

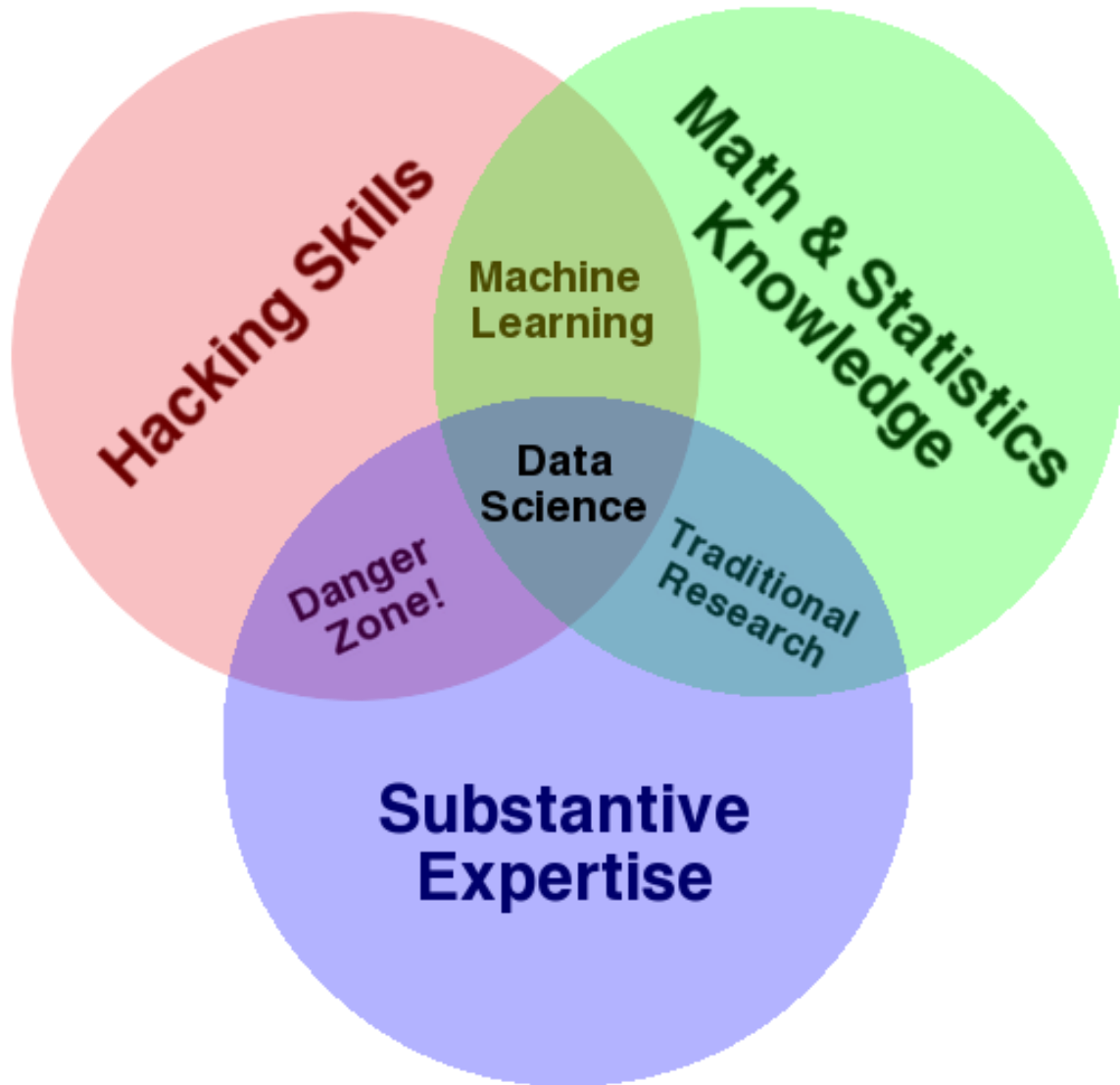


Cognitive collaboration

why effective data science needs human-centered design

James Guszcza
CAS Underwriting Collaboration
New Orleans
June 26, 2018

Data science

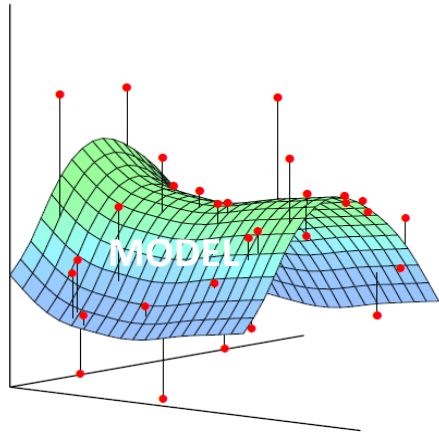


“... It is unclear how to distinguish among hackers, statisticians, subject matter experts, their overlaps and where data science fits ...”

