Twitter Thread by Surya Mattu



Surya Mattu
@suryamattu

We've spent the last ten months building #CitizenBrowser, a project that aims to peek inside the Black Box of social media algorithms, by building a nationwide panel to share data with us. Today, we are publishing our first story from the project. /1

.@corintxt crunched the numbers and found that after Facebook flipped the switch for political ads, partisan content elbowed out reputable news outlets in our panelists' news feeds. https://t.co/Z0kibSBeQZ /2

You can learn more in our methodology, where we describe how we did this and what steps we took to ensure that we preserved the panelists' privacy. https://t.co/UYbTXAjy5i /3

Personally, this project is the culmination of years of experiments trying to figure out how to collect data from social media platforms in a way that can lead to meaningful reporting. I've described a couple of highlights below ■ /4

My first attempt was in 2016 at Propublica, when I was working with @JuliaAngwin . We were interested in seeing if there was a difference in the Ad interests FB disclosed to users in their settings and the interests they showed to marketers. /5

We collected the data set of groups FB showed marketers through the FB Ad portal quite easily
but collecting what they showed to users was a non-trivial problem. /6

In order to collect this data I built a browser extension that let readers share their data with us. We released this as a part of @propublica's Breaking The Black Box series. https://t.co/OD4A61n8WN /7

Sadly, this extension no longer works, but the 52k categories we collected from volunteers are still available to download. https://t.co/qslgqE5nJt /8

It was during this investigation that we first learned about the now infamous 'ethnic affinity' ad categories on Facebook. https://t.co/C0A1I0pMG1 /9

The second attempt was with @kashhill at Gizmodo investigating how Facebook's People You May Know algorithm suggested people to befriend. /10

Since FB doesn’t have an API for this service. I built a tool that allowed people to collect all the friends' suggestions Facebook made to them on a daily basis. If they found something interesting they could share it with us. https://t.co/d2e3PheugZ /11

Facebook wasn’t happy about our tool and asked us to take it down. To this day there is no easy way for users to get their friend suggestions data from Facebook. https://t.co/n83p5zS7SM /12

#CitizenBrowser is the next step in this journey. We built an app that we paid panelists to install, which collects data from their Facebook feeds at regular intervals. /13

It’s an ambitious experiment. But given that there is no other way to see what is happening on this important platform, we gave it our best shot. Grateful for @JuliaAngwin and @nabihasyed's fearless leadership that allowed us to push the limits. /14

Also grateful for our team, who just ran a marathon at sprint speed to get this project done! @angiewaller @leonyin @corintxt @jonkeegan @sammorrisdesign @mynameisfiber @magiccia @cortesi @jeffcrouse @iaaaan @rinapalta @simplysecureorg /eof