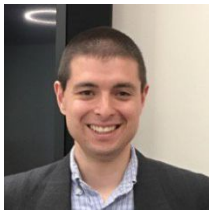
The Upworthy Research Archive
A Time Series of Experiments in U.S. Advocacy

Nov 2, 2019

CODE@MIT Conference



J. Nathan Matias
Cornell University
@natematias
natematias.com

Kevin Munger
Penn State
@kmmunger
kevinmunger.com

Andy Morris
GOOD / Upworthy
andy@goodinc.com

Goooooooooogle
Egypt
• Crisis in Egypt
• Protests of 2011
• Lara Logan
• Travel, Vacations
• Egypt Daily News
• CIA World Factbook
TED

AWESOME
MEANINGFUL
UPWORTHY
SHARE BETTER
VISUAL

FAST COMPANY

Sign up for our daily email. Enter your email address

08.07.13

How Upworthy Used Emotional Data To Become The Fastest Growing Media Site of All Time

If it makes an editor cry, it's Upworthy.

> The most important impacts of digital technology lie not in the capacity of disorganized masses to more easily speak, but the capacity of new civil society organizations to more effectively listen

Karpf, D. (2016). **Analytic activism: Digital listening and the new political strategy**. Oxford University Press.

Upworthy cofounder Peter Koechley apologizes for the clickbait 'monster' his site unleashed on the internet

Lara O'Reilly Mar 19, 2015, 8:27 AM

Upworthy's cofounder Peter Koechley has an apology to make.

Upworthy's cofounder spoke at The Guardian's Changing Media Summit in London. Business Insider/Lara O'Reilly

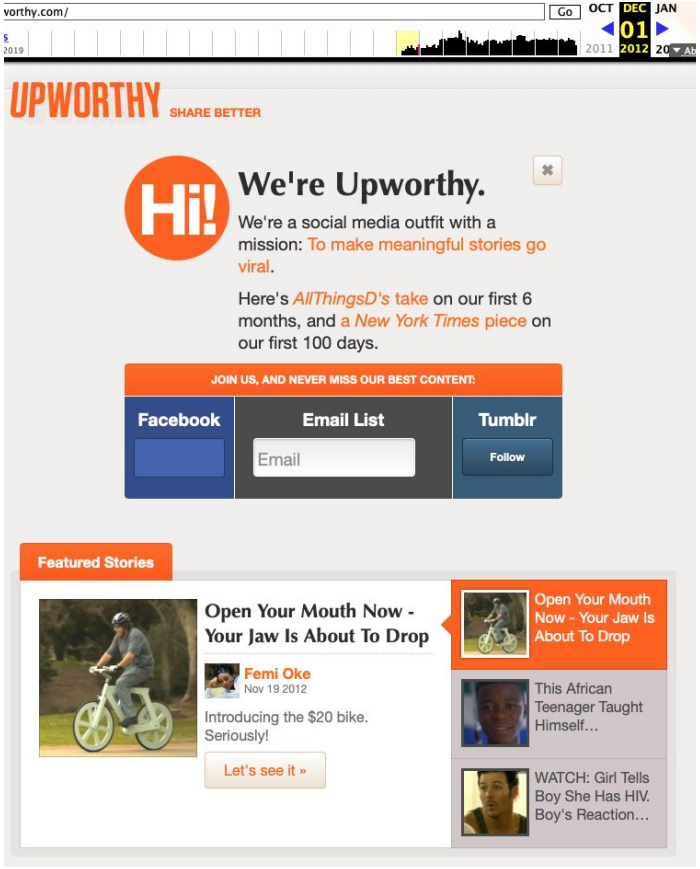
Speaking at The Guardian's Changing Media Summit in London on Thursday, Koechley apologized for what made the three-year-old site one of the fastest growing of all time: its bombastic, sensationalized, clickbait headlines such as "This Amazing Kid Just Died, What He Left Behind Was Wondtacular" and "His First 4 Sentences Are Interesting. The 5th Blew My Mind. And Made Me A Little Sick."