-- Drew Conway

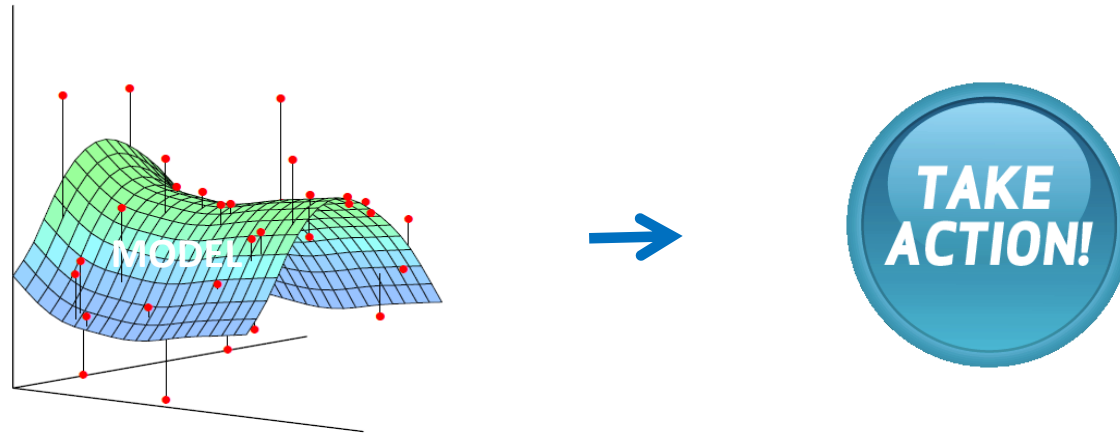
Data science's "last mile problem"

Algorithms can point us in the right direction, but are not a complete solution.



Data science's "last mile problem"

Algorithms can point us in the right direction, but are not a complete solution.



They must be followed by the right judgments, decisions, or behavior change.

'AI IS THE NEW ELECTRICITY'



“Just as electricity transformed almost everything 100 years ago, today I actually have a hard time thinking of an industry that I don’t think AI will transform in the next several years.”

Andrew Ng

Former chief scientist at Baidu, Co-founder at Coursera

Smart *algorithms* are unlikely to engender smart *outcomes* unless they are designed to promote smart *adoption* on the part of human end users.

Smart *algorithms* are unlikely to engender smart *outcomes* unless they are designed to promote smart *adoption* on the part of human end users.



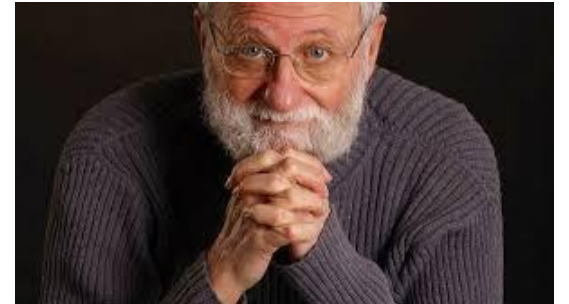
Effective AI needs human-centered design

The AI revolution needs a design revolution

The problem with the designs of most engineers is that they are too logical.

We have to accept human behavior the way it is, not the way we would wish it to be.

— Don Norman, The Design of Everyday Things

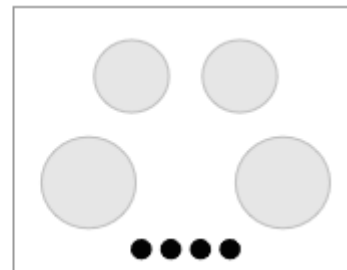
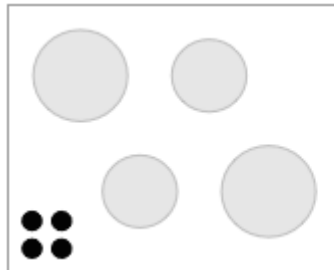


Human-centricity: understanding the user

AI and other data products will yield better outcomes if they are designed to go with the grain of human psychology.

Data science teams must think like designers ...

... not just "engineers".



