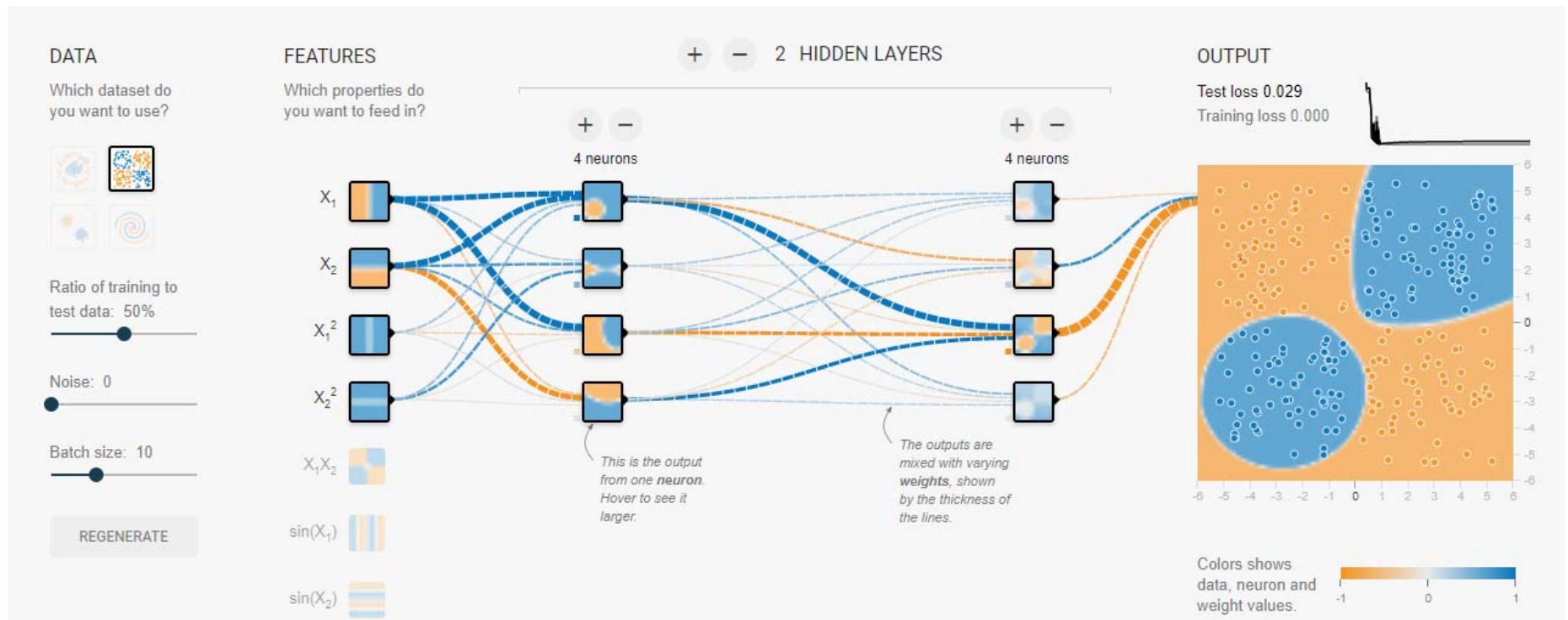
Fancy Algorithms
– thanks to high school math,
calculus 101 and linear algebra,
and an army of data scientists at
Google, Facebook, and LinkedIn



Neural Networks – How do They Work?



Neural Networks – How do They Work?

