Audio file

2020-06-20_Training_NarrativeIdentification_[Name redacted] 1.m4a

Transcript

00:00:37 Speaker 1

My quick intro would just be that, uh, we're going to be running through some scenarios. I was going to do some agenda setting and protocol thing stuff, but I think mostly it's just [Name redacted] kind of talking through this process at a high level then walking through a scenario. I think with a group this size, feel free to like.

00:00:57 Speaker 1

You know, jump in and ask a question if you want, but we'll have an open discussion after [Name redacted]'s done walking through the scenario to kind of think about, you know, a couple of key questions he's considering, like how do we help to better automate these things or how can we help people come into this work more easily as they join triage and things like that?

00:01:17 Speaker 1

And then kind of maybe dig into one of the newer options, which it sounds like might be a continuation of the example that we'll be walking through. So I tossed to you [Name redacted], it's all you.

00:01:29 Speaker 2

All right, so.

00:01:30 Speaker 2

As [Name redacted] said.

00:01:31 Speaker 2

Here's kind of the agenda pretty much as.

00:01:34 Speaker 2

You laid out.

00:01:35 Speaker 2

The only the only adjustment is the a lot of the COVID 9G stuff has been kind of sanitized and pulled down, or at least not easily queueable. So we're probably going to walk through.

00:01:50 Speaker 2

Corona palooza. I played around with that a little bit yesterday and it was pretty fun, especially with kind of the the excellent memes going back and forth kept me laughing for for a little bit of time.

00:02:03 Speaker 2

So again, this is this is less of a me kind of briefing and telling you about stuff. I want this to be a little bit more interactive. This is kind of what I do for a living is threat analysis. So a lot of the, a lot of this stuff kind of jumps out at me because.

00:02:23 Speaker 2

Consistently looking through feeds and and able.

00:02:26 Speaker 2

To kind of.

00:02:26 Speaker 2

Pick out the the disparate or the the outliers. Just because I've been following geopolitics and and nation state adversaries for quite some time. So if there's anything that that you would like me to dive a little deeper.

00:02:42 Speaker 2

That's kind of why I left the slides fairly bland. I like a more back and forth than kind of me just talking to you, so feel free to jump in, cut me off otherwise I'll talk into a wall so.

00:03:00 Speaker 2

Get into it.

00:03:02 Speaker 2

So the first thing of of threat analysis is for a threat the the definition of a threat is is capabilities and intent. And the way I like to describe it is if an adversary has capabilities but known.

00:03:20 Speaker 2

But no intent, it's it's a.

00:03:25 Speaker 2

It's not that big.

00:03:26 Speaker 2

Of a it's not as big of.

00:03:27 Speaker 2

A threat as if they both align. They have the capabilities and the intent. Then that's an operational threat. If they have the intent and no capabilities, that's an aspirational threat.

00:03:39 Speaker 2

So before you kind of can do that, when I bring out any new analysts it, there's there's kind of a.

00:03:46 Speaker 2

A kind of walk through and read and establish the baseline. So you always have to establish what what the current is, right? So without that you don't know. You can't know how to identify outliers. This goes from cyber. This goes to narratives. This goes to.

00:04:07 Speaker 2

This applies to pretty much anything.

00:04:10 Speaker 2

Go to your car.

00:04:11 Speaker 2

To know that your car is not working right, you can't just walk in and and start driving it on day one and be like yeah, this isn't working right. If you've never driven a car before, you have to have that baseline of knowledge. So any new analysts coming on, they have to kind of go through the steps, read the news.

00:04:31 Speaker 2

And then you kind of ask.

00:04:33 Speaker 2

Some of the more senior ones kind of asked the questions. Right. OK. So what are you seeing? Well, here's kind of the some of the assumptions that I've made here's kind of some of the some of the leaps and and things that I've.

00:04:43 Speaker 2

Drawn out and then.

00:04:45 Speaker 2

It's a follow on for from some of the more master ones is to walk them.

00:04:49 Speaker 2

Through. OK, that's good.

00:04:52 Speaker 2

But it it it doesn't fit. Here's why. Right? So there's biases. You have to look at the source. Is it? Is it a an opinion piece and?

00:05:01 Speaker 2

We'll go through.

00:05:02 Speaker 2

Later, so kind of the things that I drew out from what we're doing right now is for COVID is we have our master narrative list and then we also have our persistent threats, so known bots, known sources of disinfo that's kind of the infrastructure that they use and that kind of ties it from from the.

00:05:22 Speaker 2

Fiber to the to the Miss Info is is the infrastructure they're using. The domains that they're pushing it out. So we had what? What is it? National news then? Corona 2 plus. Then you have the Canaries, the the accounts, the hashtags, the the groups.

00:05:43 Speaker 2

I like to call them factions that are distributing it or carrying the the the disinformation right, because that gives you the intent. If you watch those over time, you have the baseline and then all of a sudden something.

00:05:58 Speaker 2

Changes or merges with two separate factions.

00:06:04 Speaker 2

Then that gives you an outlier, right? So why are these two factions that never push the same the same narratives? Why they're now all of a sudden pushing the same narrative? What is the event that caused that? So that's the outliers bring additional analysis.

00:06:20 Speaker 2

So it helps you.

00:06:23 Speaker 2

Whatever you want to call it, separate the wheat.

00:06:25 Speaker 2

From the chaff.

00:06:26 Speaker 2

Noise from noise from the actual threats and then established. So once you have, once you have that the known right so you know no narratives. You know who's pushing them.

00:06:40 Speaker 2

You have established that.

00:06:42 Speaker 2

Baseline. Then you need to the regular threat streams. Is kind of your collection, so you have to establish a known collection that if something gets through and you miss it, you can go back and you.

00:06:57 Speaker 2

Can say OK it.

00:06:59 Speaker 2

Came through. Here's the artifacts that we've seen. It didn't.

00:07:02 Speaker 2

Had our our net, it did it, it got through our net. So how do we close that gap and then you adjust your feeds, your collection or kind of your sources that you go to. So right now a lot of this is manual for me.

00:07:19 Speaker 2

At least, so I use things like. I find a new a new source, right? That is always giving me good tips and 90% of the time they have a subscription. So subscribe to a daily feed and subscribe to a weekly feed so I don't have those come in. So I don't have to go searching for.

