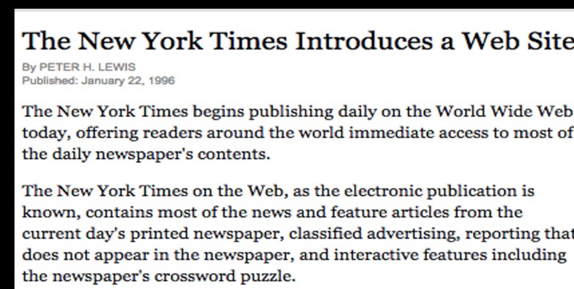
data science @ The New York Times

chris.wiggins@gmail.com

UX: 1851 vs. 1996

New-York Daily Times.

1851



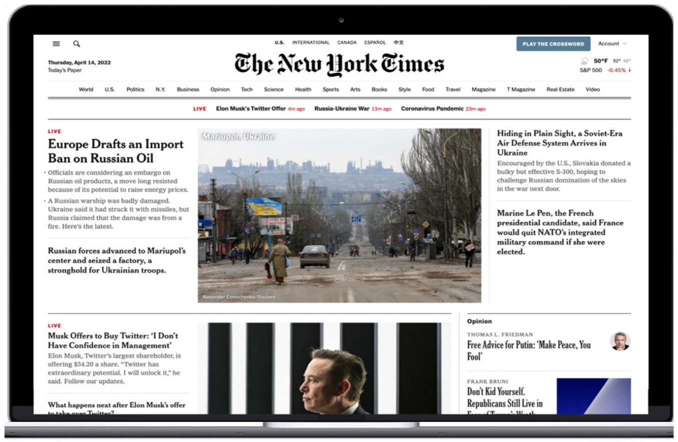
The New York Times Introduces a Web Site

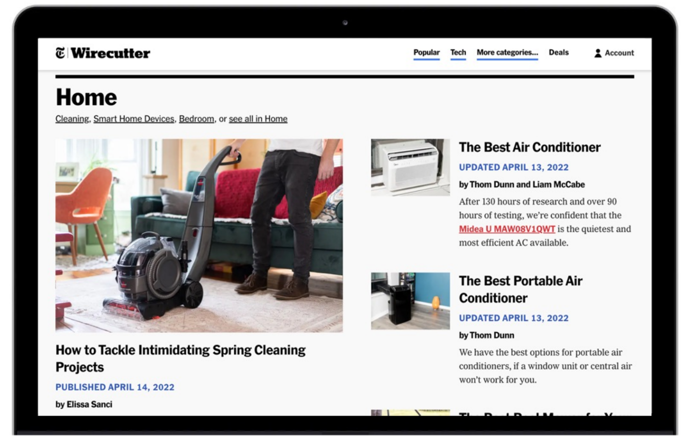
By PETER H. LEWIS
Published: January 22, 1996

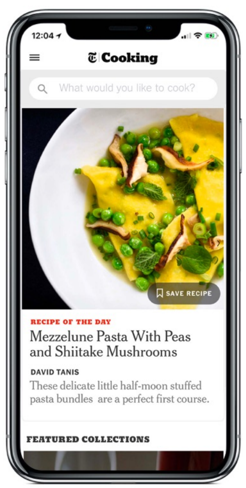
The New York Times begins publishing daily on the World Wide Web today, offering readers around the world immediate access to most of the daily newspaper's contents.

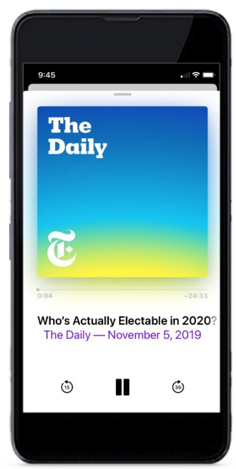
The New York Times on the Web, as the electronic publication is known, contains most of the news and feature articles from the current day's printed newspaper, classified advertising, reporting that does not appear in the newspaper, and interactive features including the newspaper's crossword puzzle.

1996


The New York Times
PLAY THE CROSSWORD
Europe Drafts an Import Ban on Russian Oil
Hiding in Plain Sight, a Soviet-Era Air Defense System Arrives in Ukraine
Marine Le Pen, the French presidential candidate, said France would quit NATO's integrated military command if she were elected.
Russian forces advanced to Mariupol's center and seized a factory, a stronghold for Ukrainian troops.
Musk Offers to Buy Twitter: 'I Don't Have Confidence in Management'
Opinion
Free Advice for Putin: 'Make Peace, You Fool'


Wirecutter
Popular
Tech
More categories...
Deals
Account
Home
How to Tackle Intimidating Spring Cleaning Projects
PUBLISHED APRIL 14, 2022
by Elissa Sanci
The Best Air Conditioner
UPDATED APRIL 13, 2022
by Thom Dunn and Liam McCabe
The Best Portable Air Conditioner
UPDATED APRIL 13, 2022
by Thom Dunn

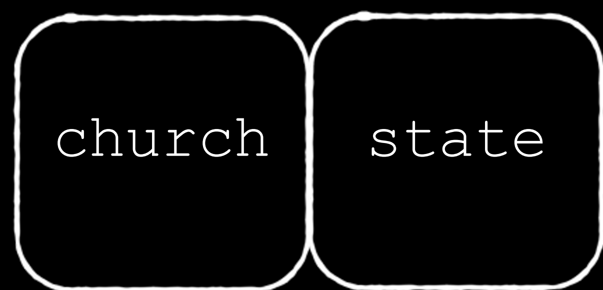

Cooking
What would you like to cook?
SAVE RECIPE
RECIPE OF THE DAY
Mezzelune Pasta With Peas and Shiitake Mushrooms
DAVID TANIS
These delicate little half-moon stuffed pasta bundles are a perfect first course.
FEATURED COLLECTIONS


The Daily
Who's Actually Electable in 2020?
The Daily — November 5, 2019


GUESS
WHICH
CRAZE
JOINS
TIMES
GAMES

2022

news: organizational chart



Data Science @ The Times

We develop and deploy
machine learning solutions
for business and newsroom problems
(using first party data)

Algorithmic Recommendations

Advertising

Media Innovations Team (MIT)

Algorithmic Targeting (ALTA)

Data Science @ The Times

We develop and deploy
machine learning solutions
for business and newsroom problems
(using first party data)

Algorithmic Recommendations

Advertising

Media Innovations Team (MIT)

Algorithmic Targeting (ALTA)