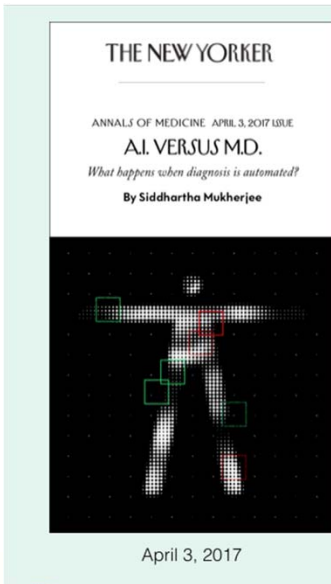


Where is Big Data Going in Health care? Artificial Intelligence/Machine Learning



Photo: eyediagnostics.net

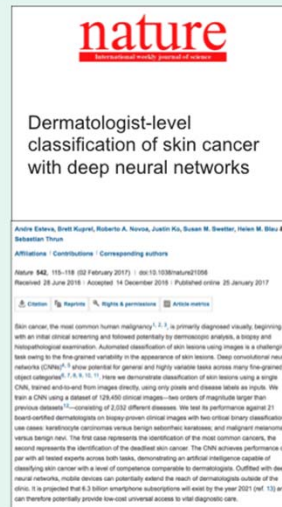
- Image-based specialties will see the most use
- Ophthalmology, dermatology, radiology
- AI will work alongside specialists, allowing MDs to work more effectively, and on more challenging cases.



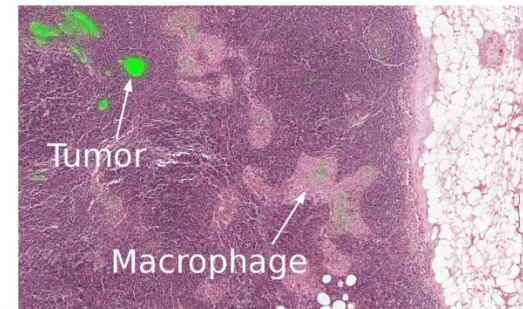
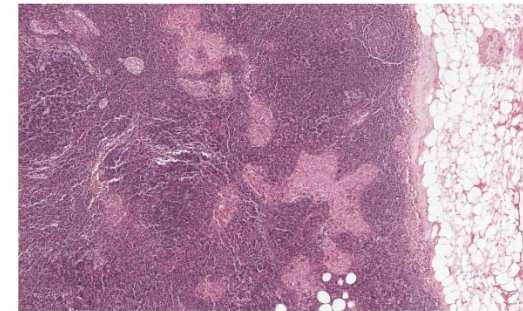
April 3, 2017

25 dermatologists
2000 biopsy-proven
images

“In every test, the
network
outperformed expert
dermatologists.”

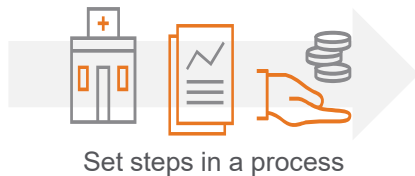


January 25, 2017



OptumLabs Innovation: Artificial Intelligence in Administrative Processes

KEY INGREDIENTS OF ADMINISTRATIVE WORKFLOWS



TYPES OF DATA MANUAL PROCESS DECISIONS

Many administrative processes for claims rely on analyzing text data

- Charts
- Notes
- Comments

Often, sequences of events are critical to determine an outcome

- Groups of claims
- A progression of care

The results of these decisions are well-suited to train a neural network

Applying Deep Learning Neural Networks(DLNNs)

Use cases underway



Avoidable ED visits	Use labeled ED visits and EHR data /claims data to predict patients at risk for these visits	Proactively reach out to priority patients to educate them about alternative care options	Audit/monitor outcomes of patients contacted to identify appropriate ED visits
Unnecessary medical claim reviews	Use labeled claims data to predict claims that should be automatically paid	Change claims review process to include automated approvals and review claims flagged for review	<ul style="list-style-type: none"> • Audit/monitor false positives • Evaluate policy change
Untimely prior authorizations	Use historical prior authorization data to predict which requests should be automatically approved	Change prior authorization process to include automated approvals and requests requiring review	<ul style="list-style-type: none"> • Audit/monitor false positives • Evaluate policy change

