Using Data Analytics in Onboarding and Employee Performance Measurement

Littler®

HR-Going For The GOLD NYS SHRM

Presented by

Jacqueline Phipps Polito
Shareholder, Littler
Rochester, NY
jpolito@littler.com

Pamela S.C. Reynolds
Associate, Littler
Rochester, NY
preynolds@littler.com

Overview

1. A brief history of prediction modeling (from 2000 BCE to 2016)
2. A briefer explanation of data science
3. An overview of data scientific applications in HR
4. A few examples of those applications
5. Values and risks of data science approach
6. Recommendations

2000 BCE

- Available data are VERY low quality; analytic tools are… not great

1956

Available data are not terrible quality; analytic tools are…

Better, but still not great

2016

- Available data are VERY high quality; analytic tools are… much much better than even 5 years ago

The Basic Progression

Expertise, knowledge, & intuition

Statistics and econometrics

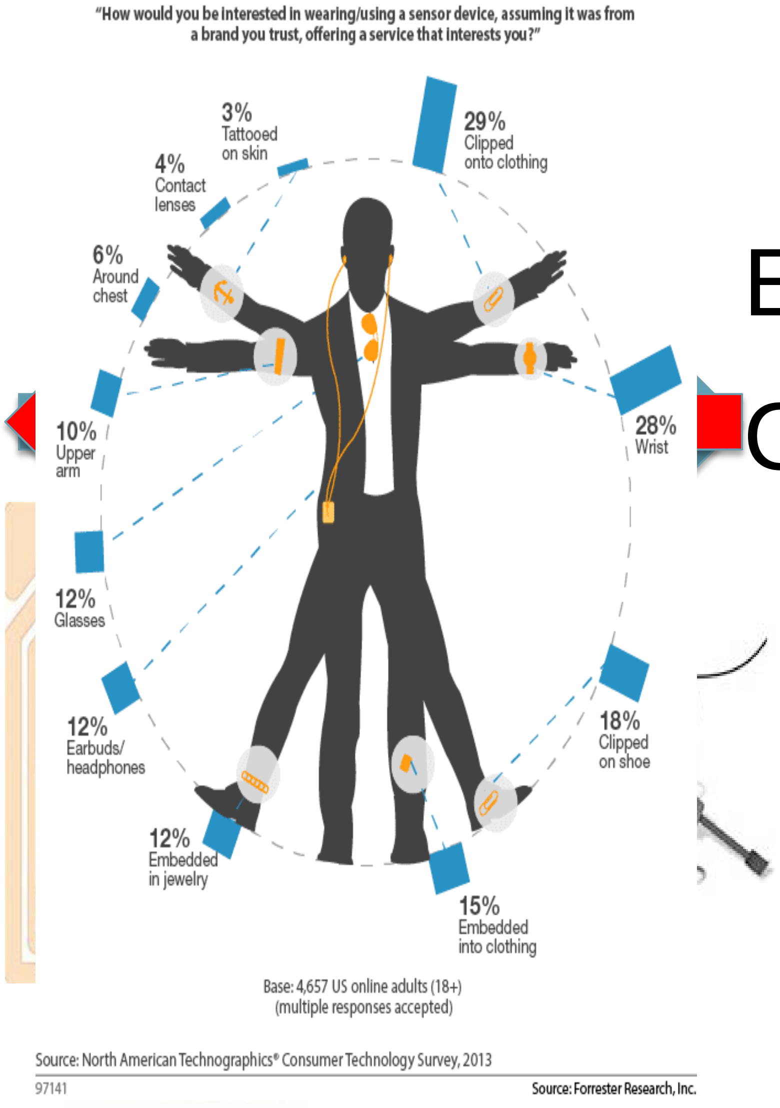
Data science, machine learning and cognitive computing

New Model of Data Flow

Employer

Employees

Consumers



Predictive Analytics are Profitable

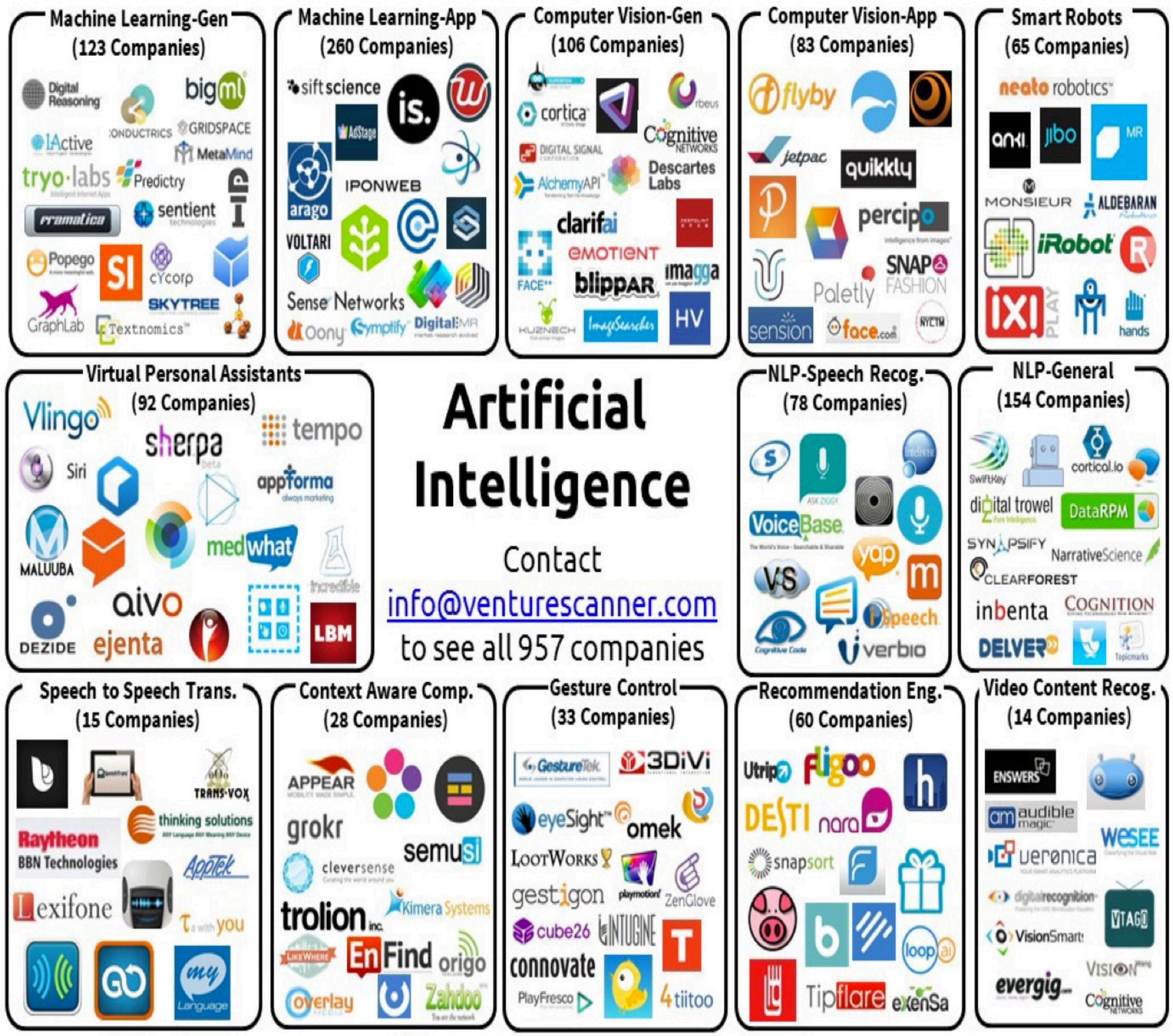
250% ROI

from business analytics solutions that incorporate predictive analytics

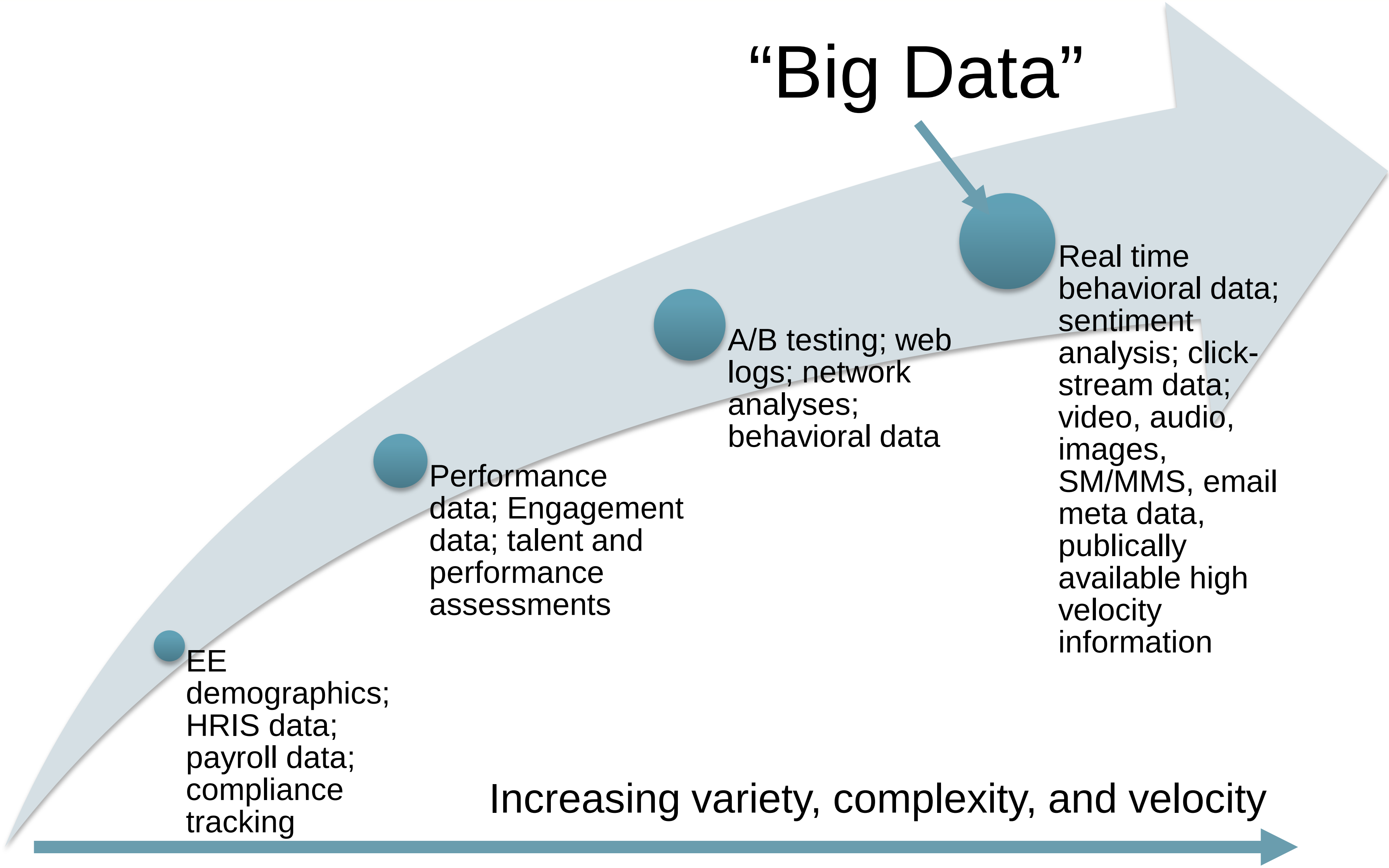
—triple the ROI that other information access and internal productivity projects offer.

Source: IDC, The Business Value of Predictive Analytics, June 2011.

Venture Scanner found 957 AI companies across 13 categories, with a combined funding of $4.8 Billion



“Big Data” are Sources of Information (Transactions, Observations, Interactions) Too Big or Fast for Traditional Analysis



4 Distinct Data Driven Services

1. Data used to <u>benchmark</u> (Why is A performing better than B?)
2. Data used for <u>recommendation</u> and <u>filter</u> systems
   - Content based recommendation engines
   - Collaborative filtering
   - Hybrid
3. Data used for <u>predictions</u>
4. Data used to describe and understand statistical relationships

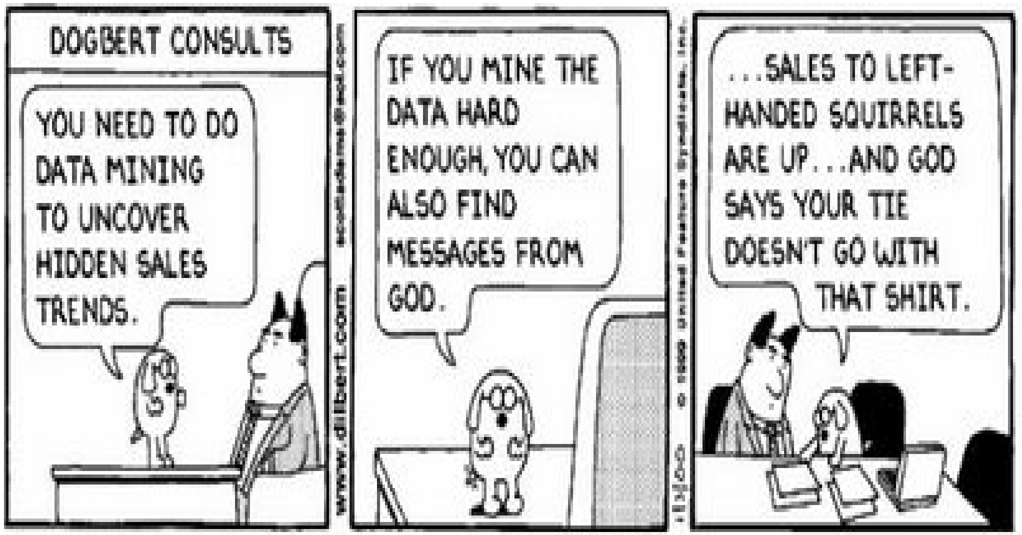
Data Mining Tasks

- **Descriptive methods**
  - Find human-interpretable patterns that describe the data
    - **Example:** Clustering

- **Predictive methods**
  - Use some variables to predict unknown or future values of other variables
    - **Example:** Recommender systems

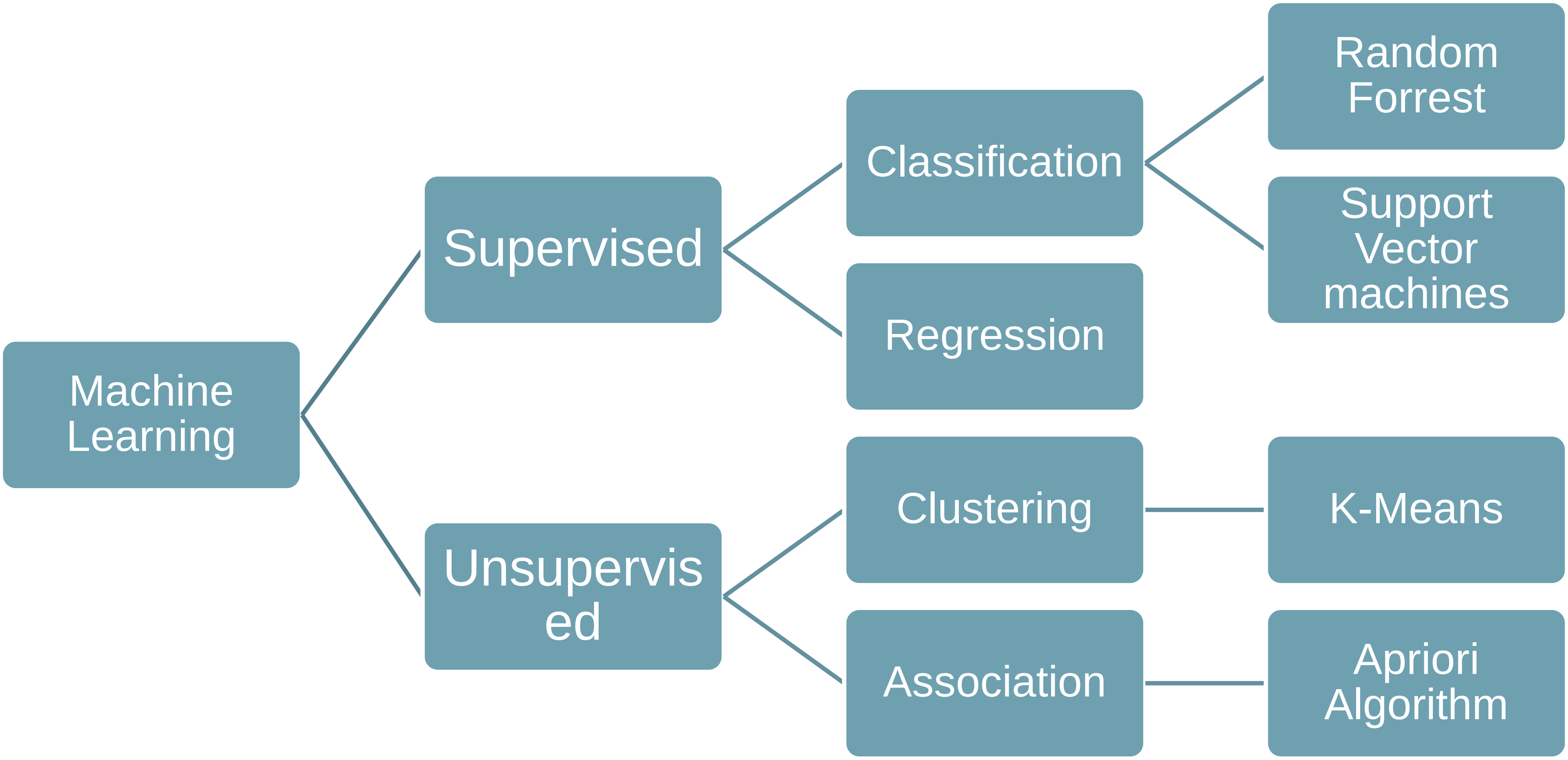
Meaningfulness of Analytic Answers

- **A risk with "Data mining" is that an analyst can "discover" patterns that are meaningless**
- Statisticians call it **Bonferroni's principle**:
  – Roughly, if you look in more places for interesting patterns than your amount of data will support, you are bound to find crap

