**John Torpey** 00:03

Hi, my name is John Torpey, and I'm director of the Ralph Bunche Institute for International Studies at the Graduate Center of the City University of New York. Welcome to International Horizons, a podcast of the Ralph Bunche Institute that brings scholarly expertise to bear on timely international issues. Today's topic is Russian disinformation campaigns through social media and their interference in US elections. We're fortunate to have with us today John W. Kelly, PhD, who is the founder and CEO of Graphika, which is a data analysis firm. The firm was founded on technology that he invented blending social network analysis, content analysis, and statistics to make complex online networks understandable. John has provided expert testimony on foreign interference in the US presidential election before the Senate Select Committee on Intelligence in the news today. He is also an affiliate at the Berkman Klein Center for Internet and Society at Harvard, where he works with leading academics to design and implement empirical studies of the internet's role in business, culture, and politics around the world. A quantitative social scientist by training, John Kelly earned his PhD in communications from Columbia and has also studied at Stanford, and at Oxford's Internet Institute. Thanks so much for joining us today, John Kelly.

**John Kelly** 01:40

My pleasure, John, thanks for having me.

**John Torpey** 01:42

Great to have you here. So I understand you were part of the team that first detected Russian disinformation campaigns on social media and its interference in the 2016 election. Can you tell us about that and explain how it came about?

**John Kelly** 01:59

Sure. Well, I guess there are a lot of ways to tell this story. I've got a lot of points in the timeline to drop in. Let me start with the period around the election itself, kind of leading up to it and through it, and then what happened after. So back in 2015, we started working with a research unit at Google called Jigsaw on algorithms that can detect basically what was real versus fake online. So when you have some cascade of information or adoption of a hashtag, or some kind of mass behavior online, could you detect whether it was happening organically? Or could you detect if it was being manipulated or sort of boosted in some way by coordinated activity. And so this was a scientific research project. And in executing this, we had to train the algorithm. So we had a number of examples of known disinformation operations or information operations, kind of manipulation of social media, things that were to train for the positive case. And then we had a number of known organic, things going viral, kind of just spontaneous cultural activities, on the other hand, and then we were busy working to construct an algorithm to tell the difference between these two things. And that was really just amazing kind of science research project that we were working on. And some of those cases of manipulation were Russian information operations - were known to be such, although it wasn't necessarily clear which Russian threat actor was responsible, because there are several different entities in Russia that do this kind of stuff. But nonetheless, here, we knew that some of them were Russian. And this was just sort of data we have been working with. And we are so we're well aware that this kind of manipulation was in their toolkit, and that some of it was being done around the election. And this is kind of in the lead up to the election. And remember, at this time nobody understood that this sort of thing was happening, it was only a kind of small group of researchers and academics that were tuned into this kind of activity. And then the election happened. And I still remember being invited about two or three weeks after the election to a think tank in DC to sit around a table and hear people mash their teeth over the outcome, and I remember piping into the conversation at one moment and saying "well, you guys realize there is a massive Russian disinformation or propaganda campaign kind of throughout that election" people looked at me like I had three heads. I mean, these were professional journalists and kind of folks in the DC communication space. And then fast forward a little bit and everyone understood that that had been happening, and so that shift is sort interesting. And so I wouldn't say it's that so much we were the first ones detecting. I'd say we were among a small community of researchers that were clued into this and looking for it and studying this kind of activity before the election, and then in the ramp up to kind of the publicizing of it, we wound up with a interesting role. And maybe I can tell the story of how that evolves.

**John Torpey** 05:26

Sure, please do.