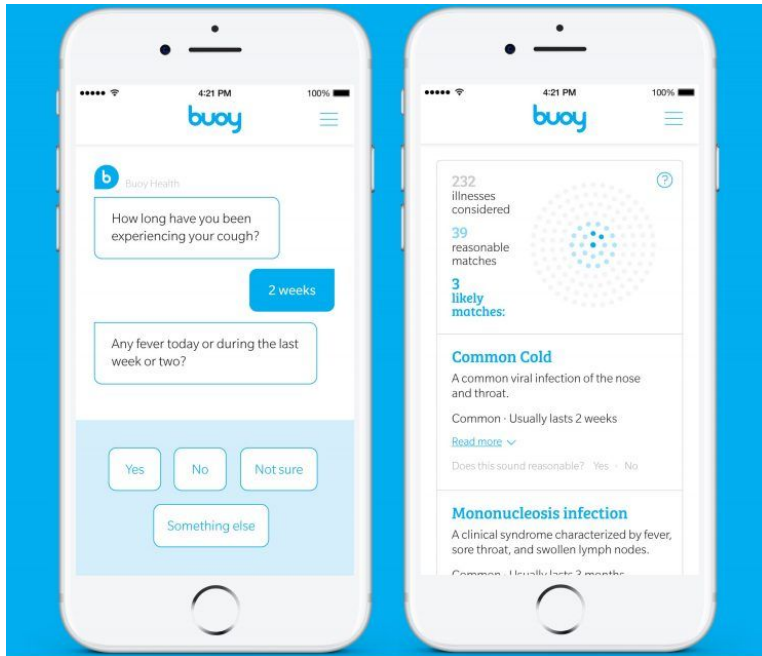
AI Promise ... and Risks



- Powerful for images and text
- May require much less data preparation
- TODAY, requires large training sets
- New techniques allow much smaller data
- AlphaGo shows the possibility that AI can surpass human performance in some cases (GO, Chess)
- Some tasks aren't amenable to AI (eg. crossword puzzles)

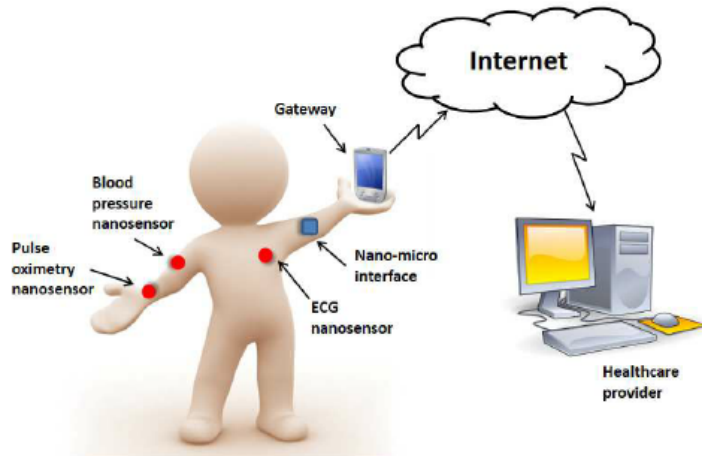
- AI doesn't replicate human thinking or human mistakes
- The system is often a black box, which may vary from human judgement occasionally or drift over time.
- Biased data will lead to incorporation of bias in decision making with major implications
- Need a "human in the loop" system.

Chatbots/Voice – Powered by AI

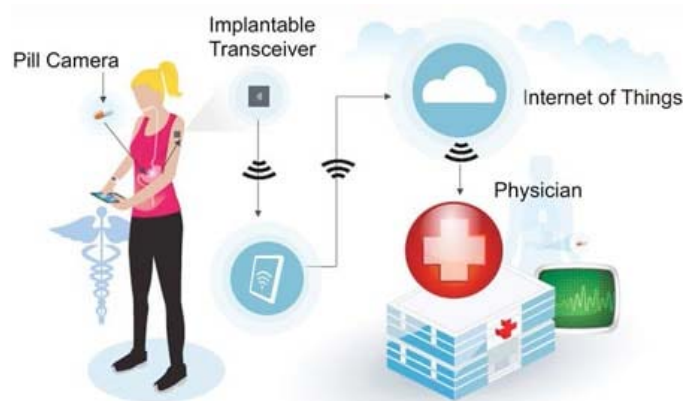


- Chatbots and voice (Alexa) are becoming a major part of consumer's lives
- Younger consumers prefer efficient, online interactions for things like:
 - Simple health questions
 - Scheduling
 - Test results
 - Ordering of medicines and medical goods
- Powered by AI, chatbots and voice bots can provide a satisfactory, if not preferred experience in these situations.