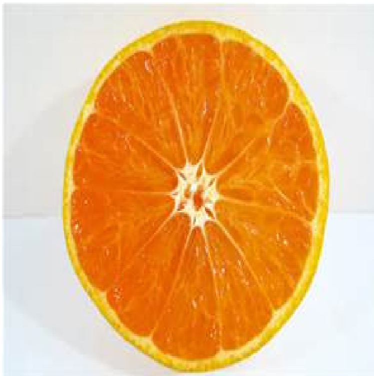


Categories
(Source: Venture Scanner)

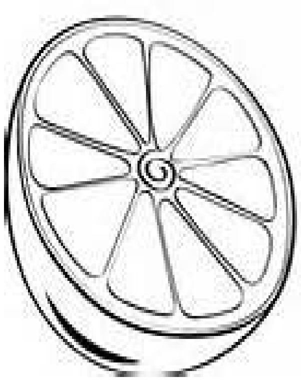
1. **Deep learning / Machine Learning (general / applications)**
2. **Natural language processing (general / speech recognition)**
3. Computer vision/image recognition (general / applications)
4. Gesture control
5. Virtual personal assistants
6. Smart robots
7. **Recommendation engines and collaborative filtering**
8. Context aware computing
9. Speech to speech translation
10. Video automatic content recognition

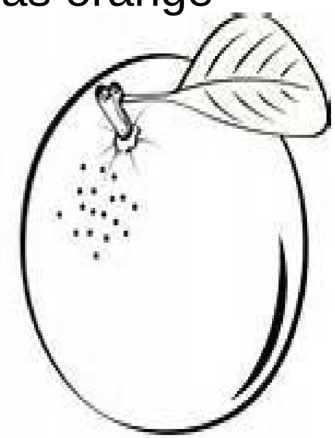

Machine Learning
Supervised
Unsupervised
Classification
Regression
Clustering
Association
Random Forrest
Support Vector machines
K-Means
Apriori Algorithm

A Seemingly Simple Example: Predict Whether the NEXT [Unseen] Image Should be Classified as an Orange or an Apple



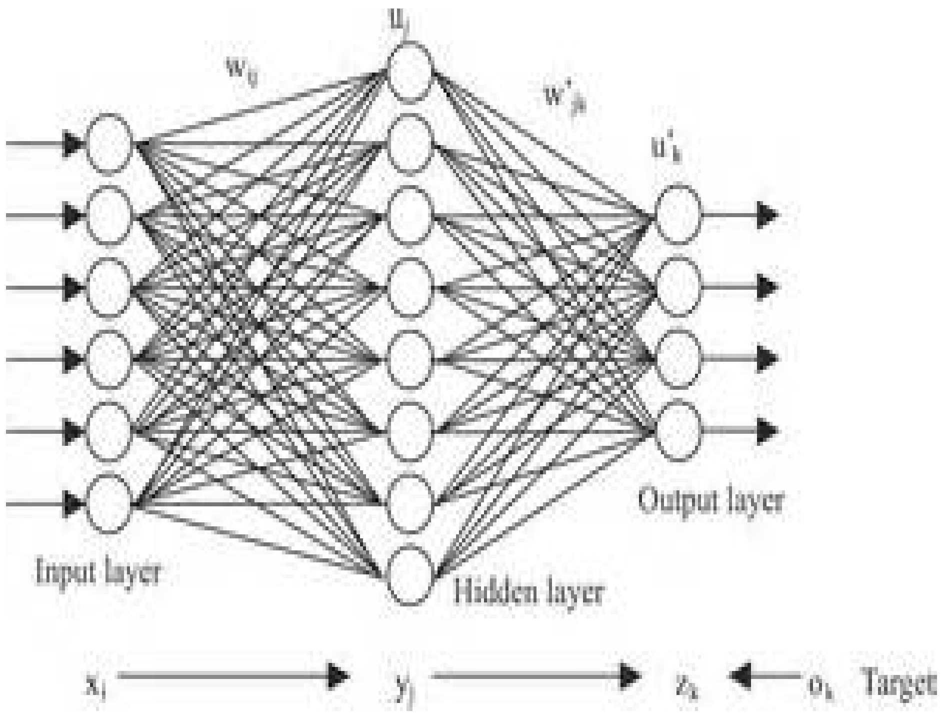
If 60% orange pixels, then classify as orange





What’s a Neural Network?

- Neuron is a rough abstraction of a human brain cell:
  - Receives input (signal)
  - Sums weighted inputs
  - Amplifiers and inhibitors
  - Pattern recognition
- Neural network: interlaced web of neurons
- “Feedfoward network” – neurons are organized into layers, with connections only between subsequent layers

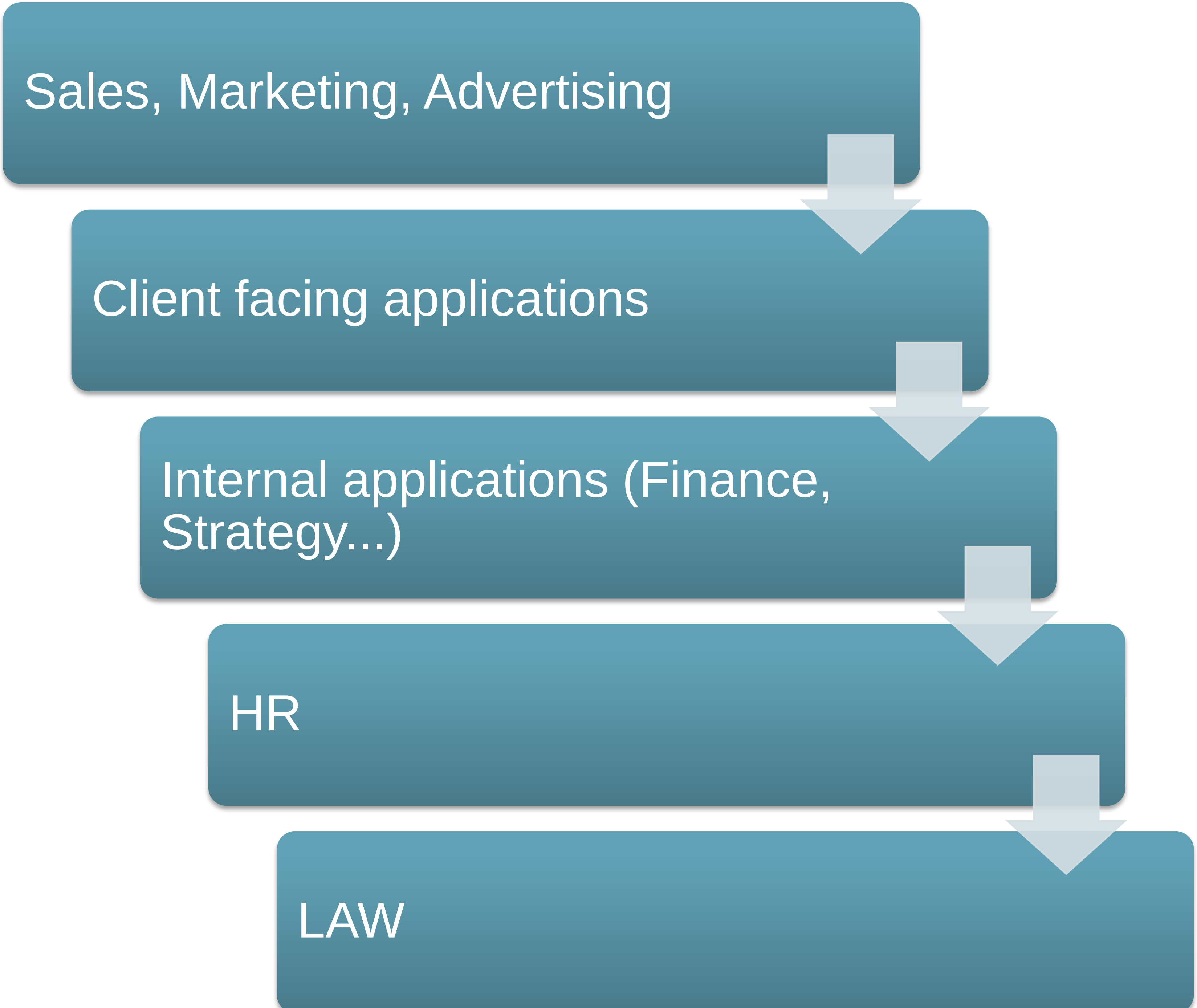


Data Scientists “Train” Neural Networks

- Forward pass: get the current estimate of the TARGET
- Backward pass: correct weights to reduce error

- $s_j^1 = \sum_i w_{i,j}^1 x_i + b_j^1;\ h_j^1 = f(s_j^1)$
- $s_k^2 = \sum_j w_{j,k}^2 h_j^1 + b_k^2;\ h_k^2 = f(s_k^2)$
- …
- $s_l^O = \sum_k w_{k,l}^{N+1} h_k^N + b_l^O;\ y_l = f(s_l^O)$

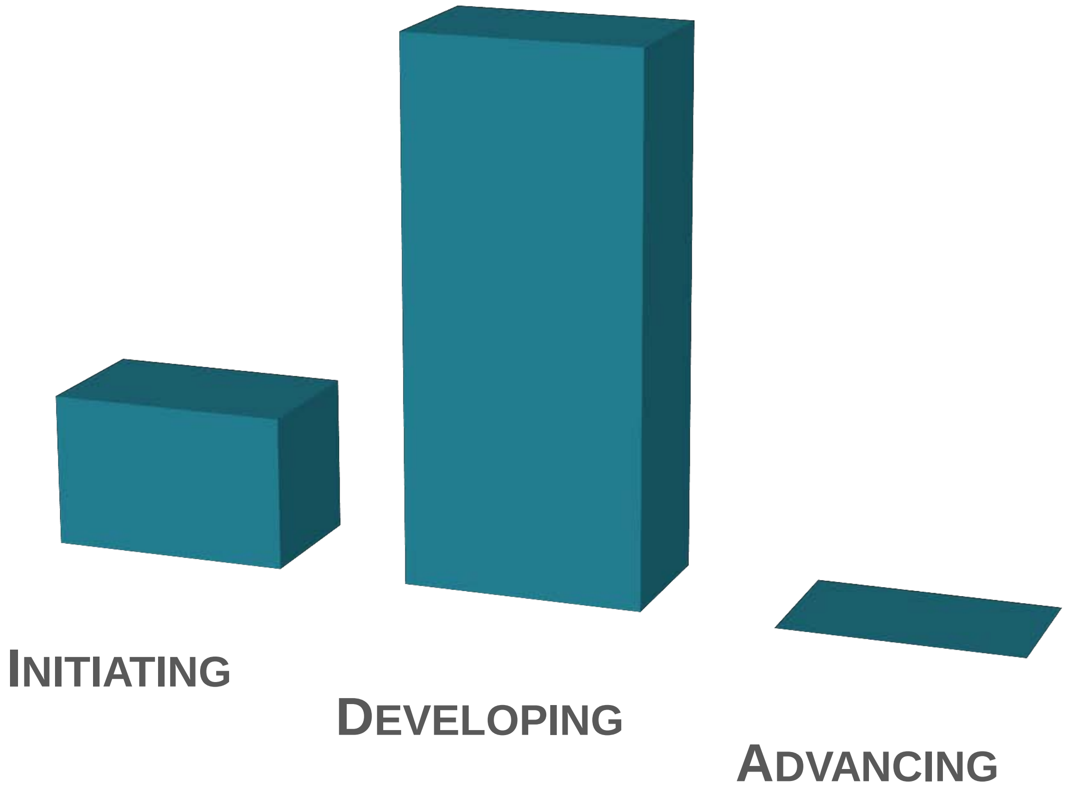
Layer	Error	Gradient (w.r.t. weights between current and prev. layer)
Output	Defined loss (e.g. $L = \sum_{i=1}^{N_o}(y_i - \hat{y}_i)^2$)	$\frac{\partial L}{\partial w_{j,l}^{(N+1)}} = \frac{\partial L}{\partial y_l} * \frac{\partial y_l}{\partial s_l^O} * \frac{\partial s_l^O}{\partial w_{j,l}^{N+1}} = \frac{\partial L}{\partial y_l} f'(s_l^O) h_j^N$
N^{th} hidden	$\delta_l^N = \frac{\partial L}{\partial y_l} * \frac{\partial y_l}{\partial s_l^O}$	$\frac{\partial L}{\partial w_{k,j}^N} = \sum_l \frac{\partial L}{\partial y_l} * \frac{\partial y_l}{\partial s_l^O} * \frac{\partial s_l^O}{\partial h_j^N} * \frac{\partial h_j^N}{\partial s_j^N} * \frac{\partial s_j^N}{\partial w_{j,l}^N} = \sum_l \delta_l^N w_{l,j}^{N+1} f'(s_j^N) h_k^{N-1}$
$(N -$	$\delta_j^{N-1} = \sum_l \delta_l^N w_{l,j}^{N+1} f'(s_j^N)$	$\frac{\partial L}{\partial w_{l,k}^{N-1}} = \sum_j \delta_j^{N-1} w_{j,k}^N f'(s_k^{N-1}) h_l^{N-2}$
…		
1^{st} hidden	δ_k^1	$\frac{\partial L}{\partial w_{i,j}^1} = \sum_k \delta_k^1 w_{k,l}^2 f'(s_j^1) x_i$


Sales, Marketing, Advertising
Client facing applications
Internal applications (Finance, Strategy...)
HR
LAW

HR Analytics are Profitable

HR organizations that “regularly use data to make talent and HR strategy decisions” generate **30% greater stock returns** than the S&P 500 over the last 3 years

78% of HR Directors Rate People Analytics in Their Organization as "Developing." The Others (22%) are in the Early Stages.



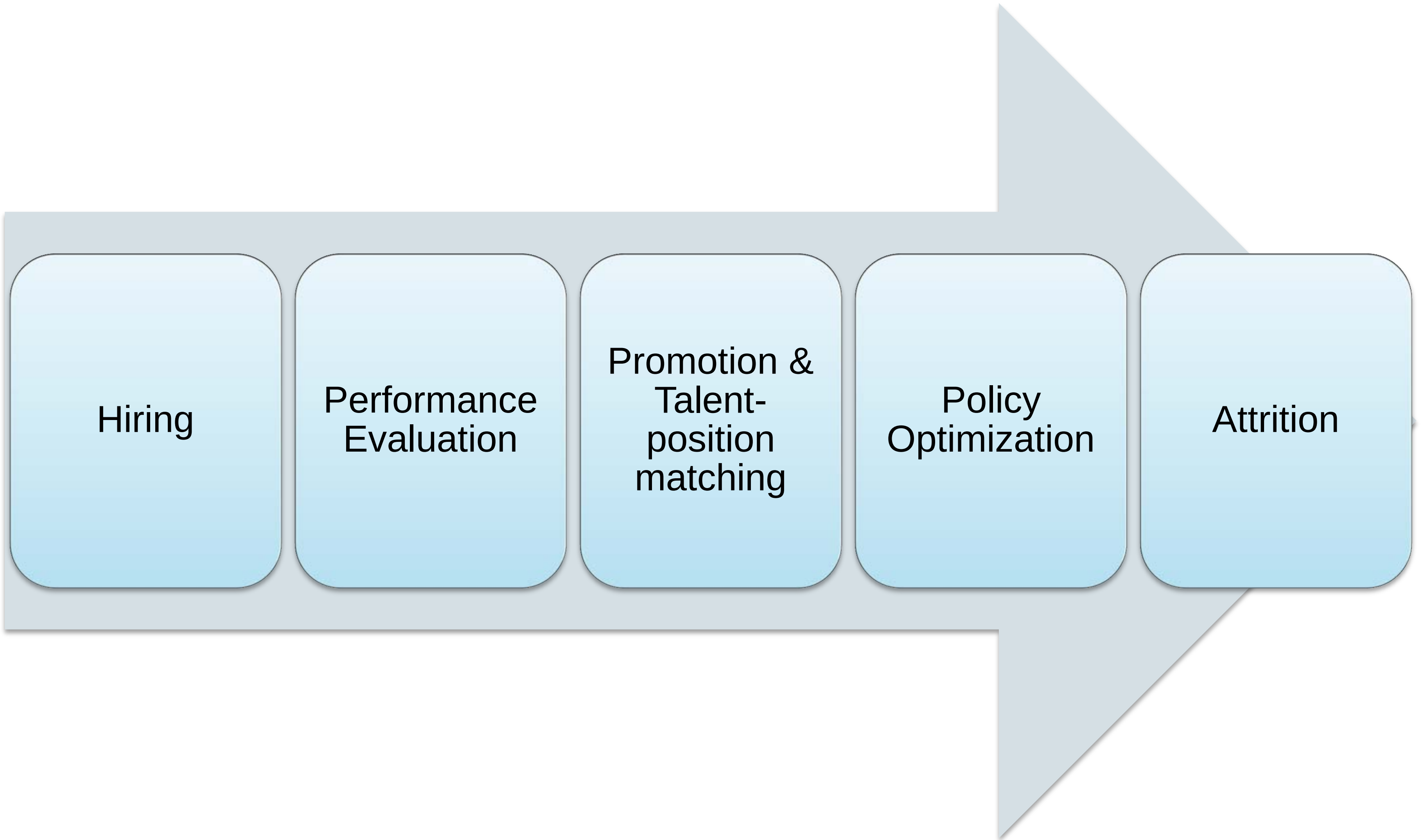
High Quality Data Sources with Predictive Power are Proliferating

Administrative & Compliance Data	Talent Management Data	Social & Behavioral Data
		Real time behavioral data Passive candidate employment and personal preferences
	Performance evaluations EE engagement results Talent assessment results	Past experiences, skills and languages
	Learning effectiveness	Professional & social networks
EE Demographics HRIS / Payroll data Compliance tracking		Email meta data; publically available EE data

Data volume & predictive power

1990 2016

5 Areas in Which Data Science Have Been Applied in HR Decision-making



Examples of Predictive Model Questions

- Which new locations are most likely to result in the ability to hire and retain the most talented employees?
- Which of 3 contemplated policies is optimal to reduce turnover, reduce workplace conflict, and improve morale?
- Which applicants pose the lowest risk of theft or abuse of paid time off?
- Which employees are likely to voluntarily quit in the next 6 months?

Hiring:
Whom Should the Employer Hire?