The drive to automation

THE FUTURE OF EMPLOYMENT: HOW SUSCEPTIBLE ARE JOBS TO COMPUTERISATION?*

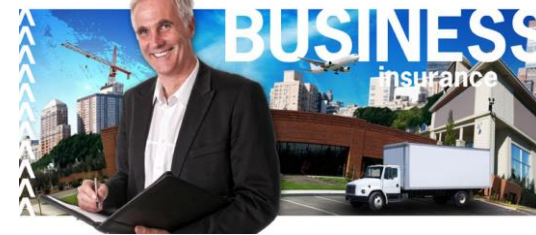
Carl Benedikt Frey[†] and Michael A. Osborne[‡]

September 17, 2013

Computerisable				
Rank	Probability	Label	SOC code	Occupation
1.	0.0028		29-1125	Recreational Therapists
2.	0.003		49-1011	First-Line Supervisors of Mechanics, Installers, and Repairers
3.	0.003		11-9161	Emergency Management Directors
4.	0.0031		21-1023	Mental Health and Substance Abuse Social Workers
5.	0.0033		29-1181	Audiologists
6.	0.0035		29-1122	Occupational Therapists
7.	0.0035		29-2091	Orthotists and Prosthetists
8.	0.0035		21-1022	Healthcare Social Workers
9.	0.0036		29-1022	Oral and Maxillofacial Surgeons
10.	0.0036		33-1021	First-Line Supervisors of Fire Fighting and Prevention Workers
11.	0.0039		29-1031	Dietitians and Nutritionists
12.	0.0039		11-9081	Lodging Managers
13.	0.004		27-2032	Choreographers
695.	0.99		13-2082	Tax Preparers
696.	0.99		43-5011	Cargo and Freight Agents
697.	0.99		49-9064	Watch Repairers
698.	0.99	.	13-2053	Insurance Underwriters
699.	0.99		15-2091	Mathematical Technicians
700.	0.99		51-6051	Sewers, Hand
701.	0.99		23-2093	Title Examiners, Abstractors, and Searchers
702.	0.99		41-9041	Telemarketers



What about underwriting more complex risks?



Issues

- Large number of data points
- Reasonably homogenous risks
- Many data sources

Innovation

- Better risk selection through “big data”: credit scores, telematics, wearables, ...

Issues

- Relatively few data points
- Heterogeneous risks
- Few common risk factors

Innovation

- Better risk selection through “playing Moneyball”
- Rather than replace underwriters, we build algorithms to help underwriters with tasks the human mind is poor at.



Actuarial vs clinical prediction – the motion picture

*Human judges are not merely worse than optimal regression equations;
they are worse than almost any regression equation.*

— Richard Nisbett and Lee Ross

Science 31 March 1989:
Vol. 243 no. 4899 pp. 1668-1674
DOI: 10.1126/science.2648573

Clinical versus actuarial judgment

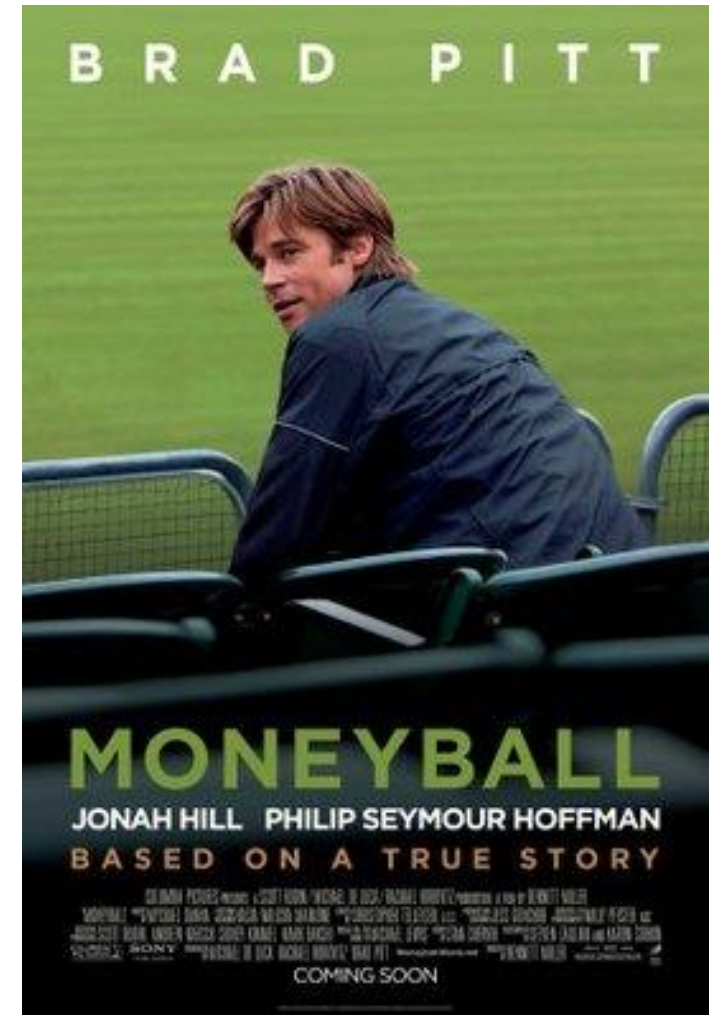
RM Dawes, D Faust and PE Meehl

[+](#) Author Affiliations



ABSTRACT

Professionals are frequently consulted to diagnose and predict human behavior; optimal treatment and planning often hinge on the consultant's judgmental accuracy. The consultant may rely on one of two contrasting approaches to decision-making—the clinical and actuarial methods. Research comparing these two approaches shows the actuarial method to be superior. Factors underlying the greater accuracy of actuarial methods, sources of resistance to the scientific findings, and the benefits of increased reliance on actuarial approaches are discussed.



