Lindsay Goddard

Topic: An idea for a new Twitter product targeted at newsrooms to help these organizations with News Discovery

News Accuracy Checker
Last updated 1/27/2019

Background

Eric Tucker, cofounder of a marketing company in Austin, TX with just 40 Twitter followers, tweeted a picture of coach busses with the tagline: "Anti-Trump protestors in Austin today are not as organic as they seem. Here are the busses they came in. #fakeprotests #trump2016 #austin."[1]

The problem was, he got it wrong. But it was too late.

The news story was shared at least 16,000 times on Twitter and more than 350,000 times on Facebook."[2] It was published on the conservative blog Gateway Pundit as "Figures. Anti-Trump Protesters Were Bussed in to Austin #FakeProtests" and shared on prominent Facebook pages, such as Robertson Family Values and Donald Trump Commander in Chief 2020.

Reporters spent valuable time investigating the accuracy of the source and only a few reached out to the bus company directly for confirmation.

False news spreads fast. "False news stories are 70% more likely to be retweeted on Twitter than true stories."[3]

Problem

Newsrooms rely on Twitter as an essential newsgathering platform for breaking news. Without the right guardrails in place to curb fake news and misinformation, a newsroom can spend precious time investigating the reliability of a source; or worse, include inaccurate info within publication.

Solution

[1] https://www.nytimes.com/2016/11/20/business/media/how-fake-news-spreads.html?_r=0
[2] https://www.nytimes.com/2016/11/20/business/media/how-fake-news-spreads.html?_r=0
[3] https://www.marketwatch.com/story/fake-news-spreads-more-quickly-on-twitter-than-real-news-2018-03-08

Just in time for the 2020 election, we want to give publishers a dashboard tool that analyzes and surfaces fake news and misinformation.

What user needs are we solving for?

Our publishers use Twitter because they want to get quickly caught up on real-time news events throughout the day and gather information for potential stories. As they come across breaking news events, they want to be certain the information is accurate -- and they want this in a quick, digestible manner.

Business goals

We aim to increase publisher's daily habituation with Twitter. To establish a relationship of trust with our publishers; increase their engagement with the overall Twitter brand; and allow us to remain competitive with other social media organizations.

What are we building?

We're building a tool for publishers that monitors and highlights the accuracy of tweets. This tool is available on a dashboard within [Twitter Analytics](). We plan to use the algorithms that power the dashboard to mark tweets themselves with a visual "accuracy" indicator.

Dashboard view for monitoring
- **Table of trending tweets**
  - Tweet
  - Tweet category (Entertainment, Politics, etc)
  - Username (displays verified account badge when available)
  - Username responsible for the original post
  - User credibility score % - based on an algorithm that crawls a user's past behavior
  - Start date and duration of the Tweet
  - Number of retweets and shares
  - Audience size
  - Truth-o-meter % - an algorithm that assembles the expected accuracy of the tweet
- **Search engine field** for users to paste specific tweets into
- **Map chart showing location and scale of shares**

Visual "accuracy" stamp on Tweets

- **Tweet symbol** - a tweet that is 99% inaccurate will be stamped with a symbol
- **User report button** - an option for marking a tweet as inaccurate. A notification is sent to a certified, third party fact checking service to investigate.

The data monitoring algorithms we use must be carefully vetted to avoid factual news from being marked as inaccurate. When confirming technical approach, some consideration should be given to consulting and potentially integrating with experienced third party data companies so algorithms have a lesser risk of bias.

We must also consider that not all fake news is intentional; some is simply misinformation by innocent individuals -- as was the case with Eric Tucker's tweet about political protests.

Competitive analysis

Within the market, there are several low-fi news monitoring sites and browser extensions that are limited in their features: analyzing full websites or pages rather than individual tweets; visual representations of rapidly spreading news with no accuracy score measurement; and manual fact-checking services that do not happen in real-time.

Additionally, no well-known third party companies or media organizations have developed robust news accuracy tools: CrowdTangle started testing the ability for users to flag tweets as inaccurate; Facebook does not offer a dashboard monitoring view but does give users the ability to mark posts as inaccurate.

If news accuracy tools become more sophisticated and competitive in the future, Twitter's point of differentiation will always be strong due to its access to internal information for robust data views (not as customizable when funneled through third parties); its ability to embed visual indicators directly within the user experience (on Tweets, etc); and its well-known brand prominence, which establishes more trust than smaller scale sites.

- **University of Michigan's Iffy Quotient** allows users to track an overall news accuracy score for both Twitter and Facebook. Not useful for individual tweets.
- **Trive**, a browser extension that alerts users when they're on a webpage with unreliable news sources. Not useful for individual tweets.
- **Botometer**, a more widely used Twitter tool that gives usernames a score based on how likely it is to be a bot.
- **Hoaxy**, a small scale site that allows users to search for a URL and see a visual of a viral news article spreading. It also shows which claims have been fact-checked.
- **PolitiFact**, a site that manually fact checks political statements and assigns a truth-o-meter score based on accuracy. Manual fact checking only.
- **Snopes**, a well-known site that fact-checks and investigates potential statements of misinformation. Manual fact checking only.

- **CrowdTangle,** a popular third party social media gathering tool used amongst publishers, recently rolled out a small [test feature](#) that allows users to mark posts as fake new. They do not offer a dashboard view for monitoring potential accuracies of posts.
- **Facebook** gives users the ability to flag posts that may be inaccurate and work with [third-party fact checking services](#) in a number of countries. They do not offer a dashboard view for monitoring potential accuracies of posts.

What is our scope?

We'll focus the test to a small percentage of the US audience. Pending technical estimation and user feedback, we may break this down into separate launches for the dashboard view and tweet visuals.

How will we measure success?

Quantitative
- Primary: Usability of the dashboard tool based on time spent, views and clicks
- Secondary: At least 10% of users flag a tweet for misinformation

Qualitative
- Overall positive satisfaction score (4/5) from stakeholders with the new tool (survey)

How will we validate our assumptions?

- Prototype various solutions, small and large, and get qualitative feedback from users
- Understand current user base in terms of usage and baseline metrics/cohorts
- Build out more robust tests to A/B test any hypotheses where we feel quantitative data is needed in order to gain additional confidence
- Perform more in-depth competitive analysis to understand what others are doing and assess what we like/dislike

What are our key milestones?

Tentative:

- Analysis, competitive analysis, hypothesis gathering: April 1 - May 1 (4 weeks)
- User research prep: May 3 - 19 (2 weeks)
- User research: May 20 - June 3 (2 weeks)
- Iterative testing and technical exploration: June 4 - July 16 (6 weeks)

- MVP definition: July 16 - 30 (2 weeks)
- MVP Build-out: begin September 1
- Expansion: TBD

How will we communicate project updates/status?

- Bi-weekly meetings with key stakeholders to discuss latest updates and blockers
- Monthly email update to all stakeholders on RACI list
- Slack channel for any questions / feedback

What is our plan to iterate?

- Experiment with allowing third party integration. The majority of publishers use CrowdTangle for social media gathering, so usage may be higher on third party platforms
- Implement email notifications/digest to alert publisher teams for more convenient access (and foster more habituation)
- Investigate the value in setting a monthly subscription fee
- Roll out fact-checking services in non-US countries

Who makes up the core team and key stakeholders?

Responsible	Lindsay Goddard (Product) Open (Design) Open (Engineering) Open (Project)
Accountable	Open (Product)
Consulted	Open (Product) Open (News Curation) Open (Advertising) Open (Data) Open (Legal) Open (Marketing)

Informed	Open (Customer Care)

Resources

TweetDeck
Twitter Analytics