Data Source	Best at predicting:	Risk(s):
Video data from interviews	Ruling out bad candidates (reliability)	Replication of bias in favor of homogeneity
Unstructured text data (resumes and writing samples, etc.)	“Fit” matching; field competence;	Failure to identify value from exogenous sources/dissimilarity
Behavioral indicators (public data like LinkedIn, Twitter, Reddit)	“Fit” matching, reliability, ruling out bad candidates	Cohort effects, type I errors
Behavioral data (video game play)	Risk-taking, impulsivity, conscientiousness, patience	Unstable, expertise “bias”
User entered data (surveys, psychometrics, demographic info, etc.)	Conformity to existing benchmarks	Self-report bias

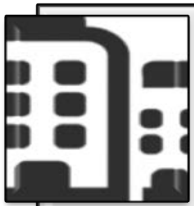
Public Data Sources for Predictive Models

Individual Data

- Demographic information
- Geographic history
- Social network data

Company Data

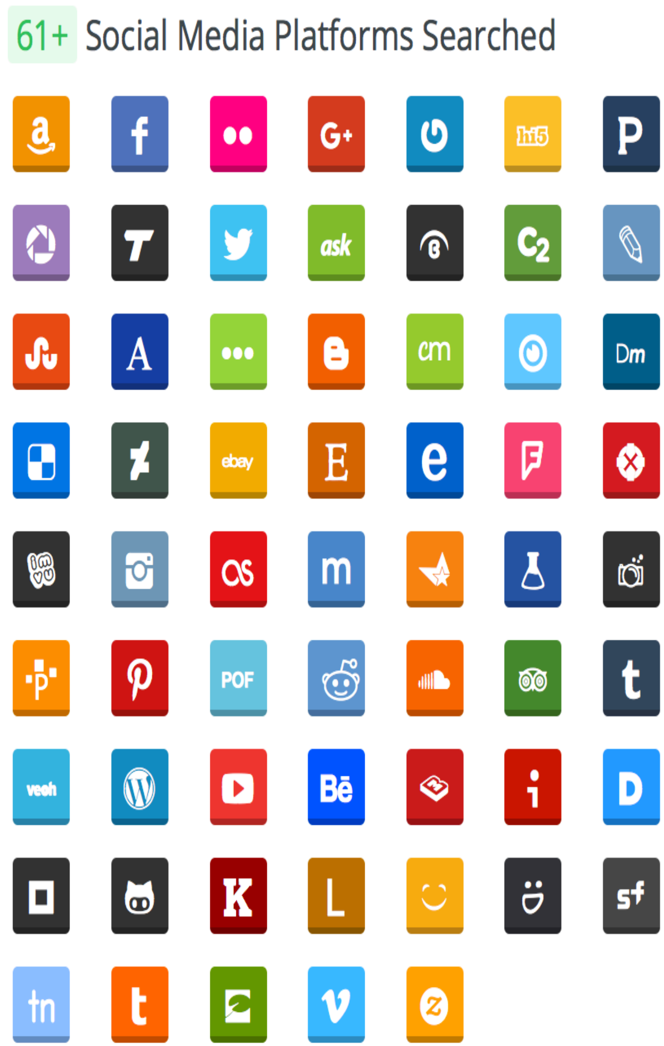
- High level indicators
- Work environment
- Status



Job Market Data

- Competition in local markets
- Labor supply/demand ratio
- Available alternatives to job

Public Data Sources Abound



I ❤ Data Sources That are Hard to Fake or System-game

"You can lie to me, you can lie to your trainer, you can even lie to yourself, but you can't lie to your Fitbit."

Public Data are Often Extremely Useful…

Jennifer Lawrence: Actors Respond to Actress's Essay on Gender Wage Gap in Hollywood

"O Jennifer Lawrence I love you so," Emma Watson shared on social media. Other actors also expressed their support for Lawrence, including Elizabeth Banks, Jessica Chastain and Bradley Cooper.

Involved in This Story

Jennifer Lawrence

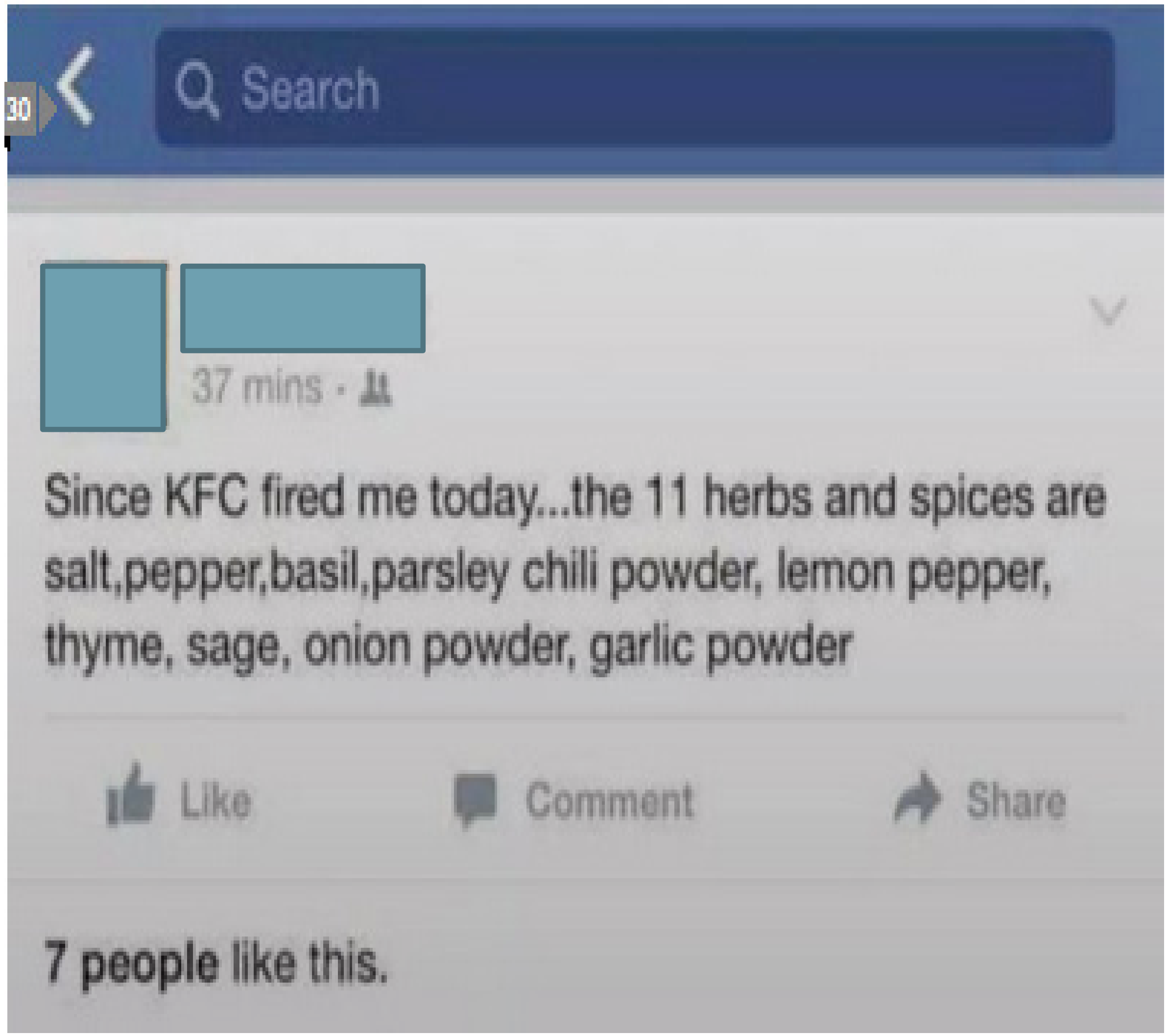
Actress wrote essay on male-female wage gap [?] · October 13 at 9:40am · Edited ·

Why Do I Make Less Than My Male Co-Stars?
(subscribe at www.lennyletter.com to read more)

When Lena first brought up the idea of Lenny to me, I was excited. Excited to speak to Lena, who I think is a genius, and excited to start thinking about what to complain about (that's not what she pitched me, it's just what I'm gonna do). When it comes to the subject of feminism, I've remained ever-so-slightly quiet. I don't like joining conversations that feel like they're "trending."...

Continue Reading

Social Media Offers a Readily Available Stage for Work-Related Expression



Hiring: Using Data Science to Improve Hiring Processes & Outcomes

Forbes / Entrepreneurs

Free Webcast: 8 Billionaire Predictions for 2016

JAN 14, 2014 @ 10:28 AM 59,584 VIEWS

Reaping The Benefits Of Diversity For Modern Business Innovation

Ekaterina Walter, CONTRIBUTOR

I write about leadership, business culture, and marketing innovation

FOLLOW ON FORBES

Opinions expressed by Forbes Contributors are their own.

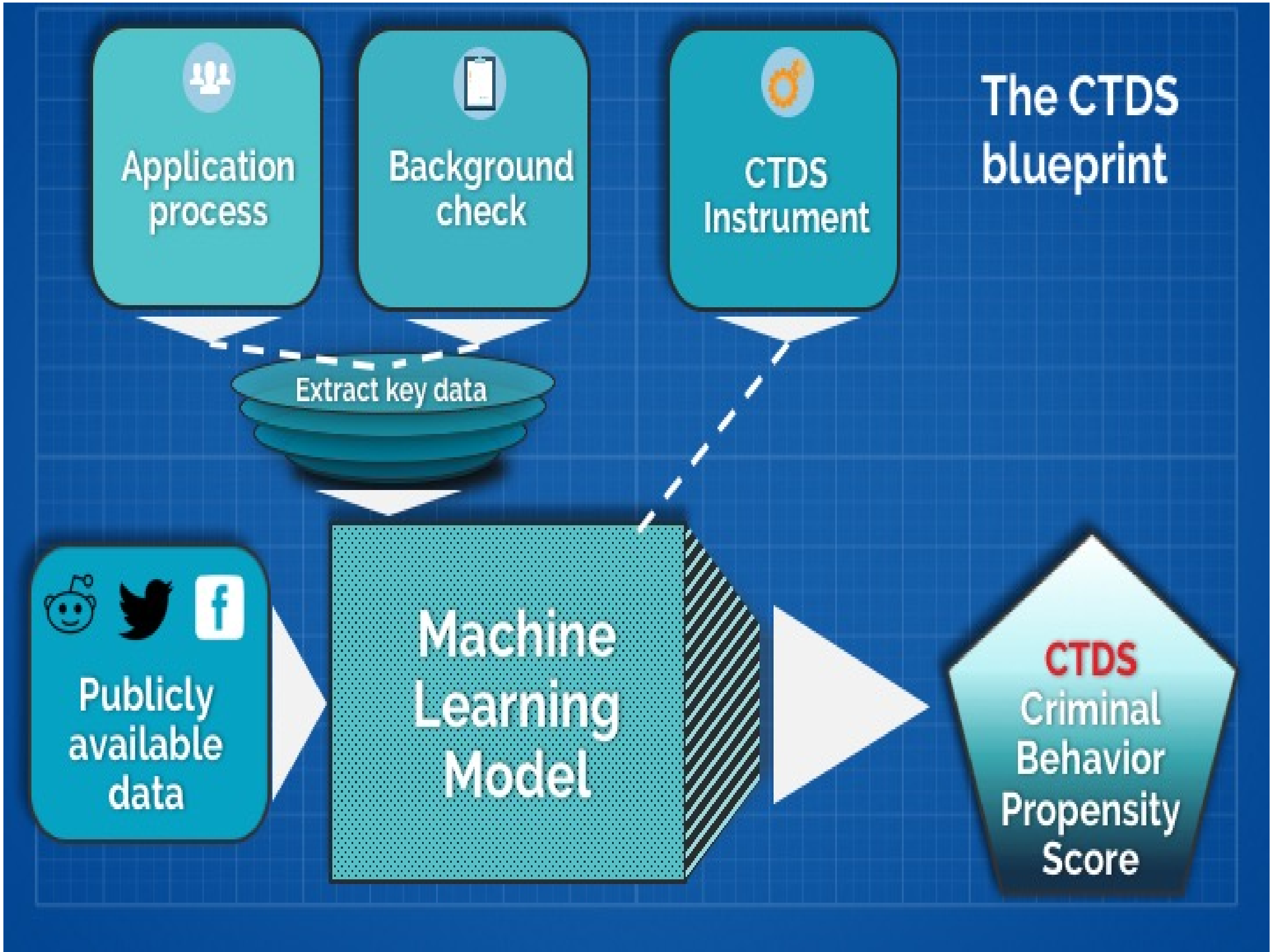
FULL BIO

Stephen Covey once said: "Strength lies in differences, not in similarities."

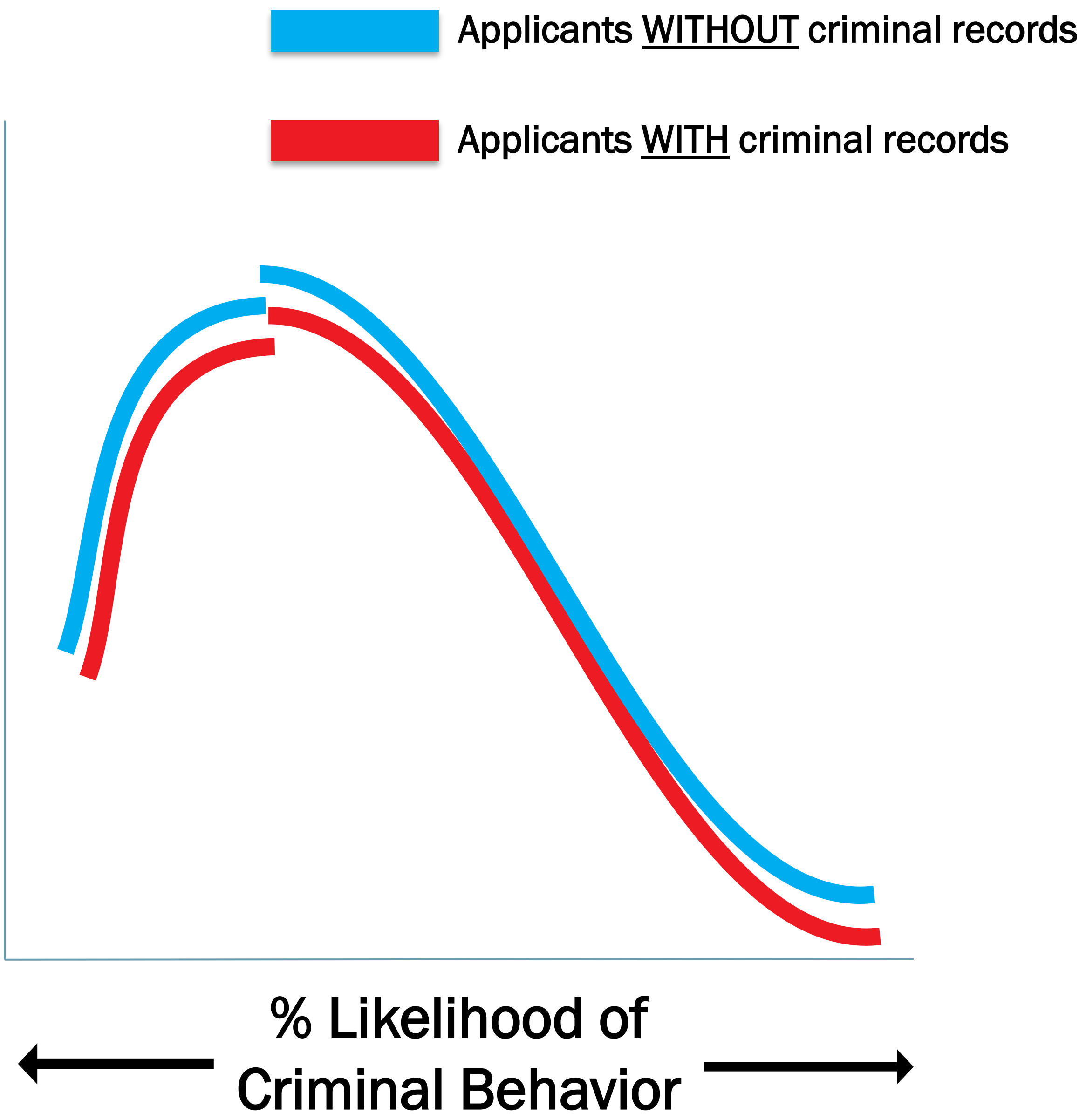
Diversity is critical for organization's ability to innovate and adapt in a fast-changing environment. Some of the most successful entrepreneurs and most admired leaders will tell you the same thing. Diversity is essential to growth and prosperity of any company: diversity of perspectives, experiences, cultures, genders, and age. Why? Because diversity breeds innovation. And innovation breeds business success. Don't believe me? Take a look:

- Forbes study has identified workforce diversity and inclusion as a key driver of internal innovation and business growth.
- Lu Hong and Scott Page showed that groups of diverse problem solvers can outperform groups of high-ability problem solvers.
- According to McKinsey, companies with diverse executive boards enjoy significantly higher earnings and returns on equity.

