

Opportunities to Collaborate in Data Science + Data Engineering

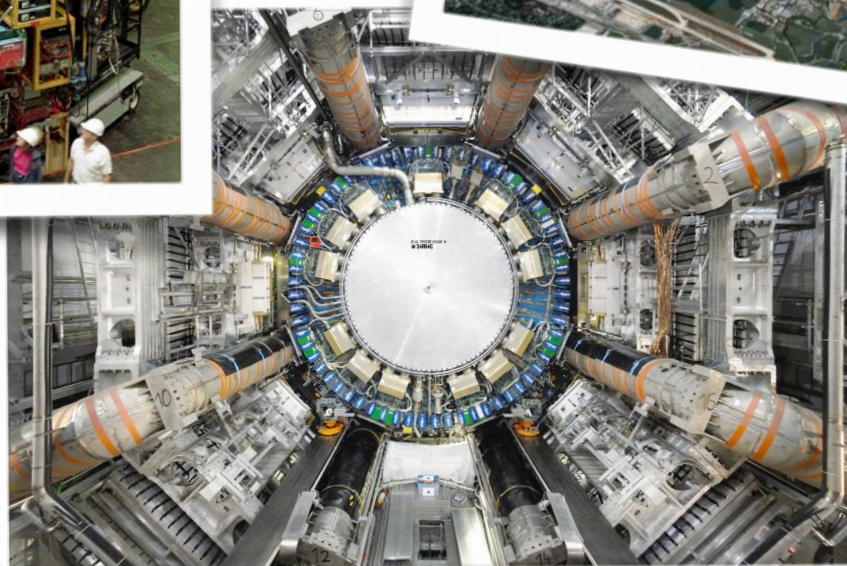
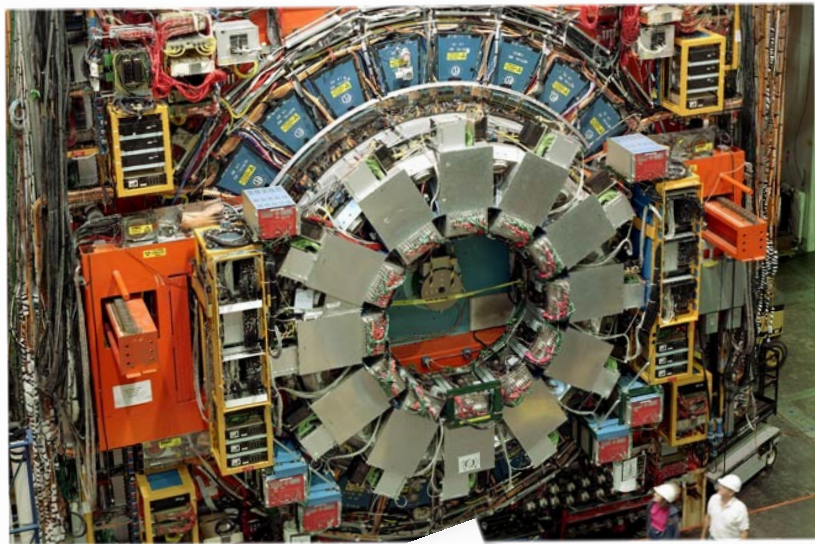
Kathy Copic

kathy@insightdatascience.com

The Insight Fellows Program

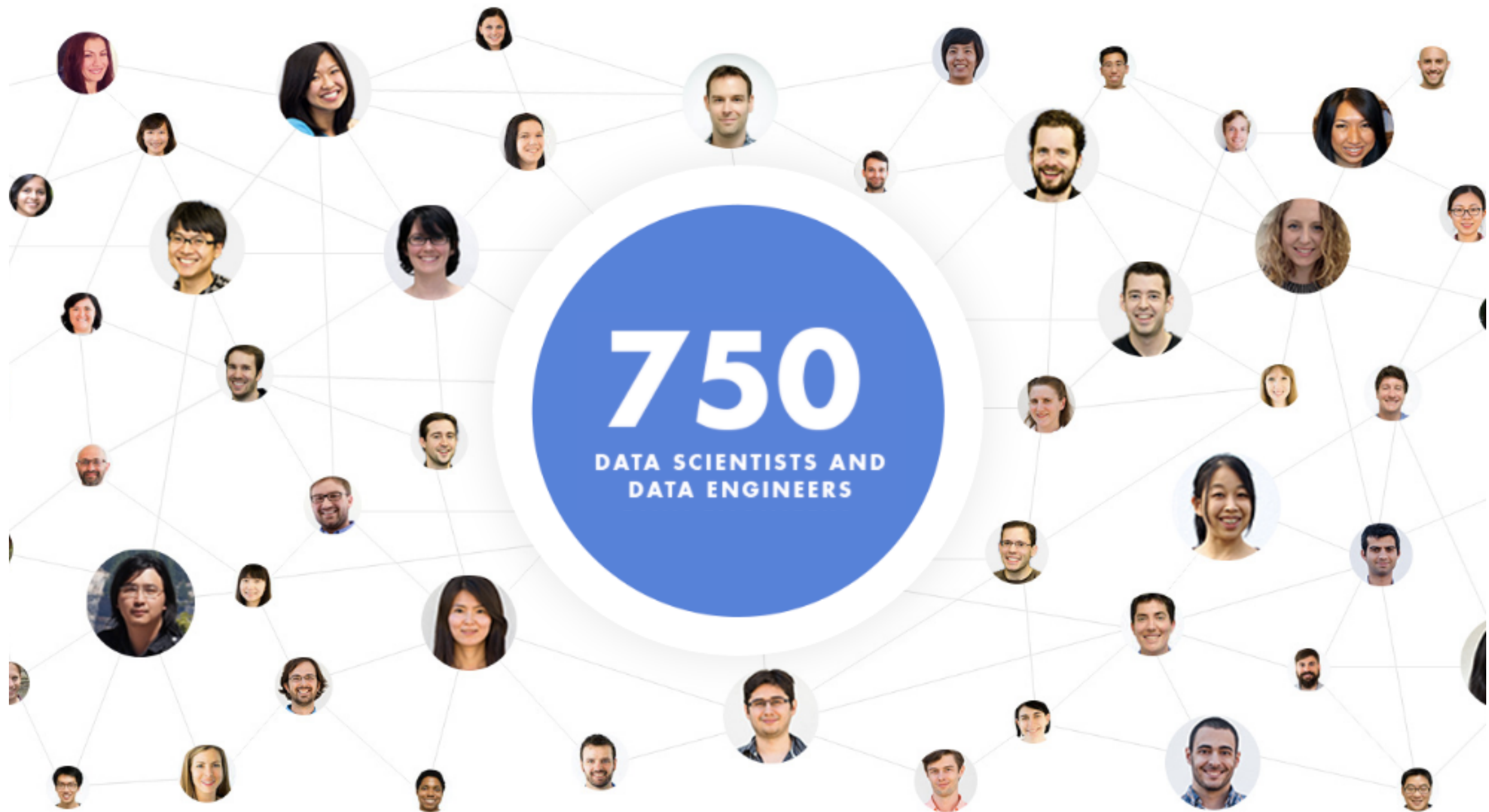
insightdatascience.com | insightdataengineering.com