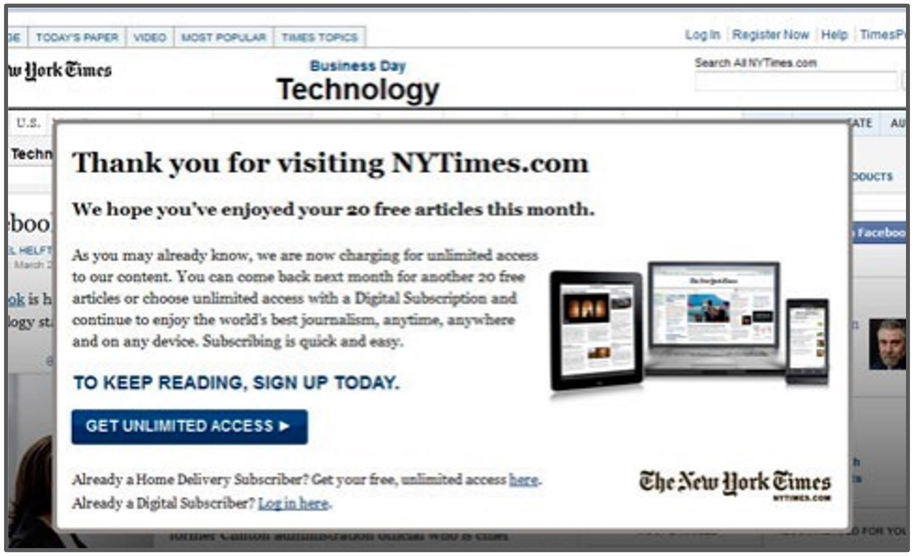
Published in NYT Open

Rohit Supekar

Aug 10 · 8 min read · Listen

How The New York Times Uses Machine Learning To Make Its Paywall Smarter

history: 2011



prognosis: doom

THE NEW YORK TIMES PAYWALL IS DESTINED FOR FAILURE

The New York Times online subscription service was a long time in the making. Originally expected to roll out in January, the plans for the so-called paywall weren't announced until last Thursday. Unfortunately, the long-awaited service is destined for failure.

PC By PCMag Staff March 24, 2011

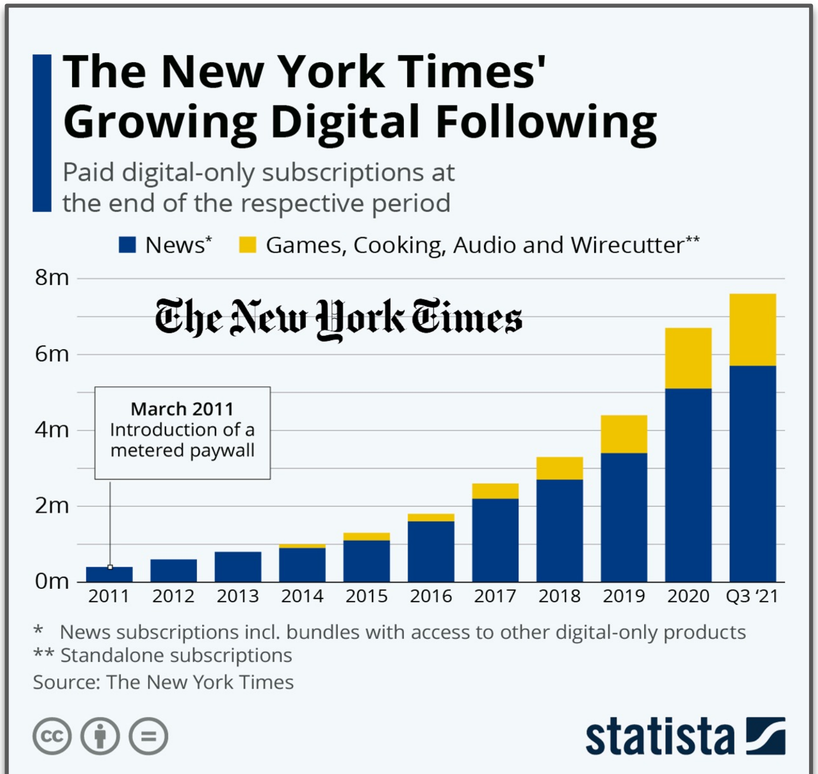
DigitalTonto
At the Crossroads of Media, Marketing and Technology...

Home About Greg Sat

5 Reasons Why The New York Times Paywall Will Fail (And Why It's Really Dumb)

2011 MARCH 23

current: not doom

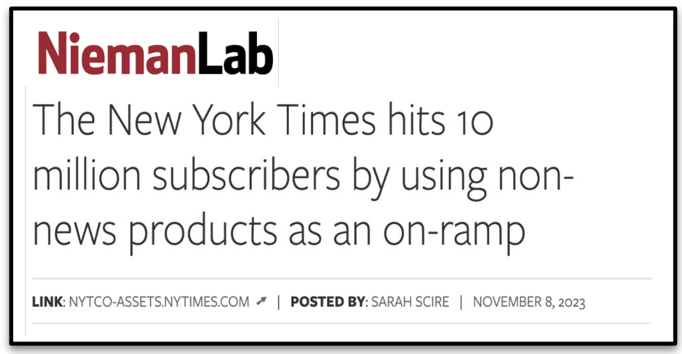


THE WALL STREET JOURNAL.

New York Times Tops 10 Million Subscriptions as Profit Soars

[Feb 2, 2022]

News publisher increases dividend, sets new target of 15 million total subscribers by end of 2027