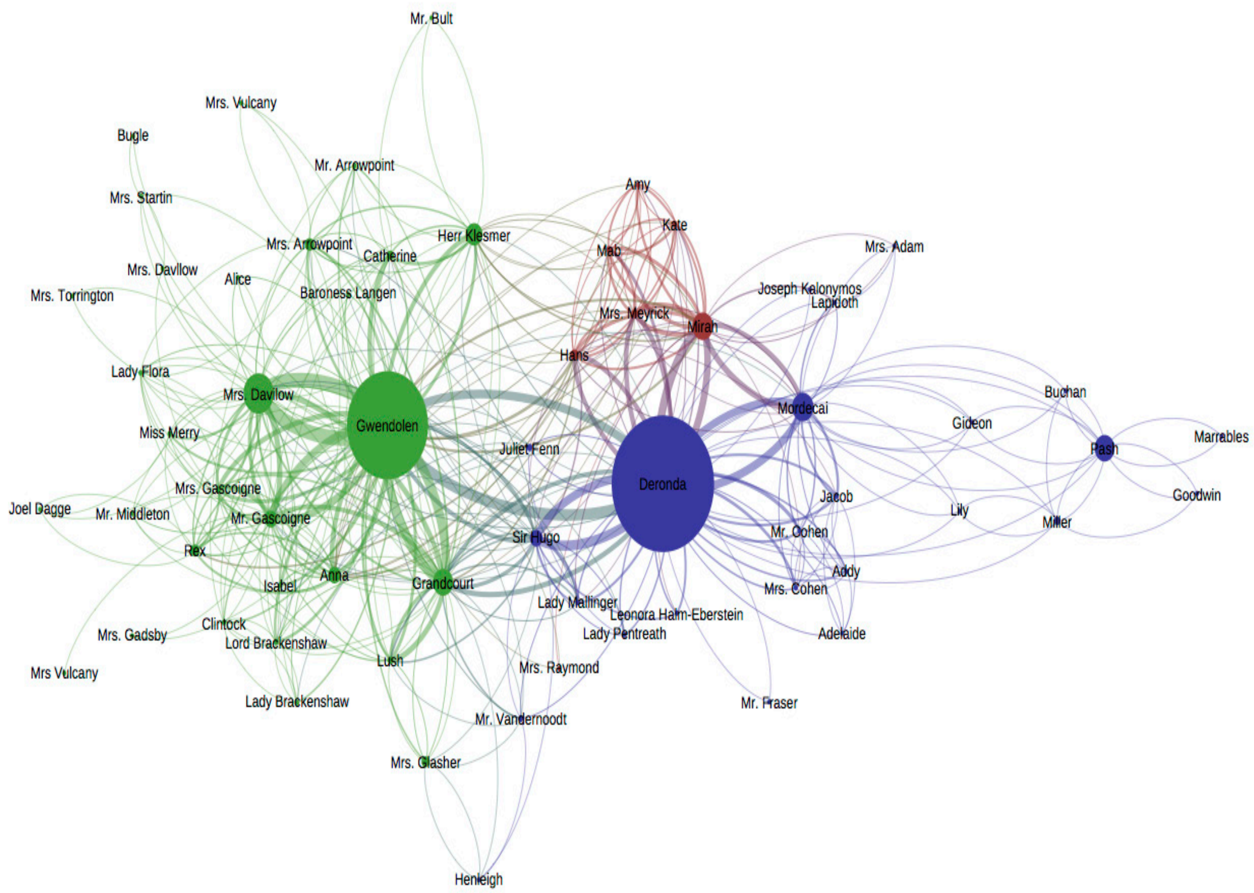
An Example: Cherry Tree Data Science



CTDS shows employers which applicants with criminal records are no more "risky" than applicants already deemed acceptable to hire.



Using Social Network Analysis ("SNA") as a Way of Measuring Employee Performance



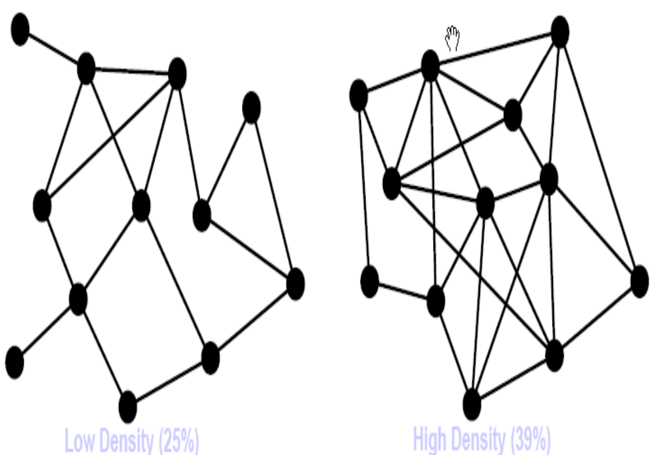
SNA Metrics Relevant to HR

Type	Description	Implications for HR
Degree centrality	# of links going into (“in degree”) or coming out of (“out degree”) a node in a network.	Id central or critical people in the knowledge flow of the org.
Density	Ratio of # of actual information ties in a network to the maximum possible ties.	Id the degree of collaborative knowledge sharing.
Cohesion	The distance, or the # of links to reach nodes in a network	Id how well distributed knowledge is in an organization.
E-I Index	A ratio of the external and internal links for particular subgroups in the network	Determines if knowledge is insular or being shared across subgroups
Ego brokerage	Measures the degree of brokerage that is occurring for different network configurations	Determines if knowledge is being disseminated in a specific manner or configuration

Network Density

$$density = \frac{l}{n(n-1)/2}$$

where *n* = number of nodes
l = number of lines (ties)

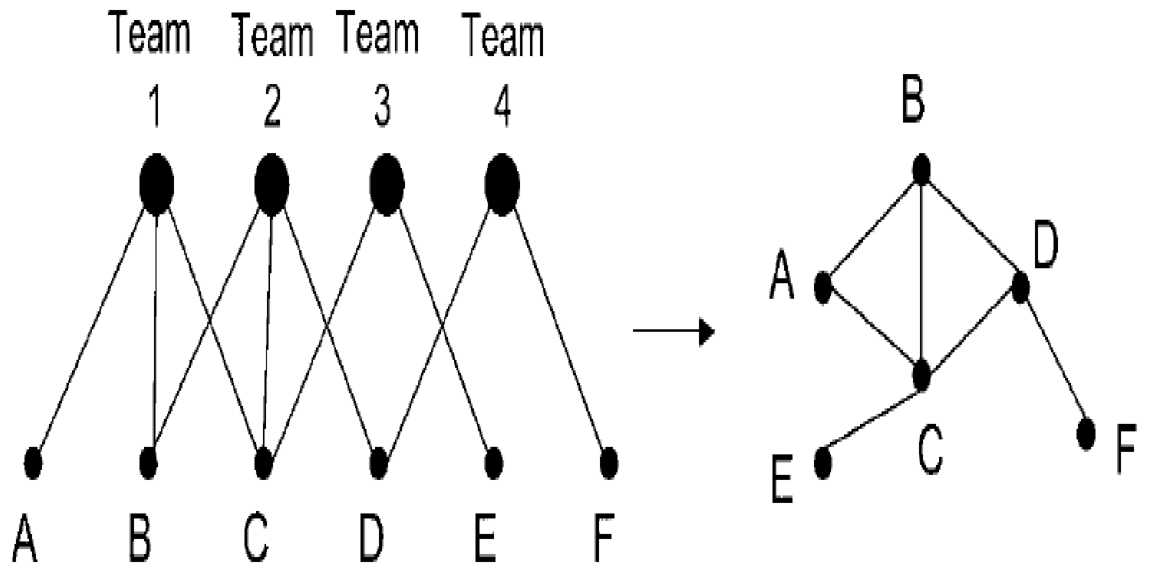


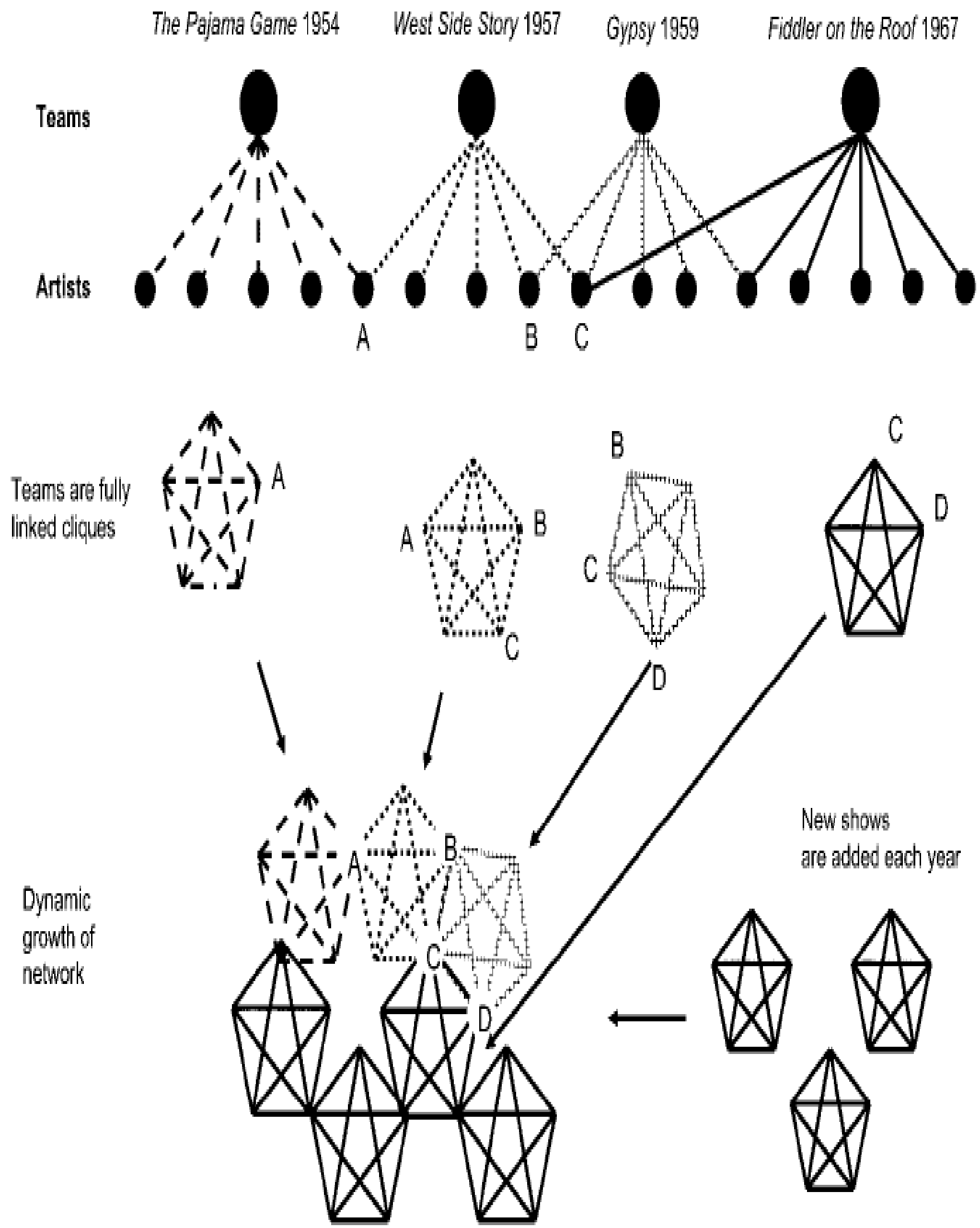
Low High

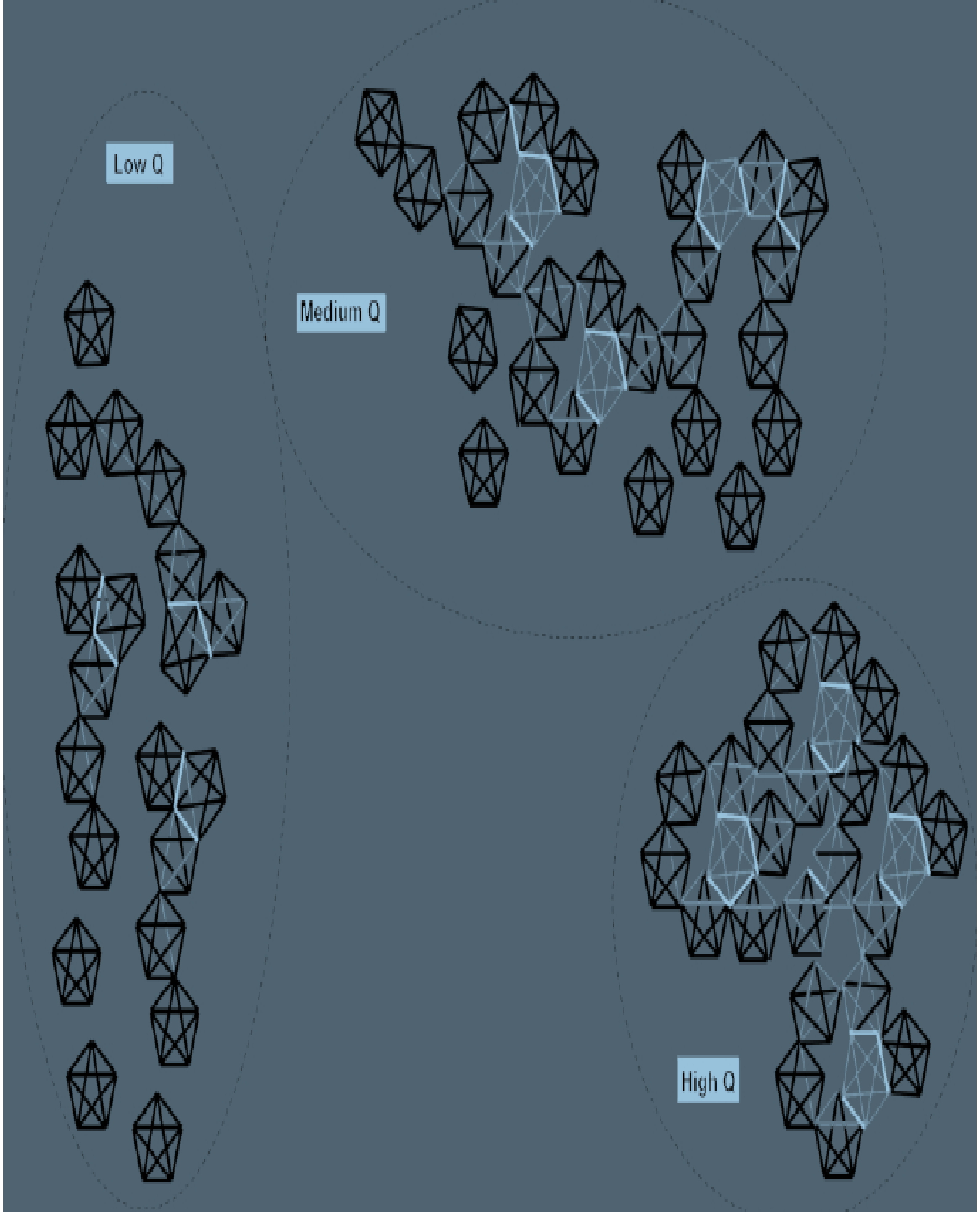
The actual number of connections in the network as a proportion of the total possible number of connections.
0 > Density ≦ 1

What's the Relationship Between Network Density and Performance Measures?

- Data: 2,092 people who worked on 474 musicals between 1945 and 1989.




The Pajama Game 1954
West Side Story 1957
Gypsy 1959
Fiddler on the Roof 1967
Teams
Artists
A
B
C
Teams are fully linked cliques
A
A
B
C
B
C
D
C
D
New shows are added each year
Dynamic growth of network
A
B
C
D


Low Q
Medium Q
High Q

Micro-Level Example: Policy Optimization

MISQ Archivist

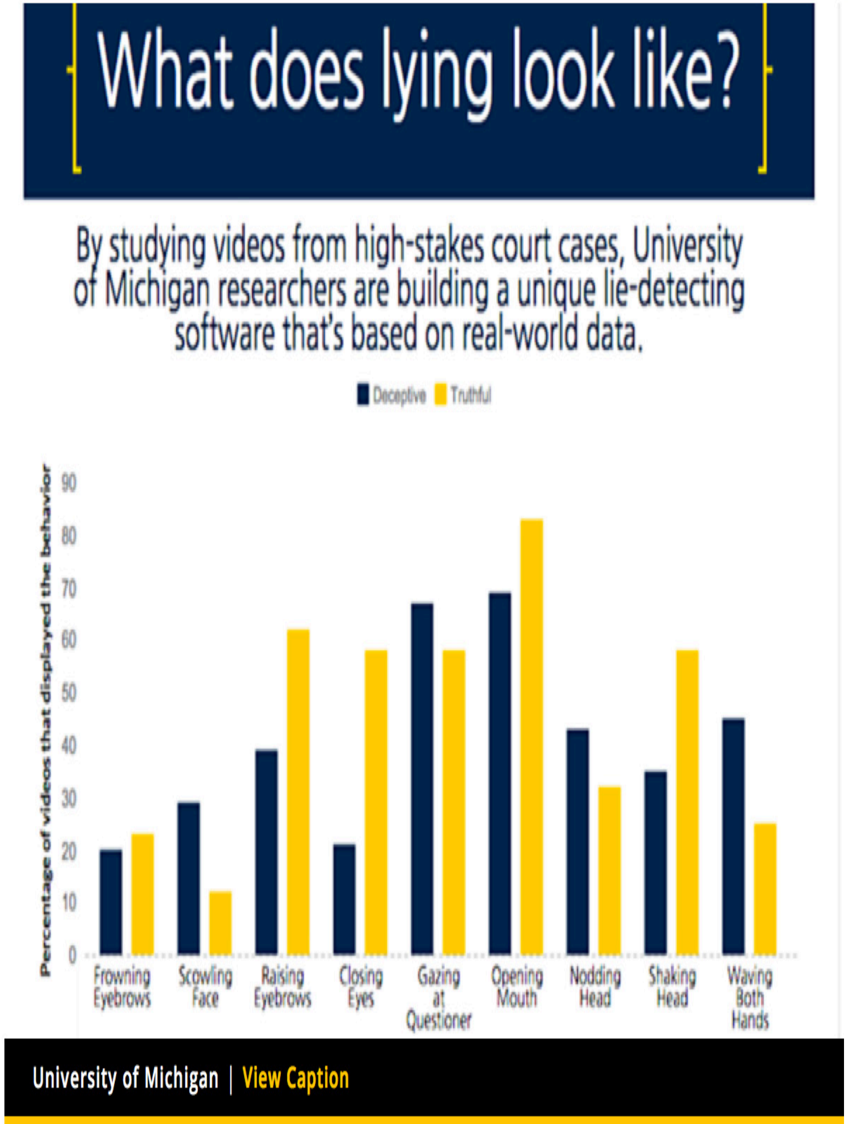
Inferring Negative Emotion from Mouse Cursor Movements

*Martin Hibbeln, Jeffrey L. Jenkins, Christoph Schneider,
Joseph S. Valacich, and Markus Weinmann*

Abstract

Experiencing negative emotion during system use can adversely influence important user behaviors, including purchasing decisions, technology use, and customer loyalty. The ability to easily assess users' negative emotions during live system use, therefore, has practical significance for the design and improvement of information systems. We utilize attentional control theory to explain how mouse cursor movements can be a real-time indicator of negative emotion. We report three studies. In Study 1, an experiment with 65 participants from Amazon's Mechanical Turk, we randomly manipulated negative emotion and then monitored participants' mouse cursor movements as they completed a number-ordering task. We found that negative emotion increases the distance and reduces the speed of mouse cursor movements during the task. In Study 2, an experiment with 126 participants from a U.S. university, we randomly manipulated negative emotion and then monitored participants' mouse cursor movements while they interacted with a mock e-commerce site. We found that mouse cursor distance and speed can be used to infer the *presence* of negative emotion with an overall accuracy rate of 81.7 percent. In Study 3, an observational study with 80 participants from universities in Germany and Hong Kong, we monitored mouse cursor movements while participants interacted with an online product configurator. Participants reported their level of emotion after each step in the configuration process. We found that mouse cursor distance and speed can be used to infer the *level* of negative emotion with an out-of-sample R^2 of 0.17. The results enable researchers to assess negative emotional reactions during live system use, examine emotional reactions with more temporal precision, conduct multimethod emotion research, and create more unobtrusive affective and adaptive systems.