Internet of Things/Wearables - Health Care Insurance 3.0



- Consumer wearables will continue to be popular for motivated healthy individuals.
- Medical grade wearables will become more available and useful for those with chronic disease.
- The focus will shift from the motivated consumer or tech savvy physician to payer plan design.
- Examples:
 - “steps” to reduce insurance premiums
 - motion sensors to reduce rehab time in hip and knee surgery
 - Continuous glucose monitor based coaching



Individual Health Records – Ontology Mapping

Individual Health Record™

Patient Summary | Provider Directory | Tasks

Theodore S. Smith Gender: **Male** Birthdate: **November 12, 1938** Age: **79 yr** [Select another individual](#)

Symptoms and Conditions

Problem	Onset	Data Source Type
Adenomatous Polyp of the Colon	Aug 23, 2018	U
Fall Risk	Jun 4, 2018	U
Rheumatoid Arthritis	Jun 4, 2018	U
Diabetes Mellitus Type 2	Apr 18, 2010	U
Essential Hypertension	Apr 18, 2010	U

Medications

Medication	Data Source Type
Methotrexate Sodium Oral Tablet 5 mg	U
Xeljanz Oral Tablet 5 mg	U
Actos Oral Tablet 30 mg (inactive)	U
Glucophage Oral Tablet 500 mg	U
Zantac Oral Tablet 150 mg	U

Health Goal Tasks

Task	Priority
Schedule patient for pneumococcal vaccine within next 2 months	●
Schedule retinopathy screening within next 10 months	●
Schedule microalbuminuria testing within the next 10 months	●
Educate patient about methods for lowering cholesterol levels during next visit	●
Schedule patient for medication reconciliation within the next 2 months	●
Educate patient about stress relief methods during next visit	●
Educate patient about diet and exercise tips for lowering blood pressure during next visit	●
Educate patient about diet and exercise tips for managing weight during next visit	●

Allergies

Allergen	Reaction	Data Source Type
Sulfamethoxazole	Muscle Pain Nausea	U
Penicillin G	Allergic Urticaria	U
Nonsteroidal Anti-Inflammatory Drug (NSAID)	Nausea	U

Tests and Exams

Start Date	Procedure Information	Indicator	Data Source Type
Sep 4, 2018 2:25 PM	Glucose Assay in Blood	●	U
Sept 4, 2018 2:25 PM	Lipid Profile in Blood	●	U
Sep 4, 2018 2:25 PM	Hemoglobin A1c Assay in Blood	●	U
Sep 4, 2018 9:30 AM	X-ray of Knee	●	U
Sep 3, 2018 11:17 AM	Glucose Assay in Blood	●	U

Vital Signs

Two line graphs showing vital signs over time. The left graph shows blood pressure (BP) and the right graph shows heart rate (HR).

Care Team

Attending

Name: Butler, Alfred	Specialty: Physician: Internal Medicine
----------------------	---

Attributed Practitioner

Name: Butler, Alfred	Specialty: Physician: Internal Medicine
----------------------	---

