Careers in Data Science

Transitioning from Academia to Industry

Zach Gazak, Wafa Soofi
Insight Data Science

PIEP Workshop
October 8, 2015

INSIGHT

RICE

Georgia Institute of Technology

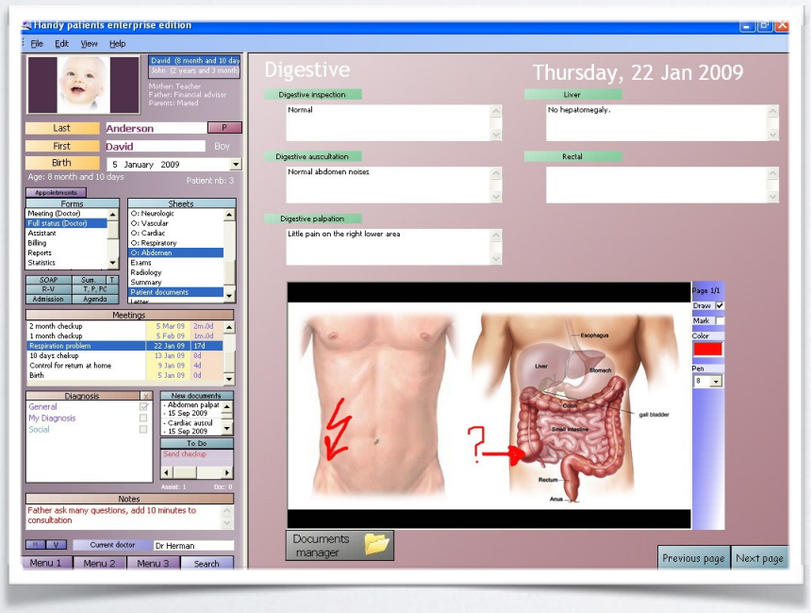
INSIGHT

Massachusetts Institute of Technology

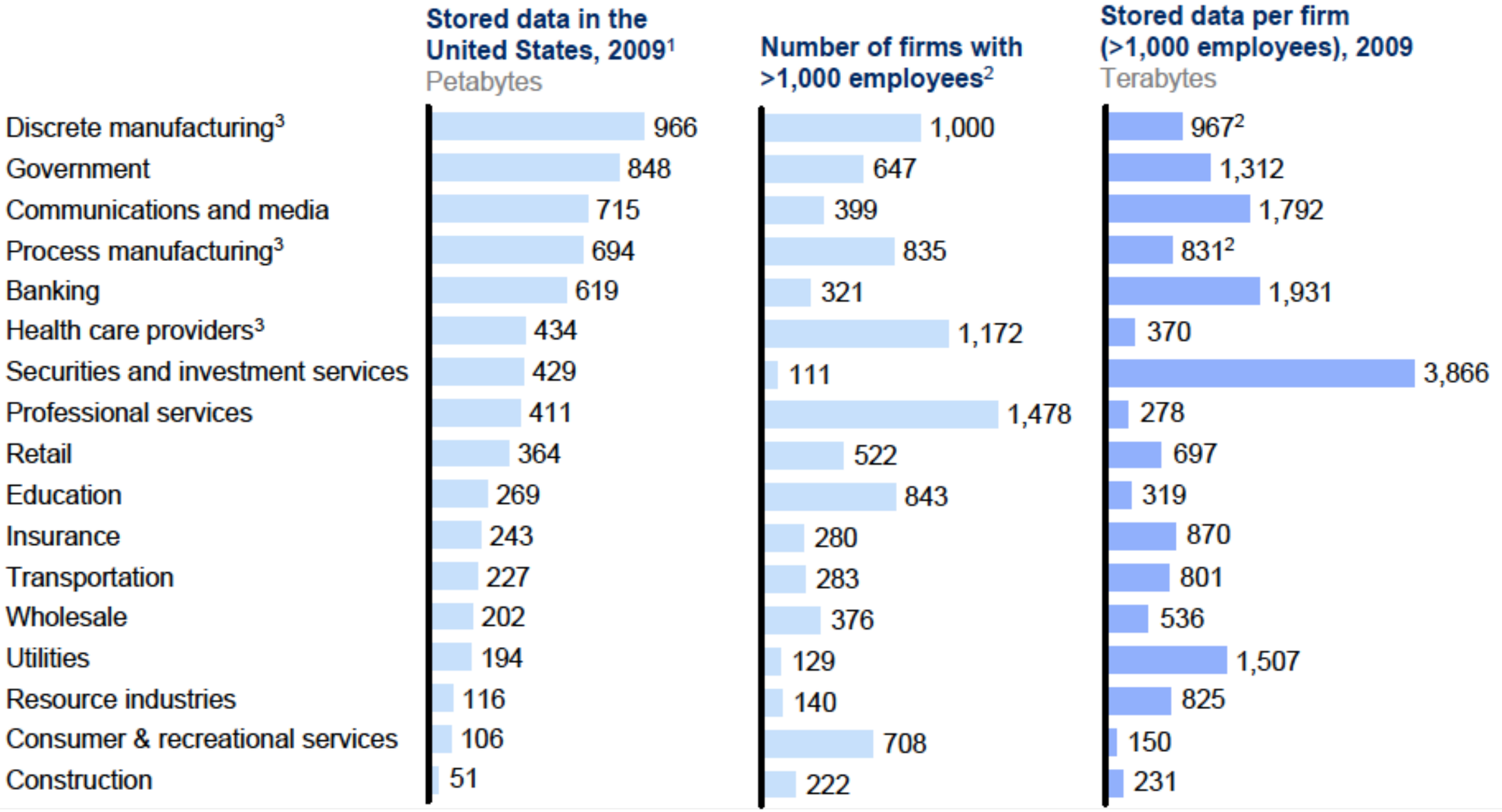
UNIVERSITY of HAWAI'I® MĀNOA

INSIGHT

Data Is Growing Exponentially



Companies in all sectors have at least 100 terabytes of stored data in the United States; many have more than 1 petabyte



SOURCE: IDC; US Bureau of Labor Statistics; McKinsey Global Institute analysis, 2011