00:07:42 Speaker 2

It makes it a lot better when you can have them coming to you. I do it the old fashioned way because not good with development, but it the the it's there so you identify the platforms, then you evaluate the platforms that you're using.

00:08:01 Speaker 2

Upfront. So establish the biases that they may have established that OK, this is social media. This is going to be a.

00:08:08 Speaker 2

Lot of opinion, this is Main Street media. There's probably an agenda. This is state media, etcetera. So once you establish those document those. So again then you can come back and you can improve those.

00:08:22 Speaker 2

If something gets through and units see it, so this is kind of a a moving, evolving piece so that you can.

00:08:31 Speaker 2

And you don't want to filter down too far, because if you get it too filtered, a lot of stuff's going to come through from from additional sources that you may not. So cast your net wide and then kind of be able to filter from the wide net, be able to evaluate.

00:08:52 Speaker 2

And then pick out those outliers.

00:08:55 Speaker 2

So that goes into kind of the persistent repeatable monitoring. So I do it on a daily basis, same time if you could set parameters in, in the data sets that you are that you're that you're searching and can save those, that's very good. So if you can go from.

00:09:15 Speaker 2

Hey, every day I'm going to look from.

00:09:18 Speaker 2

07 to 007 the day before 2:07 today, based on this keyword search or this this parameter, if you can do that, then the next day you know you're not missing anything kind of in your call. So you're evaluating the same data set.

00:09:39 Speaker 2

Consistently. And then you're growing to that so that you can you can have a plan.

00:09:46 Speaker 2

So some of the things that I use because they're mobile and a lot of my stuff is mobile is Google News, very good just because of the fact that they're, they're ULA. It gets you through the ability to see a lot of this stuff, even if even if it's subscription. So like.

00:10:05 Speaker 2

Some of the things I've had an issue with is like Bloomberg or Financial Times or some of the other stuff that that has subscription. They'll always give you a few, a few free before you have to verify your account.

00:10:20 Speaker 2

Or other news aggregation sites. So I've used things like Flipboard, tweet, deck.

00:10:26 Speaker 2

Where you can set kind of parameters that you can go back to and just.

00:10:31 Speaker 2

Kind of continue to filter them and grow them.

00:10:34 Speaker 2

And and and and again. This has got to be repeatable so that you can train new people coming on if you can't point back to I I hate sops and work instructions. So when you're training, when you're training a.

00:10:48 Speaker 2

Enterprise or or new analysts, and you're bringing them on. If you don't have something to give them baseline, it's hard to get them up and running.

00:11:03 Speaker 2

So that's so Google. I don't know if anybody knows like the the Google dorks or the Google hacks.

00:11:11 Speaker 2

So kind of the the Boolean parameters that you can set. So you can say I just want PDS. If you're looking at at finished reporting. I just want CSV's or or if you're looking for feeds and then you can set. I just want stufffrom.govor.eduor.com or so you can kind of.

00:11:30 Speaker 2

Go through that.

00:11:32 Speaker 2

And then you could set the time parameter as well.

00:11:36 Speaker 1

Maybe this is a?

00:11:37 Speaker 1

Good point to let [Name redacted] jump in for a second because she had a few different ideas.

00:11:44 Speaker 3

I I was just putting notes in as as we were going. So in terms of tips on threat streams, we also have tips coming in.

00:11:51 Speaker 3

Yeah. And that's something's been fed in a lot and that you're talking about information streams, but we've also got the other groups that are monitoring disinformation streams, things like the.

00:12:02 Speaker 3

University of Indiana dashboards and photometer.

00:12:06 Speaker 3

So we can look at those.

00:12:08 Speaker 2

And I think we have a lot of.

00:12:10 Speaker 2

That stuff in.

00:12:10 Speaker 2

It we have a lot of.

00:12:11 Speaker 2

That stuff in the.

00:12:12 Speaker 2

Data science. So I was looking through the data science training and a lot of those are are in there. So I think as we as we start.

00:12:19 Speaker 2

To document our.

00:12:21 Speaker 2

Data streams. They can go OK, open news partnerships and and tips.

00:12:28 Speaker 2

Social media, etc. And then we.

00:12:32 Speaker 2

Could also go about.

00:12:34 Speaker 2

Outside of those actions to take or.

00:12:35 Speaker 3

Whatever. Yeah. And I gave you an early view of the data science notes. I'm I'm writing. So yeah, we'll get.

00:12:42 Speaker 3

That done. Awesome. No, that's great.

00:12:46 Speaker 1

And this I can't remember was.

00:12:47 Speaker 1

This your last slide.

00:12:49 Speaker 1

Before jumping into.

00:12:49 Speaker 1

The example or did you have one more?

00:12:53 Speaker 2

Two more.

00:12:54 Speaker 1

OK, awesome. Keep doing great.

00:12:54 Speaker 2

Going. Yeah. So when you're when you're.

00:12:57 Speaker 4

Asking a question about.

00:13:01 Speaker 4

Sorry, can I ask a question about crowd, crowd tangle. There's two ways of using it. One is the extension and also there are the dashboards that those with the access to the.

00:13:18 Speaker 4

To the full.

00:13:21 Speaker 4

Database do which I found. I found the the dashboard that someone on the group made was very effective.

00:13:36 Speaker 4

How how would you use the the?

00:13:40 Speaker 4

How would you use crowd angle?

00:13:42 Speaker 4

To like to to do your monitoring.

00:13:47 Speaker 2

So I would go from. I would go to probably [Name redacted]. [Name redacted] is the one that, that that has the one for us. I am very, very new to the crowd tangle.

00:14:00 Speaker 2

And I have seen it used so one time we used.

00:14:03 Speaker 2

It for like.

00:14:04 Speaker 2

There was so, so much disinformation kind of feeding everybody down. We looked at just like COVID good news stories, right, so that there, there was the ability to set up dashboards. So from this we could do it regionally.

00:14:20 Speaker 2

So I I.

00:14:21 Speaker 2

Believe you could put in put in sources.

00:14:24 Speaker 2

I'm not sure if you could put in hashtags or you could put in like persona.

00:14:32 Speaker 2

And users and and stuff in there. And then you just build that out and you keep, you keep growing that and then you can have the different, the different fees, the different dashboards that you can subscribe to. And however you want to get that in. If we get the subscription or just do API polls, I'm not sure how the back end of that works.