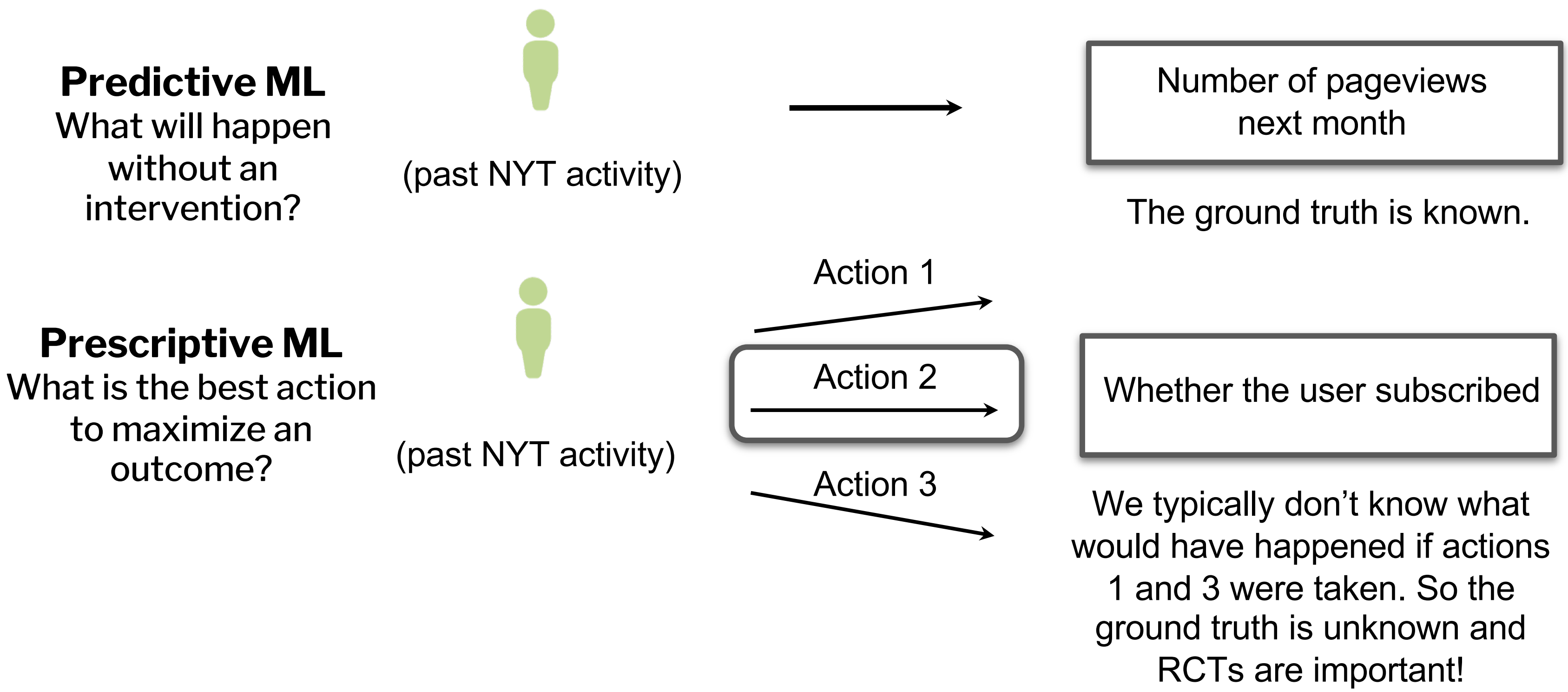


NiemanLab

The New York Times hits 10 million subscribers by using non-news products as an on-ramp

LINK: NYTCO-ASSETS.NYTIMES.COM | POSTED BY: SARAH SCIRE | NOVEMBER 8, 2023

Predictive vs Prescriptive Machine Learning



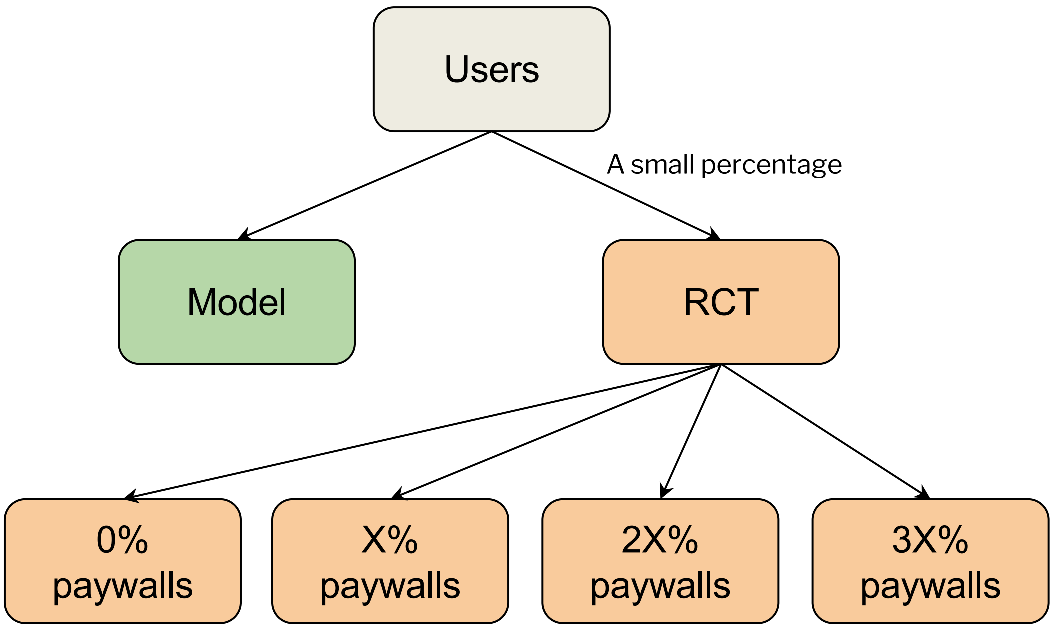
Business trade-offs

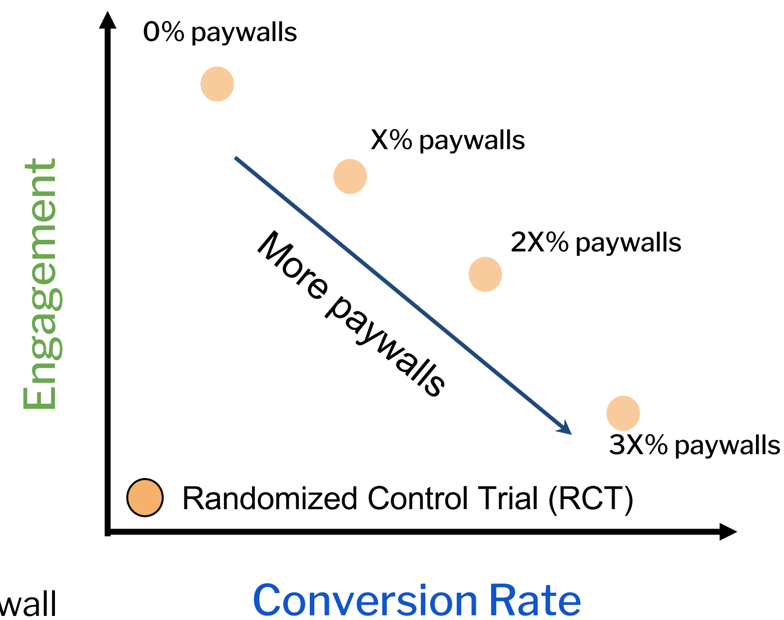
- NYT mission and business goal:
  - Engagement: We seek the truth and help people understand the world
  - Conversion: 15 million subscribers by 2027

- Engagement and Conversion have a inherent trade-off that is controlled by paywall rate.
  - More paywalls give us more subscribers but hurt engagement and reader habituation.

- We use Randomized Control Trials (RCTs) or A/B tests to understand the relationship between the engagement and conversion tradeoff.

Learning trade-offs from RCTs



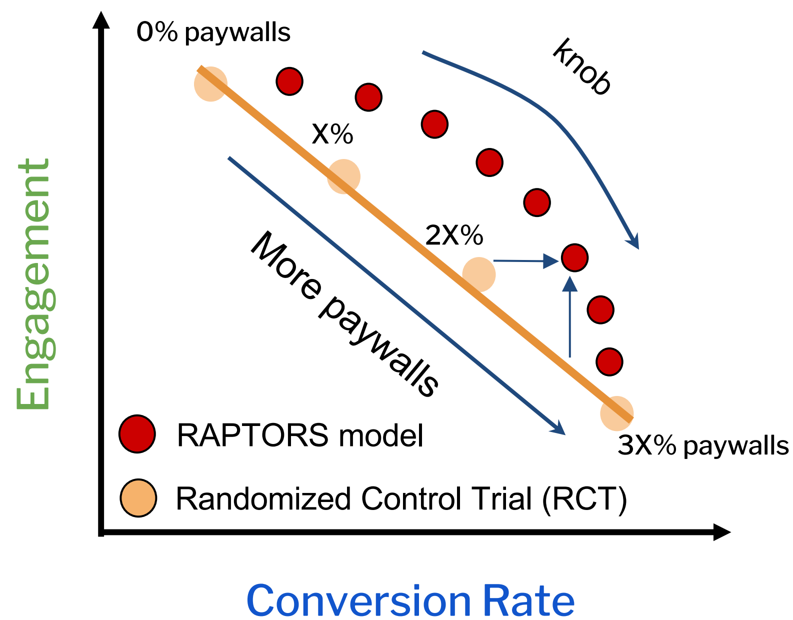


- Users in these variants have a random chance of seeing a paywall
- RCT is constantly active to capture any time-varying effects