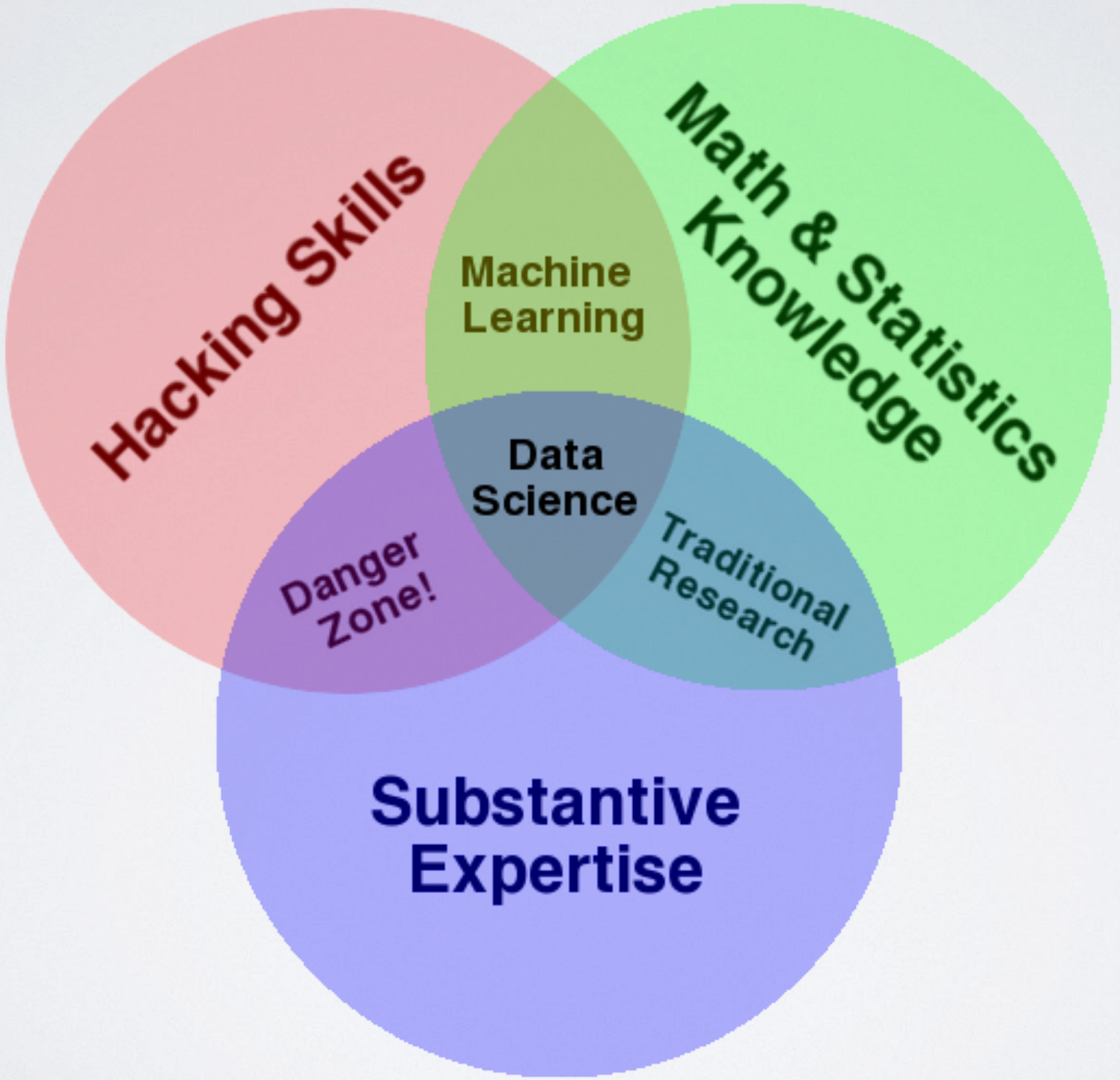
Job Trends from Indeed.com
"data science"
Percentage of Matching Job Postings
0.04
0.03
0.02
0.01
0
Jan '06
Jan '07
Jan '08
Jan '09
Jan '10
Jan '11
Jan '12
Jan '13
Jan '14
Jan '1
indeed
TM
one search. all jobs.

Source: McKinsey Global Institute analysis, 2011

"The best data scientists tend to be 'hard scientists,' rather than computer science majors."

– DJ Patil

WHAT IS DATA SCIENCE?



Drew Conway

THE GAP

1. Tools
2. Context
3. Pace
4. Priorities

BRIDGING THE GAP

PIEP POSTDOC INDUSTRY EXPLORATION PROGRAM

INSIGHT

EVIDENCE OF INTEREST

Build a data project.

Show off your:

Initiative

Technical Ability

Product Sense

GATHER INTEL

Do some informational interviews.

Ask questions like:

What is your data science stack/toolset?

Do you work on production-level tools?

How big is the data science team at your company?

CONNECT

Stay informed.

Check out:

Meetup groups (or start your own)