“The mind is a machine for jumping to conclusions”

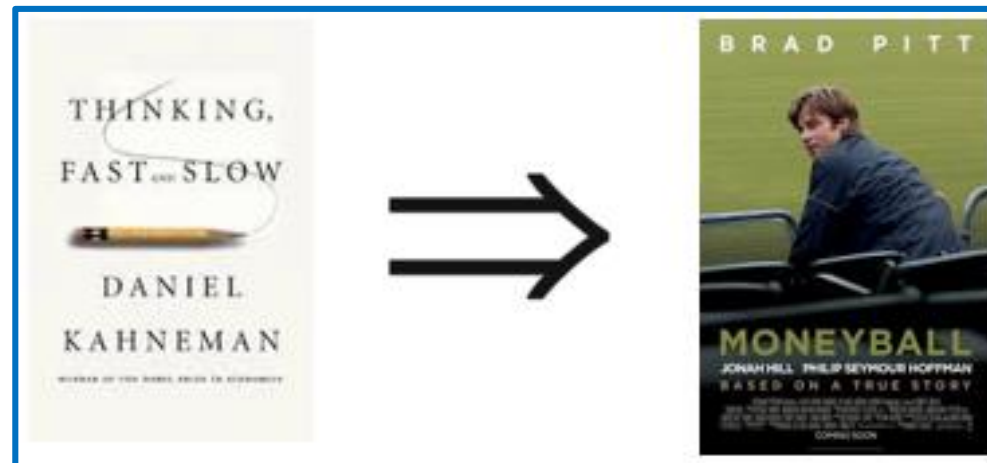
Daniel Kahneman: there are two types of mental operations.

- System 1: automatic, effortless, **associatively** coherent.
- System 2: controlled, effortful, **logically** coherent.



Most of our mental operations are “System 1” in nature.

And “System 1” has a lot of trouble with statistics.



There's something about Linda

bias

Think about this person:

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.



There's something about Linda

bias

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

Which is more likely?

- a) Linda is a bank teller
- b) Linda is a bank teller active in the feminist movement



A sample of cognitive biases

- **Law of small numbers:** we overgeneralize from personal experience
- **“What you see is all there is” [WYSIATI]:** we assume the evidence before our eyes is all that’s relevant
- **Base rate neglect:** we tend to ignore base rates (Bayesian priors) in favor of case-specific information
- **Representativeness heuristic (“Linda”):** when asked about a probability we instead consider how similar the case is to other cases.
- **Confirmation bias:** we selectively process information that confirms our “gut feel” / first impression
- **Availability heuristic:** our estimates of probabilities are affected by how easily an example comes to mind
- **The halo effect:** we form an overall impression based on specific factors we like or dislike
- **The affect heuristic:** we selectively process information about a risk depending on whether we basically “like” or “dislike” it. “The emotional tail wags the rational dog”
- **The Dunning-Kruger effect:** “Ignorance frequently begets more confidence than knowledge”