RCT data power ML model training

- Use of RCT data:
    a. the model learns the effect of paywalls applied to access requests on the conversion/engagement tradeoff
    b. model performance can be measured

- The goal of the model is to
    - improve the trade-off by only paywalling the most worthwhile clicks
    - provide a “knob” to change the level of the trade-off



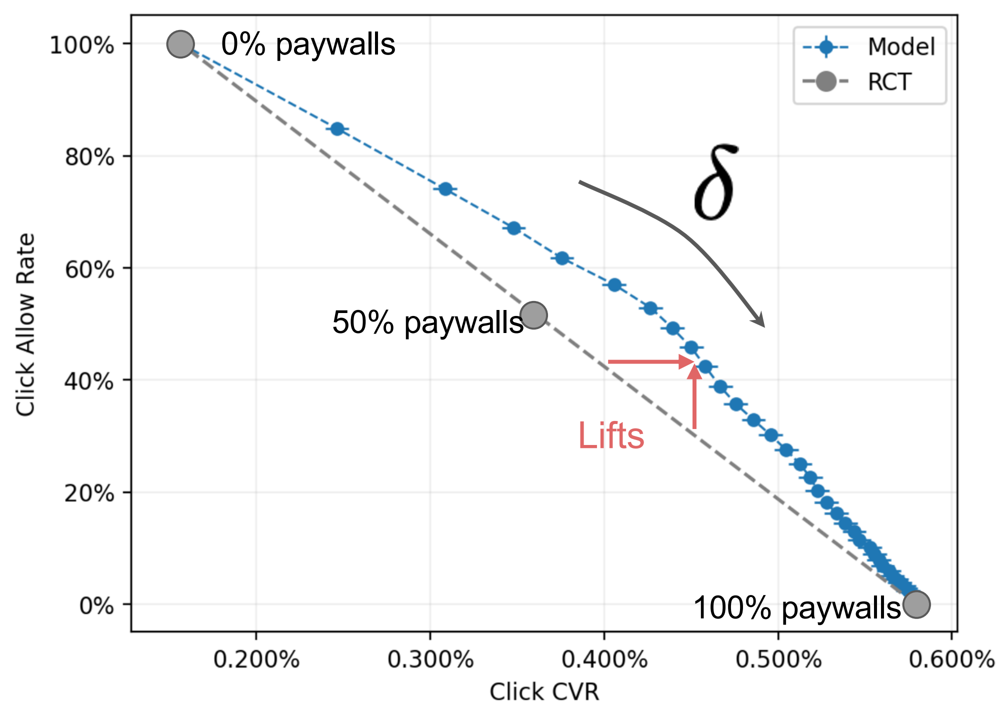
Backtesting: Offline model evaluation

How would the model have performed in the past?

Click #	1	2	3	4	5	6	...
RCT	✅	❌	❌	❌	❌	✅	...
Model	✅	❌	✅	❌	✅	✅	...

Estimate model performance

*totally fake numbers



Data Science @ The Times

We develop and deploy
machine learning solutions
for business and newsroom problems
(using first party data)

Algorithmic Recommendations

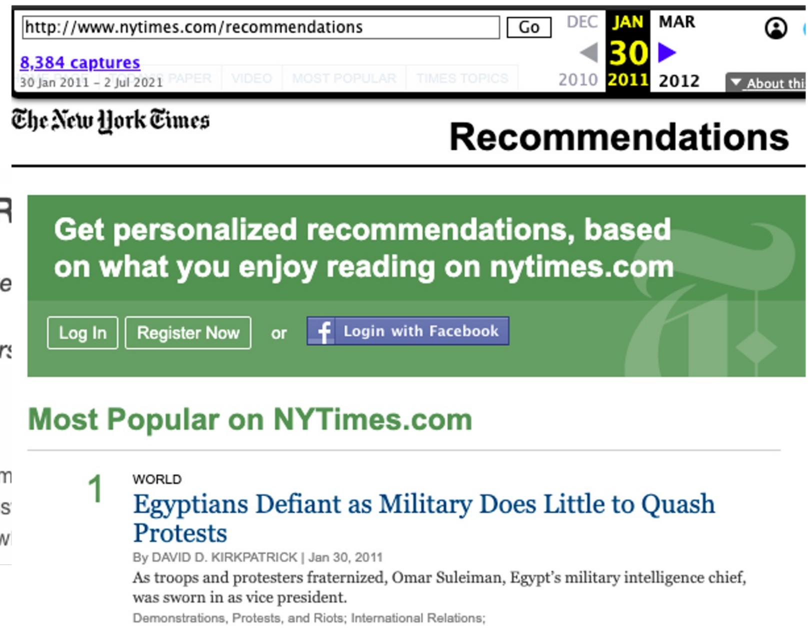
Advertising

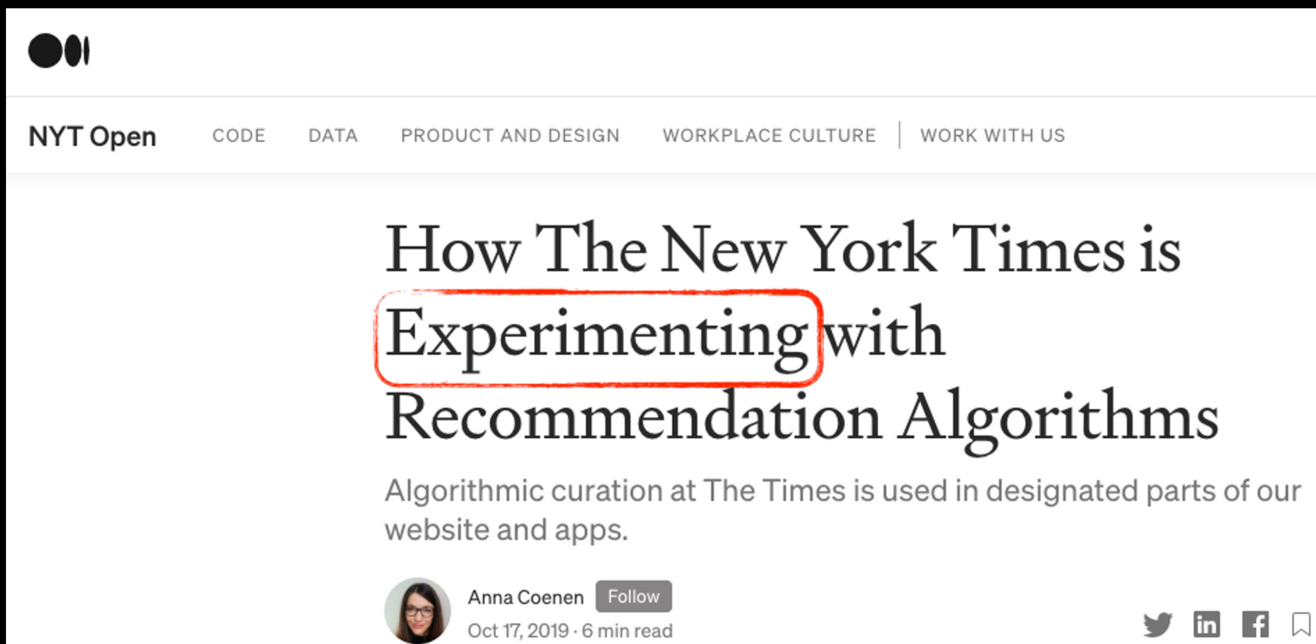
Media Innovations Team (MIT)