00:14:50 Speaker 3

You you can have a PR pull off the back of it. I I've done that. The problem is is.

00:14:56 Speaker 3

That's it.

00:14:57 Speaker 3

Yeah. So.

00:14:58 Speaker 3

If you want to get a feel for what dashboards do what, you tend to get is. So tweet deck is a free version not of this one, but it's a free dashboard that you basically put a bunch of queries in.

00:15:10 Speaker 3

And then it streams from those queries for.

00:15:13 Speaker 2

You. So I have. I have I.

00:15:15 Speaker 2

Just started messing around with tweet that I was kind of old.

00:15:18 Speaker 2

People just still using Twitter, but yeah, so I have that pulled up and that will go through that and kind of work through that.

00:15:25 Speaker 2

In the example.

00:15:26 Speaker 3

In the big thing we get out of the.

00:15:32 Speaker 3

The other dashboard, whose name I forget, is that we get access to the Facebook streams.

00:15:38 Speaker 1

Yeah, the ground title, yeah.

00:15:39 Speaker 3

Yeah, crab.

00:15:40 Speaker 3

Tangle Facebook is hard, which I still owe you. The finishing that Facebook groups graper.

00:15:46 Speaker 1

Awesome. I'm going to have you finish up for app with the last two so that we can jump into the examples because I think a couple of these will kind of come out in as you show people.

00:15:55 Speaker 2

So when you're looking for outliers and new narratives, the the big thing to look for is emerging or re emerging narratives. So narratives that have kind of fallen dead. So we had to reopen as we reopen for a second time. If we go into in into a second wave or however you want to.

00:16:15 Speaker 2

On it, those are likely going to reemerge. Then you have kind of some of the the COVID naive G Those will probably pick up overtime or kind of the we see the the emerging narratives, right, so.

00:16:30 Speaker 2

Russia and China, they all of a sudden started their their groups and their allies started pushing or their state media started pushing kind of some of the.

00:16:40 Speaker 2

Same narratives.

00:16:43 Speaker 2

That is something that you you definitely want to dig in deeper and then kind of local and world events that we know are getting a lot of.

00:16:53 Speaker 2

Media attention, especially Main Street media, because then we'll start to see it trickle down in social media and then start getting.

00:17:02 Speaker 2

Twisted pushed and new narratives kind of emerge out of that. So we've seen the protests. We've seen kind of George Floyd's George Floyd's campaign or the campaign around the riots and and things like that. You had the reopen, you had gridlock. And then.

00:17:22 Speaker 2

Some anonymous, anonymous or significant size online activity. So for this one I usually watch for like I'll check sporadically for trending hashtags.

00:17:34 Speaker 2

Throughout the day.

00:17:36 Speaker 2

And then kind of so the way that Twitter does it is if it's an actual hashtag, it'll show you kind of what the other hashtags to the what the other hashtags that are getting associated with that. And I think there's a lot we can do on the back end with that throwing those in and then analyzing that.

00:17:54 Speaker 2

And then kind of.

00:17:55 Speaker 2

Depicting that graphically, so we can kind of pick out the the groups or whatever.

00:18:00 Speaker 2

So when you're, when you're analyzing the outlier narratives, so you, you definitely first thing is you have to evaluate the source biases.

00:18:09 Speaker 2

So knowing the source and kind of doing some digging, if it's past, if it's. If it's Chinese, Chinese media, if it's Russian media. So RT right? So that that comes with an agenda, if it's so if it's an.

00:18:28 Speaker 2

Opinion piece out of the Washington Journal or an opinion piece out of what have you. They're now doing a better job of kind of bottom line of.

00:18:38 Speaker 2

Front and putting those in, but those are the things you need to you need to identify and kind of establish when you're making your assessments and your assumptions. So find additional sources. So as soon as you find a narrative, just use anything unique in identifying in there.

00:18:59 Speaker 2

String of characters plug it.

00:19:01 Speaker 2

And plug it into Google and see what other sources are using those same that same line, right? And then you can start to see, OK, this is all US that comes with its own bias, right? These are all the same or these are competing narratives.

00:19:22 Speaker 2

Why are they competing? Why are they the same as? The fact is it opinion? What is different? Why is it different and some of these you have to hypothesize about and make assumptions, but then later you check those assumptions before you make a final analysis.

00:19:37 Speaker 2

This and then try to establish the intent or the agenda behind it. Is it political? Is it influences it to cause disruption, distraction, confuse harm? And then how could this narrative be used for bad? And this is kind of when you put on?

00:19:56 Speaker 2

Your red Team hat, right?

00:19:58 Speaker 2

If I was back, here's how I could spin this if I wanted to. If I wanted to push this again.

00:20:03 Speaker 2

The and then what? What would the the what? What could be the impact of the narrative if leveraged for bad and I kind of left out who, who who may use that right. So we know some of the bad actors, we know their agendas, we know their goals, their strategic objectives.

00:20:23 Speaker 2

And kind of mapping those mapping those to that and the the hypothesis we came up in the last one.

00:20:33 Speaker 2

Now it's play.

00:20:33 Speaker 2

Time. So if there's any questions, let me stop sharing here.

00:20:43 Speaker 2

Let me check this all right.

00:20:47 Speaker 5

Not a question, but a comment. You know how we have the T-shirts that say [Name redacted] is my IDs.

00:20:53 Speaker 2

Yes, yes.

00:20:53 Speaker 5

I feel like we need a T-shirt that says [Name redacted] is my threat intelligence dashboard.

00:20:59 Speaker 2

Absolutely. We should definitely do that.

00:21:02 Speaker 2

I'm down with that.

00:21:04

Let me see.

00:21:06 Speaker 2

Let's do this. Hit full screen and pull up here.

00:21:19 Speaker 2

There you go.

00:21:19 Speaker 1

So graph you're going to run us through the example of Corona Palooza.

00:21:24 Speaker 1

Yeah. Awesome. So.

00:21:29 Speaker 2

Kind of where this came in. Right. So yesterday kind of a little background. So there was the, the Juneteenth and and Trump was kind of going to he was doing his first post.

00:21:44 Speaker 2

Post in quotes rally in Tulsa.

00:21:49 Speaker 2

On the 19th, the 19th was actually the.

00:21:52 Speaker 2

Last day or.

00:21:53 Speaker 2

Was the day that the information got.

00:21:56 Speaker 2

All the way to Texas.