Hacker News

Data Science Weekly

LinkedIn

EVIDENCE OF INTEREST

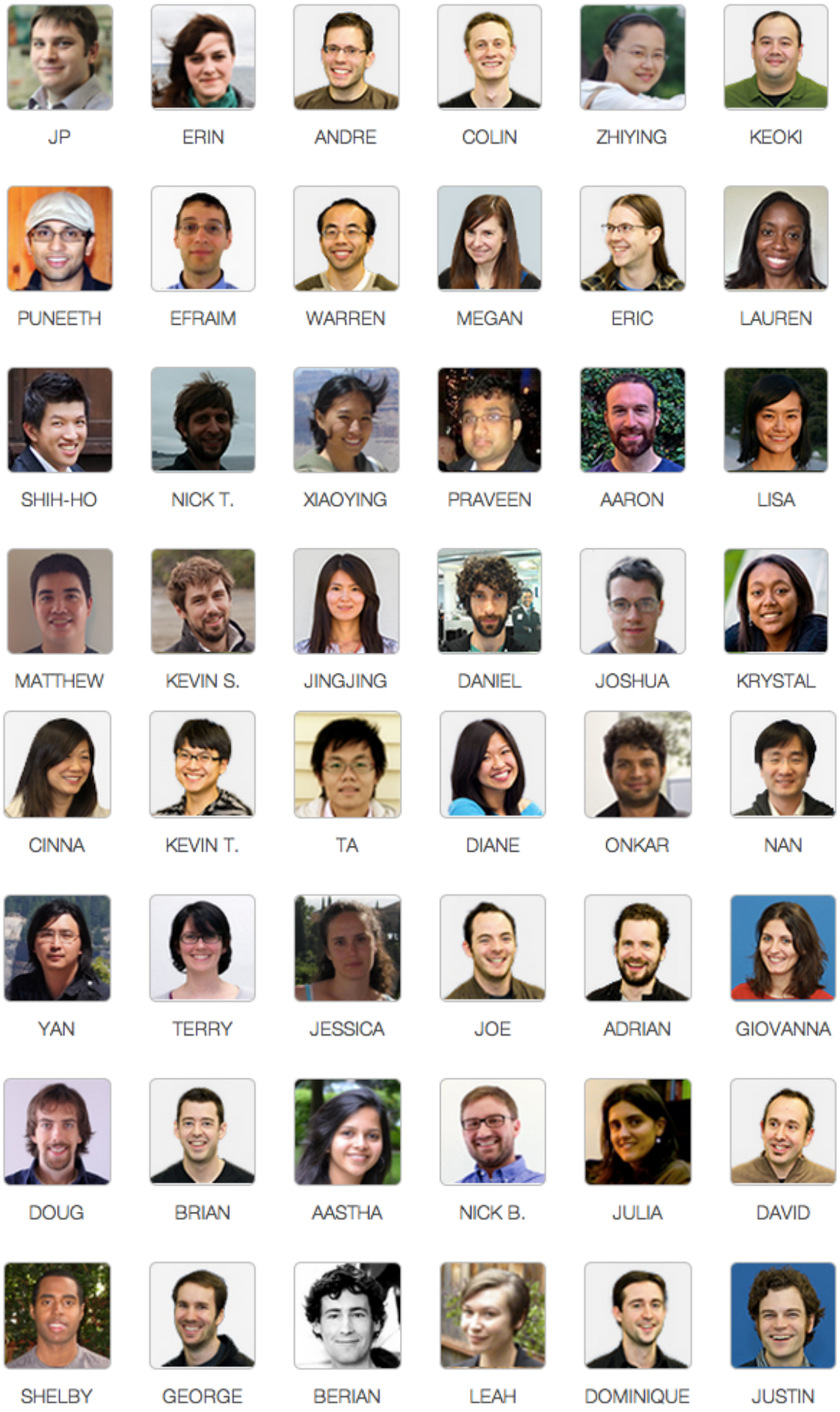
Build a data project

GATHER INTEL

Do some informational interviews

CONNECT

LinkedIn, Twitter, meetups



400+

Insight Fellows working as Data Scientists and Data Engineers

Insight Fellows are Data Scientists and Data Engineers at:

LinkedIn

facebook

twitter

Memorial Sloan-Kettering Cancer Center
The Best Cancer Care. Anywhere.