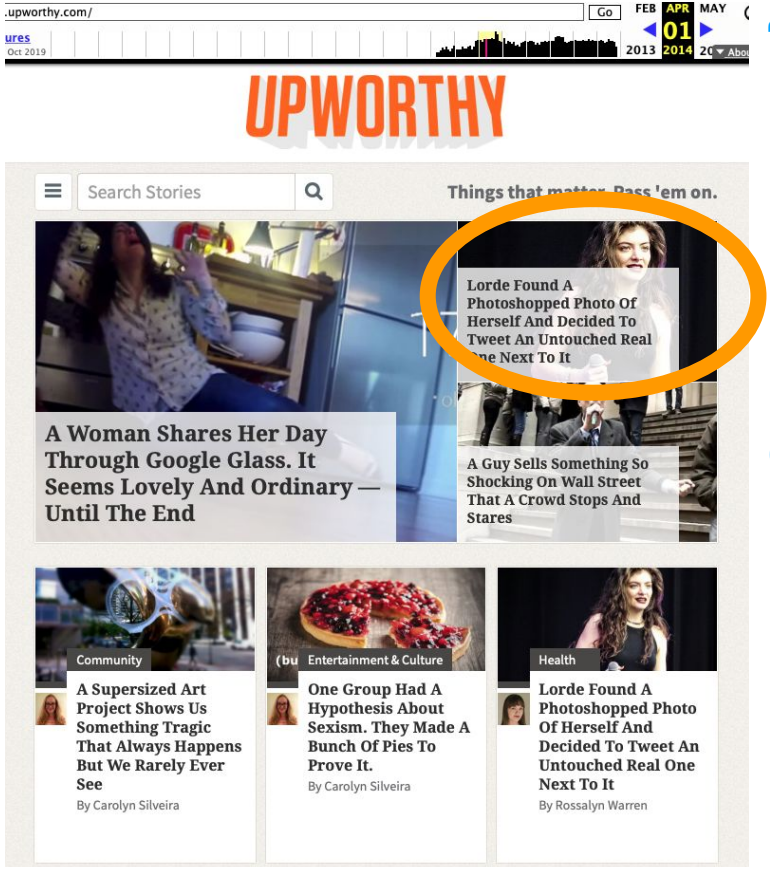
Those types of headlines were punished by a change in Facebook's algorithm last year, which sought to clean up the social-media site's News Feed to surface more "high quality" stories, resulting in huge drops of traffic for Upworthy and similar sites.

Koechley said: "We sort of unleashed a monster. Sorry for that. Sorry we kind of broke the internet last year. I'm excited going forward to say goodbye to clickbait."

The Upworthy Research Archive

Agreement to Study, Publish, and Release the Full Archive for Research and Educational Purposes





"Package"

- Headline
- Subhead
- Thumbnail
- Social Media summary (sometimes)

Outcome information:

- Impressions
- Clicks
- "Winner"

Some Jerk Corporation Patented Your Genes. Literally. I'm Not Joking.

Your Body Was Just Patented By A Corporation. You. Your Specific Body. I'm Not Joking.

Creepy Corporation Does Something So Cartoonishly Evil That We Have To Make A Dr. Evil Reference

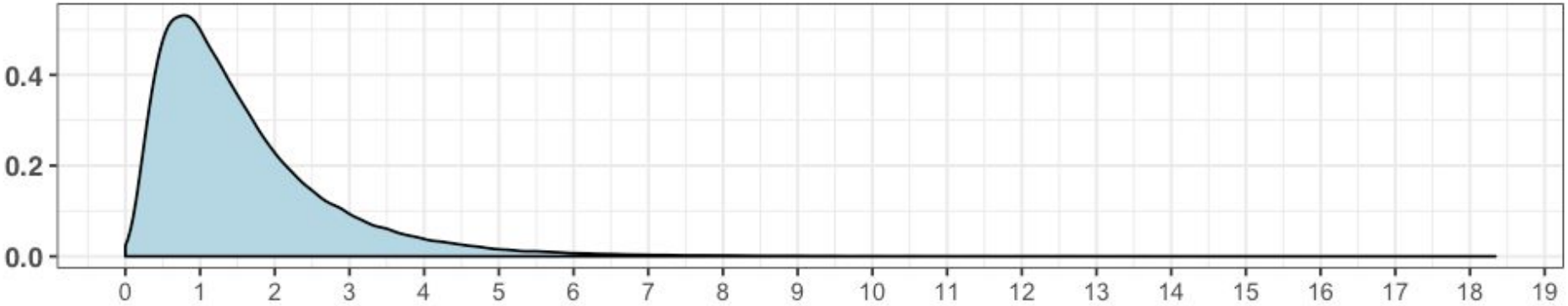
The One Where A Corporation Thinks They Can Control What You Do With Your Own Body