insightdata.ai | insighthealthdata.com



INSIGHT

Insight Fellows across the US

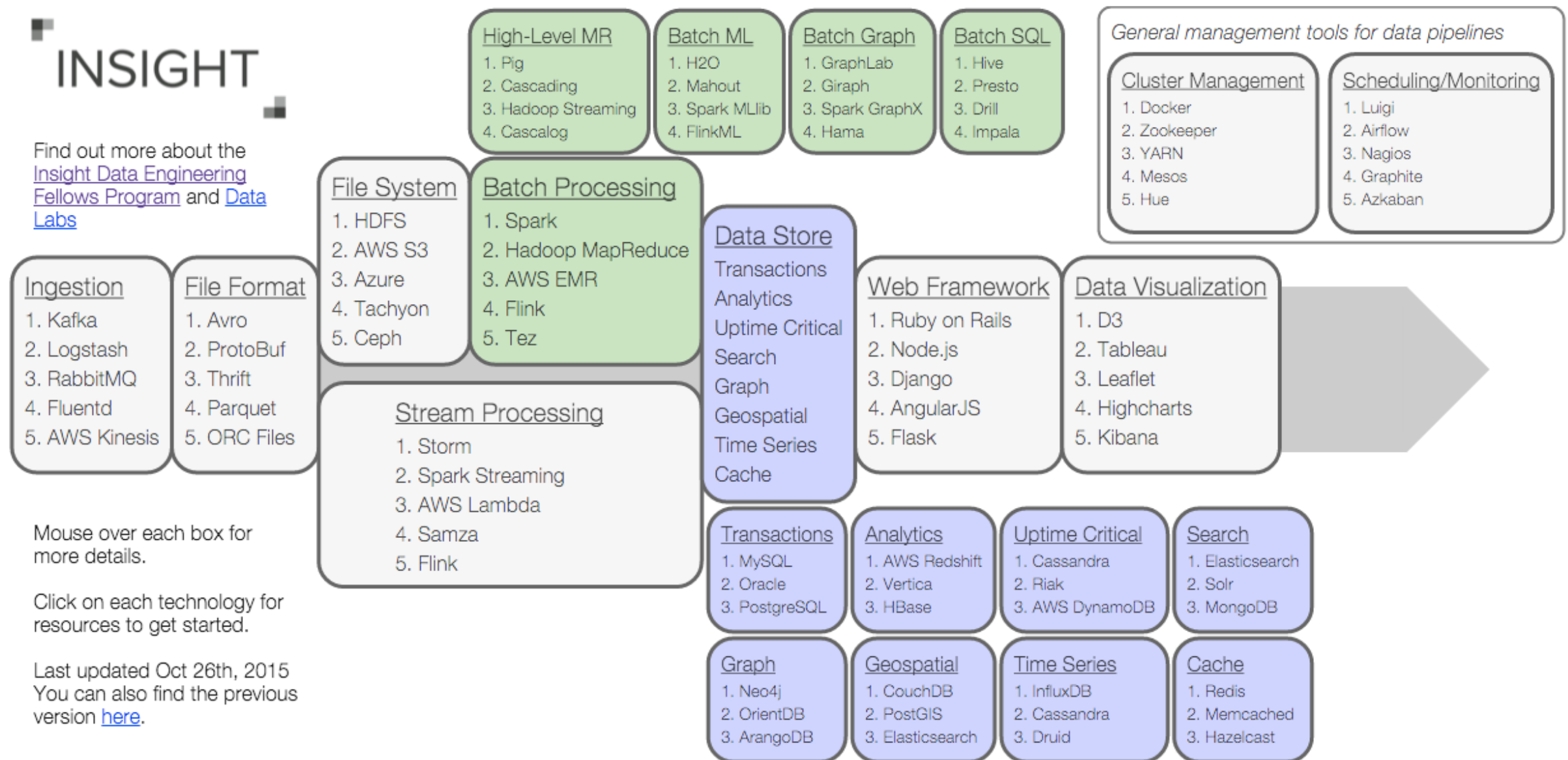


Insight Fellows are Data Scientists and Data Engineers at:



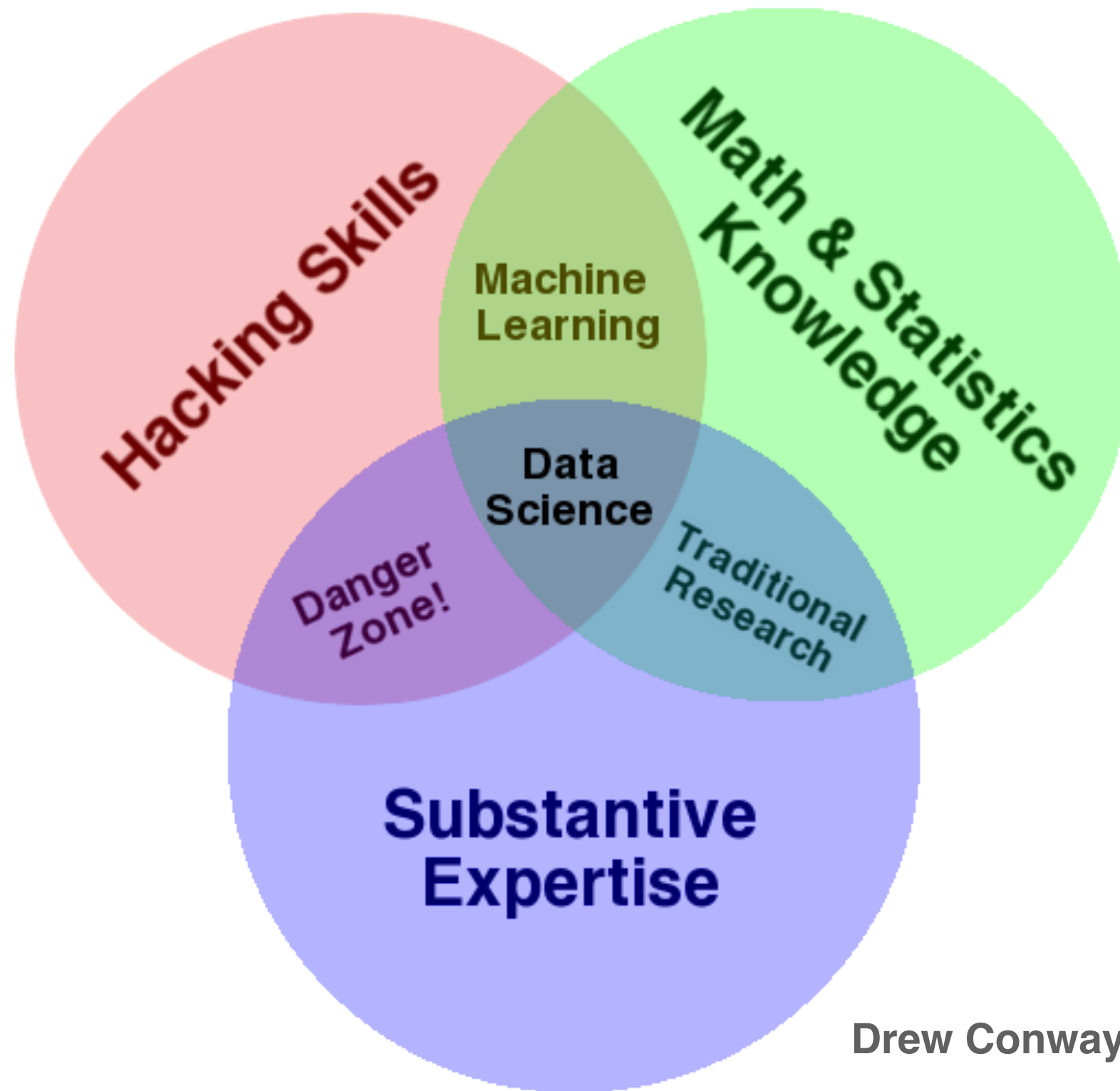
Silicon Valley • New York • Boston • Los Angeles • Seattle
+ many others...