00:21:58 Speaker 2

That the that slaves had been freed after the the Civil War. And I actually think it was something like two years after that that the some of the slaves in in Texas had realized or found out that they were actually that they were actually free.

00:22:16 Speaker 2

And then the the Tulsa aspect was kind of spun because of the Tulsa massacre in 1921, there was, I mean, they came through, they blew up the entire city. You had KKK members that were keeping the fire departments out.

00:22:37 Speaker 2

While the city was burning, people were being kind of just just massacred from from the air. What?

00:22:45 Speaker 2

So those two kind of grew A narrative which made Trump change his the the date of his rally to actually today. But that didn't really change a whole lot with respect to crowds and whatever. So then you had the COVID piece.

00:23:05 Speaker 2

Kind of rolled in there and the narrative narrative that emerged was.

00:23:13 Speaker 2

So this is tweet deck. Sorry, this is tweet deck so.

00:23:16 Speaker 2

This is basically Twitter.

00:23:17 Speaker 3

We can't see it.

00:23:19 Speaker 2

I'm not sharing. I thought I was sharing.

00:23:22 Speaker 2

All right, so screen one, sharing now.

00:23:28 Speaker 2

Was wondering why there was a one and a two on my screens. Apparently I just hadn't picked something.

00:23:33 Speaker 2

So this is this is tweet deck so you log in with your Twitter. Obviously this is one that I use just for this info J French 13. So this is kind of where you can you can look at what's trending currently and we'll just go.

00:23:52 Speaker 2

Corona Palooza started started trending. There's some. There's some awesome memes coming out of here, but when you're when you're kind of drilling down to this right you you'd be keeping like analysis notes. So you find something that that is COVID. Let's say when I when when we found COVID N.

00:24:14 Speaker 2

It was covered 5G. Then we started digging in seeing the associated accounts, right? So the associated hashtags, if there was any any media images, we were identifying those so.

00:24:30 Speaker 2

You can kind of look through here and you can just kind of scroll and see the different perspectives, right? So pro Trump, anti Trump kind of so kind of look through these and then ones that you see I do this a lot.

00:24:49 Speaker 2

Noble, right? So I'll just throw them into a one note. Just kind of copy copy link to this tweet, throw it in there so you can do deeper analysis on it.

00:25:00 Speaker 2

And kind of bring out trends in, in patterns to identify the narratives. Then you kind of look at. OK, here's the source. They're they're all pushing against, let's say, George Takei and.

00:25:15 Speaker 2

And and it's it's very, very interesting when you start digging into this and you can see COVID, it's so basically the narrative around Corona Palooza is that.

00:25:32 Speaker 2

Is that it's bringing everybody together when the.

00:25:39 Speaker 2

The numbers are just going up so that that everybody that's going here has to sign a a disclaimer against against the the fact that they could be exposed to Corona, it's inside, so it's being pushed by by people.

00:25:59 Speaker 2

And try to get people to not.

00:26:00 Speaker 2

Go to the not go.

00:26:01 Speaker 2

To the the rally.

00:26:04 Speaker 2

To build a narrative saying that hey, in two weeks this is the reason why the numbers are going up. So there's kind of a fall back, but we've seen the same thing with the reopen Memorial Day than you had kind of June 15th as everything kind of started walking through and.

00:26:23 Speaker 2

Going to for the most part, a lot of them going to phase two.

00:26:28 Speaker 2

So really just diving into this.

00:26:32 Speaker 2

And so this is from the social media aspect. I'm not. I'm not very, very apt on on the Facebook side, I'm kind of a A.

00:26:43 Speaker 2

Twitter born seller so now kind of jumping over into the the kind of the.

00:26:54 Speaker 2

Actual media, right? So I go to to Google News and you can build safe searches so you can set parameter.

00:27:04 Speaker 2

There's for time and then you use Boolean searches, so this is what I usually go to at least once a day just to kind of see all of the different. They do a pretty good job of of breaking up the narratives and if something catches my eye, come dig into it, see the sources.

00:27:24 Speaker 2

But this gives you a good data stream to kind of go through. That's consistent, right? So if I check this every day at 7:00, I know for one day I'm going to get everything corona at this information.

00:27:43 Speaker 2

So let's just give this information and and another good one that I've that I've found is on on Twitter a a lot of people.

00:27:53 Speaker 2

Have been calling out a lot of this information stuff and it's brought it's brought a a deeper look into some of this.

00:28:03 Speaker 2

So I think it's doing a good job of of bringing kind of.

00:28:07 Speaker 2

Awareness it is being used obviously both ways, but for now I think it's being used kind of for the good.

00:28:18 Speaker 2

So you can go save searches Corona disinformation, and then you'll go through. OK, so let me jump here and then obviously, so evaluating the the source, right. So it's Al Jazeera. You obviously know that that.

00:28:36 Speaker 2

They are known to.

00:28:39 Speaker 2

Certain political agendas for certain countries.

00:28:43 Speaker 2

There. So evaluate it with. For what it's worth and look OK COVID-19 and Russia, fake news forced concessions. So then what you would do is.

00:28:58 Speaker 2

What I would do is.

00:29:01 Speaker 2

Go back here.

00:29:08 Speaker 2

And go Russia and force.

00:29:18 Speaker 2

So then now you can start to see, OK and I would go, let's say that just happened today.

00:29:27 Speaker 2

So when go four days.

00:29:32 Speaker 2

So that way you can capture.

00:29:33 Speaker 2

Kind of both sides of.

00:29:34 Speaker 2

It so that's the only one. So you can kind of widen that net seven days.

00:29:44 Speaker 2

Let's go. Russia. Fake news. There you go. So, like you can start to see and put these together and then evaluate. OK, Financial Times. You got Washington Post and then start drawing.

00:30:02 Speaker 2

Drawing observables out of these right and additional resources.

00:30:09 Speaker 2

So you have the. This is the Mueller report.

00:30:15 Speaker 2

But so you kind of jump.

00:30:17 Speaker 2

Through and it gets to be.

00:30:21 Speaker 2

It gets to be a lot. So definitely when you're doing this, definitely make sure that you're you're kind of keeping track and it does a pretty good job of keeping your history.

00:30:31 Speaker 2

Go Google.

00:30:33 Speaker 2

Does anybody have any kind of emerging things that you see on the horizon that you would like me to try to run through kind of the process?