**John Kelly** 05:27

Yep. So I think we first wound up in the public eye for this around the work with the Senate Select Committee on Intelligence or what I'll just call SSCI. And so in the aftermath of the election, SSCI was interviewing American experts and some foreign experts, folks like Phil Howard at Oxford, on foreign influence and information operations. And so they're had a, just a lot of folks coming through that they were interviewing and kind of asking questions and trying to build their own - this is the SSCI staffers - understanding how it worked. And so I was just on their list and started talking to them. And then because I had this experience, from our work training these algorithms, had some hard experience with international data, they just started inviting me back. And that began a conversation with SSCI, that ultimately led to the research projects that wound up being made public as a as a report. And so, you know, they had been given all the data on the Internet Research Agency's activity: the Russian troll farm and St. Petersburg Internet Research Agency, part of the [Yevgeny] Prigozhin network. He's also known as Putin's Chef, Prigozhin is a Russian oligarch who sits on top of all this sort of dark arts, online manipulation capabilities. And so SSCI had all the data that Google Facebook and Twitter had attributed to the Russians, and it turned over to them. And they wanted to do something with it. But they don't have open source analysts on staff. And it turned out that they were interested to put together a research project. And they allowed me to be involved in that. And we helped design a project to turn over all that data to two different teams, so that then the teams weren't supposed to know who each other were, or have any contact with each other so that the reports and the findings that came out would be verifiable against one another. And so we wound up on one team with Oxford Internet Institute, Phil Howard's shop, their computational propaganda program there, plus Graphika, was one of the teams. And then the other team was a company called New Knowledge along with a scholar from Columbia's Tow Center, and a couple other folks. And so they messengered an encrypted hard drive up to New York, with all that data from the platforms on it. And we spent about seven months looking at that, and just a fantastically interesting and rewarding experience to dive through all that stuff and reconstruct what the Russians had been up to during that time, or at least what the Internet Research Agency had been up to at that time. Because one of the really, key findings for us at the end of that project was that of the known Russian information operations we knew about and had studied in our research work, the IRA was only responsible for a few of them, which meant there was another Russian actor out there and other organizations that had done the other ones.

**John Torpey** 08:48

Fascinating story. Maybe you could tell us a little bit more about how your software works. And how you track down bots and that sort of thing. And what more broadly has been done since 2016, to try to mitigate the consequences of Russian and perhaps other disinformation. We'll get into that next.

**John Kelly** 09:17

So our software, which is essentially my dissertation kind of turned into a software stack. I did a lot of work in network sociology at Columbia at my PhD program; I was fascinated by network analytic techniques and their overlap with communications between political communications. And that led to creating a machine learning based network analytic stack that would understand online conversations not as a kind of river of content, which is the way that most systems understand it. So take the Twitter firehose and put a listening platform on top to understand how things are trending in that firehose of content. But rather, I was interested in understanding online communications as a network with relationships between authors and kind of understanding the pattern of information flow in that complex network. So I invented a technique for mapping out and then using machine learning to cluster or segment, very complex online discussion networks into different communities and then understand things about each of those communities, who and what made them tick, what kinds of information resonated where. And that was what my dissertation was on web blogs. And then, over the years, started developing that technology, which eventually became the basis of Graphika's software platform. Now, what we're then doing is - allow me a digression, if you will, to discuss the way we look at the world because it's, I think, critical to understanding the technology, what it means, which is that humanity evolved in a geo terrain, you know, the planet Earth, the 2D planes wrapped around sphere, and all human collective activity happens in that space and for thousands of years, as we developed as a species, and then as a civilization, so civilizations, all human collective activity requiring communications happened in that space, and the speed of those human communications was the speed of a horse, or someone running down, down across the ground, or the speed of a boat was limited by distance. And then you get to this magical moment with the invention of the telegraph, where you kind of punch a wormhole through geo terrain, and communicate at planetary scale in real time that radically rewired all the world's major organizations, how empires were run, how armies fought, how markets worked, all that stuff, was radically transformed by the telegraph, even though nobody but Abraham Lincoln would have had these things in their house. And then fast forward to today, where we're all connected at planetary scale in real time, via devices that we carry around in our pockets, and social networks, and email and all these different things. And so now human communication, which is what forms the lifeblood of human collective action, and civilization, and nation states, and everything we do, is primarily not mediated in geo terrain anymore, it's mediated in what we call a cyber-social-terrain. And that is constituted by the fabric of connections that arise from all the little choices people make about who to follow, and what to like. And all these things create relationships and condition algorithms that determine, you know, how communities form and what and who is visible to who and how that information travels around. So we think about that as cyber social terrain. And we think of ourselves as cartographers of cyber social trends, we're using large scale, graph analysis, network analytics machine learning to map out these complex online topologies, and understand what flows with them. So that's what the technology does. And then how it becomes useful in understanding information operations online is that because we've been mapping out organic, real communities and how they form, we've got a great deal of understanding about that, and we know what the patterns look like. And likewise, we know what the patterns of things that are fake look like. So, you know, there's a community that is, you know, 10,000 accounts being run by some small number of entities that has certain telltale markers in terms of how it's constructed as a network in terms of how it behaves semantically, all that sort of stuff. So that's the work we were doing in our research project back when the Senate found us - was trying to understand how we could algorithmically instrument, the measurement of that.