Tremendous innovation and production of tools



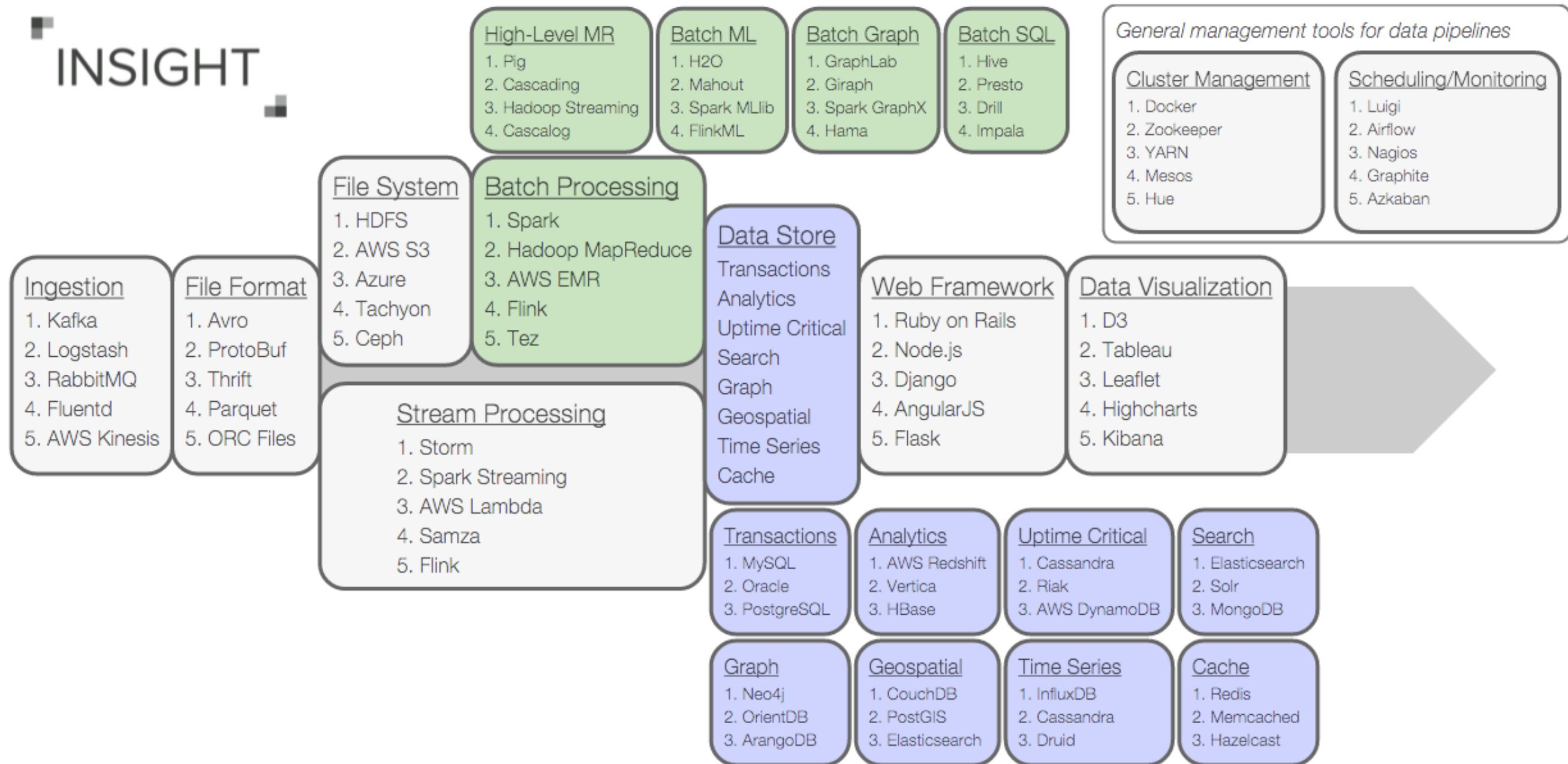
... and Open Source!

What is Data Science?