Micro-Behavioral Example: Predicting Lying



Attrition: Who is Likely to Voluntarily Quit?

Data Source	Analytic Concern(s):	Legal Risk(s):
Behavioral data (email, RFID, etc.)	Undetectable systematic variation	Discrimination, privacy
Behavioral indicators (public data like LinkedIn, Twitter, Reddit)	Cohort effects; interpretation of missing data; system-gamable(?)	Discrimination, NLRA issues
User entered data (surveys, psychometrics, demographic info, etc.)	System-gamable (type I and type II errors); validity, failure to detect mediating effects, endogeneity	Discrimination
Social Network Analysis Data	Unreliability of self-reported data; methodological problems	Discrimination

An Example: HiQ Labs

What is the Value of a Data-Analytic Approach to Organizational Decision-Making?

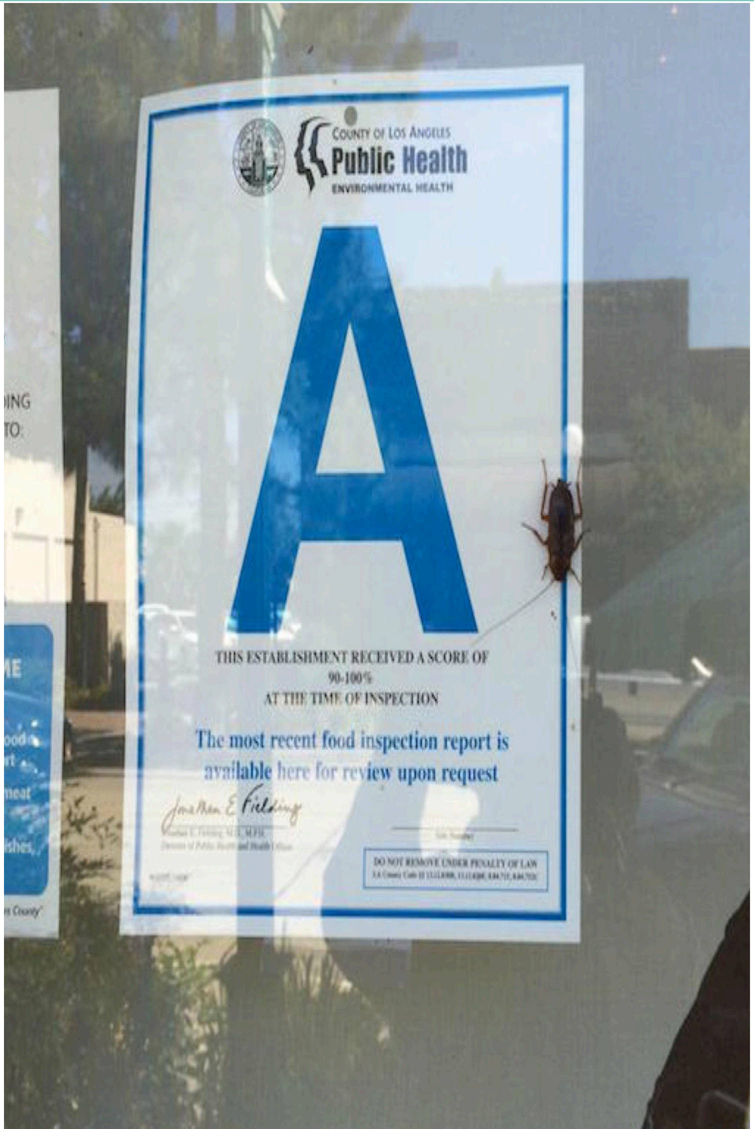
1. Increases the <u>availability</u> of information on which to rely for decisions
2. Increases the <u>reliability</u> of information drawn from diffuse sources
3. Increases the <u>accuracy</u> of decision-making criteria (by applying algorithms to tether hiring criteria to objective performance measures)
4. **<u>Reduces bias</u> (both illegal bias and bias that is not illegal but is inefficient)**

What are the risks of a Data-Analytic Approach to HR Decision-Making?

1. Analytics done incorrectly yield sub-optimal, erroneous, and costly decisions:
   - Type-I errors (erroneous conclusion that risk exists when it doesn't)
   - Type-II errors (failure to properly detect risk)
2. Reports not protected by attorney-client privilege may result in discoverable "smoking gun" evidence in litigation.
3. "Insider" analytics may yield erroneous or self-interested results.

The Drawback of “Insider” Analytics

- “4 out of 5” dentists problem.
- Agency costs.
- Difficulty seeing problems in one’s own organization.
  - Example: hidden bias



Analytics is an Excellent Tool for Uncovering Bias

"But the biggest problem isn't their policies, it's their managers' unwitting preferences. Can any company be immune?"

The War Over Unconscious Bias

Wal-Mart and others are facing class actions for job discrimination. But the biggest problem isn't their policies, it's their managers' unwitting preferences. Can any company be immune? *BY ROGER PARLOFF*

Last February a federal appeals court panel in San Francisco decided, 2-1, to allow the largest class action employment discrimination case ever convened to go forward against Wal-Mart Stores. The class includes the more than two million women who have worked at any of the company's more than 4,000 retail stores nationwide since Dec. 26, 1998.

The case, known as *Dukes v. Wal-Mart*, accuses the retailer of discouraging the promotion of women store employees to managerial positions and of paying them less than men across all job positions. The suit seeks changes in the company's internal procedures, more than $1 billion in backpay, and punitive damages.

Wal-Mart denies any wrongdoing and asserts that it has put "enormous resources in seeking to have a diverse workplace and to make sure that women and

Do managers favor employees who look like them and ignore others?

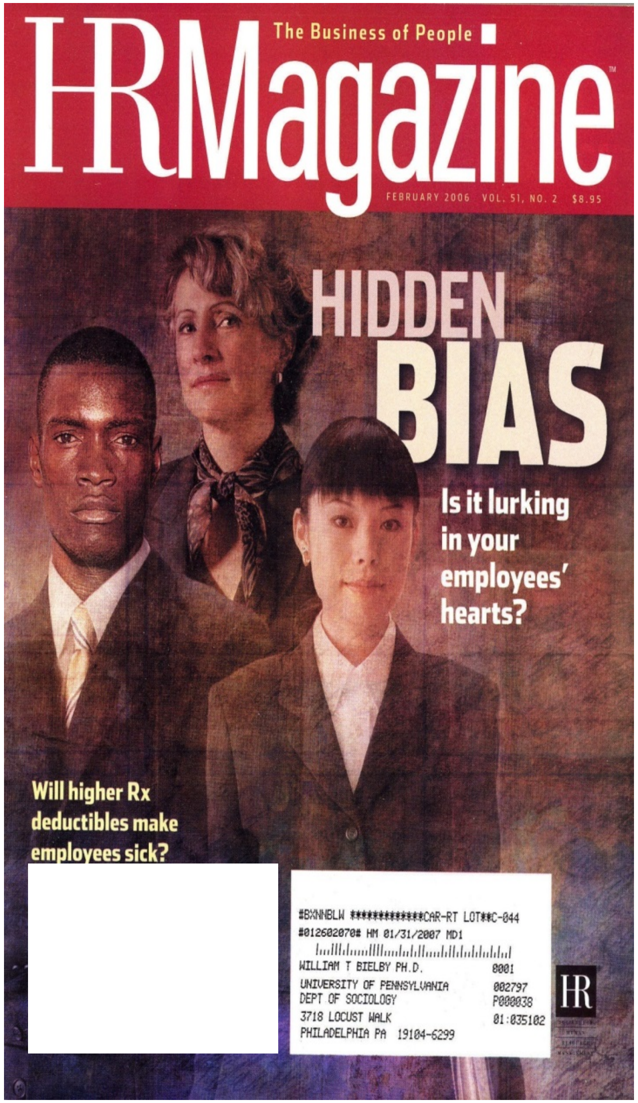
90 • FORTUNE

ILLUSTRATIONS BY BRAD HOLLAND

Analytics is an Excellent Tool for Uncovering Bias

“you may not see it, but it’s probably lurking among your managers – and perhaps even you.”

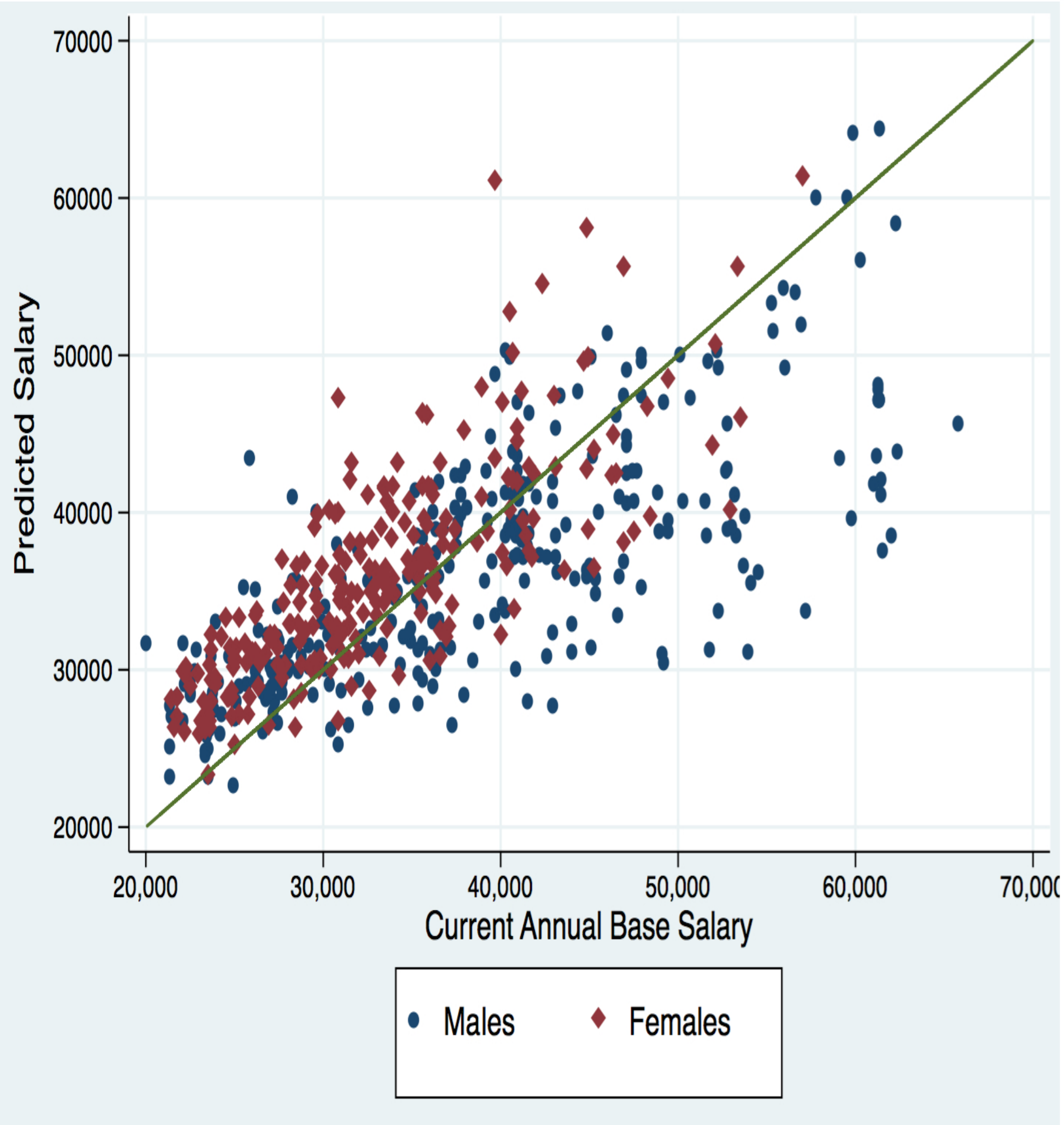
“Hidden bias can negatively affect your employees, your customers, and your business.”



Big Data Offers Great Value, But May Pose Huge Risks

- FTC advises that companies “should review these laws and take steps to ensure their use of big data analytics complies with the discrimination prohibitions that may apply.”
  - ECOA
  - Title VII
  - ADA
  - ADEA
  - FHA
  - GINA

Managing Uncovered Risk



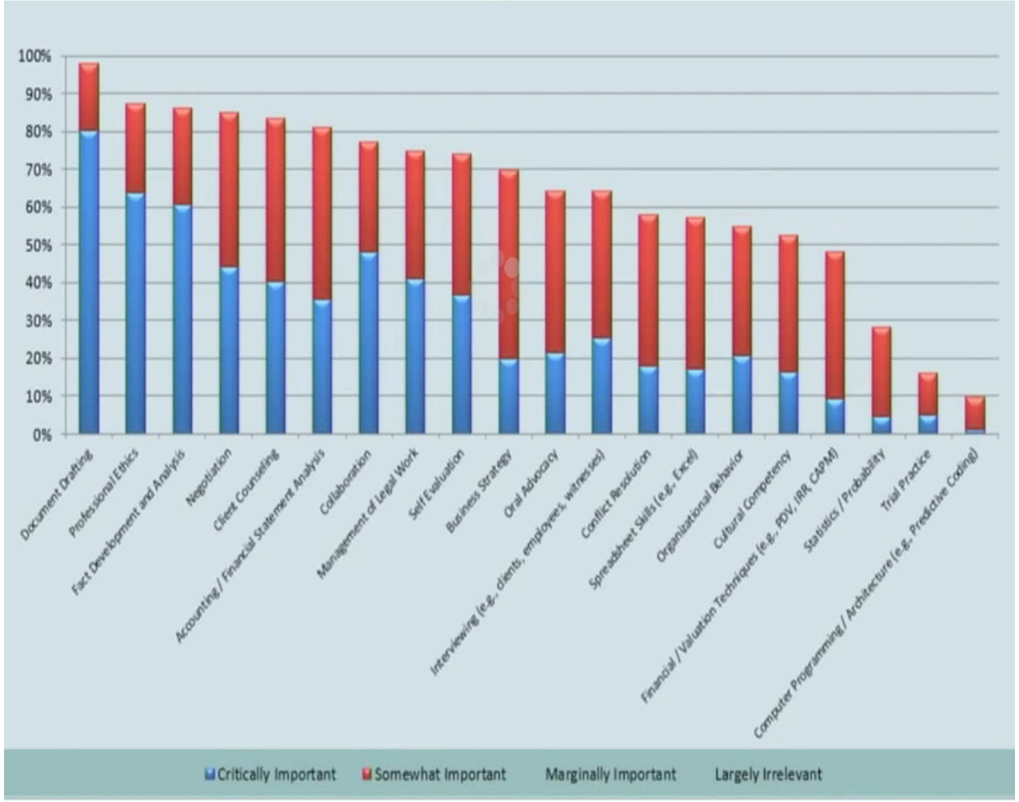
Privacy Issues

- “Big data will face huge challenges around privacy, especially with the new privacy regulation by the European Union. Companies will be forced to address the “elephant in the room” around their privacy controls and procedures.” (Forbes, 3/15/2016)
- Gartner Inc. predicts that by 2018, 50% of business ethics violations will be related to data.
- We predict that by 2020, most privacy causes of action will be related to data.

Recommendations

- Ensure that multiple data sources are examined and cross-validated.
- Don’t replicate one uniformly dictated approach with another. Test and re-test.
- Integrate data scientific approaches into decision making processes, but avoid replacing discretion with algorithms wholesale.
- Legal risk is VERY costly, and often personally costly. Avoid/reduce it by ensuring legal risk evaluation at 2 process points:
  - Feature identification / Model building
  - Report generation: don’t create “smoking gun” documents
  - Taking action on recommendations/ output
- **<u>Remember: Data are NEVER protected by attorney-client privilege. Reports generated internally in the ordinary course of business are often NOT protected by attorney-client privilege.</u>**

- In light of your area of expertise, please evaluate the following list of skills/competencies for new lawyers. How important is it for a lawyer to have achieved core competency in each skill in her/her first 2-3 years.

