Media and Artificial Intelligence

Matthew Gentzkow
Rite Aid posts 3Q profit

Associated Press January 3, 2018

CAMP HILL, Pa. (AP) _ Rite Aid Corp. (RAD) on Wednesday reported fiscal third-quarter net income of $81 million.

The Camp Hill, Pennsylvania-based company said it had profit of 8 cents per share. Earnings, adjusted to account for discontinued operations, came to less than 1 cent on a per-share basis.

The drugstore chain posted revenue of $5.35 billion in the period.

In the final minutes of trading on Wednesday, the company's shares hit $2.11. A year ago, they were trading at $8.18.

This story was generated by Automated Insights (http://automatedinsights.com/ap) using data from Zacks Investment Research. Access a Zacks stock report on RAD at https://www.zacks.com/ap/RAD
Game Summary:

The Yorktown Patriots triumphed over the visiting Wilson Tigers in a close game on Thursday, 20-14.

The game began with a scoreless first quarter.

In the second quarter, The Patriots' Paul Dalzell was the first to put points on the board with a two-yard touchdown reception off a pass from quarterback William Porter.

Wilson was behind Yorktown 7-0 heading into the second half. Wilson's Anton Reed tied the score with a two-yard touchdown run. The Patriots took the lead from Wilson with a two-yard touchdown run by Tanner Wall. The Patriots scored again on Adam Lunker's 29-yard field goal.
Today’s focus: **Demand side**

*AI helping consumers…*

Find what they want to see

Find what others (society, gov’t, advertisers) want them to see
Outline

1. Framework
2. AI & Matching
3. AI & Bias
4. AI & Capture
1

Framework
Media Goods

- Heterogeneous tastes (over consumers and over time)
- Quality learned through experience; sometimes not even then (e.g., truth)
- Many goods, low marginal costs

- Key problem is matching consumers to content
Sources of inefficiency

1. Consumers can’t find what they want
   - Imperfect search, matching, recommendations

2. What consumers want isn’t what’s good for society
   - Fake news, bias, Kardashians, violence

3. Gov’t, firms, etc. have other ideas
   - Censorship, capture, foreign manipulation, persuasive ads
Artificial Intelligence

Potential to dramatically improve decisions when we have:

• *Lots of data on prior cases*

• *Tightly specified decision problem*

• *Measurable, clearly defined objectives*
2

AI & Matching
Problem:

Choose Netflix movie recommendations, Google search results, Facebook newsfeed content, or informative ads to maximize consumer utility
Lots of data on prior cases?

Tightly specified decision problem?

Measurable, clearly defined objectives?
Yes

Loads of data on historical viewing, searches, browsing

Clicks, likes, time spent provide easily quantifiable objectives
Matching content to consumers is a canonical high-impact application for AI
Yet...
Hannak et al. (2017) “Measuring Personalization of Web Search”

• 90% of Google search results not personalized
• 10% that are only due to location and login status
Recommended for you, Matthew

Buy It Again in Health and Personal Care
6 ITEMS

Buy It Again in Office
5 ITEMS

Buy It Again in Other Categories
16 ITEMS

Science Fiction & Fantasy Books
99 ITEMS
Bakshy et al. (2015) “Exposure to ideologically diverse news and opinion on Facebook”

Conclusions:
• Most differences in what consumers see driven by what their friends share
• Role of algorithmic ranking beyond this relatively small
The Personalization Paradox

• In practice, personalization of search results, recommendations, etc. remains remarkably limited

• Where it’s tried, quality is mostly poor
Why?

I don’t know

But here is one hypothesis…
Possible tasks for AI

1. Rank products by average utility
2. Parse queries (learn what consumers want right now)
3. Personalization (learn stable differences between consumers)

Which is most valuable?
Figure 1: Estimated Site Utility: Conservatives vs. Liberals

- Mean utility: conservatives
- Mean utility: liberals

Notes: Figure plots estimated mean utility for conservatives \((a_j + g_j)\) on the y-axis against estimated mean utility for liberals \((a_j - g_j)\) on the x-axis. The unit of observation is an individual Internet news site. See text for details.
3

AI & Bias
Problem:

Choose accurate, informative, socially valuable content instead of what consumers appear to want
Why might private and social utility diverge?

Externalities (Becker, Downs)

- Too little demand for information
- Demand for the wrong kind of information (biased, false, violent, etc.)

“Internalities”

- E.g., partisan content as a temptation good
Nothing can now be believed which is seen in a newspaper. Truth itself becomes suspicious by being put into that polluted vehicle. The real extent of this state of misinformation is known only to those who are in situations to confront facts… with the lies of the day.

Thomas Jefferson
Lots of data on prior cases?

Tightly specified decision problem?