Algorithmic Targeting (ALTA)

2011: “Recommended for you”

- Limited to one location
- Simple algos




NYT Open

CODE DATA PRODUCT AND DESIGN WORKPLACE CULTURE | WORK WITH US

How The New York Times is Experimenting with Recommendation Algorithms

Algorithmic curation at The Times is used in designated parts of our website and apps.

Anna Coenen Follow

Oct 17, 2019 · 6 min read

critical difference: observation vs. intervention https://open.nytimes.com/how-the-new-york-times-is-experimenting-with-recommendation-algorithms-562f78624d2

NYT Open

How we design and build digital products at The New York Times

Follow

655

A contextual recommendation approach

One recommendation approach we have taken uses a class of al
called contextual multi-armed bandits. Contextual bandits lear
how people engage with particular articles. They then recomme
that they predict will garner higher engagement from readers.
contextual part means that these bandits can use additional info
get a better estimate of how engaging an article might be to a pa
reader. For example, they can take into account a reader's geog
region (like country or state) or reading history to decide if a pa
article would be relevant to that reader.

- “bandit”: Bush, Robert R., and Frederick Mosteller. "Stochastic models for learning." (1955).

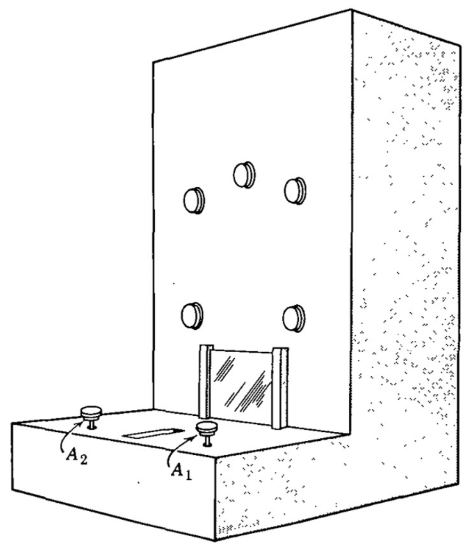
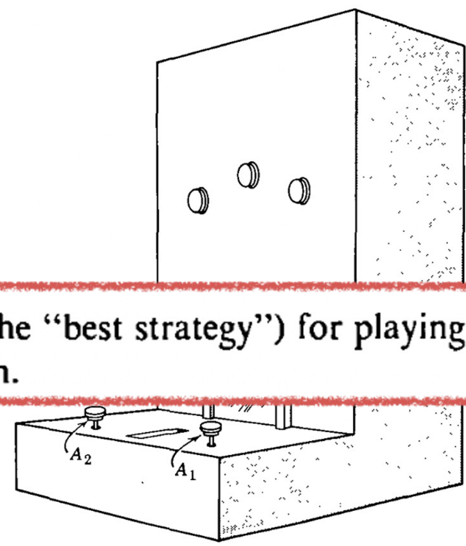


Fig. 13.2. Sketch of the two-armed bandit used by Goodnow, Robillard, and the authors. When the machine is in operation, the three upper lights are on and a soft buzz may be heard. The subject pushes button A_1 or button A_2, and this turns on the light directly above the button pushed. Poker chips are dropped into the region behind the glass.

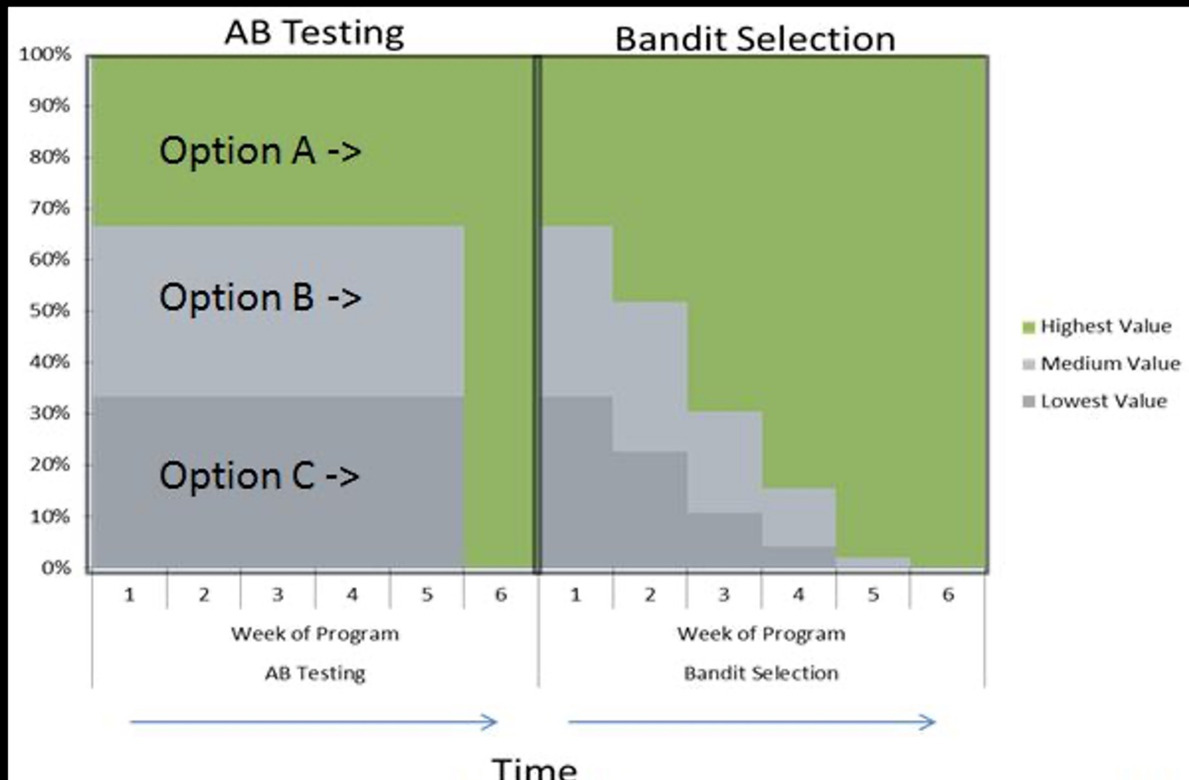
* The optimum decision rule (the “best strategy”) for playing the two-armed bandit for n trials seems not to be known.