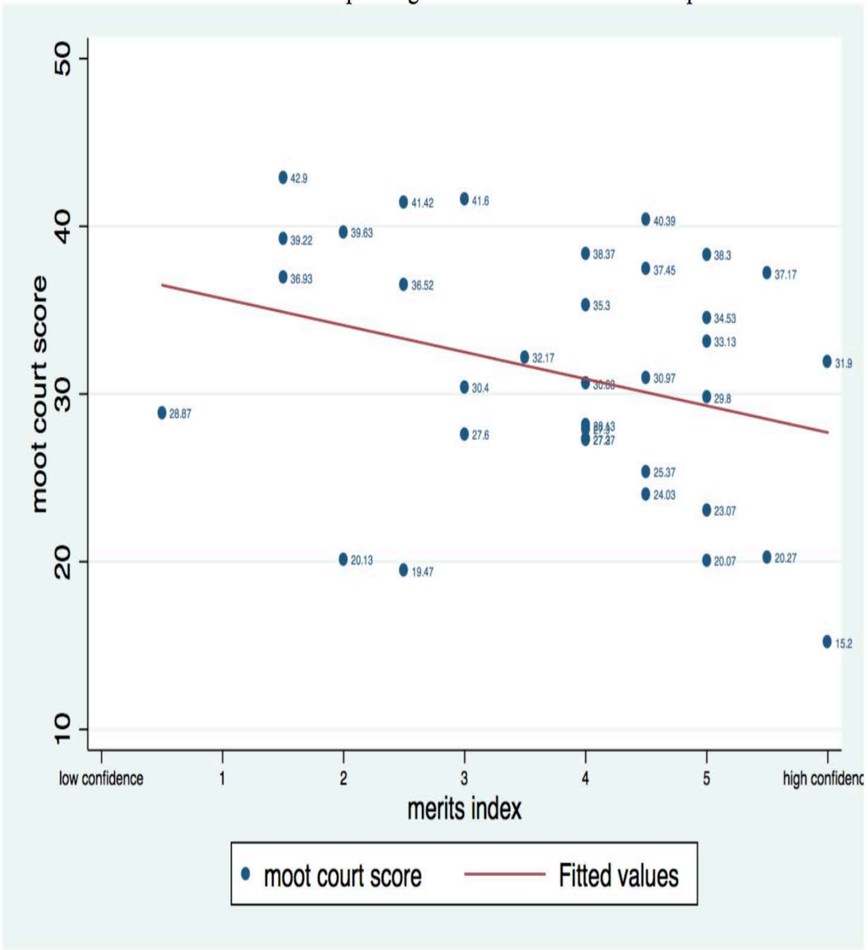


Lawyers' Over-Optimism Can Get Them in Trouble

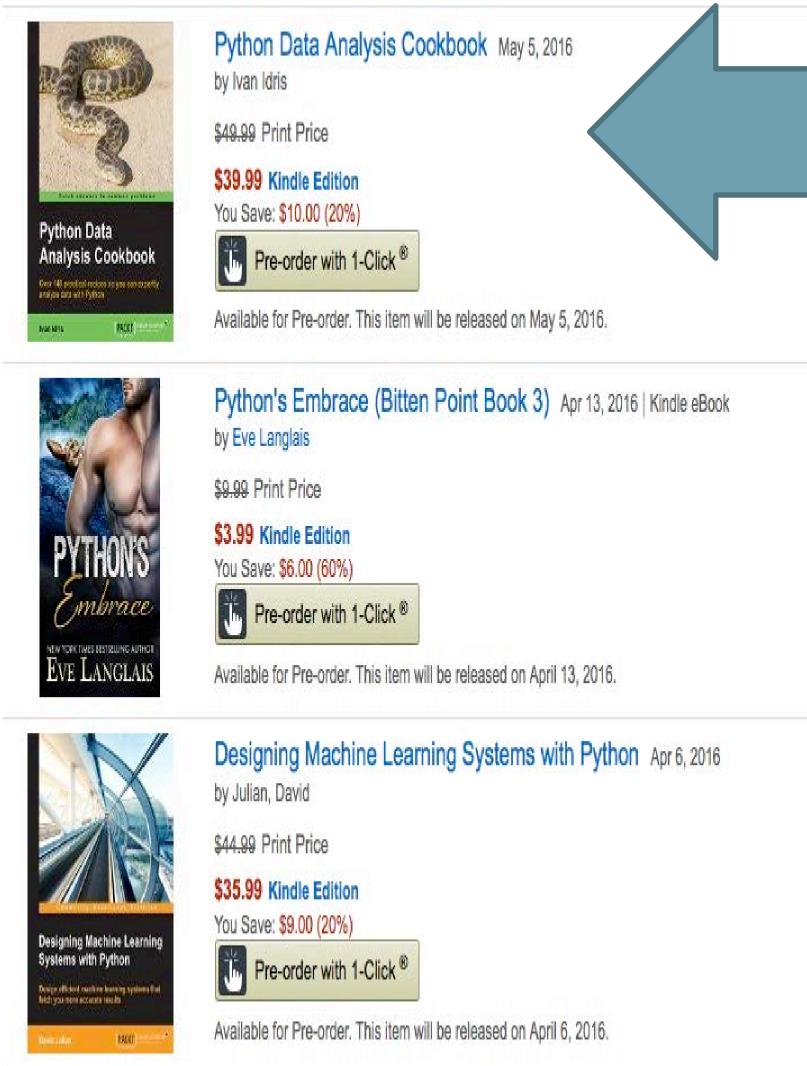
- Do Lawyers Really Believe Their Own Hype and Should They?: A Natural Experiment

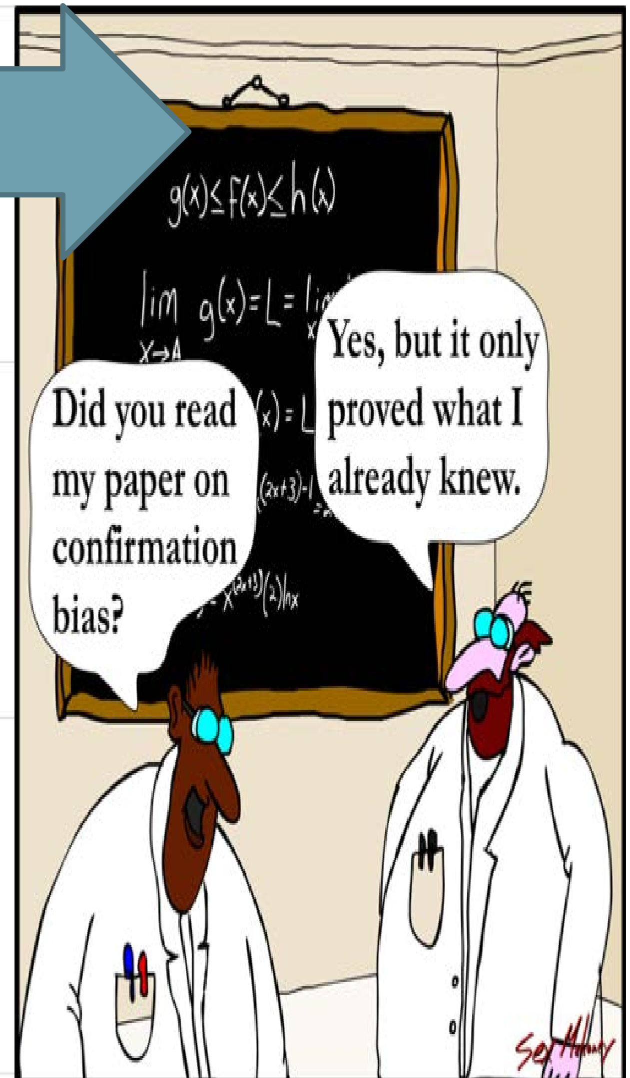
(Eigen & Listokin, JLS 2012)

FIGURE 6: Moot court scores and corresponding merits confidence index with prediction line



Balance “Data Science Evangelism” with “Data Science Doom and Gloom”

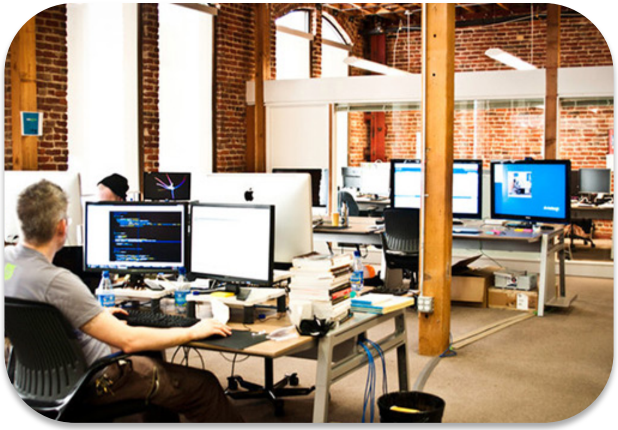




Beware of False Promises; Strive to Understand How to Consume Data Science Output

Data Science for Legal Decision Making

Inside the firm
(obo the firm as client)



Outside the firm
(obo clients)

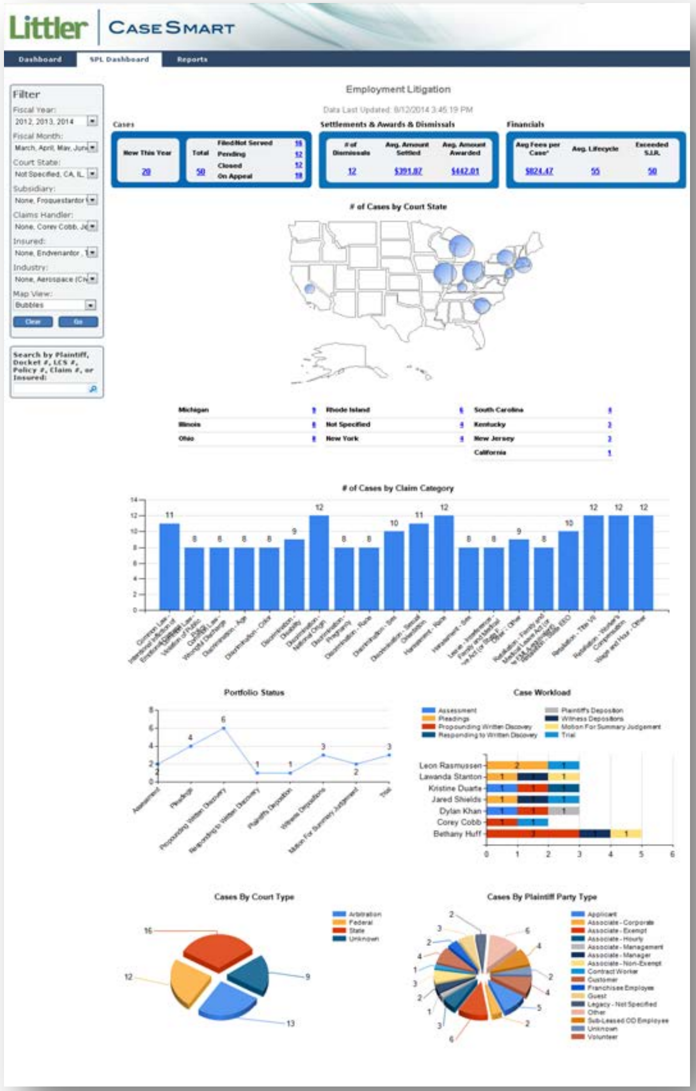
4 Applications

- Staffing / Pricing
- Rule Interpretation/ decision automation
- Task Automation
- Outcome & Cost Prediction Modeling

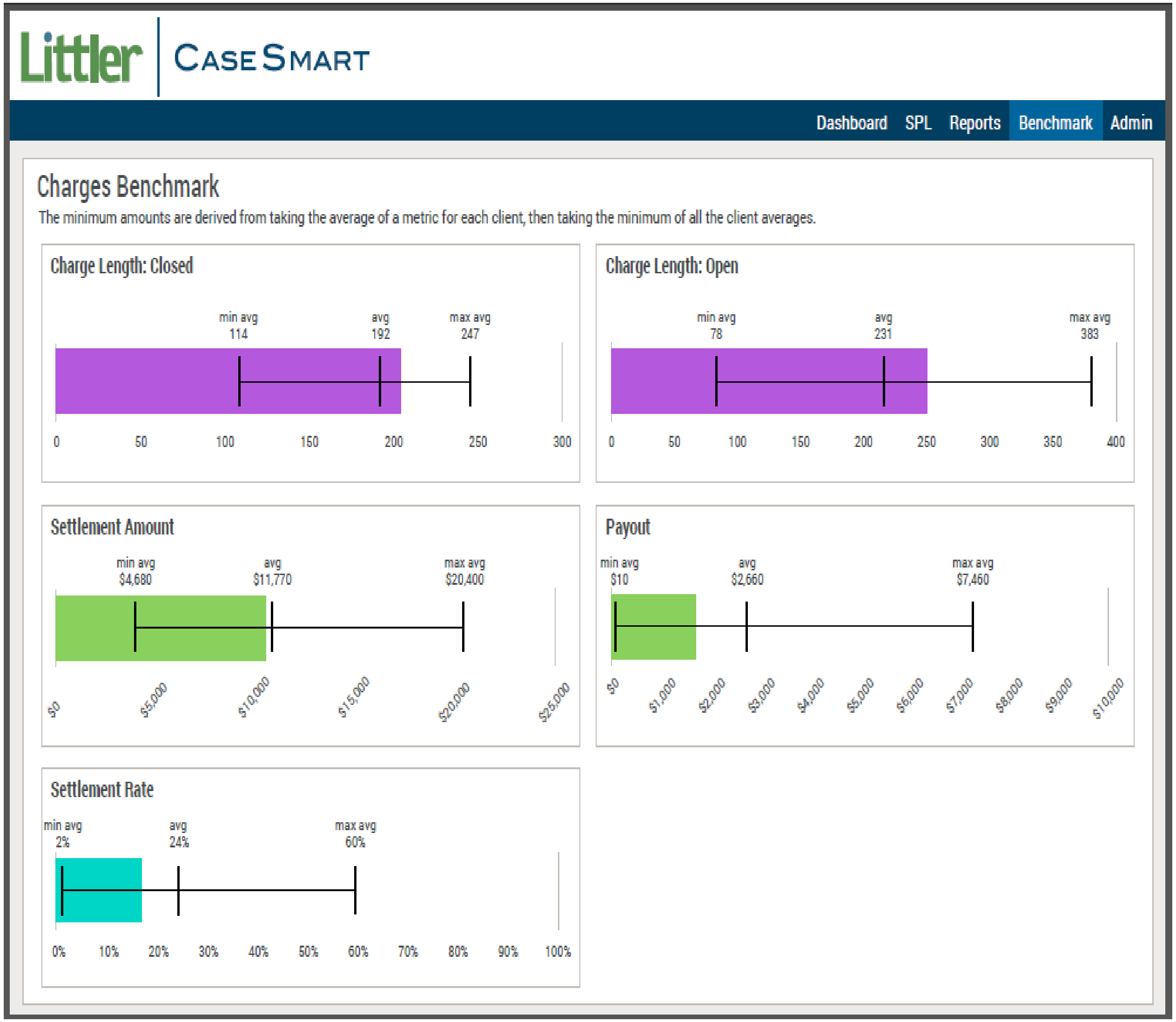
Examples in the 8 “Boxes” of Innovation

	Staffing/ Pricing	Rule Interpretation / decision automation	Task automation	Outcome & Cost Prediction Modeling
Inside the Firm (Firm as client)	Hiring lawyers, rate setting		Citation, opinion analysis	Marketing, rate setting
Outside the Firm (Clients as client)	Optimizing selection processes	Interpreting IC/Employee distinction	Contract analysis, E-discovery	Improved legal strategic decision-making

Staffing / Pricing

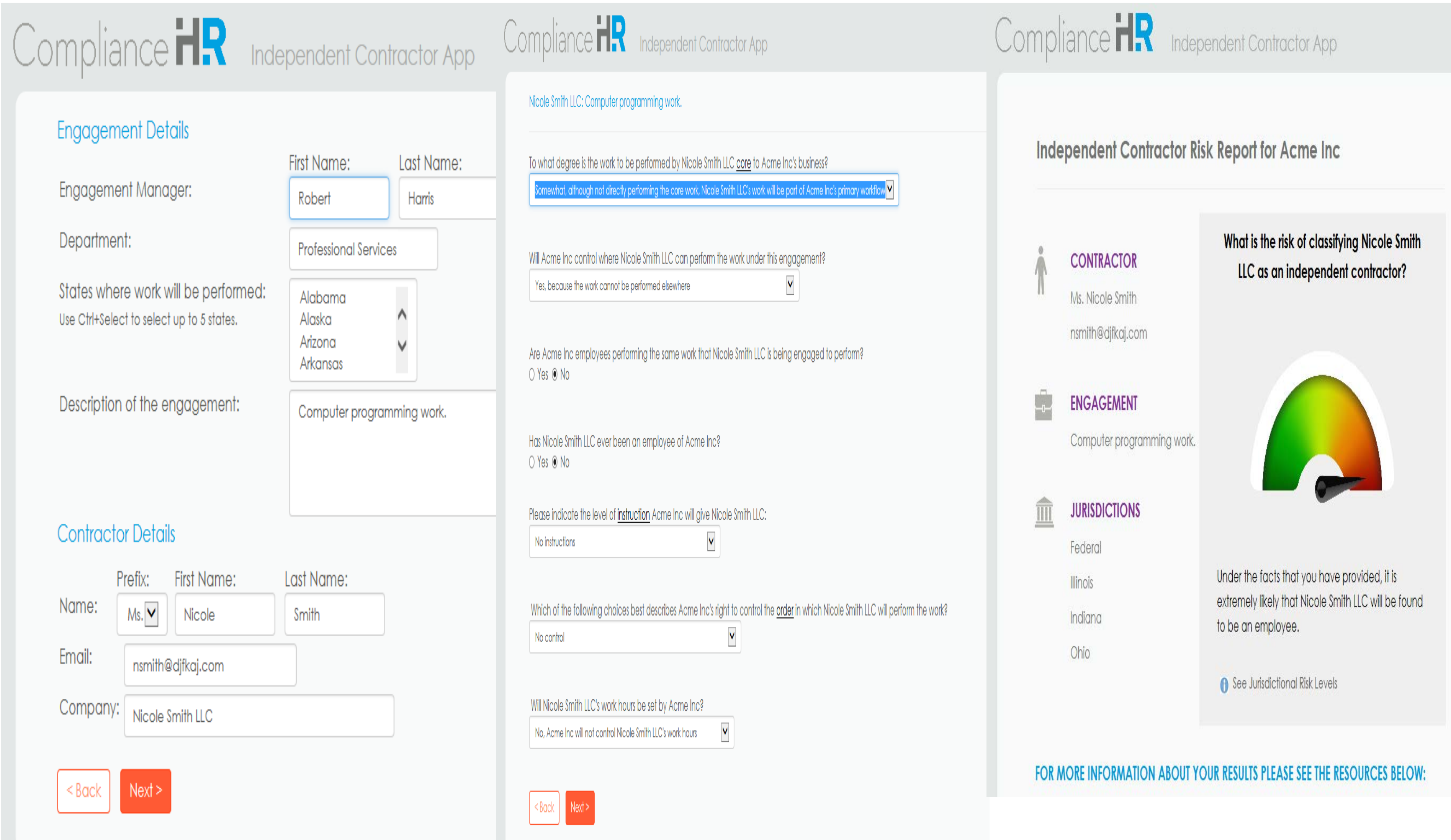


CaseSmart



Littler
CASESMART
Dashboard
SPL
Reports
Benchmark
Admin
Charges Benchmark
The minimum amounts are derived from taking the average of a metric for each client, then taking the minimum of all the client averages.
Charge Length: Closed
min avg 114
avg 192
max avg 247
0 50 100 150 200 250 300
Charge Length: Open
min avg 78
avg 231
max avg 383
0 50 100 150 200 250 300 350 400
Settlement Amount
min avg $4,680
avg $11,770
max avg $20,400
$0 $5,000 $10,000 $15,000 $20,000 $25,000
Payout
min avg $10
avg $2,660
max avg $7,460
$0 $1,000 $2,000 $3,000 $4,000 $5,000 $6,000 $7,000 $8,000 $9,000 $10,000
Settlement Rate
min avg 2%
avg 24%
max avg 60%
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Rule Interpretation Example: Compliance HR