00:30:49 Speaker 5

Why don't we take a look at something related? You see, COVID was it?

00:30:53 Speaker 5

COVID palooza.

00:30:55 Speaker 5

Yeah, like COVID jella.

00:30:58 Speaker 2

Oh, that's a good one.

00:30:59 Speaker 2

Too. So you go in here and and you don't even have to put the hashtag because some.

00:31:05 Speaker 2

People don't CHELL.

00:31:10 Speaker 2

A or LA.

00:31:12 Speaker 5

I think it's double L.

00:31:14 Speaker 5

It's like Coachella, but with.

00:31:18 Speaker 3

Good grief.

00:31:21 Speaker 2

So yeah, so.

00:31:24 Speaker 2

Then you obviously you'll start to see the narratives coming out. You'll start to see, should we watch it so COVID, that's the new one. COVID stock, COVID palooza. What else do we have?

00:31:39 Speaker 1

Like it so it pulls out.

00:31:40

The people who like.

00:31:41 Speaker 1

Different types of music based on.

00:31:42 Speaker 1

The type of COVID rally they would go to.

00:31:44 Speaker 2

Yeah, absolutely. And then you have stuff from Trump genocide in there too, you know.

00:31:49 Speaker 2

Because that doesn't show your bias at all.

00:31:54 Speaker 2

And I I.

00:31:55 Speaker 2

Love when they do this and they.

00:31:58 Speaker 2

They throw a whole bunch of just.

00:32:02 Speaker 2

Hashtags in there.

00:32:04 Speaker 2

For for a couple of reasons. Because sometimes you'll see like you'll see a whole bunch of stuff that's super unrelated and you're like, OK, so you're just trying to jump on the.

00:32:16 Speaker 2

And to bolster whatever you're trying to so you can kind of count.

00:32:21 Speaker 2

That out or or look.

00:32:23 Speaker 2

At that right. So you.

00:32:24 Speaker 2

Can do some quick.

00:32:25 Speaker 2

Analysis on OK. Are these all? Are these all related? Yeah, probably. But if you start seeing if you start seeing Corona Palooza with, I don't know, the Incredibles 2.

00:32:38 Speaker 2

It's like, OK, so.

00:32:40 Speaker 2

It it gives you that quick.

00:32:42 Speaker 2

That quick triage right?

00:32:44 Speaker 2

But they can. You can start looking at these and kind of documenting these.

00:32:48 Speaker 2

Down in a note dig.

00:32:50 Speaker 2

Into them and then say, hey, related interesting facts. Here's some of the artifacts that I'm finding out here. Here's some of the narratives that I'm pulling out so.

00:33:03

Let's see.

00:33:06 Speaker 5

That's interesting the the the hashtag.

00:33:12 Speaker 2

Oh yeah, that's.

00:33:13 Speaker 5

That's something I noticed, and this is a long time ago, but back back in 2010, 2011 in Wisconsin, when the legislature was pushing through the Act 10 for busting public unions, people would be basically trying to hashtag or hijack the hashtags.

00:33:33 Speaker 5

Of the other.

00:33:34 Speaker 5

Quote UN quote.

00:33:36 Speaker 5

Yeah. So you see lots of.

00:33:37 Speaker 5

Tea party people trying to hijack union hashtags and vice versa, and.

00:33:41 Speaker 5

So that gives.

00:33:42 Speaker 5

You a really interesting perspective on, yeah.

00:33:44 Speaker 2

Into your intent, right? So it gives you it gives you insight into their intent whether they.

00:33:50 Speaker 2

Realize it or.

00:33:51 Speaker 2

Not I, I would say 90% of the time they don't. But if you look and and it's just.

00:33:57 Speaker 2

It's not even.

00:33:58 Speaker 2

Like the first thing I do is like I I'll put in the I'll put in the search parameters and then I'll just start.

00:34:04 Speaker 2

Scrolling through, I'm like, OK, well this video.

00:34:06 Speaker 2

Is showing up a lot, right?

00:34:08 Speaker 2

Or or this article showing up.

00:34:10 Speaker 2

Let me let me tag that put it into a bookmark and I'll go back into it later and kind of dig into it, but it's the it's kind of that, that wide swath of being able to look at everything and kind of an even stream.

00:34:26 Speaker 2

To be able to identify the outliers if everything's ohk. OK. Yeah. Corona Palooza, dead, corona, whatever.

00:34:35 Speaker 2

It it allows you, or if they're all the same, the same wording, the same phrasing, the same hashtags, and you're seeing it a lot, right? So that gives you a reason to think it's a little more coordinated, a little more inauthentic. But one of the things that I see.

00:34:53 Speaker 2

A lot too.

00:34:54 Speaker 2

On Twitter is.

00:34:59 Speaker 2

To try to draw their.

00:35:00 Speaker 2

Hashtag into the into the mainstream. They'll just start adding OK at CNN's at at.

00:35:09 Speaker 2

At Fox News, whatever. And they'll just start trying to. And sometimes it works because it'll it'll draw it into whoever's following that person, right? It'll draw.

00:35:22 Speaker 2

That it it'll.

00:35:23 Speaker 2

Draw them into that stream and then maybe they'll dig into it and then maybe they'll post something, right?

00:35:29 Speaker 2

So it's bringing people from both sides from from kind of grassroots.

00:35:37 Speaker 2

In the Main Street, I mean, you see it all the time with with at the Real Donald Trump, right?

00:35:45 Speaker 2

They'll just add him. He's not really going to respond, but it's usually a pretty good memetic warfare going on. It's kind of fun.

00:35:53 Speaker 2

To watch.

00:35:55 Speaker 1

So do you have?

00:35:58 Speaker 1

Any ideas of ways that you wanted to kind of get input from people about automating some of these things? I think that was one of your questions after kind of.

00:36:06 Speaker 1

The walk through was how the yeah.

00:36:07 Speaker 2

Yeah. So, so kind of.

00:36:10 Speaker 2

Some of the.

00:36:10 Speaker 2

Things that I that that would be.

00:36:12 Speaker 2

Nice for me. Right is if we could.

00:36:16 Speaker 2

Automate some of this collection right? So if we find something and trending trending hashtags.

00:36:24 Speaker 2

That we want to.

00:36:25 Speaker 2

Be able to pull.

00:36:29 Speaker 2

Kind of pull a feet of this.