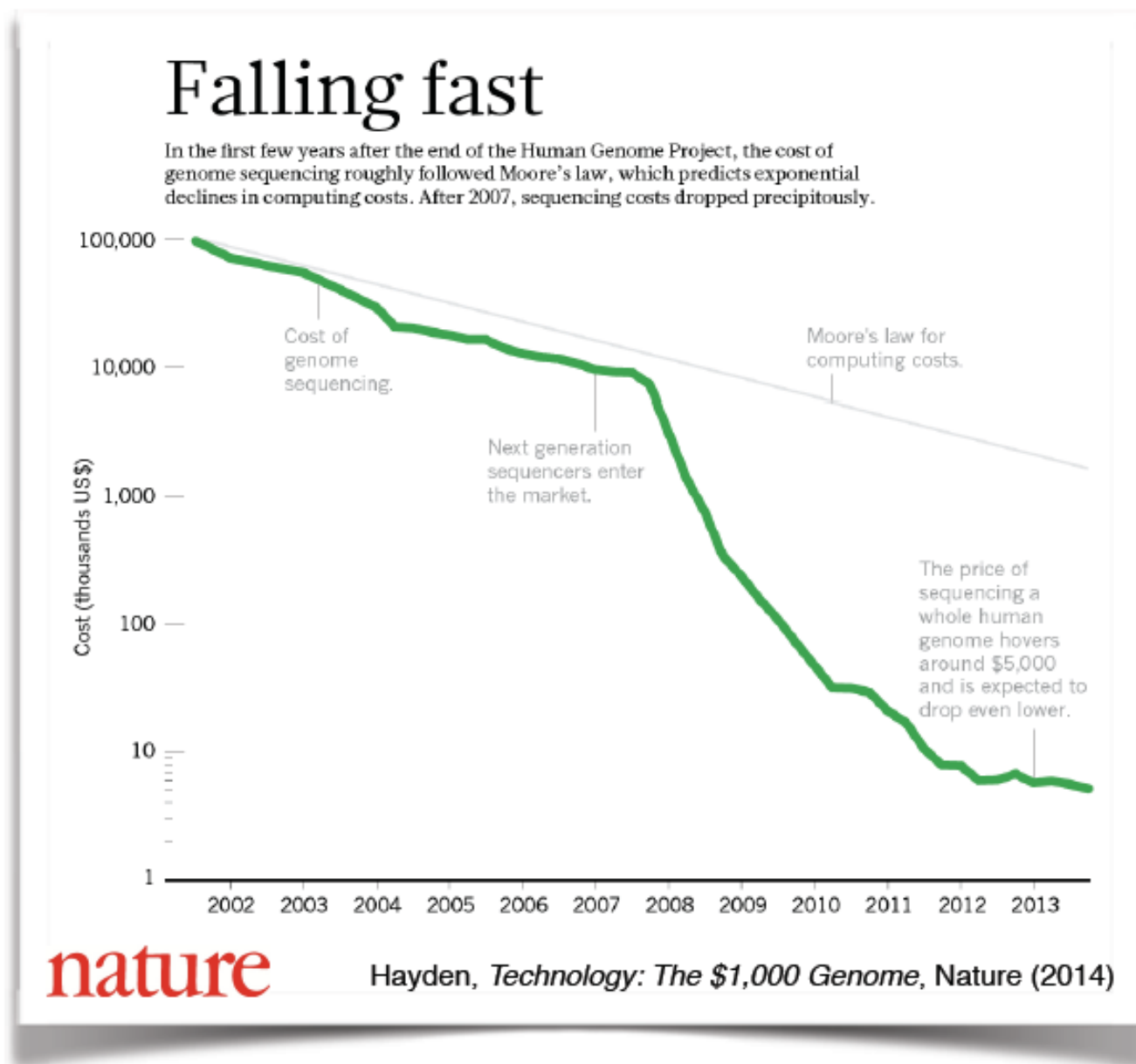


Drew Conway

What is Data Engineering?



Health Data: Explosion!



- Genome sequences
- Electronic medical records
- Claims data
- Quantified-self data
- Longitudinal studies
- Drug screens
- Clinical trial data

Segmentation in the Field

Segmentation in the Field

Data Analytics / Decision Science
Analysis to drive choices

Segmentation in the Field

Data Analytics / Decision Science
Analysis to drive choices

Data Products
Building Products used by people

Segmentation in the Field

Data Analytics / Decision Science
Analysis to drive choices

Data Products
Building Products used by people

AI / ML / Deep Learning Engineer
Next-level, “Reinforcement learning”

Segmentation in the Field

Data Analytics / Decision Science
Analysis to drive choices

Data Products
Building Products used by people

AI / ML / Deep Learning Engineer
Next-level, “Reinforcement learning”

Data Engineers
Deal with size, speed, quality
Big difference in Big vs Small Companies

How to connect:

How to connect:

Go to **Meet-ups** or watch online.

How to connect:

Go to **Meet-ups** or watch online.

Subscribe to **Data Science Weekly**.

How to connect:

Go to **Meet-ups** or watch online.

Subscribe to **Data Science Weekly**.

Join **LinkedIn!** Meet up with people.

How to connect:

Go to **Meet-ups** or watch online.

Subscribe to **Data Science Weekly**.

Join **LinkedIn!** Meet up with people.

Podcasts: Software Engineering Daily

How to connect:

Go to **Meet-ups** or watch online.

Subscribe to **Data Science Weekly**.

Join **LinkedIn!** Meet up with people.

Podcasts: Software Engineering Daily

Read **Twitter!**

How to connect:

Go to **Meet-ups** or watch online.

Subscribe to **Data Science Weekly**.

Join **LinkedIn!** Meet up with people.

Podcasts: Software Engineering Daily

Read **Twitter!**

Do nearby universities have a **center for data science?**

How to connect:

Go to **Meet-ups** or watch online.

Subscribe to **Data Science Weekly**.

Join **LinkedIn!** Meet up with people.

Podcasts: Software Engineering Daily

Read **Twitter!**

Do nearby universities have a **center for data science?**

Make **open source contributions**, get involved!

How to connect:

Go to **Meet-ups** or watch online.

Subscribe to **Data Science Weekly**.

Join **LinkedIn!** Meet up with people.

Podcasts: Software Engineering Daily

Read **Twitter!**

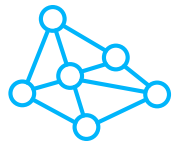
Do nearby universities have a **center for data science?**

Make **open source contributions**, get involved!

Read our **Insight Blog! :)**

INSIGHT

PROGRAMS



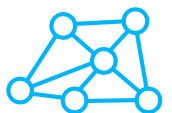
DATA SCIENCE
FOR PHDS



**DATA
ENGINEERING**

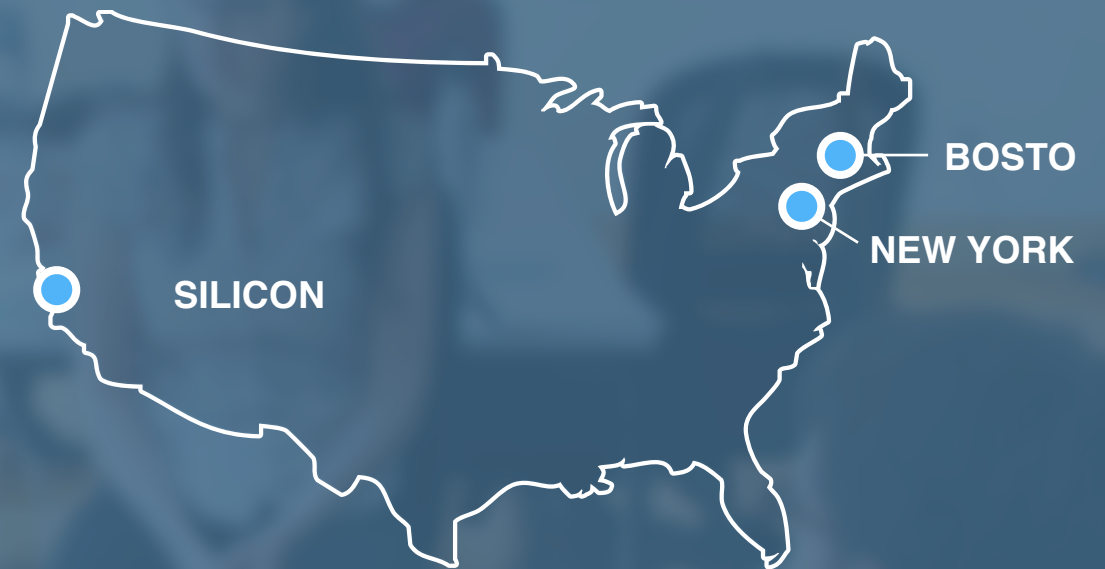


HEALTH ANALYTICS
FOR PHDS + MDS



**ARTIFICIAL
INTELLIGENCE**

LOCATIONS



+ REMOTE

Upcoming Session: January 17, 2017

Application Deadline: October 24, 2016

Visit our website to apply or sign up for notifications

Contact me!

kathy@insightdatascience.com

Kathy Copic: LinkedIn!