32,487
Valid Deployed Tests

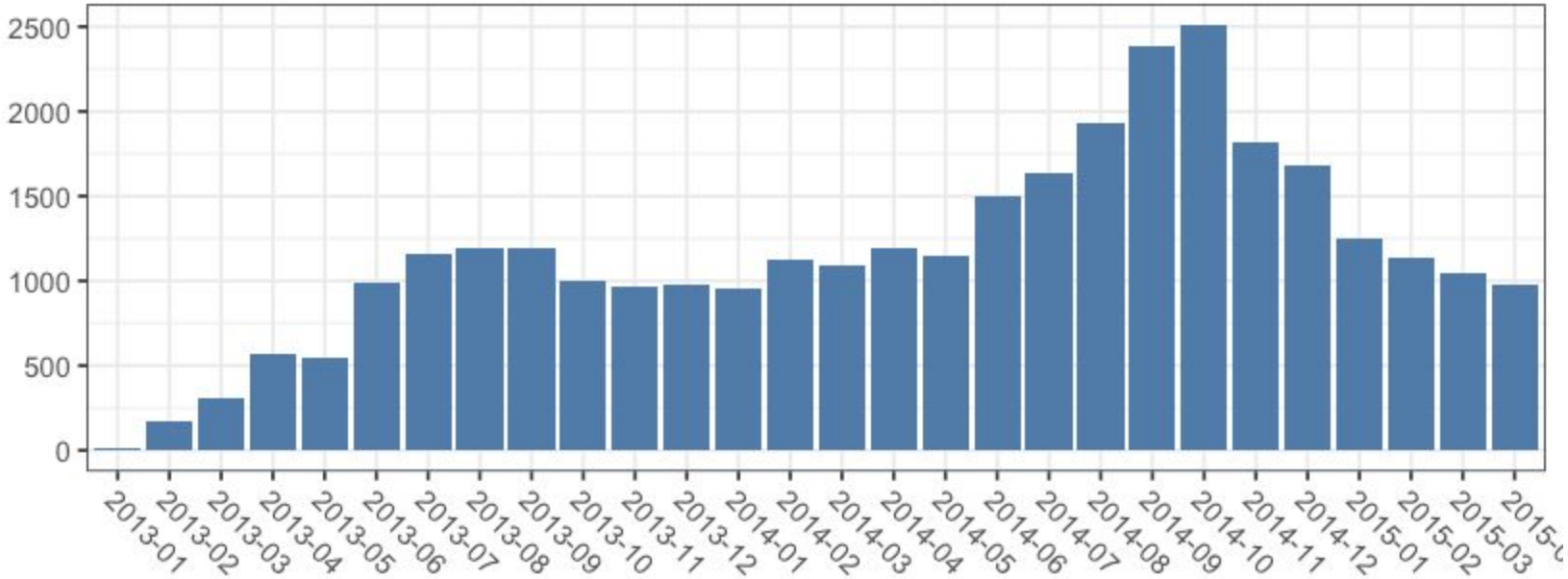
25,566
Tests That Compare Headlines

9,158
Recorded Decisions

150,817
Total Tested Packages



Clickthrough rate on Upworthy.com across 150,817 packages from 2013-01-23 to 2015-11-02.
Data analysis by J. Nathan Matias and Marianne Aubin Le Quere



A/B tests per month conducted by Upworthy between 2013-01-24 and 2015-04-30.
Analysis by J. Nathan Matias and Marianne Aubin Le Quere