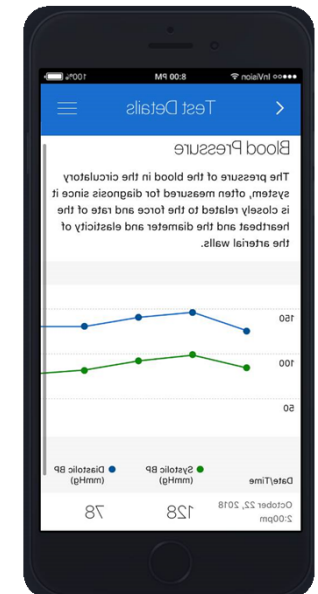
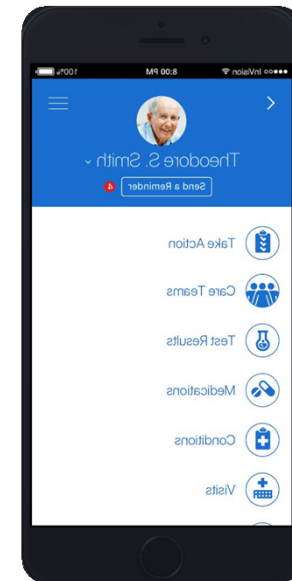
Managing Clinician

Name: Butler, Alfred	Specialty: Physician: Internal Medicine
Opal, Edmund	Physician: Gastroenterology, Internal Medicine
Bones, Douglas	Physician: Orthopedic Surgery

Consultant

Name: [Redacted]	Specialty: [Redacted]
------------------	-----------------------

- EHR Interoperability is a long way from happening
- An emerging approach is longitudinal mapping into a single ontology
- Provider has a complete record regardless of where they practice
- Patient has access to useful data
- Apple has a simple version.



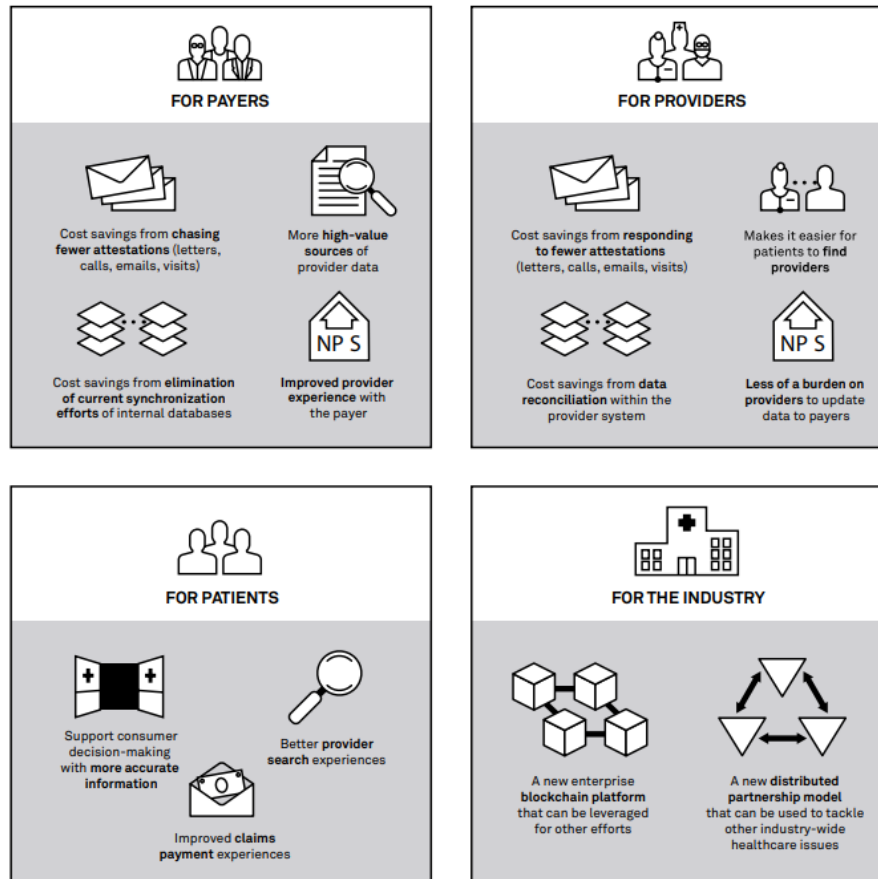
Blockchain



- Distributed, shared, tamper-free “ledger” of chronological transactions.
- Allows transactions between strangers and eliminates the need for a central authority that records and stores transactions (eg. the Registry of Deeds or a bank).
- Most obvious use is to create a usable currency without an issuing authority (eg. Bitcoin).
- 2019 view – nobody has come up with the “killer application” for non-currency uses of blockchain.