@KathyScientist



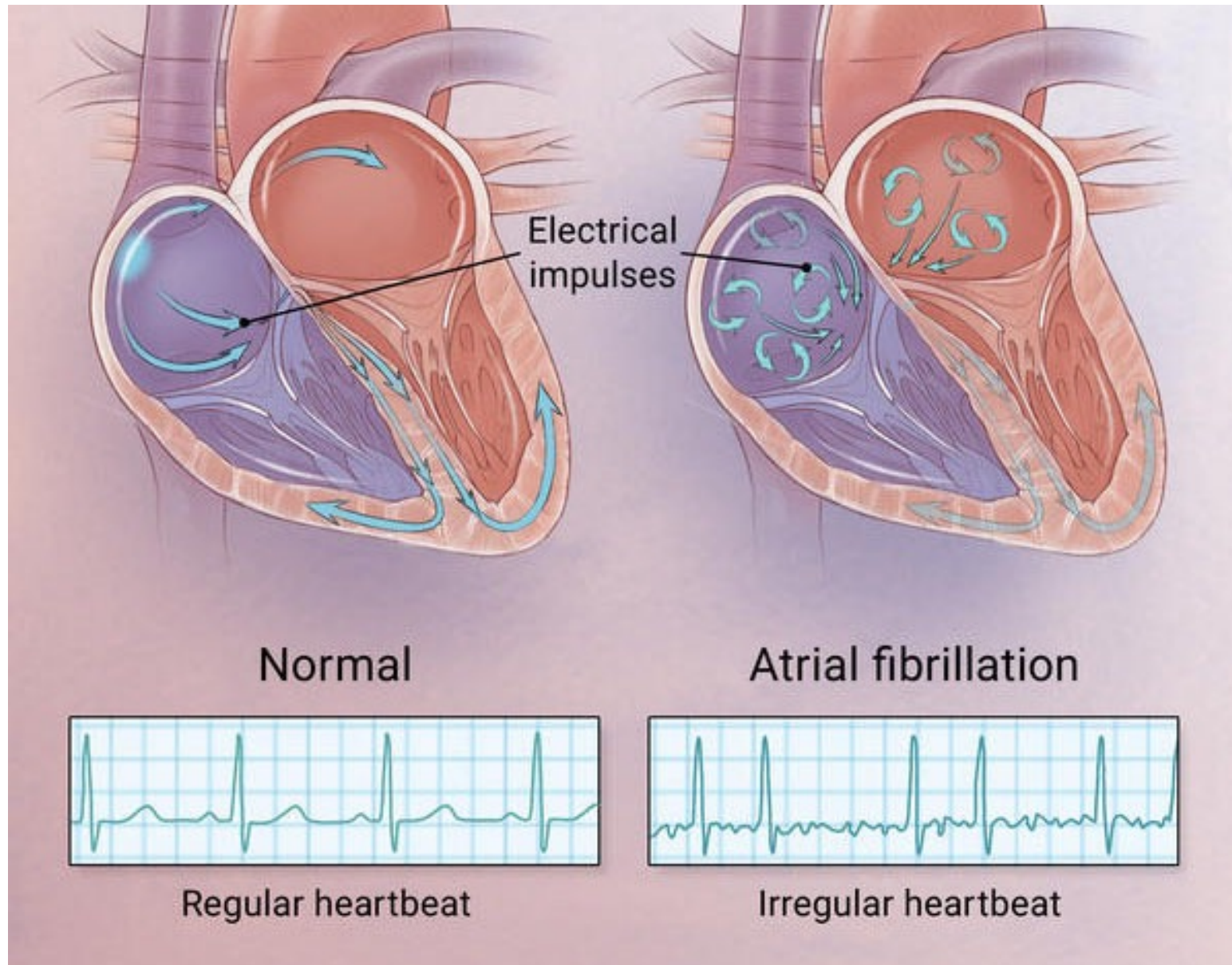


Data Science & Data Engineering programs



Bay Area, NYC, Boston and more!