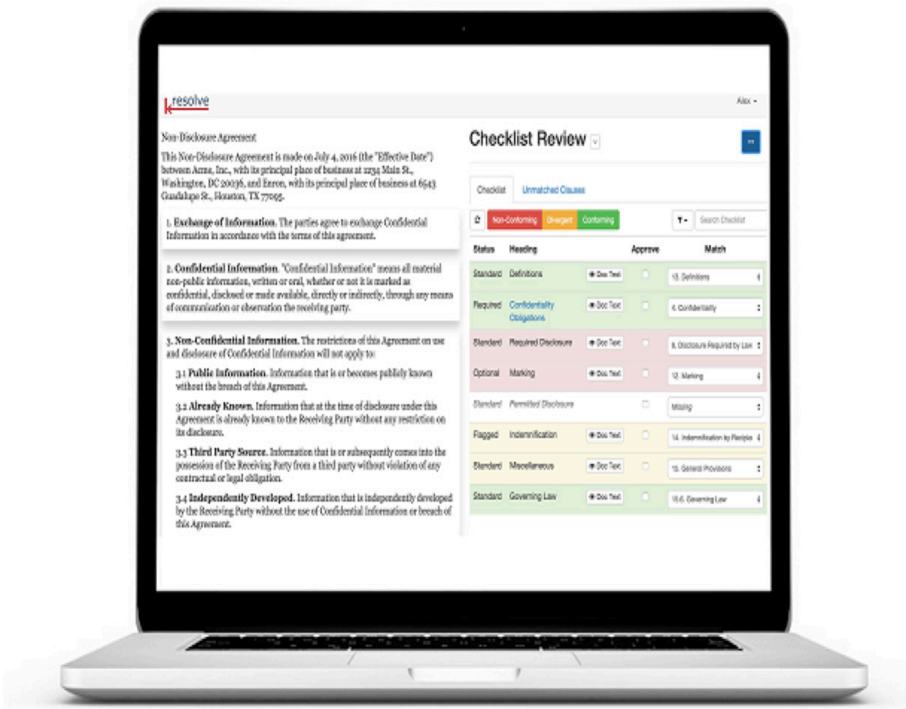
Task Automation Examples

- Citation/ opinion analysis
  – Hierarchical clustering / topic models; arguments acknowledged; rationales offered for decisions
- Legal document analysis and comparison
  – E-discovery / document tagging; analysis of judicial opinions
  – Cross-sectional / time series classification of statutes; regulations; public filings; contracts; law firm effects

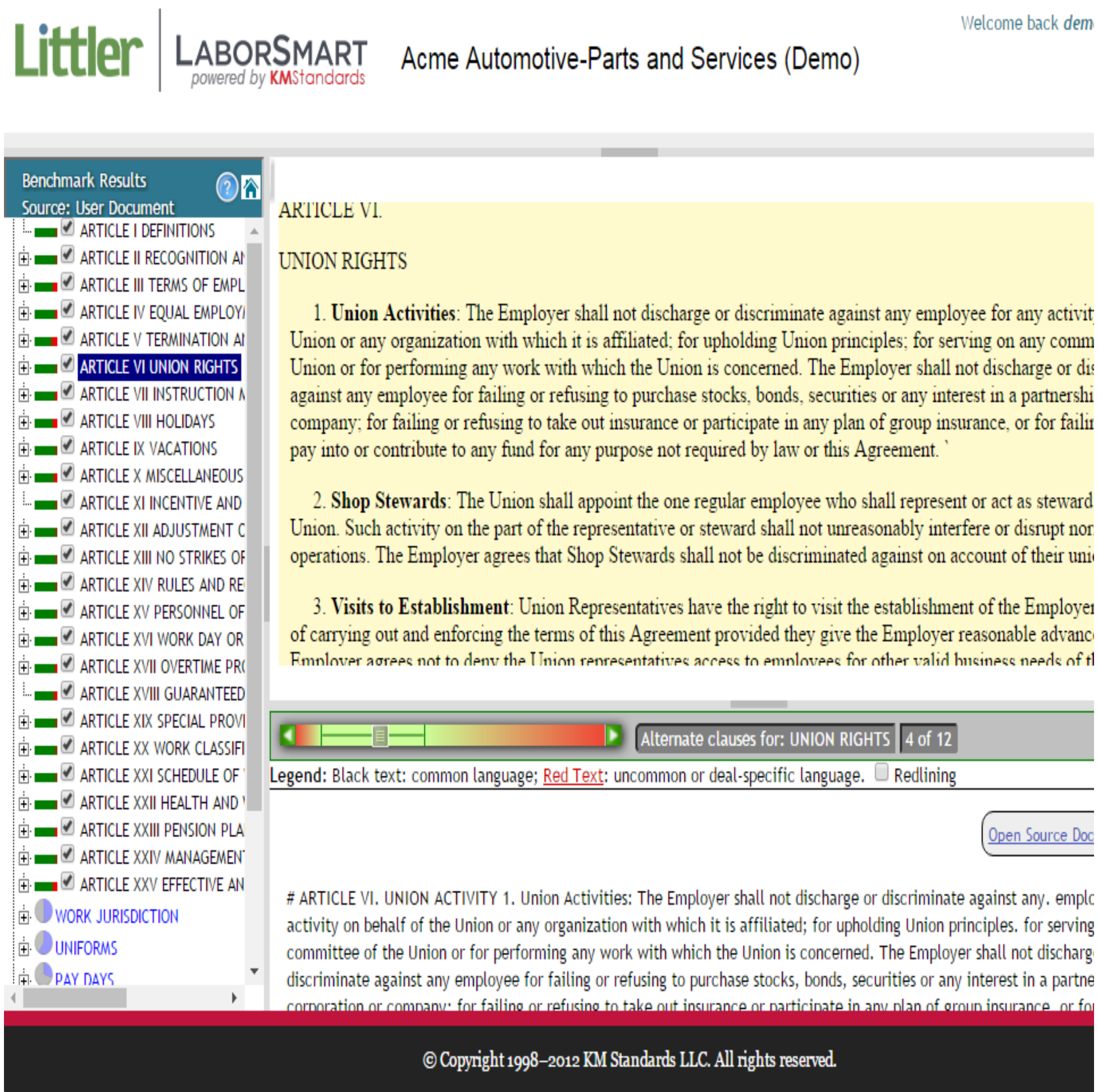
Task Automation Example: Contract Standards