Measurable, clearly defined objectives?
No

Externalities are not easily measurable

Truth is not easily measurable

Both are contested
Two Worlds

1. Consumers want to know the truth, but can’t tell what sources or information they can trust, or are tempted by partisan content

2. Consumers don’t care about the truth, and just want to be told they’re right
Two Worlds

1. Consumers *want to know the truth*, but can’t tell what sources or information they can trust, or are tempted by partisan content

2. Consumers *don’t care about the truth*, and just want to be told they’re right

• The positive impact of AI likely bigger in World 1 than in World 2
Trends in the Diffusion of Misinformation on Social Media

Hunt Allcott, New York University, Microsoft Research, and NBER*
Matthew Gentzkow, Stanford University and NBER
Chuan Yu, Stanford University

September 2018

* E-mail: hunt.allcott@nyu.edu, gentzkow@stanford.edu, chuanyu@stanford.edu. We thank the Stanford Institute for Economic Policy Research (SIEPR), the Stanford Cyber Initiative, the Toulouse Network for Information Technology, the Knight Foundation, and the Alfred P. Sloan Foundation for generous financial support. We thank David Lazer, Brendan Nyhan, David Rand, David Rothschild, Jesse Shapiro, and Nils Wernerfelt for helpful comments and suggestions. We also thank our dedicated research assistants for their contributions to this project.
• Sample of 570 sites identified by others as sources of false stories

• Measure Facebook engagements and Twitter shares Jan 2015 – July 2018

• Major news, small news, and business / culture sites as comparison groups
Figure 1: Engagement on Facebook and Twitter

Panel A: Facebook Engagements

Number of sites: 38

Major News Sites

Number of sites: 78

Small News Sites

Number of sites: 54

Business and Culture Sites

Number of sites: 570

Fake News Sites

Number of sites: 570

Notes:
This figure shows monthly Facebook engagements and Twitter shares of all articles published on sites in different categories averaged by quarter. Data comes from BuzzSumo. Major News Sites include 38 sites selected from the top 100 sites in Alexa's News category. Small News Sites include 78 sites selected from the sites ranking 401-500 in the News category. Business and Culture Sites include 54 sites selected from the top 50 sites in each of the Arts, Business, Health, Recreation, and Sports categories. Fake News Sites include 570 sites assembled from five lists. The complete lists can be found in the appendix.
Figure 1: Engagement on Facebook and Twitter

Panel A: Facebook Engagements

Panel B: Twitter Shares

Notes: This figure shows monthly Facebook engagements and Twitter shares of all articles published on sites in different categories averaged by quarter. Data comes from BuzzSumo. Major News Sites include 38 sites selected from the top 100 sites in Alexa’s News category. Small News Sites include 78 sites selected from the sites ranking 401-500 in the News category. Business and Culture Sites include 54 sites selected from the top 50 sites in each of the Arts, Business, Health, Recreation, and Sports categories. Fake News Sites include 570 sites assembled from five lists. The complete lists can be found in the appendix.
Figure 2: Relative Engagement on Facebook

Major News Sites: 38 sites
Small News Sites: 78 sites
Business and Culture Sites: 54 sites
Fake News Sites: 570 sites

Notes:
This figure shows the ratio of monthly Facebook engagements over Twitter shares of all articles published on sites in different categories averaged by quarter. Data comes from BuzzSumo. Major News Sites include 38 sites selected from the top 100 sites in Alexa's News category. Small News Sites include 78 sites selected from sites ranking 401-500 in the News category. Business and Culture Sites include 54 sites selected from the top 50 sites in each of the Arts, Business, Health, Recreation, and Sports categories. Fake News Sites include 570 sites assembled from five lists. The complete lists can be found in the appendix.
This figure shows the ratio of monthly Facebook engagements over Twitter shares of all articles published on sites in different categories averaged by quarter. Data comes from BuzzSumo. Major News Sites include 38 sites selected from the top 100 sites in Alexa’s News category. Small News Sites include 78 sites selected from the sites ranking 401-500 in the News category. Business and Culture Sites include 54 sites selected from the top 50 sites in each of the Arts, Business, Health, Recreation, and Sports categories. Fake News Sites include 570 sites assembled from five lists. The complete lists can be found in the appendix.
• Consistent with the view that FB’s efforts since the election have been at least somewhat successful

• However
  • Absolute level remains high
  • Most effective strategies not mainly about AI
AI & Capture
Problem:

Governments, firms, and other third parties want to use media to *manipulate* consumers

Can AI make them more effective?  
Can AI be used to stop them?
Lots of data on prior cases?

Tightly specified decision problem?

Measurable, clearly defined objectives?
Yes, but on both sides

• Targeting ads and propaganda, surveillance, censorship

• Filtering out ads and propaganda, evading surveillance and censorship, detecting Russian robots
Around midnight on March 29, 2014, some Chinese internet night owls noticed that the hazard factor of P-Xylene (PX) had been changed from “low” to “high” on Baidu Encyclopedia—the Chinese equivalent to Wikipedia. The next morning, hundreds of protestors assembled in Maoming—a city in southern China’s industrial heartland—where a large-scale PX plant was planned. At 8:38 am, a message with pictures of the protest was posted on Sina Weibo—the Chinese equivalent to Twitter. Tens of thousands of people joined the protests, demanding responses from local officials, burning a car, and throwing bottles until police dispersed the protestors with tear gas and batons. The next day, pictures of bloodied protestors circulated online. Thousands of posts debating the PX project and condemning the government’s action appeared on various social media platforms.

In the era of advanced information technology, social media can in some cases provide a huge information shock to a country like China, in which information and public communication has been limited by government control. How does such an information shock generated by social media affect the participation of Chinese citizens in political events? And how does the Chinese government respond?

Why Does China Allow Freer Social Media? Protests versus Surveillance and Propaganda

Bei Qin, David Strömberg, and Yanhui Wu
On one hand: Discussion of sensitive topics, collective action, remains common and difficult to suppress

On the other hand: AI/ML applied to social media content can be a very effective surveillance tool
5

Conclusion
The impact of AI on media markets will be large
It will be most effective at helping match consumers to the content they (really) want

Broader impact depends on the extent to which this is also what’s good for society