Maybe it's time we break for lunch

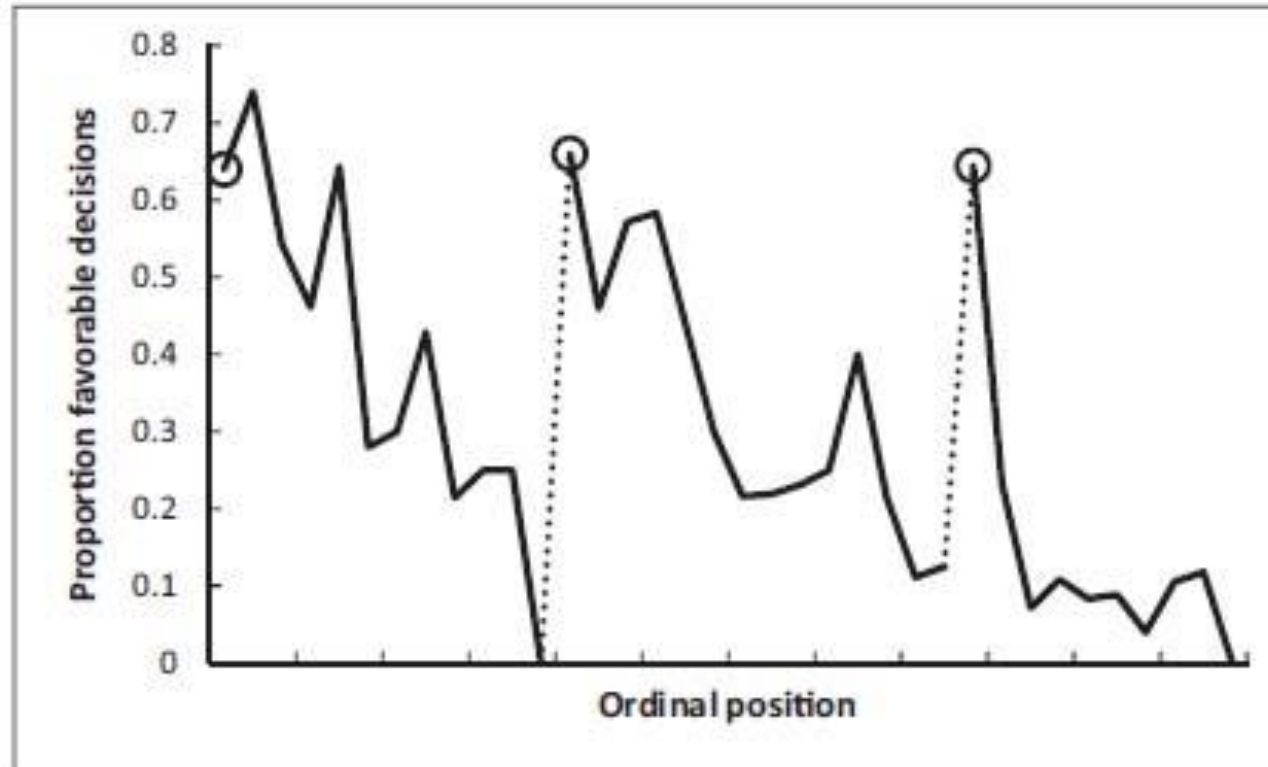


Fig. 1. Proportion of rulings in favor of the prisoners by ordinal position. Circled points indicate the first decision in each of the three decision sessions; tick marks on x axis denote every third case; dotted line denotes food break. Because unequal session lengths resulted in a low number of cases for some of the later ordinal positions, the graph is based on the first 95% of the data from each session.

ENSE

ΣΚΕΨΟΥ

THINK

DENKE

SMACHINIS

PENSE



\$24,000

\$77,147

\$21,600

Who is Stoker?
(FOR ONE WELCOME OUR
NEW COMPUTER OVERLORDS)
\$1,000

Who is Bram
Stoker?
\$ 17,973

WHO IS
BRAM STOKER?
\$5600

AI = Augmented Intelligence

Marvin [Minsky] was advocating what's called "commonsense reasoning".

Machines have shown essentially no examples of doing that.

Therefore, they are complements to people. People are actually not so bad at that.

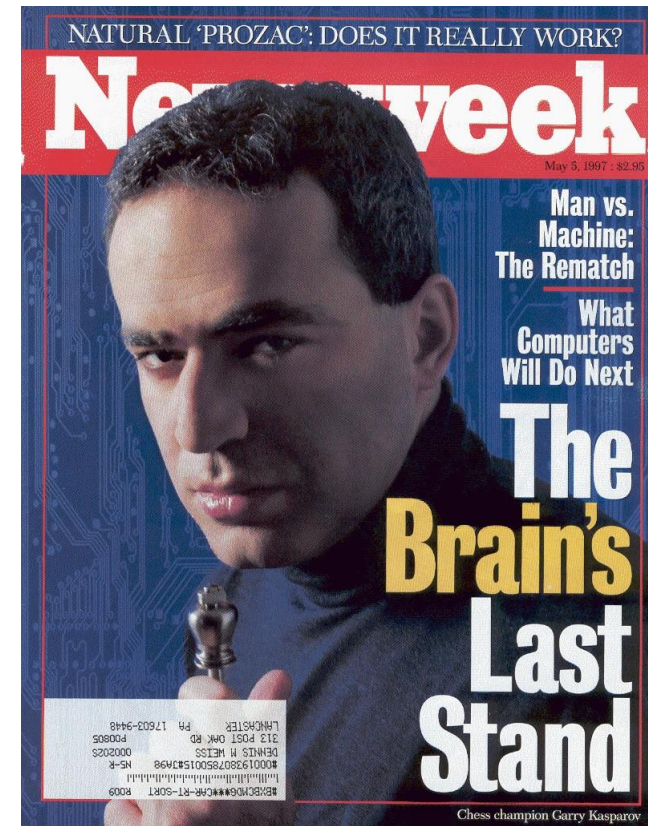
However, they are somewhat lousy at tuning things and keeping exact accounts of stuff. Machines are good at that.

That gives the idea that there could be a human-machine partnership...