**John Torpey** 13:50

Got it. So do you think that defenses that you're involved in creating and monitoring are those doing what we need them to do, or is this a never ending battle that we will always be playing cat and mouse?

**John Kelly** 14:14

Yeah, well, I've got a lot of thoughts on that. I think on the one hand, it's a constant game of whack-a-mole, you're discovering assets that are controlled by some threat actor that are doing some kind of information operation or another and you want to discover those as soon as you can, before they've been able to build up an audience and have whatever impact they're there to have, and take them down. So we work very closely with those social media platforms. So let me back up. If you think of what our technology does, its sort of like radar, but right radar is only part of the defense. You need a lot of other components to it. And what we specialize in is, is discovery and profiling of online manipulation networks across platforms. So if you're if Facebook or Google or Twitter, or a Pinterest or a Reddit, you know, you've got responsibility to look at your own platform and try and discover and mitigate any coordinated activity, inappropriate coordinated activity, that could be happening on your platform. But your responsibility kind of stops at your own borders, and any sophisticated contemporary information operation is cross platform and multi platform. I mean, there's a Russian operation called Secondary Infection that we discovered where we detailed this past year, which manipulated hundreds of websites and platforms, and most of them will manipulate at least three or four or five. So you need someone who's taking a look at what's transiting not just within a platform, but across them. That's what we specialize in. But you have to have the cooperation of all those entities. And I think the good news since 2016 - there's are a number of bullet points in the good news - one is people know what's happening now. And you know, they're not naive about it anymore. The second is that the platforms - those that can afford it - have invested a great deal in discovery and mitigation threat intelligence capabilities. And so they've got really talented teams that this is their job 24/7. And then you've got a kind of research community that's growing and becoming more sophisticated, including academics, folks like the Alex Stamos's Research Center at Stanford, as well as companies, folks like us, and we all communicate with each other and work in a kind of coalition manner. So that you've got a lot of these different pieces, which are operating together to try and tease these things out, and whack the moles faster than they were whacked before.

**John Torpey** 17:15

Right. But there's do seem to be more moles in the field now. I mean, we've just learned, it seems that Iran is a major player in this kind of activity. Maybe you could speak to that. Is that really news to you? Are there other players? And I definitely want to ask you about China's involvement in these kinds of activities, since that seems never really to come up for discussion.

**John Kelly** 17:40

Yeah. So I think the good news I just listed is somewhat offset by the bad news, the bad news is that even though Russia gets all the press, Iran has been active in online manipulation on American platforms for about as long as Russia has, and has been far more active than most people realize. And the news today isn't really news to us. I mean, the specifics of it are that they're responsible for that particular operation. But the Iranian capability is significant, and it's been in operation for decades. The Chinese are more recent to the game of international Western platform-based information operations, and their techniques are somewhat different. And their approaches and their objectives are somewhat different, but they're evolving quite rapidly. But beyond just the kind of the headline nation states, there's a vast increase in other types of malicious actors out there manipulating the Internet, and that's smaller nation states, like the Saudis are doing this. And there's a great report out from Oxford Internet Institute on what they call cyber troops. This is something that dozens and dozens of countries invest in, people that are members of their military or intelligence services or proxies, that their job is to manipulate online conversations and to hack stuff and do other kinds of operations. But this is becoming a facility that a lot of states have, but also a lot of non state actors. So you have anything from dark arts PR firms for hire to individuals that do this manipulation. So the manipulation of online conversations and platforms is proliferated. And the techniques for doing that have gotten somewhat more sophisticated, even as the techniques for discovery have gotten more sophisticated. And this is rippled out into a number of kind of interesting changes in the landscape.

**John Torpey** 19:51

Interesting, I wonder to some extent what exactly are these other countries - insofar as we're talking about countries - what exactly are they trying to achieve? I mean, it seemed clear in 2016 that Russia was intervening on behalf of the election of Donald Trump, it was basically trying to support Trump's effort to get elected. But it also seems that a certain amount of the activity really is oriented simply to reducing people's trust in the legitimacy of the system. And one might argue that that's the main problem that exists in political life in the United States at least that trust in government has been in decline at the very least since Watergate and to some degree since Reagan declared government to be part of the problem, not the solution. So I wonder if you could talk a little bit about the aims and objectives of these countries when they're engaged in this activity?