00:36:30 Speaker 2

Right, for every. I don't know if maybe 3 hours.

00:36:33 Speaker 2

Every right. Whatever.

00:36:35 Speaker 2

Where it's where it's doable that we could just look at it and not actually have to go into the platform.

00:36:41 Speaker 2

Because if it's dropped into CSV, Excel, spreadsheet, database, whatever, you could start to manipulate it.

00:36:48 Speaker 2

And then you could start to.

00:36:49 Speaker 2

Visualize it. So I think a good way to go ahead.

00:36:55 Speaker 3

We have the pieces to do that. So you've seen the.

00:37:01 Speaker 3

The GFCI visualizations. We've got this code in the GitHub called Anti Patel scraper or something. It's under collection and that pulls off a hashtag.

00:37:15 Speaker 3

In fact, it pulls from a set of hashtags off Twitter.

00:37:20 Speaker 2

So how do we?

00:37:21 Speaker 2

How do we integrate that into?

00:37:24 Speaker 2

I don't think we want. We would necessarily want that integrated into like the the 4 Chan.

00:37:29 Speaker 2

But we may want that integrated into the triage channel into like a bot or something, right? So.

00:37:37 Speaker 3

That that's why I just checked in the feed, we've got. [Name redacted] has been doing scrapes straight off Twitter for us. So he's been pulling data and pushing it into. We have three Twitter.

00:37:51 Speaker 3

I think the channel which we have where he's been hunting by state.

00:37:54 Speaker 1

OK.

00:37:57 Speaker 3

Three Twitter disinformation.

00:38:00 Speaker 3

Yeah, five. I mean, string feed that he's.

00:38:02 Speaker 3

He's put together for us.

00:38:04 Speaker 5

And tossing stuff into five disinformation data also is another option.

00:38:08 Speaker 3

Yeah, Twitter streaming is is a.

00:38:13 Speaker 3

Good place to put streams.

00:38:17 Speaker 3

We at the moment it it's I think it's just me using the [Name redacted].

00:38:23 Speaker 3

So I just run the code, get some CSV's, do a quick analysis. So I look at the top hashtags, look at top Co hashtags and look at the.

00:38:31 Speaker 3

I throw it into [Name redacted] and look at the GFCI graphs.

00:38:35 Speaker 3

[Name redacted] you can you can run as a web app.

00:38:40 Speaker 3

Has an API we can push over to that the.

00:38:44

So So what the code has?

00:38:46 Speaker 2

To be run the.

00:38:47 Speaker 2

Code has to be run in command line, right?

00:38:50 Speaker 3

The [Name redacted] stuff? Yeah, I run in command line just because that's where it is. We can we can put an app around that we've got.

00:38:58 Speaker 3

A hackathon coming up.

00:39:00 Speaker 3

So ask for what you want and.

00:39:02 Speaker 1

Yeah, if you could show us how to do that, that would be amazing because.

00:39:05 Speaker 1

It sounds like that.

00:39:05 Speaker 1

Would be super.

00:39:06 Speaker 3

Helpful. Basically just say Python [Name redacted] gets scraper dot pie and then the hashtags you want or the phrases you want in quotes.

00:39:15 Speaker 3

It's written up in the big book, I think.

00:39:19 Speaker 3

If it isn't.

00:39:20 Speaker 2

Wait one I will pull.

00:39:22 Speaker 2

The I think I have the.

00:39:23

Big book up.

00:39:26 Speaker 3

If it isn't, I will write it in there.

00:39:29 Speaker 2

It is. I was actually just looking at it all right, now you're good. So here's the here's the [Name redacted]. So it's in six, three.

00:39:40 Speaker 2

My big thing is my home system is the majority of Windows, right? So so it's it's a little harder when we're trying to execute, so I don't know if we could maybe during the hackathon figure out figure.

00:39:56 Speaker 2

Out a way to bought that.

00:39:59 Speaker 3

Yeah, we can.

00:40:00 Speaker 2

Because that way that way.

00:40:02 Speaker 2

Yeah. So that way you could just, hey, run this scrape and then have the output of the the data file that it that it goes to or into the database and then yeah. So I think that would be good, right? So getting the.

00:40:21 Speaker 2

Getting kind of some of the some of the newer people or some of the people that don't necessarily have the the infrastructure to do this. If we can do it within the slack channel and then have it put in and then have some of the have the leads kind of go through that on a regular basis.

00:40:38 Speaker 2

And start pulling that out and and throwing it into triage, say hey, this is what came in today. There were three submissions, all regarding this. Let's look into that a little deeper and then we kind of we kind of farm that out and say hey is.

00:40:54 Speaker 2

Anybody willing to?

00:40:56 Speaker 2

Kind of dive into this.

00:40:57 Speaker 2

This data set.

00:41:00 Speaker 2

And getting getting getting more of that because I think I think.

00:41:04 Speaker 2

I think I think we're doing a good job of of of finding this stuff, but it is a a smaller set of people, so if we can get four kind of more involved, which is why I try to do the the the post that I did the other I think it was yesterday.

00:41:25 Speaker 2

UM.

00:41:28 Speaker 2

Yeah, this one. So kind of.

00:41:31 Speaker 2

With regards to 2nd wave, what to watch for?

00:41:36 Speaker 2

And and kind of so that way we can go back to that.

00:41:42 Speaker 2

But get get.

00:41:43 Speaker 2

Everybody more involved. We got a lot of people in here and I think just kind of managing managing resources and that kind of falls down on me getting to know some of the more the people that we're bringing to triage. So as we get them onboarded.

00:41:57 Speaker 2

Maybe we do a regular meeting, like a resources meeting or experience meeting with the leads of hey, we got these these people over the last week onboarded. Here's kind of what they've established, that they're that they're good at and what they want to do.

00:42:13 Speaker 3

You're showing your thingy window.

00:42:18 Speaker 2

Yes. Oh yeah.

00:42:21 Speaker 2

Yeah. Yeah. So, so getting them more.

00:42:26 Speaker 2

More involved and I think that's just a a management thing, right? So here's the people we've on boarded in in the last week and that may be a Wednesday meeting.

00:42:34 Speaker 2

That we can have.

00:42:36 Speaker 3

Yeah, have a a newbies meeting. Just get him set up. But we need to. The other thing is that Facebook post scrapers, there are plenty of those out. The thing I've fighting is group scraping so you can do an equivalent on Facebook too.