— Sandy Pentland, Deloitte Review 2017



The prequel to "Jeopardy" ...



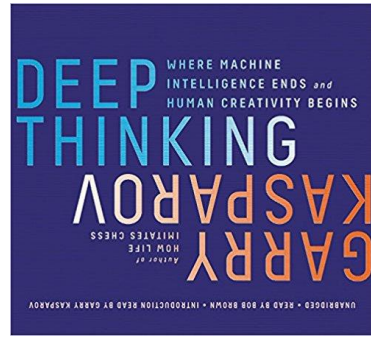
Human-computer collective intelligence

Dark horse ZackS wins Freestyle Chess Tournament

6/19/2005 – The computer-assisted PAL/CSS Freestyle Chess Tournament, staged on Playchess.com, ended with a shock win by two amateurs: Steven Cramton, 1685 USCF and Zackary Stephen, 1398 USCF, using three computers for analysis, defeated teams of strong grandmasters all the way to victory in the finals. We bring you a first flash report with games and results.



Human-computer collective intelligence



Dark horse ZackS wins Freestyle Chess Tournament

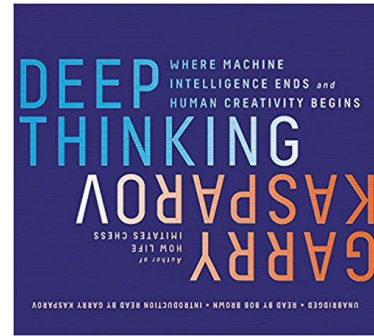
6/19/2005 – The computer-assisted PAL/CSS Freestyle Chess Tournament, staged on Playchess.com, ended with a shock win by two amateurs: Steven Cramton, 1685 USCF and Zackary Stephen, 1398 USCF, using three computers for analysis, defeated teams of strong grandmasters all the way to victory in the finals. We bring you a first flash report with games and results.

Their skill at manipulating and “coaching” their computers to look very deeply into positions effectively counteracted the superior chess understanding of their grandmaster opponents and the greater computational power of other participants.

Weak human + machine + better process was superior to a strong computer alone and, more remarkably, superior to a strong human + machine + inferior process.

— Garry Kasparov

Human-computer collective intelligence



Dark horse ZackS wins Freestyle Chess Tournament

6/19/2005 – The computer-assisted PAL/CSS Freestyle Chess Tournament, staged on Playchess.com, ended with a shock win by two amateurs: Steven Cramton, 1685 USCF and Zackary Stephen, 1398 USCF, using three computers for analysis, defeated teams of strong grandmasters all the way to victory in the finals. We bring you a first flash report with games and results.

Their skill at manipulating and “coaching” their computers to look very deeply into positions effectively counteracted the superior chess understanding of their grandmaster opponents and the greater computational power of other participants.

Weak human + machine + better process was superior to a strong computer alone and, more remarkably, superior to a strong human + machine + inferior process.

— Garry Kasparov

Creating such processes goes beyond data science - psychology and design thinking are needed.

Experts need equations...