Potential Uses of Blockchain in Health care

■ Potential Benefits of Blockchain Technology in Healthcare



- Supply chain management
- Drug traceability
- Opioid management
- Smart contracts, including management of the right to edit data

Synaptic Alliance

- Provider data is a major issue for multiple stakeholders – payers, providers, patients
- Leads to payment issues, patient and provider frustration, and inefficiencies, and Federal Penalties
- Management of the data is done in parallel, repetitively, and by various actors without inherent trust.
- Humana, Optum, UnitedHealth care, Quest, and MultiPlan partnering on a blockchain pilot
- Creates a synchronized, shared source of high-quality provider data with a complete audit trail and inherent data integrity.

Challenges and Limitations

- Data scientists are in very short supply and health care isn't paying as well/as attractive as the FANG companies
- Untrained "consumers" of health care data analytics
- Health care data is a very attractive target for hackers
- Constant calls for "liberating" health care data is at major odds with:
 - governmental policy
 - hospital and some corporate approaches
 - societal concerns
- Health care data standardization, which might allow easy aggregation, is proceeding slowly, and incentives aren't aligned to accelerate this.

BIG DATA SKILLS SHORTAGE



By 2018, there will be a **shortage of**

1.5 million Big Data managers and analysts

1.7
million
workers

with big data skills
in the U.S. alone



140,000 – 190,000
Big Data workers

Predictions are no different than lab tests



Search 10:22 AM 75%



Tweet



Elissa Lombardo
@elissalom



[@tim_cook](#) New Apple watch saved my husbands life this week! Only two days old and it diagnosed A-Fib and 150bpm. He went to ER which he never did with same symptoms. Found major blockage in arteries as a result. Two stents later, he is as good as new! Telling the world. Thank U!

7:12 PM · 1/10/19 · [Twitter for iPad](#)

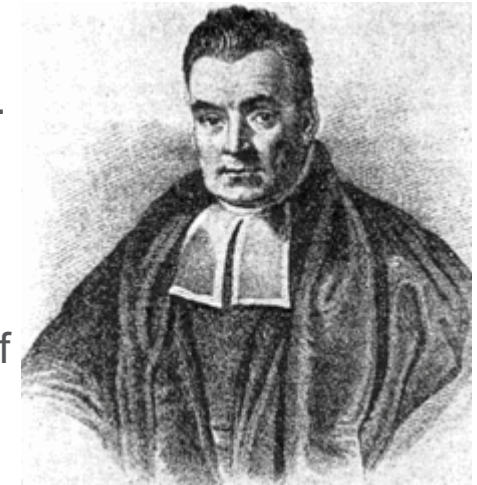
429 Retweets 3,163 Likes

Tweet your reply



Risks: The Curse of Reverend Bayes

Without screening for high risk populations, many predictive technologies will lead to increased costs and patient anxiety. The new **Apple Watch ECG** tool is a good example:



- Prevalence of Afib in 18-64 y.o. is 0.92%, of which 0.09% is undiagnosed
- The watch has an impressive sensitivity of 98.3% and specificity of 99.6%.
- For every 1,000,000 18-64 y.o.'s who use the Apple Watch for Afib detection:
 - 8,300 will have their Afib diagnosis confirmed.
 - 4,848 will be told they may have newly diagnosed Afib
 - 885 (18% of 4,848) will actually have newly diagnosed Afib
 - The remaining 3,963 will be told they might have Afib, but actually don't
- **82%** of positive "findings" will be false +'s leading to excess costs for testing, substantial anxiety, and leaving many with an ongoing belief that they have Afib, even though confirmatory tests don't show it.