00:42:50 Speaker 2

OK.

00:42:50 Speaker 3

And we could start doing this for other sources like, you know, Instagram and Reddit and places with APIs.

00:42:57 Speaker 2

Well, I think that goes into our kind of a having a.

00:43:03 Speaker 2

I I hate. I hate I hate Excel spreadsheets, but when you're talking about kind of collection platform and and identifying what you currently have the capabilities to do, you can start to identify what we need to

build the capabilities on. So maybe that's something that we need to noodle through before the hackathon, which is what? The 24th.

00:43:25 Speaker 3

I mean starting with this equivalent in Facebook so you can pull this, check this, put it up on a dashboard.

00:43:33 Speaker 3

Just the things I keep telling saying like the you.

00:43:35 Speaker 3

Know the these.

00:43:36 Speaker 3

These accounts seem to be most prominent in this. Let's go look at them.

00:43:41 Speaker 2

And even the some of the some.

00:43:43 Speaker 2

Of the the known, the known infrastructure. Right. So we have domains. If we identify that there's an event coming up.

00:43:51 Speaker 2

That they will.

00:43:52 Speaker 2

Likely get active in. Maybe we set up a monitor to say, hey, if anything new is posted to any of these or pushed out by any of these.

00:44:01 Speaker 2

I don't know if you can.

00:44:01 Speaker 2

Do that in Twitter based on based on domain.

00:44:06 Speaker 2

But if you can.

00:44:07 Speaker 2

Say any anything new, drop it in here.

00:44:10 Speaker 3

So if you're talking about like domain as in a new site, throwing out all that most of us have included on.

00:44:15 Speaker 2

So yeah, well, so all domain hosting or.

00:44:21 Speaker 2

Domain hosting, right? So they're hosting?

00:44:23 Speaker 2

Some they're hosting some disinformation messaging that is getting retweeted and kind of pushed out in the narratives. So I don't know if.

00:44:32 Speaker 2

You can do that.

00:44:33 Speaker 2

Like if you could say, hey, if there is a URL that is LinkedIn, a post with this root level domain.

00:44:42 Speaker 2

Have us pull it down. I don't know the back.

00:44:44 Speaker 2

End of the API looks like.

00:44:46 Speaker 3

Twitter has T dot codes for its URLs, so I will that goes to the Python as well. Is there some way to unpack?

00:44:49 Speaker 2

OK.

00:44:54 Speaker 2

Yeah, because that way if we could just get root level domains and we have or I mean even second level and we can start saying, hey if if these are starting to be posted, start dumping them here for follow on analysis and then we could start seeing kind of that early warning.

00:45:10 Speaker 2

Hey today it had five. Here's here's.

00:45:14 Speaker 2

The the narratives that went with it, here's who's pushing it and do some follow on analysis with that and continue to collect.

00:45:21 Speaker 2

I don't know if.

00:45:24 Speaker 1

Open this up a little bit more and see if there's anybody else who had ideas for automating or ways to kind of continue building on this or question.

00:45:42 Speaker 3

Let's see.

00:45:46 Speaker 1

OK.

00:45:46 Speaker 5

So one of the things that I find helpful.

00:45:50 Speaker 5

When I'm building automation for my own use.

00:45:56 Speaker 5

Going through and sort of dispassionately looking at what exactly I'm doing and detailing it, so like, OK, open this browser. Ohh go to this URL, set the setting.

00:46:12 Speaker 5

Export this data into CSV and then move it into this folder right? So as much of that as you can detail.

00:46:21 Speaker 5

In in sort of Amelia Bedelia.

00:46:24 Speaker 5

Safe steps if you know what I mean.

00:46:27 Speaker 5

Is incredibly helpful, especially when you've got somebody coming in for the first time to help you write something and they've got 72 hours to do it as.

00:46:35 Speaker 5

Part of a hackathon.

00:46:39 Speaker 5

That's that's incredibly helpful, especially also but.

00:46:41 Speaker 2

30 seconds.

00:46:43 Speaker 5

Yeah, go ahead.

00:46:44 Speaker 2

Now you go.

00:46:46 Speaker 5

Especially because you can also take that, then take that detailed description of what you're doing.

00:46:53 Speaker 5

Drop it into a Python document and just basically put a hash or a pound sign in front of each line. So now you've got built in comments and so basically as you go you just write a function or write the the code you need.

00:47:06 Speaker 5

For each.

00:47:07 Speaker 5

You know.

00:47:08 Speaker 5

Just step of the way.

00:47:10 Speaker 2

And don't we use Anaconda?

00:47:14 Speaker 2

Because, like we could, I mean we.

00:47:15 Speaker 2

Could you could technically modulate that right? So input like a here and then you you have your inputs and then we could grow on that. We used to do it for deploying systems, so we'd have a white.

00:47:33 Speaker 2

Like system update and download these and then execute this. So I mean doing that in some so I don't know if that would be.

00:47:42 Speaker 2

A playbook.

00:47:43 Speaker 6

So and I.

00:47:46 Speaker 6

I'm going to interrupt you. So this is.

00:47:49 Speaker 2

OK.

00:47:53 Speaker 6

Something I think as [Name redacted] and [Name redacted] have already started talking about. So I'll pass this over to [Name redacted].

00:47:58 Speaker 6

To comment but.

00:48:00 Speaker 6

One of the things we could do is.

00:48:03 Speaker 6

Build out Jupiter notebooks with kind of playbooks or like patterns of of things you want to analyze, right? And then yeah, dump your data in those and and.

00:48:13 Speaker 6

The idea is.

00:48:14 Speaker 6

You shouldn't have to. You shouldn't have to sit down and think about what you want to do every single time. You should have a set of tools in which you just plug in some variables and see like what the results are, right? And so that's really like a notebook type thing. And we can build those. We can, we can make that happen, but we need to know, you know, what do you want? You know, what are you looking at?

00:48:22 Speaker 5

Right.

00:48:35 Speaker 6

And once we list those out.

00:48:36 Speaker 2

And the variables. What are the?

00:48:38 Speaker 2

Variables that you want to paint.

00:48:39 Speaker 2

Like you want to queue off.

00:48:40 Speaker 2

Of and what are the data sources that you want us to?

00:48:43 Speaker 2

Run these against.

00:48:44 Speaker 3

Well, the nice thing about you for the notebooks is you can change almost.