resolve

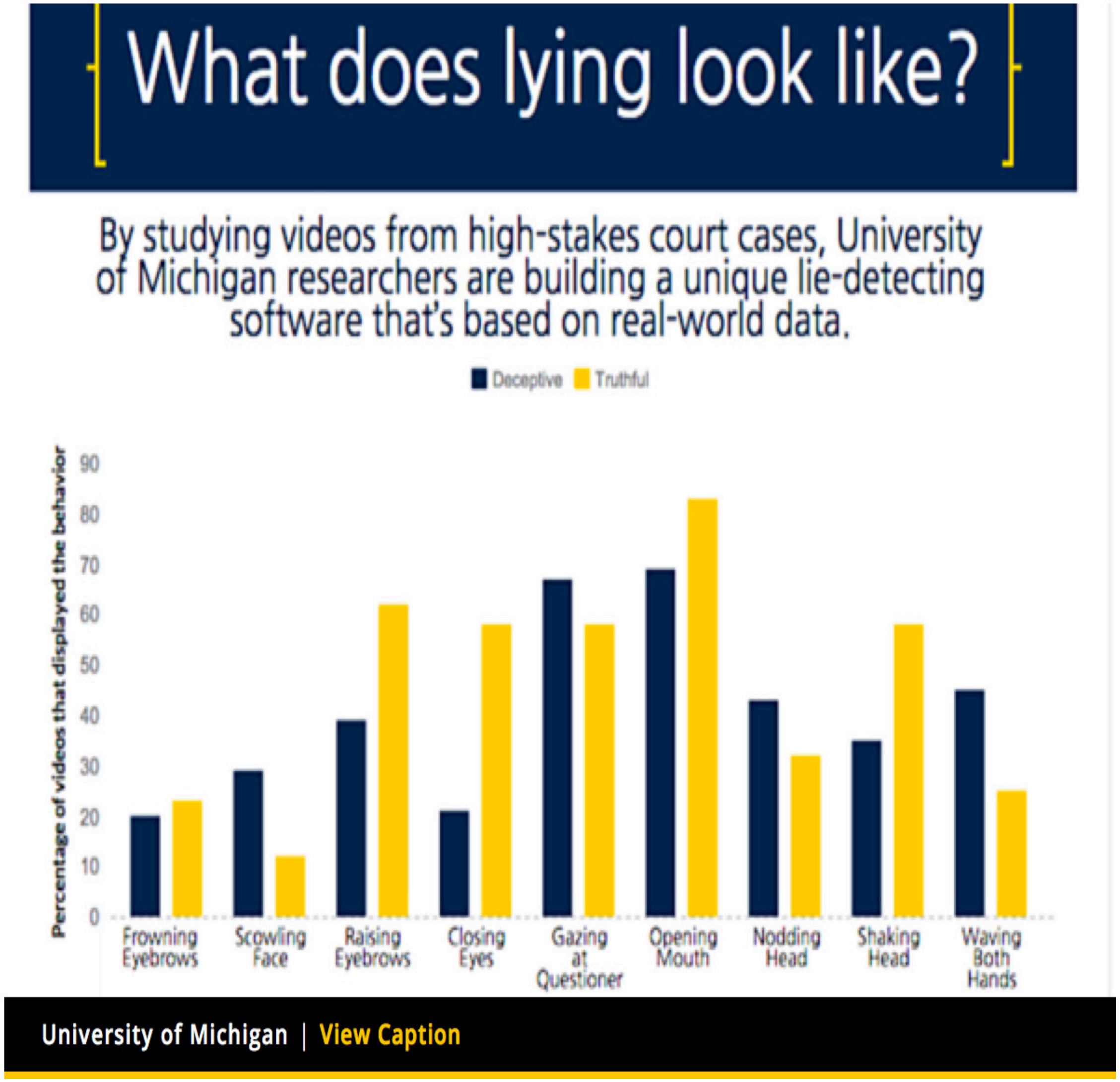
get better terms



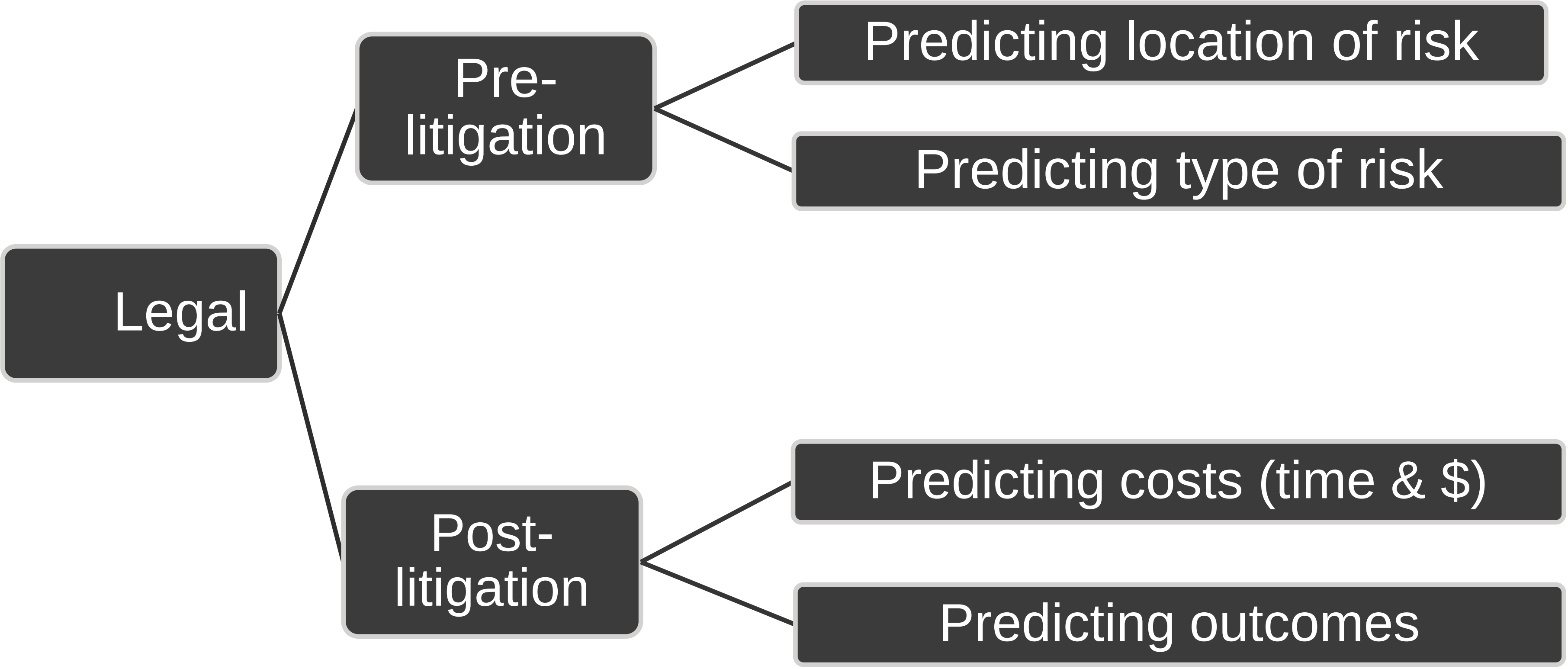
Task Automation Example: Labor Smart



Micro-Behavioral Example: Predicting Lying

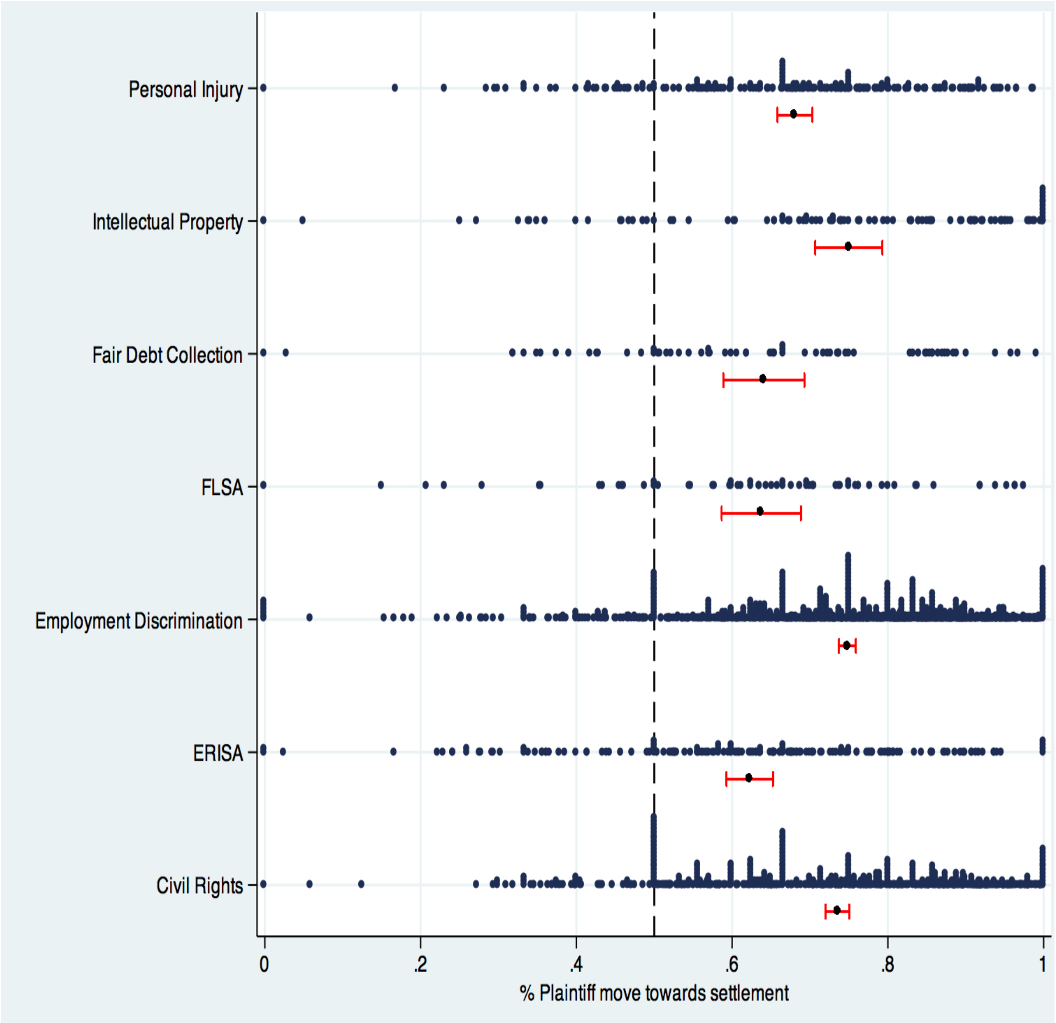


Outcome & Cost Prediction Modeling

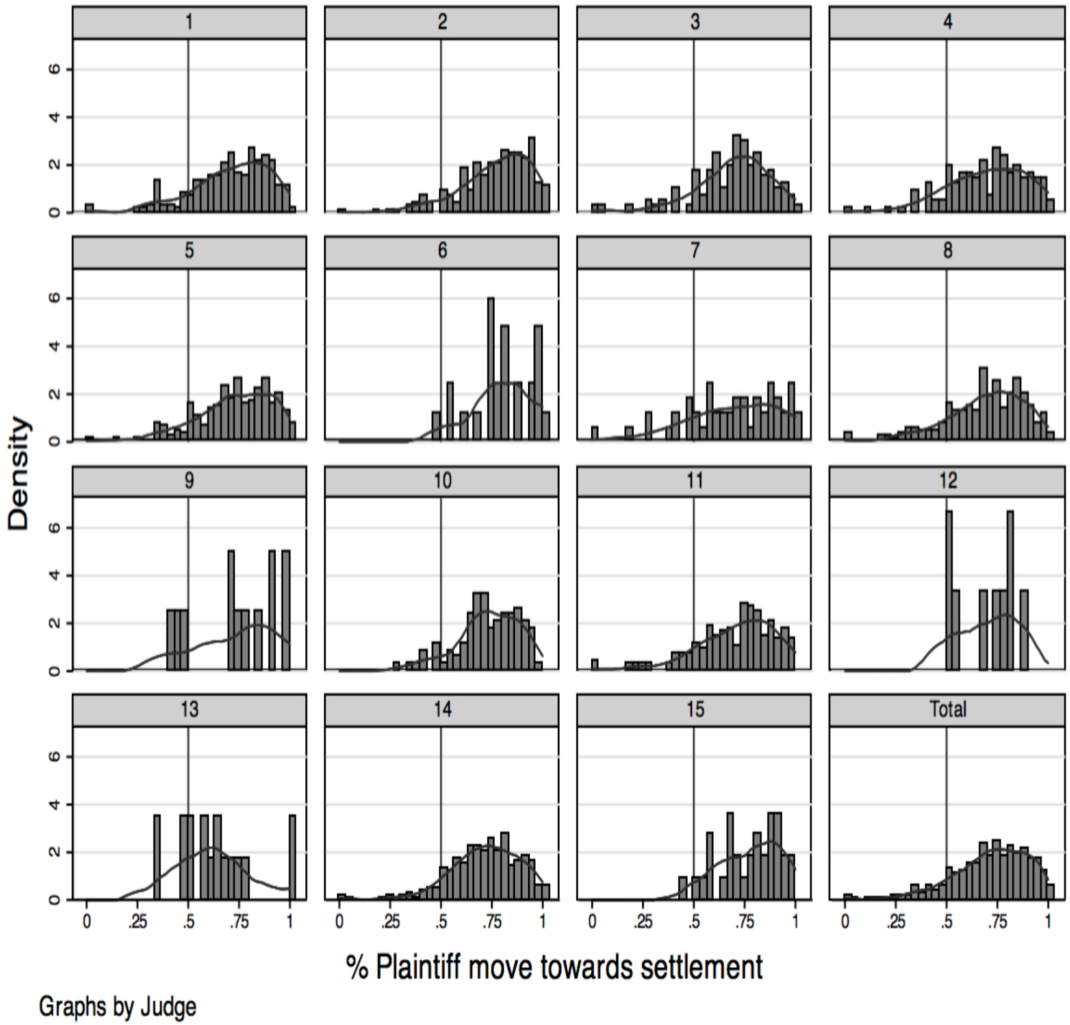


Outcome Prediction

- Likelihood of success on the merits
- Cost of defense
- Settlement value
- Timing of settlement



Example: Judicial Assignment



Some More "Predictive" Talent Acquisition Software Providers

"[Our software] uses outcomes-based modeling and machine learning to identify those applicants who are most likely to succeed"

Pre-Recorded Video Interviewing

“Wouldn’t it be great if there was a way to take a ton of information—say, data from more than 3 million interviews—and turn it into something that predicts which candidates will be your top performers?”

“Candidates can launch their interviews from a corporate career site, a LinkedIn job posting or even from Twitter! The hiring company drives the content from an approved pick-list for customization. [We] partnered with assessment vendors ___ and ___ for behavioral questions….”

HCM is Seeing Significant Activity in All Areas

More than $5B has been invested into HCM technology startups over the past 5 years helping many startups to gain traction/awareness very quickly

Talent Acquisition

- Jibe — 2014 $20M (C) / $40.9M
- HIRED — 2014 $15M (B) / $32.7M
- Jobvite — 2014 $25M(D) / $55.5M
- LEVER — 2014 $10M(A)

Learning & Development

- lynda.com — 1995 $289M $1.5B sale (LinkedIn)
- TALENTRANK — 2016 20K(incubator)
- PRO SKY — 2015 2.3M (Seed)

Planning and Analytics

- VISIER — 2014 $25.5M (C)/$46.5M
- pegged software — 2014 $7.5M (growth)
- VoloMetrix — 2014 $12M / 16.9M Acquired by Microsoft
- SMASHFLY — 2016 $22M (B) / $31M

Engagement

- slack — 2015 $160M (E)/ $340.2M
- GLINT — 2015 $15.5M (A) / $22M

Talent Management

- BetterWorks — 2014 $15.5M (A)
- TMBC — 2014 $5M (A)
- ZUGATA — 2015 $3.2M(Seed)
- Wrike.com — 2015 $5M(B) / $26 M
- Trello — 2014 $10.3M(A)

Core HR

- ZENEFITS — 2015 $500M (C) / $583.6M
- Namely — 2015 $45M (C) / $73.1M

Key: Year founded; last round VC funding; Total funding

“PUSH”

e.g., compensation, career path, bad manager, commute, and culture

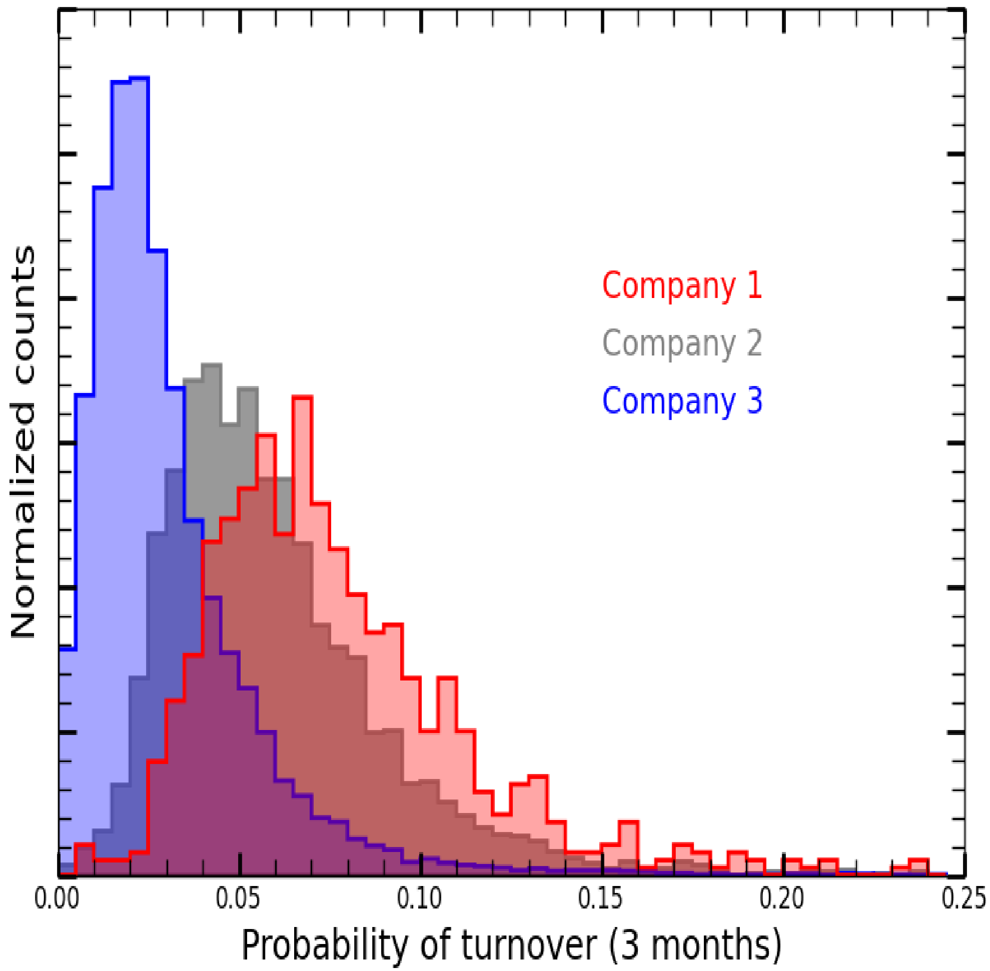
VS.

“PULL”

e.g., heavily recruited, high demand skill, career progression, growing sector, high visibility performer

The Public Data Landscape

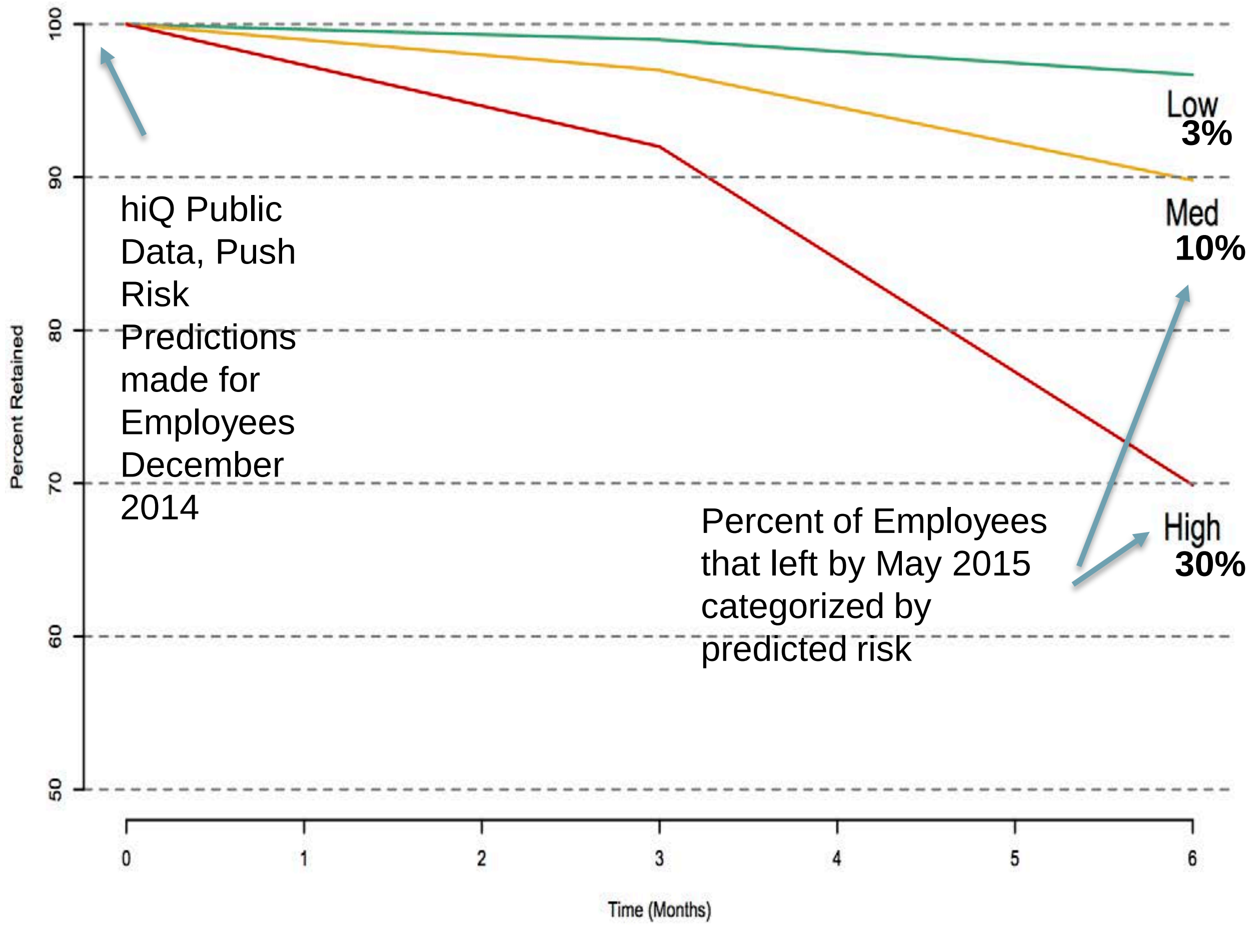
Labor Market Pull Risk Profiles



Individual Probability of Turnover

Company data December 15- May15; hiQ Static Model; Firm made no interventions

An Example of Attrition Risk Predictions: Dec 2014 – May 2015



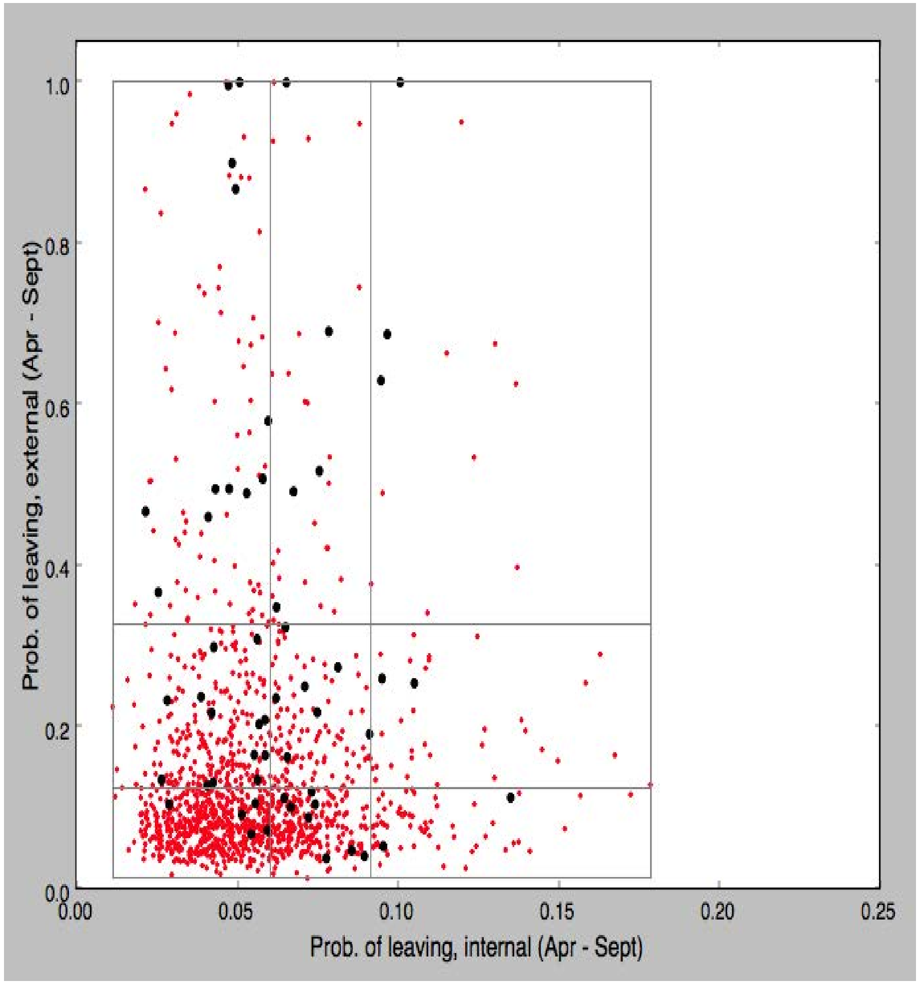
Company data December 15- May15; hiQ Static Model; Firm made no interventions

Push and Pull Predictions Together

<u>Predictions Will Be:</u>

- **More Accurate**
- **More Robust to Changing Conditions**
- **More Credible**
- **More Actionable**

Correlation of Retention Index and Pull Risk -0.0018



Push predictions use 2014 survey model applied to 2015 survey data; hiQ predictions made April 5th, 2015. Results based on actual turnover from April 20th to June 5th.

79

“*Get to the Best Candidates, Faster*”

“[Our software compares] more than 15,000 interaction, behavioral, and performance attributes such as attitude, engagement, word choice, and operational performance outcomes to predict which candidates are likely to be top performers…. **In an instant, [our software] will rank all of your interviews based on how your top managers would’ve ranked them**”

Games, Scored Interviews, Other Online Personality Assessments

- “[P]lay games to find your first job or best fit career….[We] leverage neuroscience and big data to optimize your company's recruiting and candidate sourcing process”
- “[Our software] dynamically obtains, crunches, compares, and contrasts all of the raw data and then generates a statistically significant percentage of “fit” between candidates and open positions – what we like to call a Fitness Index”

Key Potential Legal Risk Areas

- Disparate Impact
  - Sourcing v. Selection
  - Who is an “Applicant”?
  - Validation
- Disparate Treatment
- Disability Discrimination

Sourcing v. Selection

- Passive recruiting: Sourcing bleeding into selection
  – Why it matters: UGESP
  – Ok to restrict to Linked In?
- Who is an “applicant”
  – Invites…note the assumption
  – Everyone touched by an algorithm?
- Accessibility re disabilities

Big Data Validation Challenges

- Interpreting the result
  - Rationale for observed relationships
  - Rationale plus statistical meaningfulness
- Validation issues
  - Focus on job requirements vs. people characteristics
  - Test items may not be secure or “face valid”
  - Algorithms may be based on “tests” designed for very different purposes (e.g., clinical psychological diagnoses)
  - Determining “passing” score

Vendor Due Diligence is a Must

- Has the process demonstrated adverse impact?
  - What validation evidence has been collected to establish the job relatedness of the algorithm? For each job?
  - Does the validation evidence comply with the requirements of UGESP? Get a copy of the validation study.
  - What steps have been taken to ensure the security of test questions?
  - What kind of ongoing monitoring do you provide as we continue using the instrument?
  - Indemnification?

Are “Predictive” Big Data Software Solutions The “Glue”?

Wal-Mart v. Dukes

The White House Weighs In

- A White House report in 2014 recommended the federal government (DOJ, EEOC) should:
  - "…expand their technical expertise to be able to identify practices and outcomes facilitated by big data analytics that have a discriminatory impact on protected classes … The agencies may consider the classes of data, contexts of collection, and segments of the population that warrant particular attention, including for example genomic information or information about people with disabilities."

Executive Office of the President (May, 2014). Big Data: Seizing opportunities, preserving values. Washington, D.C., p. 65


Thank You!