- Bush, Robert R., and Frederick Mosteller. "Stochastic models for learning." (1955).



* The optimum decision rule (the "best strategy") for playing the two-armed bandit for *n* trials seems not to be known.

Fig. 13.2. Sketch of the two-armed bandit used by Goodnow, Robillard, and the authors. When the machine is in operation, the three upper lights are on and a soft buzz may be heard. The subject pushes button A_1 or button A_2, and this turns on the light directly above the button pushed. Poker chips are dropped into the region behind the glass.



Bandits and Adaptive Optimization, Matt Gershoff (Conductrics) 2012

How The New York Times Incorporates Editorial Judgement in Algorithms to Curate Home Screen Content

A look into how editorially-driven algorithms assist content curation on The New York Times home page.

The NYT Open Team   Follow   9 min read · Oct 2, 2024   **By Zhen Yang**

Figure 1

Figure 1. New York Times home page: algorithmically programmed modules are highlighted in different colors.

Figure 2

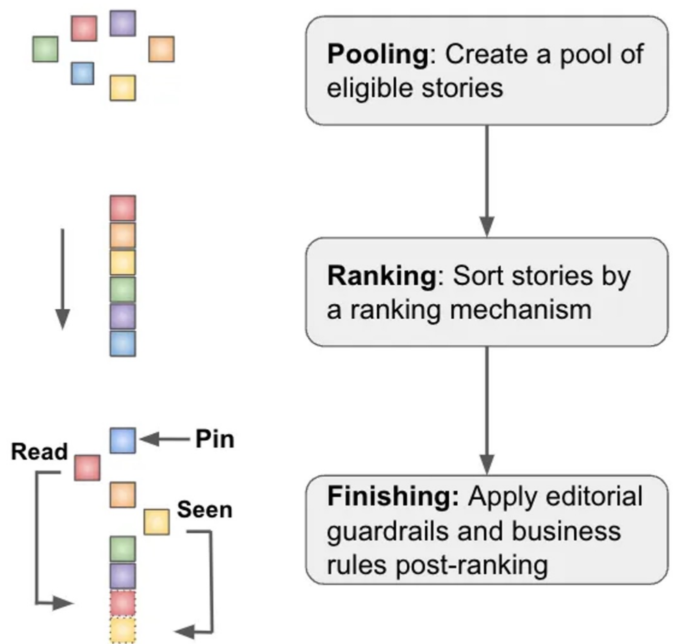


Figure 2. The process of algorithmic programming includes three steps: pooling, ranking, and finishing. Pinning an important story at the top and removing a story that has already been read or seen are examples of applying editorial guardrails and business rules at the finishing step to override the algorithmic output.