Risks: Bias

Ad related to latanya sweeney ⓘ

Latanya Sweeney Truth
www.instantcheckmate.com/
 Looking for **Latanya Sweeney**? Check **Latanya Sweeney's Arrests**.

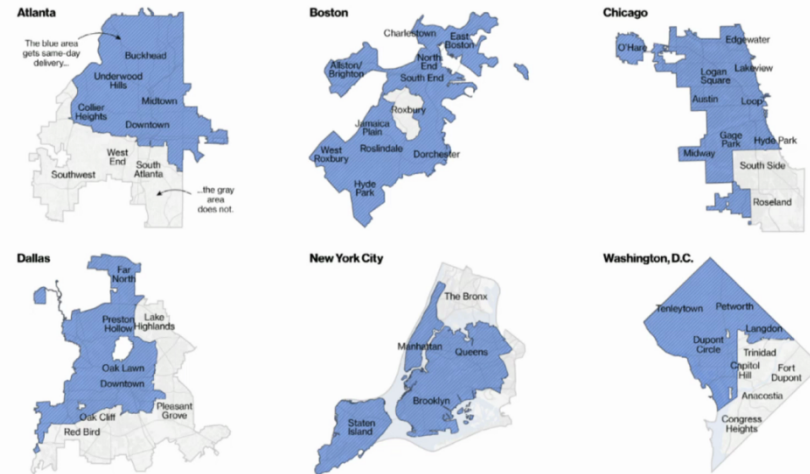
Ads by Google

Latanya Sweeney, Arrested?
 1) Enter Name and State. 2) Access Full Background Checks Instantly...
www.instantcheckmate.com/

Latanya Sweeney
 Public Records Found For: **Latanya Sweeney**. View Now.
www.publicrecords.com/

La Tanya
 Search for La Tanya Look Up Fast Results now!
www.ask.com/La+Tanya

Amazon Prime – Same Day Delivery



<http://www.bloomberg.com/graphics/2016-amazon-same-day/>

Turkish - detected ▾

o bir aşçı
 o bir mühendis
 o bir doktor
 o bir hemşire
 o bir temizlikçi
 o bir polis
 o bir asker
 o bir öğretmen

English ▾

she is a cook
 he is an engineer
 he is a doctor
 she is a nurse
 he is a cleaner
 He-she is a police
 he is a soldier
 She's a teacher

Big Health Care Data Risks: Privacy

Record	*****
Hospital	162: Sacred Heart Medical Center in Providence
Admit Type	1: Emergency
Type of Stay	6: Subacute
Length of Stay	6 days
Discharge Date	Oct-2011
Discharge Status	under the care of an health service organization
Charges	\$71708.47
Payers	1: Medicare 6: Commercial insurance 625: Other government sponsored patients
Emergency Codes	E8162: motor vehicle traffic accident due to loss of control; loss control mv-mocycl
Diagnosis Codes	80843: closed fracture of other specified part of pelvis 51851: pulmonary insufficiency following trauma & surgery 2761: hyposmolality & or hyponatremia 78057: tachycardia 2851: acute hemorrhagic anemia
Age in Years	60
Age in Months	720
Gender	Male
ZIP	98851
State Reside	WA
Race/Ethnicity	white, Non-Hispanic

MAN 60 THROWN FROM MOTORCYCLE
 A 60-year-old Soap Lake man was hospitalized Saturday afternoon after he was thrown from his motorcycle. Ronald Jameson was riding his 2003 Harley-Davidson north on Highway 25, when he failed to negotiate a curve to the left. His motorcycle became airborne before landing in a wooded area. Jameson was thrown from the bike; he was wearing a helmet during the 12:24 p.m. incident. He was taken to Sacred Heart Hospital. The police cited speed as the cause of the crash. [News Review 10/18/2011]

From: Latanya Sweeney,
<https://techscience.org/a/2015092903/>