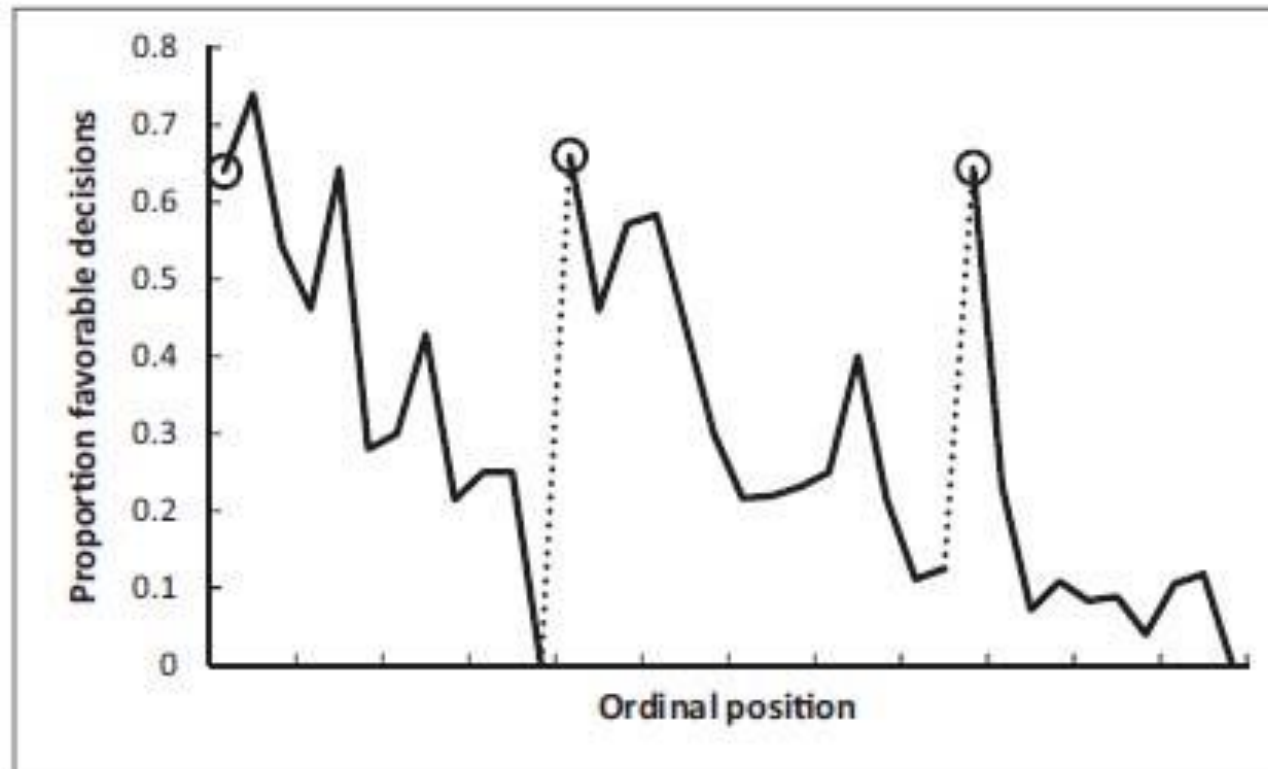


Fig. 1. Proportion of rulings in favor of the prisoners by ordinal position. Circled points indicate the first decision in each of the three decision sessions; tick marks on x axis denote every third case; dotted line denotes food break. Because unequal session lengths resulted in a low number of cases for some of the later ordinal positions, the graph is based on the first 95% of the data from each session.

... but equations can be biased too



▼ Machine Bias



Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica

May 23, 2016

Prediction Fails Differently for Black Defendants

	WHITE	AFRICAN AMERICAN
Labeled Higher Risk, But Didn't Re-Offend	23.5%	44.9%
Labeled Lower Risk, Yet Did Re-Offend	47.7%	28.0%

“Freestyle decision-making”

*Kasparov's Law: Weak human + machine + **better process** is superior to strong human + machine + inferior process*

Harvard
Business
Review

DECISION MAKING

Creating Simple Rules for Complex Decisions

by Jongbin Jung, Connor Concannon, Ravi Shroff, Sharad Goel, and Daniel G. Goldstein

APRIL 19, 2017

AGE		
RANGE	SCORE	CALCULATION
18-20	8	2
21-25	6	
26-30	4	
31-50	2	
51 and older	0	
+		
PAST COURT DATES MISSED		
RANGE	SCORE	
0	0	6
1	6	
2	8	
3	9	
4+	10	
=		
8		

SOURCE: JONGBIN JUNG ET AL. © HBR.ORG

“Freestyle decision-making”

*Kasparov's Law: Weak human + machine + **better process** is superior to strong human + machine + inferior process*

“Optimal” should be conceived from the perspective of human and organizational needs, not what's technically possible.

Harvard
Business
Review

DECISION MAKING

Creating Simple Rules for Complex Decisions

by Jongbin Jung, Connor Concannon, Ravi Shroff, Sharad Goel, and Daniel G. Goldstein

APRIL 19, 2017

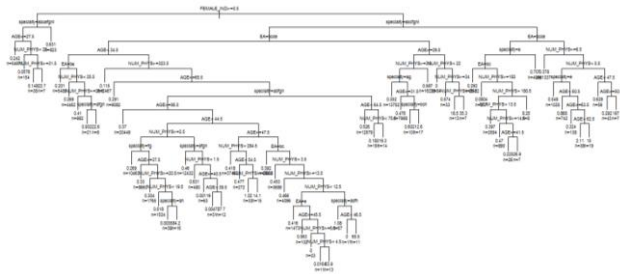
AGE		
RANGE	SCORE	CALCULATION
18-20	8	2
21-25	6	
26-30	4	
31-50	2	
51 and older	0	
+		
PAST COURT DATES MISSED		
RANGE	SCORE	
0	0	6
1	6	
2	8	
3	9	
4+	10	
=		
FLIGHT RISK		
8		