Figure 4

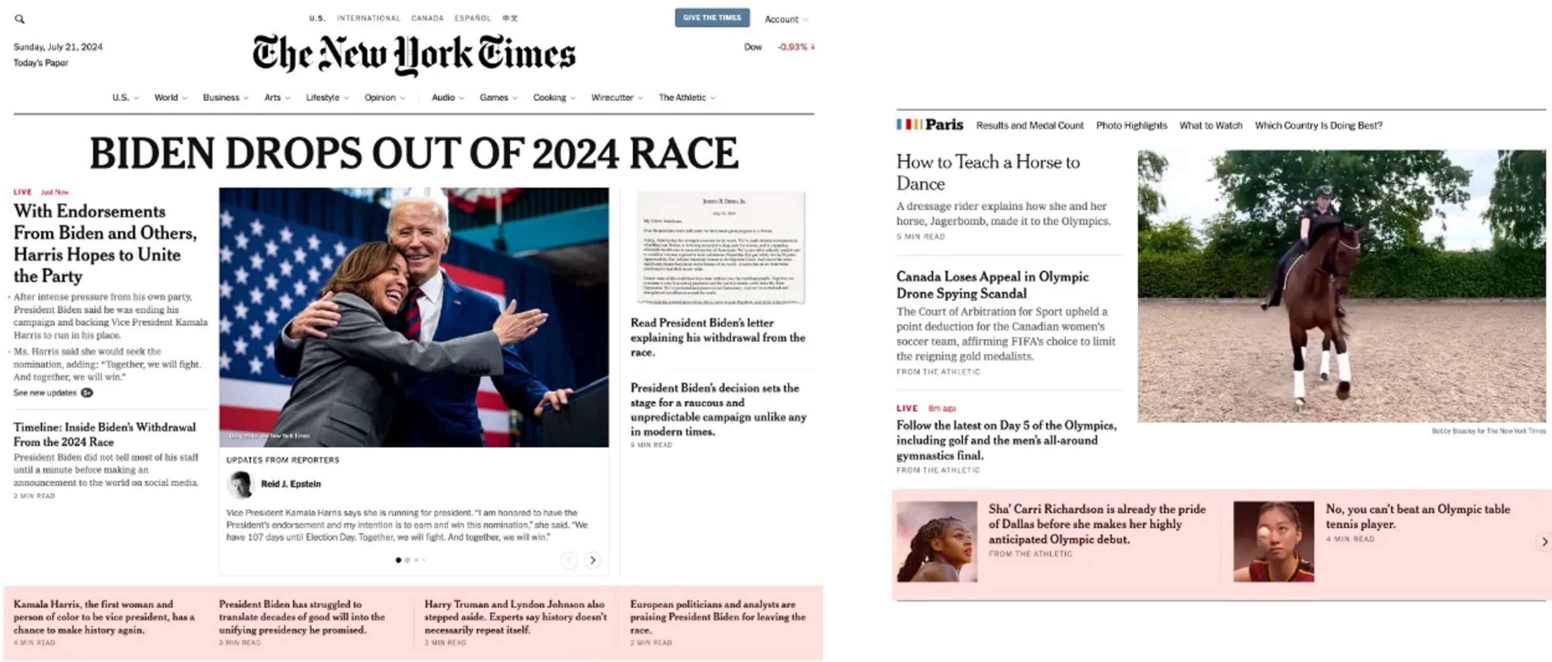
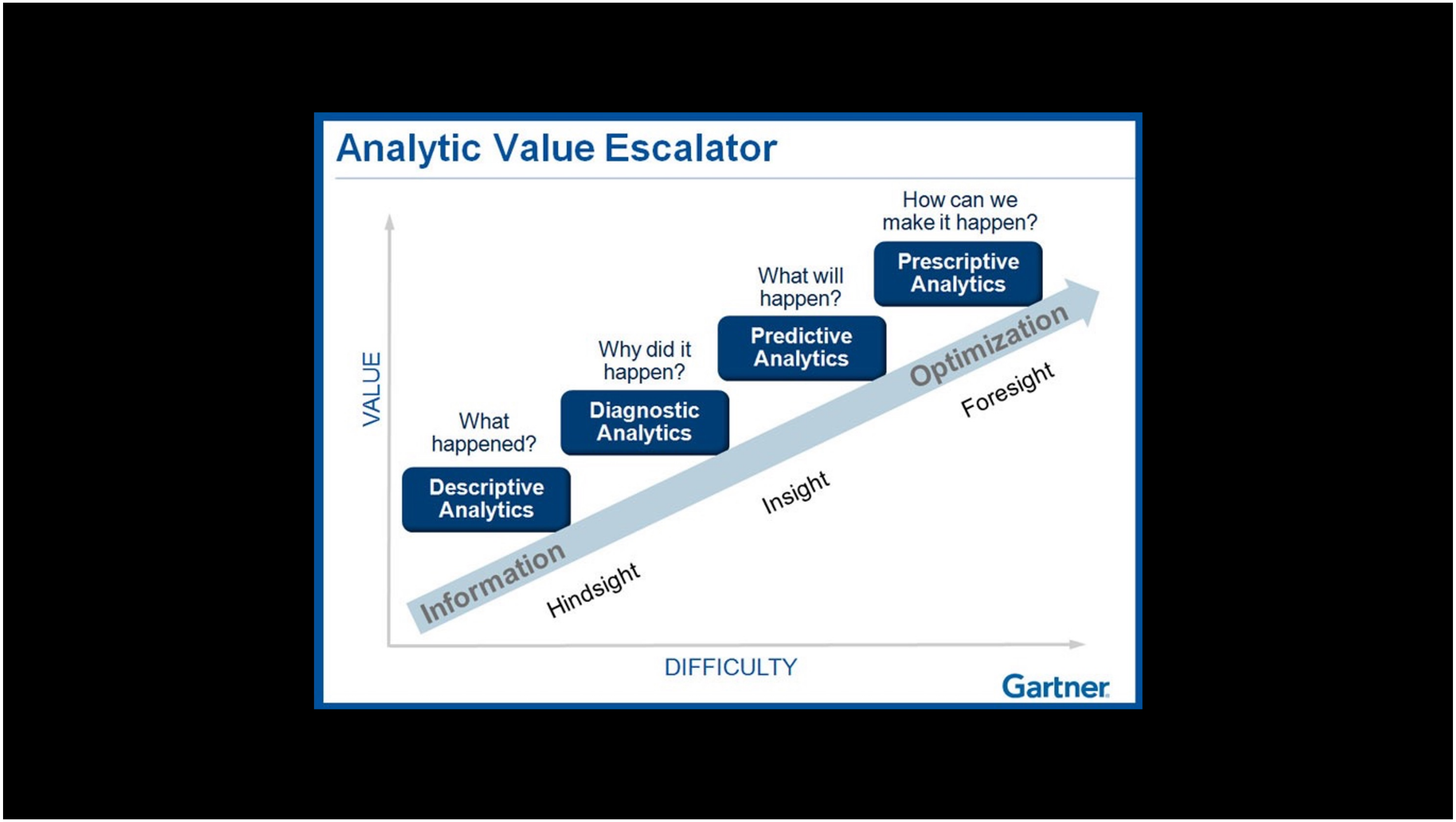
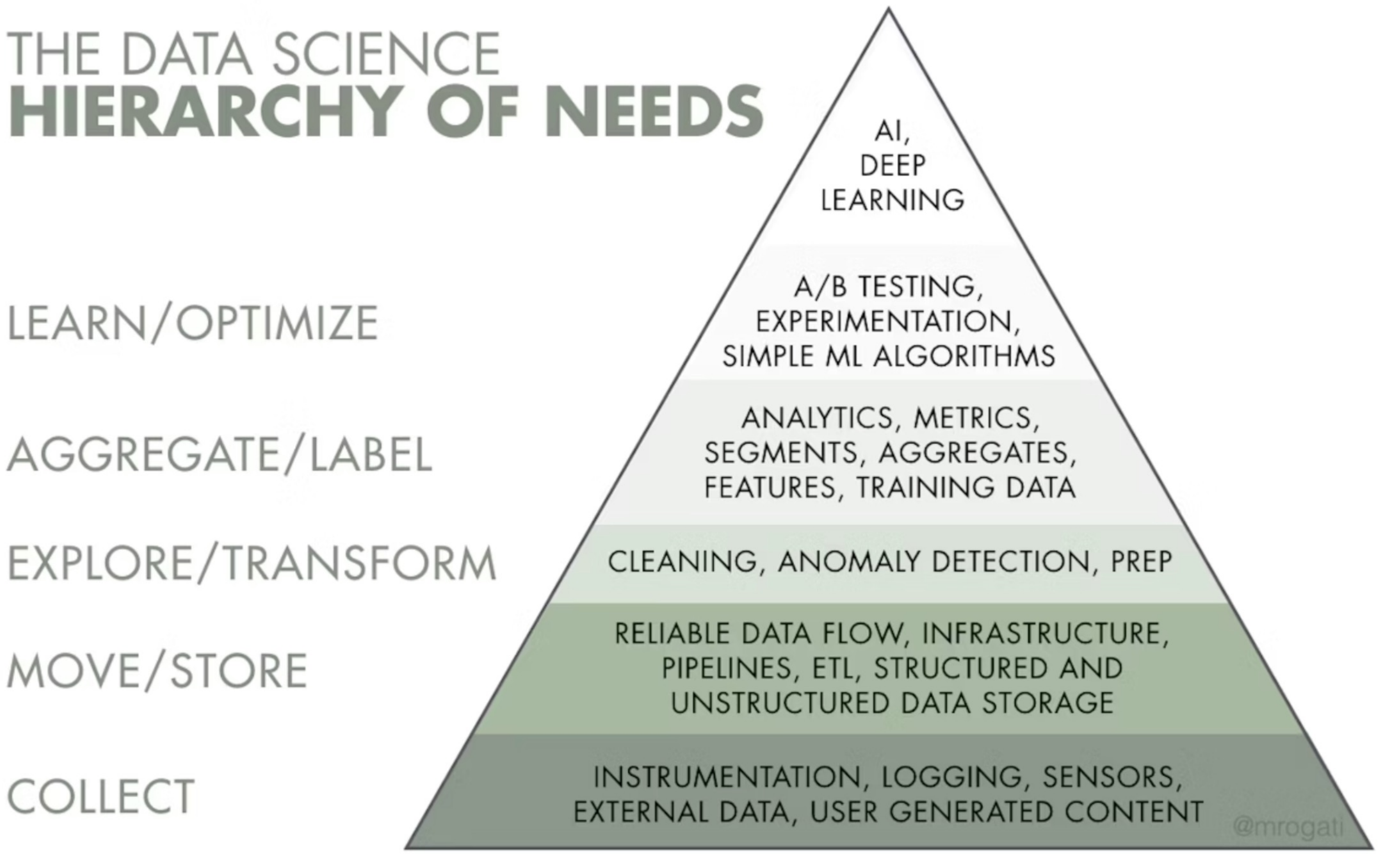