Analytics Upworthy

Upworthy

Go to report

Demographics: Overview

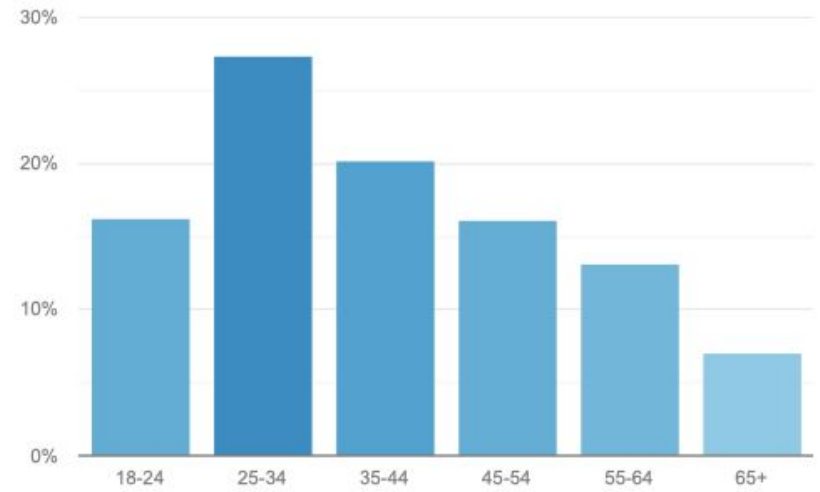
All Users
100.00% Users

Jan 1, 2013 - Dec 31, 2015

Key Metric:

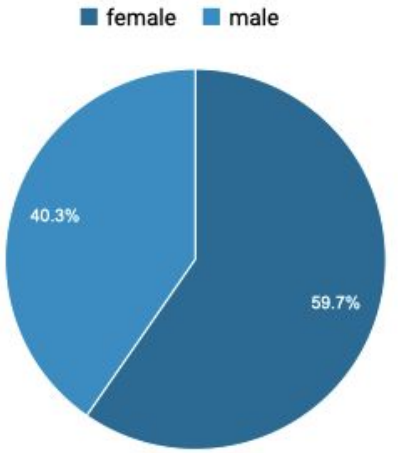
Age

15.44% of total users



Gender

14.25% of total users



The Upworthy Research Archive

Why Do This?

- RCTs produce valid *local* knowledge
- Next step for social science: aggregating this local knowledge
- Two problems:
    - Methods are still in early stages of development
    - RCTs are expensive and thus under-supplied

Ioannidis, J. P. (2018). **Randomized controlled trials: Often flawed, mostly useless, clearly indispensable: A commentary on Deaton and Cartwright.** *Social science & medicine (1982)*, *210*, 53.

Why Do This?

- “the use of multiple experiments...provides fertile ground for future work” BUT
    - There aren’t enough datasets in the public domain
    - Science advances through whoever is best at negotiating with firms, which is slow, small-scale, and inefficient

Athey, S., & Imbens, G. W. (2017). **The state of applied econometrics: Causality and policy evaluation.** *Journal of Economic Perspectives*, *31*(2), 3-32.

Scientific Advances from the Archive

- **Communication & Org Behavior:** How firms learn over time (or not) through experimentation
- **Political Science & Psychology**: Theories about kinds of headlines that people are more likely to click
- **Statistics**: Meta-analysis & other methods
- **Computer Science**: Machine learning, cybersecurity
- **Education**: Teaching with real, large datasets of experiments

Matias, J. N. (2018). T**eaching Large-Scale Digital Experimentation to Undergraduates and Graduate Students.** CODE@MIT 2018

Practical Advances from the Archive

- Evaluate strategic choices (exploration v exploitation)
- Test ways for A/B testers to incorporate information from previous tests when planning subsequent tests
- Estimate the cost/benefit of A/B testing over time
- Inspire further data sharing from other publishers

Temporal Validity

- Online media was changing rapidly in 2013-5
- Allows us to formalize “knowledge decay” due to low “temporal validity” in a context with high rate of change (Munger, 2019)
- Determine sources of bias---drift in tastes, changes in Upworthy demographics, and tweaks to Facebook algorithm

Sharing The Dataset Wisely

- Goals:
  - Support broad investigation
  - Minimize risk of data mining, p-hacking overfitting
  - Eventually publish full dataset
- Current ideas:
  - Provide training set or synthetic dataset
  - Organize a special issue & a pre-registration process
  - Inspirations: Fragile Families, Psych Science Accelerator

Questions or Ideas? Contact Us:

nathan.matias@cornell.edu

kmm7999@psu.edu

Acknowledgments

Marianne Aubin Le Quere

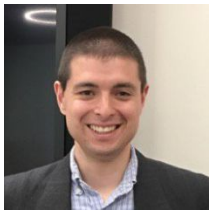
Cornell University

mariannealq.com

The Upworthy Research Archive
A Time Series of Experiments in U.S. Advocacy

Nov 2, 2019

CODE@MIT Conference



J. Nathan Matias
Cornell University
@natematias
natematias.com

Kevin Munger
Penn State
@kmmunger
kevinmunger.com

Andy Morris
GOOD / Upworthy
andy@goodinc.com