00:48:47 Speaker 3

Anything in them. So we work our way through the.

00:48:48 Speaker 5

Yeah, yeah.

00:48:51 Speaker 6

And if you want to.

00:48:51 Speaker 1

I was like.

00:48:51 Speaker 6

Do that from slack, you know then what we can do.

00:48:54 Speaker 6

Is like, you know, pipe off all the variables you want to look at, like hashtags or whatever to a Jupyter notebook. Maybe return a a URL or something, right? And then you.

00:49:03 Speaker 6

Can just like.

00:49:04 Speaker 6

Load the web app you know you don't have to.

00:49:05 Speaker 1

Sure. Like [Name redacted].

00:49:06 Speaker 6

Like have Linux.

00:49:08 Speaker 5

And the point that I was getting at is as we're trying to.

00:49:10 Speaker 5

Figure out how to automate this stuff.

00:49:12 Speaker 5

The more detailed in terms of figuring out what steps we need to take the the the easier it's going to be, especially with the hackathon coming up to to hand it off to somebody who does this sort of stuff on a daily basis.

00:49:25 Speaker 5

That's all I was getting at.

00:49:26 Speaker 2

And to identify where, where, where the commonalities are right. So we we all we all want this stuff to be run against Twitter. We all want the stuff to be run against Facebook. Let's run it against Reddit and and identify the different platforms we want to run against. And again the inputs, right. So this is where it gets down to.

00:49:48 Speaker 2

The capabilities what do we have the?

00:49:49 Speaker 2

Capabilities of now.

00:49:51 Speaker 2

To do what are the inputs?

00:49:54 Speaker 2

What are the outputs?

00:49:56 Speaker 2

And then what do we want to do? And then how do we best integrate those with the current environment and current personnel?

00:50:05

My name is [Name redacted].

00:50:07 Speaker 2

Hey, [Name redacted].

00:50:09 Speaker 5

Sorry, I was. I was.

00:50:12 Speaker 5

I just reimaged [Name redacted] and I forgot that I was unmuted.

00:50:16

That was so awesome.

00:50:20 Speaker 2

I mean you you evaluate current capabilities, you evaluate the current environment and the personnel and then you figure out how best to close the gap so that the people can use the capabilities the fastest. And then after that close the gap on or or add additional capabilities.

00:50:39 Speaker 2

But it's literally going to be decompiling what we want to do, what the inputs, what the expected outputs are, and then overlaying them. I think of everything as a map. So kind of a see through map on the old, the old projectors, right.

00:50:58 Speaker 2

So here's here's what we want for Twitter. Here's what we want, and it's kind of that that one to one data matching. Here's all the inputs. Here's the format that it's.

00:51:07 Speaker 2

Take and here's how you're going to be able to call it. Here's where you're going to be able to call it. Here's the output. Here's how you're going to be able to access the output. Here's how you going to be able to manipulate the data, and here's how you're going to have like an output. Here's the expected output for action, right?

00:51:27 Speaker 2

So like all of that, should go in before an incident gets kicked off.

00:51:33 Speaker 2

And a lot of that that pre I mean and we've talked about this multiple times, a lot of that collection can happen seamlessly with automation and then make sure that the that it's available to the majority of everybody and that there's.

00:51:52 Speaker 2

Consistent evaluation of that data. So make sure that somebody's somebody's going in there every day and looking at it and giving a rough summary and next steps or.

00:52:05 Speaker 2

Or gaps in the data, right? So the the after action reports. Hey, this was good. Let me give you some feedback on how this could be better and that's how you improve and and close those gaps.

00:52:19 Speaker 1

So Speaking of summaries and next steps, we have about 3 minutes before the happy hour starts and also the Corona Palooza starts, so bringing folks.

00:52:32 Speaker 1

Kind of to a close here. It sounds like one important next step is a conversation about Jupiter notebooks among the leads and maybe triage, but probably.

00:52:41 Speaker 1

Most of the leads.

00:52:42 Speaker 1

To talk about how we might want to use that as a way of automating some of the things we're looking at here and it sounds.

00:52:47 Speaker 1

Like, that's already kind of happening a little bit and then?

00:52:52 Speaker 1

I would say just diving in to.

00:52:56 Speaker 1

The next example or two of incidents that we see popping off and have people kind of start to use some of some more of these strategies. Were there any other things you wanted to to say before we wrap up or any questions?

00:53:08 Speaker 1

From other folks.

00:53:10 Speaker 2

No, I I really think that's good is is the big thing is taking what our current capabilities.

00:53:17 Speaker 2

There and I think we have that pretty much captured in.

00:53:20 Speaker 2

6th of the big.

00:53:21 Speaker 2

Look in Chapter 6 and then handling the artifacts like so. These can grow into building out these Jupiter notebooks 6:00 and 7:00.

00:53:37 Speaker 2

And then the analysis outputs of eight but six and seven are pretty much the the the crux of what we have now, maybe, maybe some of four, but I mean a lot of the data we've already captured. So we just need to go.

00:53:53 Speaker 2

Through it.

00:53:54 Speaker 2

And standardize it and figure out.

00:53:57 Speaker 2

OK, here's here's what we currently can do.

00:54:01 Speaker 2

And here's how we want to be able to call it and use it.

00:54:05 Speaker 2

And then just noodle through that and.

00:54:07 Speaker 2

Drill those out.

00:54:11 Speaker 2

Sorry, my word of the day is noodle.

00:54:14

It works.

00:54:18 Speaker 2

Well, yeah, that's that's.

00:54:24 Speaker 2

Yeah, that's. I mean, that's pretty much.

00:54:27 Speaker 2

The try to make the the.

00:54:32 Speaker 2

It as as.

00:54:34 Speaker 2

Useful for everyone as we can, and that's how we get more people involved. Make it simple. Make it. I mean, obviously, everyone's everyone's on the slack channel, so if we can, if we can integrate it with a bot and then just teach people how to throw stuff into the bot and then we have triage.

00:54:56 Speaker 2

To look at that data on a consistent basis.

00:54:59 Speaker 2

And then we have, I don't know if we have an additional analysis piece where people are kind of broken out into regions or focus areas or or whatever expertise to do those deeper dives and and draw those, those analytical, those analytical products out.

00:55:20 Speaker 2

For follow on.

00:55:25 Speaker 1

I will.