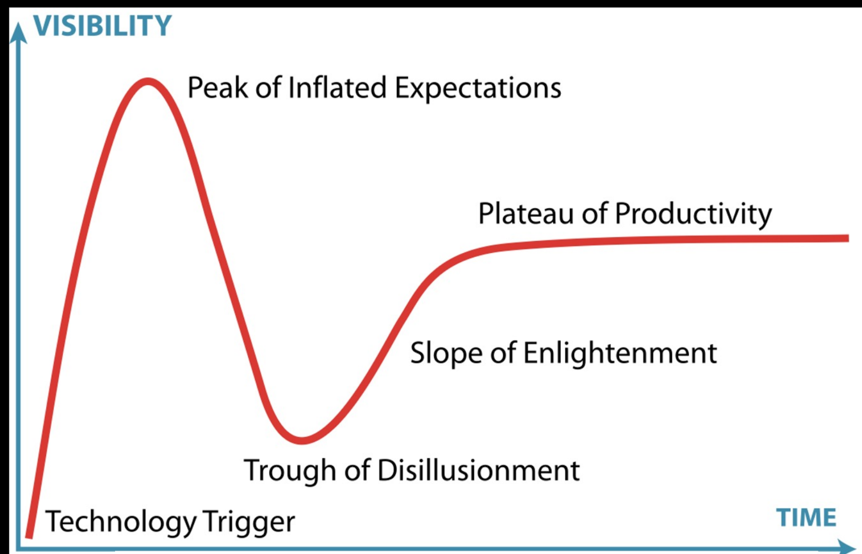
Figure 4. Self-service algorithmic modules were used high on the home page (the pink highlighted areas) during major news events: when President Biden dropped out of the 2024 race (left panel) and during the 2024 Paris Olympics (right panel).

lessons learned


Analytic Value Escalator
How can we make it happen?
Prescriptive Analytics
What will happen?
Predictive Analytics
Why did it happen?
Diagnostic Analytics
What happened?
Descriptive Analytics
VALUE
Information
Optimization
Hindsight
Insight
Foresight
DIFFICULTY
Gartner



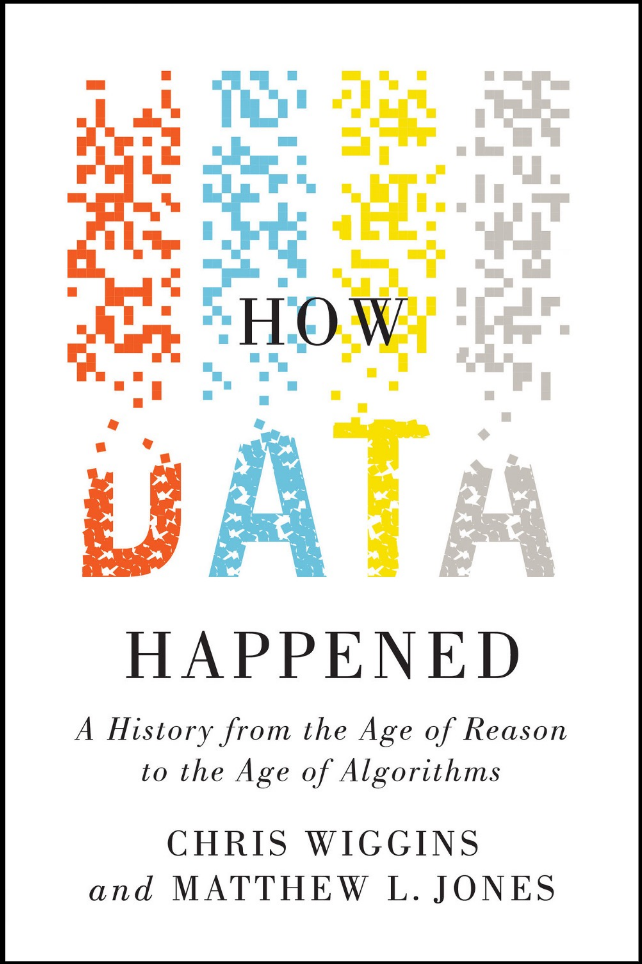
Monica Rogati June 12th, 2017 https://hackernoon.com/the-ai-hierarchy-of-needs-18f111fcc007

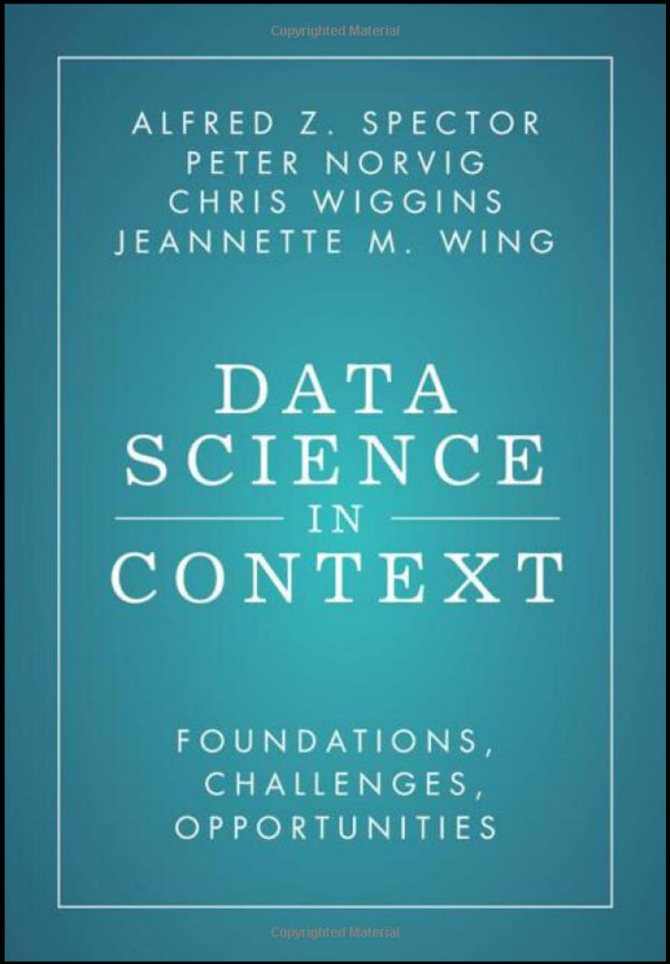


Gartner Research

Understanding Gartner's Hype Cycles

Published: 30 May 2003


HOW
DATA
HAPPENED
A History from the Age of Reason to the Age of Algorithms
CHRIS WIGGINS
and MATTHEW L. JONES


Copyrighted Material
ALFRED Z. SPECTOR
PETER NORVIG
CHRIS WIGGINS
JEANNETTE M. WING
DATA SCIENCE IN CONTEXT
FOUNDATIONS, CHALLENGES, OPPORTUNITIES
Copyrighted Material

data science @ The New York Times

chris.wiggins@gmail.com