**John Kelly** 21:01

Yeah, that's a great question because they're not really the same. I think the Russians, in my view, are the most dangerous to us, because their objective is to take our society apart at the seams, to sort of drive wedges into all the crevices that already exist and try and make those crevices unbridgeable. Their objective is really to pull our society apart, so it just falls apart. That's really different than the Iranians who, even though they do a little bit of electoral manipulation as we've just seen, it's primarily aimed at swaying an electoral outcome or kind of influencing a particular election cycle, when it is targeted at us. The vast majority of their activity is targeted in their own sphere of influence in the Middle East, so most of the Iranian kind of online propaganda covert and overt apparatus is aimed at their part of the world, and they're kind of squaring up against the Saudis, who also have capability. And they're really focused more on their own neighborhood. The Chinese are focused on promoting narratives that make them look good, and support critical foreign policy objectives, like their claims to sovereignty in South China Sea, their claims around Taiwan, so they have more of kind of a punch list of specific foreign policy objectives and trying to support the narratives that lead to acceptance of those. And then it's just a matter of kind of how far abroad are they pushing that information. And lately, they've been kind of pushing that narrative much more broadly, including into the American space, things like their stories around their interpretation of the events in Hong Kong, for instance. But they're really focused on things that are of direct interest to the Chinese state. They're not trying to do things that make one category of Americans hate another category of Americans more, which is what the Russians really doing. And then I think as you cascade into other state based actors, they're really focused on - they're kind of training those capabilities on - kind of whatever their own foreign policy objectives are, usually very close to their own countries, or they're training them on internal targets. So there's a report we just put out just in the last couple of days on the military in Myanmar, and we've reported on that before. Also what the Filipino government does in the Philippines. A lot of these other countries, that capability is mainly focused on solidifying their government's internal power and sort of attacking their internal enemies.

**John Torpey** 24:08

So, I mean, at some level, I'm inclined to say that this is just a kind of extension of warfare. And thinking of Clausewitz's his argument that war is just an extension of politics by other means. And it sounds like this sort of activity, if we want to call it cyber warfare, and it's sometimes directed at internal populations rather than other countries, it does seem to have a lot of that kind of aspect to it, that it's really about gaining advantage over these other actors with which one might have engaged, let's say diplomatically in the past, but now one can do it through these other means. Is that reasonable way to think about this?

**John Kelly** 24:57

Yeah, I think that's exactly right. And I think people don't think broadly enough about how they conceive of the problem. And the frame often ends up being stuck on "Oh, the Russians - or insert the Iranians - are trying to affect the 2020 election." And it's really a much, much bigger problem that. So first off, one of the things we discovered in our work for the Senate around the 2016 election is that across every platform that the Russians were manipulating that we had the data on - Instagram, YouTube, Twitter, Facebook - the the level of Russian activity went up after the election in 2016, by typically 50 to 60% per platform, but up by almost 240% on Instagram. So the Russian stepped on the gas after the election, it wasn't just about the election. But if you step back and think about an even bigger frame, we have to ask ourselves, what's at stake for political bodies for, let's just say, nation states. And as I mentioned earlier humanity doesn't just live in geo terrain anymore. We live in the cyber social terrain. Nations are to adopt Benedict Anderson's view on imagined communities, and in the old geo terrain, they've got physical borders and so it's easy to kind of know where they begin and end, because you've got a border, and usually an army or some defensive capability around it. But in cyber social terrain, you don't have that. And yet, all the narrative underpinning - the what are the stories about who we are and what we stand for, and how we got here, and where we're trying to go and who's the day and how do we relate to them - all these things that are the lifeblood of a body politic, now have open borders that we don't know how to defend, and we don't even know how to think about. And so I think we have to understand that, you know, to the extent that humanity is woven together in the 21st century in a cyber social terrain that is primary over geo terrain in terms of human collective activity, and that is, thus far, we haven't figured out how to defend that against manipulators who are injecting narratives that make it much harder for us to sort of feel that we're of common cause and purpose and common identity with one another in the future. It's a very hard problem.

**John Torpey** 27:36

Interesting. I mean, reminds me a little bit about Michael Mann's claims - Michale Mann the sociologist, not the filmmaker - who has argued against the whole term "society", because he thinks basically, the world consists of networks of power of different kinds, and that, despite being a sociologist, he'd like to get rid of the term "societies". But the point you're making is that we really are, as you said earlier in the discussion, we very much are connected to each other in real time globally. I had a Zoom conversation this morning with colleagues in Germany and Italy. No problem. I mean, I didn't do it from my phone, but I could have, and, you know, some kind of remarkable transformation in that sense.