Microsoft

Square

Palantir

NBC

airbnb

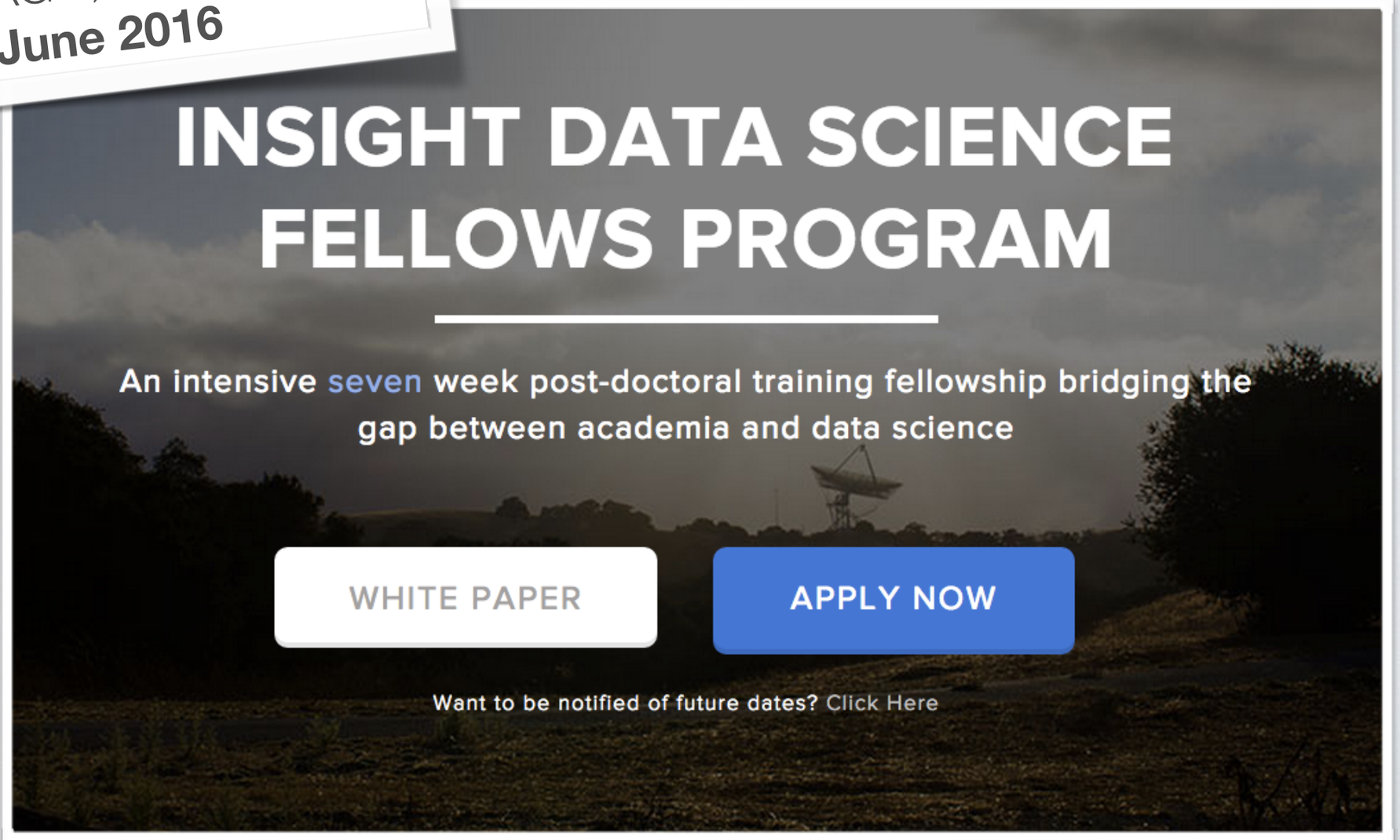
NETFLIX

intuit

tripadvisor

insightdatascience.com

Upcoming Sessions:
SILICON VALLEY
NEW YORK
SEATTLE, LA, DC, CHICAGO, & Other Cities
January 2016, June 2016



INSIGHT

Program Timeline

7 Weeks

4-8 Weeks

Week 1 — Weeks 2-4 — Weeks 5-7 → Job Interviews

Intro Week

Data Project

Company Visits

AREAS OF FOCUS

Product Analytics

Predictive Modeling (e.g. Risk, Sales)

Data Products

MANY QUESTIONS IN DATA SCIENCE

Khan Academy: Millions of math problems are attempted online per day. When is someone proficient? What is the best problem order?

The New York Times and The Wall Street Journal: What is the right payment model for journalism?

LinkedIn: Does this LinkedIn account correspond to a real person?

MANY QUESTIONS IN DATA SCIENCE

Netflix: What new shows should we make?

Memorial Sloan Kettering Cancer Center: Who is a good candidate for clinical trials?

Everywhere: What are some efficient ways in which the data can be stored, streamed, or accessed?

DATA PROJECTS

- lendinglounge.rodeo
- reviewskimmer.com
- hammerpricer.com
- beautifulcity.me
- routearoundcrime.com