SOURCE: JONGBIN JUNG ET AL. © HBR.ORG

A false comparison

Equations > experts

$$Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \dots + \beta_n * X_n$$

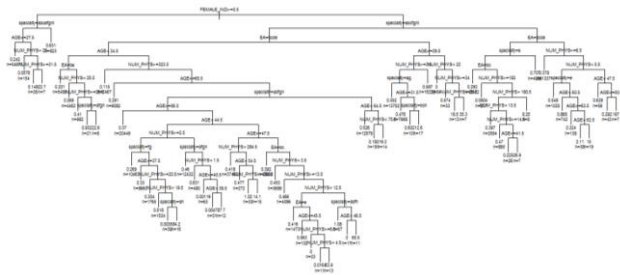


Algorithms are eyeglasses for the mind's eye

~~Equations > experts~~

(Equations + experts) > experts

$$Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \dots + \beta_n * X_n$$



Algorithms are eyeglasses for the mind's eye

~~Equations > experts~~

(Equations + experts) > experts

To achieve “freestyle underwriting”, we need models that:

- Reflect end-user needs (key HCD concept)
- End-users understand (have a clear mental model of)
- End-users trust
- Incorporate end-users' domain/institutional knowledge
- Provide “why” indications along with scores
- Designed to be compatible with business rules, over-ride mechanisms
- Are part of an iterative “test and learn” process (not “one and done”)

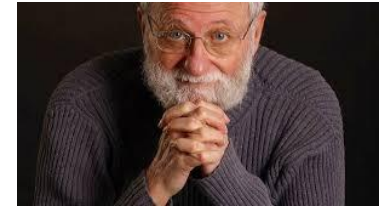
This is different from the models with the highest lift/AUC!



The future of work is Freestyle x

*The problems that we face with technology are fundamental...
We need a calmer, more reliable, more humane approach.
We need augmentation, not automation.*

- Don Norman

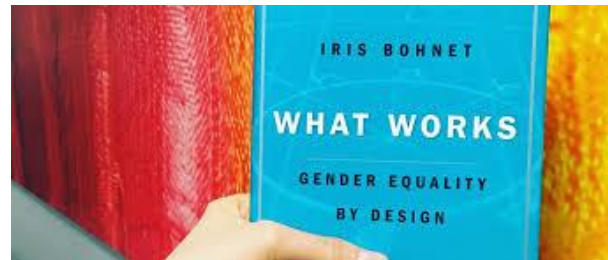


data + human judgment / empathy → decisions that are...

consistent



de-biased



informed



meaningful

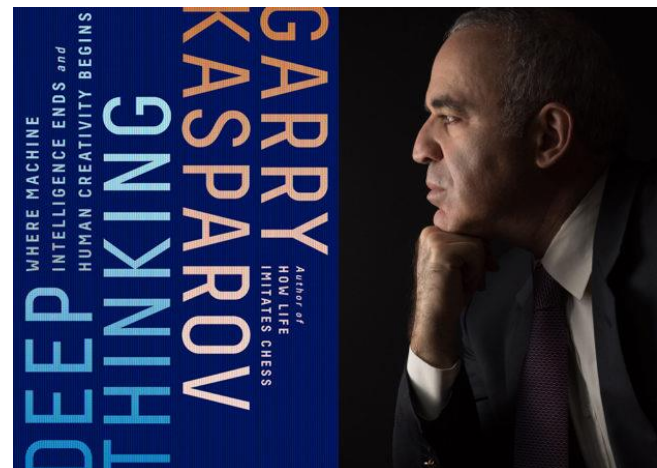


The future of work is Freestyle x

Many jobs will continue to be lost to intelligent automation...

But if you're looking for a field that will be booming for many years, get into human-machine collaboration and process architecture and design.

- Garry Kasparov, Deep Thinking



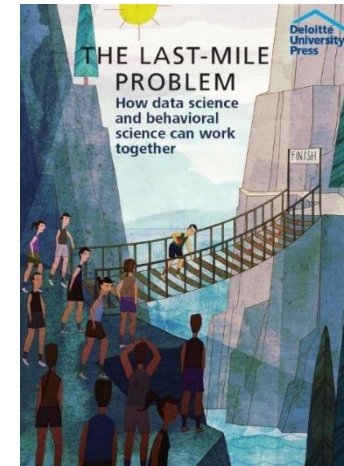
“Copies available in the lobby”

For more discussion see:

“The Last Mile Problem: how data science and behavioral science can work together”

Deloitte Review, January 2015

<http://dupress.com/articles/behavioral-economics-predictive-analytics/>



“The Importance of Misbehaving: a conversation with Richard Thaler” *Deloitte Review*, January 2016

<https://dupress.deloitte.com/dup-us-en/deloitte-review/issue-18/behavioral-economics-richard-thaler-interview.html>



“Cognitive collaboration: why humans and computers think better together”

Deloitte Review, January 2017

<https://dupress.deloitte.com/dup-us-en/deloitte-review/issue-20/augmented-intelligence-human-computer-collaboration.html>

“Smarter together: Why artificial intelligence needs human-centered design”

Deloitte Review, January 2018

<https://www2.deloitte.com/insights/us/en/deloitte-review/issue-22/artificial-intelligence-human-centric-design.html>