**John Kelly** 28:22

Yeah. But the sort of phenomenon of jumping from your locality, up to global level and how that gets mediated and formed online: a good example of this is in the height of the George Floyd protests, there was this DC blackout hashtag in hopes that spread like wildfire. This idea that the government or somebody cut the power in Washington, in the middle of these protests, and then it didn't happen, but it was trending for hours. And really made all the different sensors and needles that monitor online conversations tilt over into the red and cause kind of mass panic. And this was something that wound up printing globally, not because it was true, or because people in Washington were clamping onto it and saying it, it got boosted by a network of Anonymous accounts - you know, literally the organization Anonymous - followed by a kind of big boost from KPop fans in Asia that are extremely powerful online, picked this thing up. And you wound up with something driving the news cycle, a hoax or mistake about a blackout in Washington, DC, which happened because of international actors in the cyber social sphere. It's just a example.

**John Torpey** 29:50

Just a reminder that it's not all good. So, we're coming up now to an election. We're practically there. It's only a few days away really at this point. How would you say this kind of interference has affected the election process? And have we been able to kind of keep it, relatively speaking, under control? You listed some positives, but you've also listed some negatives. But could you speak specifically to the election?

**John Kelly** 30:30

Sure. So it's not over yet. And we're being vigilant and interesting things are happening. I think that there are a lot of differences between the, you know, the tactics and approaches of different threat actors. And so, the IRA really is most well known of the Russian state actors. And that model where you're trying to cultivate online assets that weave seamlessly into political communities, organic, real political communities, and then affect their behavior in order to, say diminish support for a more mainstream candidate - kind of attacking Hillary from the left, you see the same kind of thing attacking Biden from the left - these kinds of techniques that involve IRA-like sets of assets have been less effective this time. Both the things that they've been trying to do to hide, now that the platforms and us and others are out looking for them, have made them less effective, but generally, people are more clued in to the possibility of that kind of activity and a little more wary of it. The thing that remains dangerous, though, was more the approach of the GRU [Russian Military Intelligence], which was responsible in the 2016 cycle for hacking or acquiring the DNC [Democratic National Committee] internal data, and for releasing it via Wikileaks and other methods as well around like the Podesta emails, the DNC leaks cases and things like that. So that hack and release method is very, very dangerous. And that's more what Russian military intelligence does. We haven't seen that yet. I mean, the sort of, you know, Hunter Biden laptop story is a little bit similar to it, but we're, I think, still unpacking what's really going on with that. So I think that with respect to the election itself, so far, I'm encouraged and feel like we've done a good job. And the platforms have done a good job of kind of stepping up here. But it's a sophisticated problem. And a lot of the manipulation now is domestic and it always has been, but I'd say the biggest problem we face with trying to mitigate foreign influence campaigns online is that we tolerate them coming from Americans. And the range of domestic disinformation and call it covert online influence campaigns, that sort of stuff is really the fact that we tolerate it domestically makes it much, much harder to find out when other states are doing it. So the jury's still out. I'm encouraged, but it's not over yet.

**John Torpey** 33:20

I see. Well, that's an encouraging note, perhaps on which to end. You know, since one of the big problems with our election seems to be precisely that people are concerned that it's being manipulated by various actors, and is perhaps not as legitimate as we would like to think it is. But thank you, John Kelly. Sorry, did you want to say something?

**John Kelly** 33:45

One parting thought on that? Yeah, I think what you said is just its really important that one of the greatest dangers to this kind of activity is to delegitimize the communities, particularly vulnerable communities, that have been targeted by it. So you know, if you have the Russians' Internet Research Agency had, for instance, heavily targeted Black Americans and Black political activists. But they've also targeted, the far left and the targeted other groups as well. If it is a danger that the fact that a community has been targeted by this stuff somehow supports a narrative that they are compromised or that their views are less legitimate. And I think that's really something to guard against, is not allowing this activity to delegitimize those who are actually victims.

**John Torpey** 34:38

Great, thanks for that concluding thought, very helpful. I want to just say thanks for to John Kelly of Graphika for sharing his insights into the Russian and other disinformation campaigns that are affecting us - have been affecting us at least since 2016. That's it for today's episode of International Horizons. I also want to thank Merrill Sovner of the EU Studies Center at the Ralph Bunche Institute for helping produce this episode and Hristo Voynov for helping on the technological side. This is John Torpey of the Ralph Bunche Institute for International Studies, saying we look forward to having you with us next time on International Horizons.