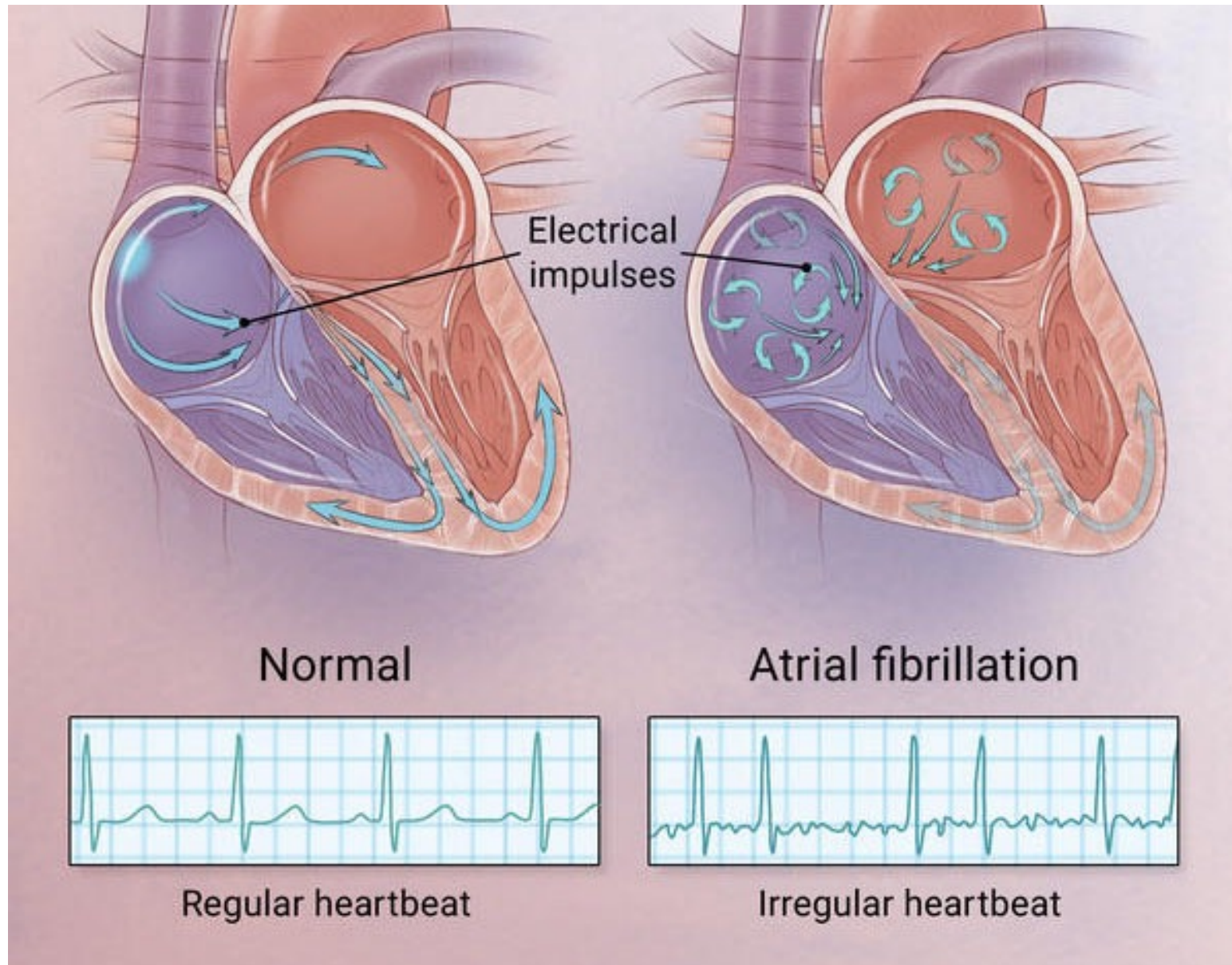
Health Data Challenges



2.7 million patients in the US

Leading cause of stroke

Health Data Challenges



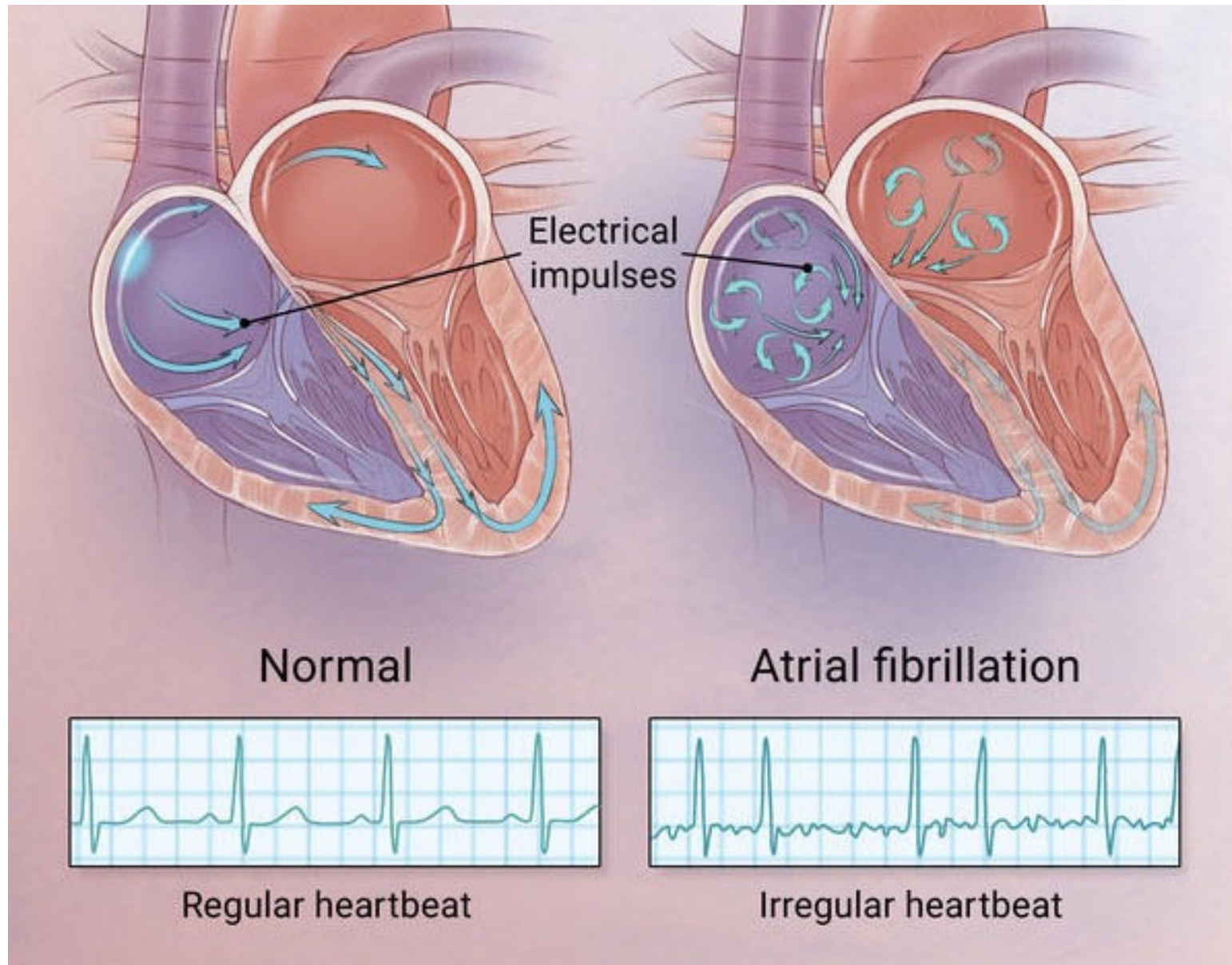
UCSF's Health eHeart Study:

13.5 million heart rate measurements from
~500 users

2.7 million patients in the US

Leading cause of stroke

Health Data Challenges



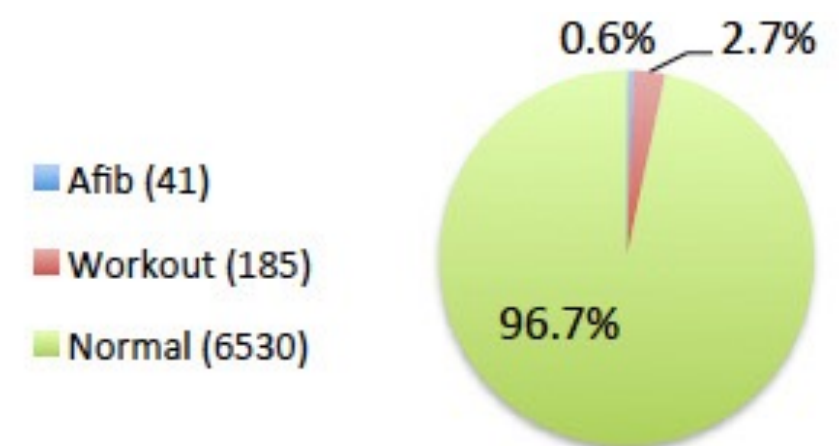
2.7 million patients in the US

Leading cause of stroke

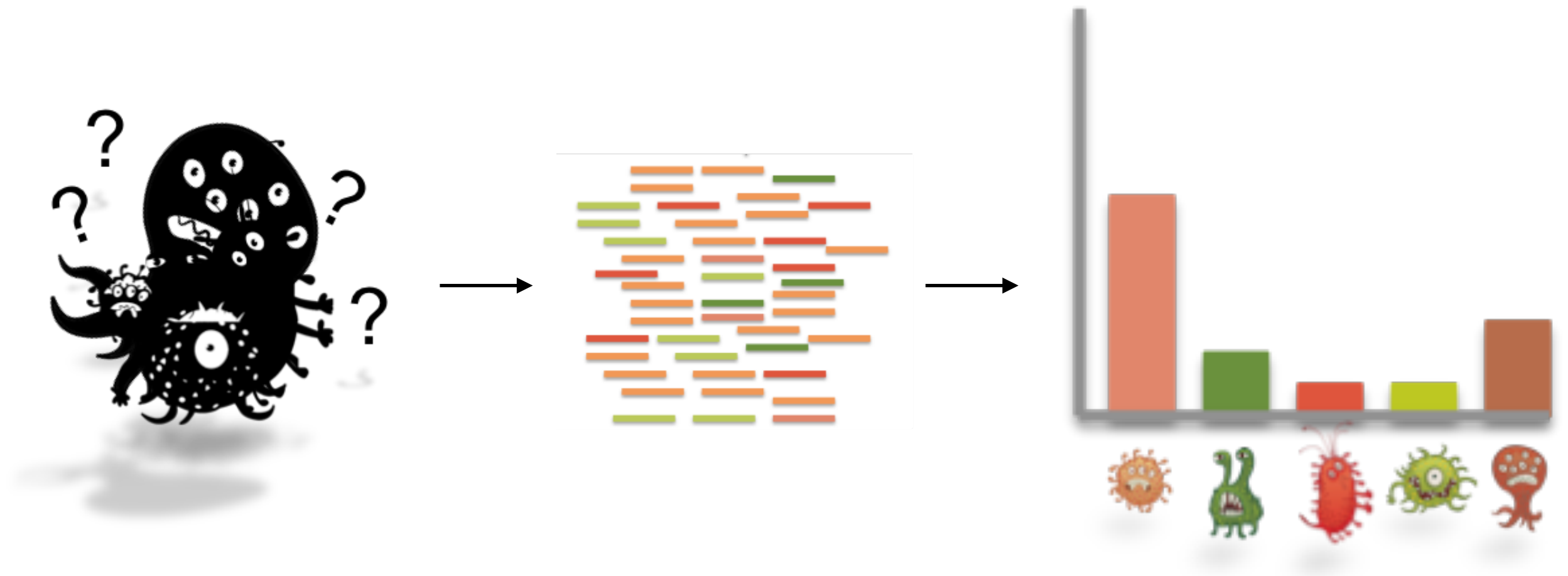
UCSF's Health eHeart Study:

13.5 million heart rate measurements from
~500 users

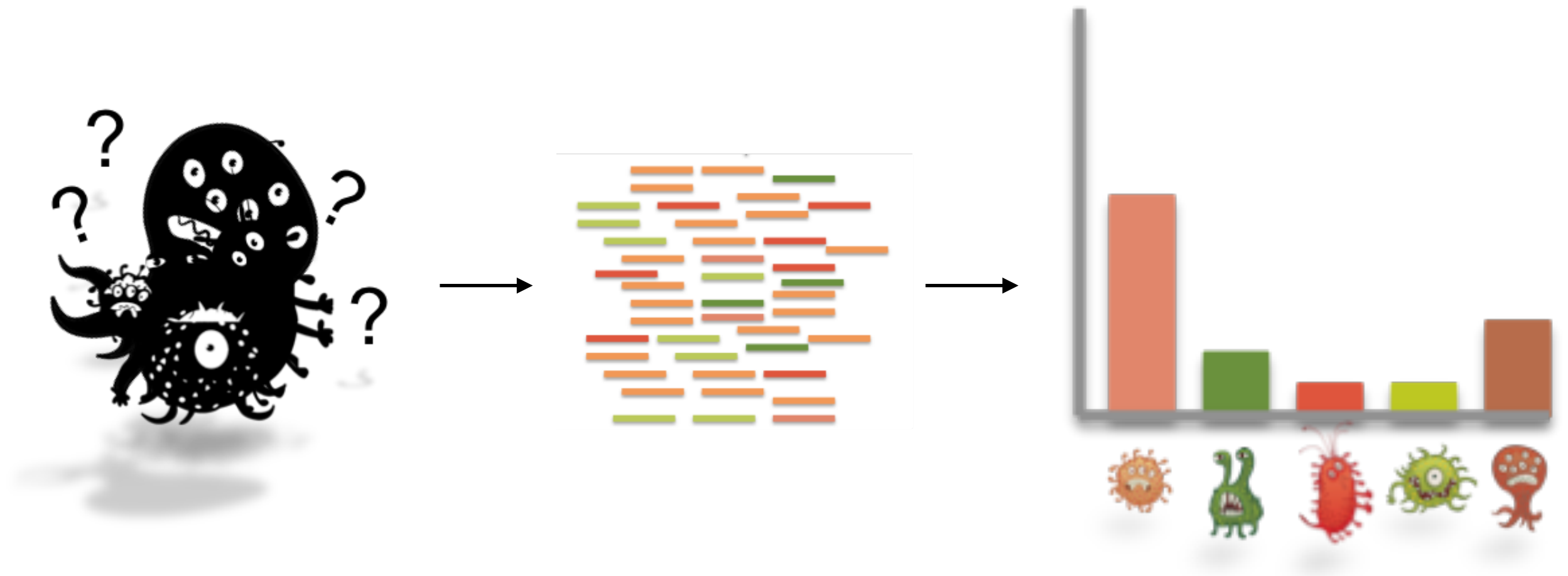
Only 41 Afib cases



Health Data Challenges

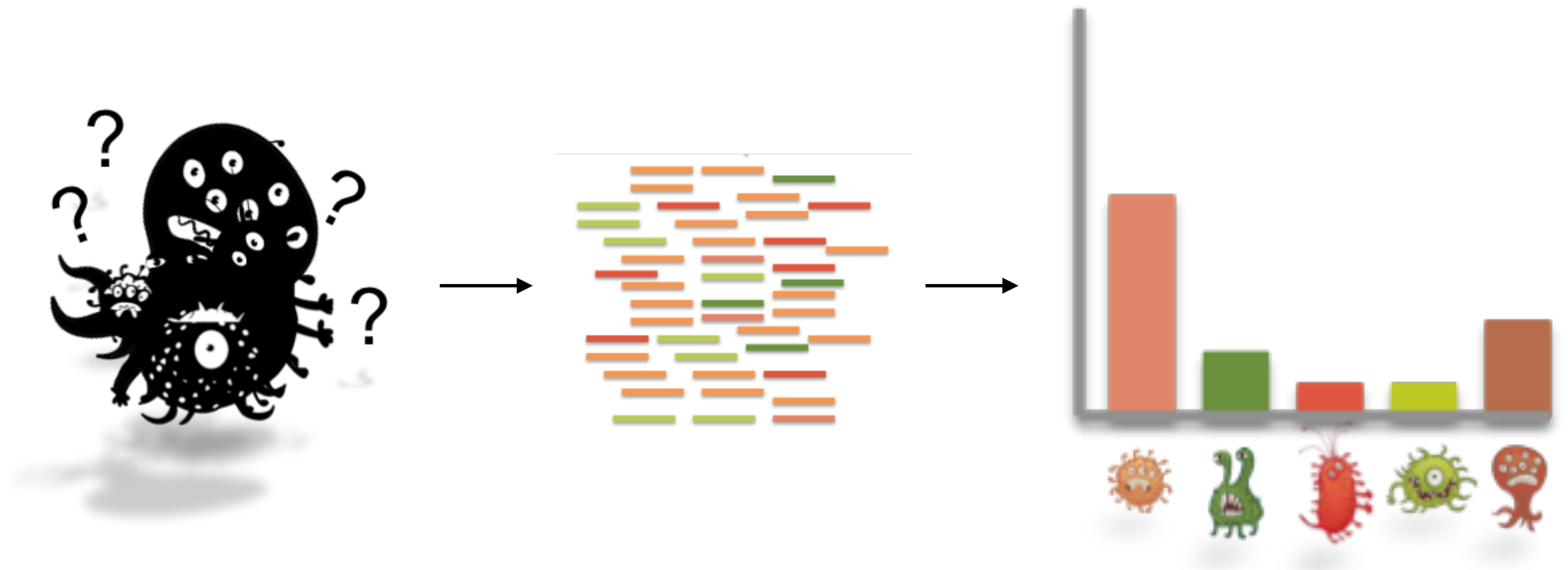


Health Data Challenges



YC startup One Codex —
500 FASTQ sequencing files, ~1M+ reads

Health Data Challenges



YC startup One Codex —
500 FASTQ sequencing files, ~1M+ reads

Applied LDA topic modeling: a technique from
Natural Language Processing

Many exciting questions:

Many exciting questions:

Memorial Sloan Kettering Cancer Center:

Who is a good candidate for clinical trials for treatments?

Many exciting questions:

Memorial Sloan Kettering Cancer Center:

Who is a good candidate for clinical trials for treatments?

Biogen: If you combine data from genomics, clinical research, environmental studies, etc, are there new approaches to treat disease?

Many exciting questions:

Memorial Sloan Kettering Cancer Center:

Who is a good candidate for clinical trials for treatments?

Biogen: If you combine data from genomics, clinical research, environmental studies, etc, are there new approaches to treat disease?

Constant Therapy: Can an iPad app deliver highly personalized, continuous therapy to patients with traumatic brain injury, stroke, etc?

Many exciting questions:

Memorial Sloan Kettering Cancer Center:

Who is a good candidate for clinical trials for treatments?

Biogen: If you combine data from genomics, clinical research, environmental studies, etc, are there new approaches to treat disease?

Constant Therapy: Can an iPad app deliver highly personalized, continuous therapy to patients with traumatic brain injury, stroke, etc?

Many exciting questions:

Memorial Sloan Kettering Cancer Center:

Who is a good candidate for clinical trials for treatments?

Biogen: If you combine data from genomics, clinical research, environmental studies, etc, are there new approaches to treat disease?

Constant Therapy: Can an iPad app deliver highly personalized, continuous therapy to patients with traumatic brain injury, stroke, etc?

Trace Genomics: Can we identify microbes and genetic mutations that cause or protect crops from disease, improving yield and shelf life?

Many exciting questions:

Memorial Sloan Kettering Cancer Center:

Who is a good candidate for clinical trials for treatments?

Biogen: If you combine data from genomics, clinical research, environmental studies, etc, are there new approaches to treat disease?

Constant Therapy: Can an iPad app deliver highly personalized, continuous therapy to patients with traumatic brain injury, stroke, etc?

Trace Genomics: Can we identify microbes and genetic mutations that cause or protect crops from disease, improving yield and shelf life?

Many exciting questions:

Memorial Sloan Kettering Cancer Center:

Who is a good candidate for clinical trials for treatments?

Biogen: If you combine data from genomics, clinical research, environmental studies, etc, are there new approaches to treat disease?

Constant Therapy: Can an iPad app deliver highly personalized, continuous therapy to patients with traumatic brain injury, stroke, etc?

Trace Genomics: Can we identify microbes and genetic mutations that cause or protect crops from disease, improving yield and shelf life?

Everywhere: Data architecture —
how can the data be stored, streamed, accessed?

Many exciting questions:

Memorial Sloan Kettering Cancer Center:

Who is a good candidate for clinical trials for treatments?

Biogen: If you combine data from genomics, clinical research, environmental studies, etc, are there new approaches to treat disease?

Constant Therapy: Can an iPad app deliver highly personalized, continuous therapy to patients with traumatic brain injury, stroke, etc?

Trace Genomics: Can we identify microbes and genetic mutations that cause or protect crops from disease, improving yield and shelf life?

Everywhere: Data architecture —
how can the data be stored, streamed, accessed?

Many exciting questions:

Memorial Sloan Kettering Cancer Center:

Who is a good candidate for clinical trials for treatments?

Biogen: If you combine data from genomics, clinical research, environmental studies, etc, are there new approaches to treat disease?

Constant Therapy: Can an iPad app deliver highly personalized, continuous therapy to patients with traumatic brain injury, stroke, etc?

Trace Genomics: Can we identify microbes and genetic mutations that cause or protect crops from disease, improving yield and shelf life?

Everywhere: Data architecture —
how can the data be stored, streamed, accessed?

Many exciting questions:

Memorial Sloan Kettering Cancer Center:

Who is a good candidate for clinical trials for treatments?

Biogen: If you combine data from genomics, clinical research, environmental studies, etc, are there new approaches to treat disease?

Constant Therapy: Can an iPad app deliver highly personalized, continuous therapy to patients with traumatic brain injury, stroke, etc?

Trace Genomics: Can we identify microbes and genetic mutations that cause or protect crops from disease, improving yield and shelf life?

Everywhere: Data architecture —
how can the data be stored, streamed, accessed?

Tremendous potential for impact!