Panel 1: Improving Trust in the Information Ecosystem
Moderators: Nadya Bliss (ASU) and Duncan Watts (UPenn)
Speakers: Claudia Deane (Pew Research), Kate Starbird (University of Washington), David Lazer (Northeastern University), Chris Wiggins (Columbia University)

138
01:06:34.350 --> 01:06:43.800
Duncan Watts: Thank you. So thanks, Beth. And thank you, Margaret and Skip, for that introduction, just like to add my thanks to the NSF for your

139
01:06:44.340 --> 01:07:04.890
Duncan Watts: enthusiasm and support throughout this process. Even as we've had to change our clients. We have a really amazing line up of speakers today in our three sessions. I'm happy to introduce the the first panel. My name is Duncan Watts. I'm a

140
01:07:06.390 --> 01:07:09.570
Duncan Watts: At the university professor at the University of Pennsylvania.

141
01:07:10.830 --> 01:07:22.920
Duncan Watts: I have appointments in the Department of Computer Science, The Annenberg School of Communication, and the Wharton School of Business. So I'm sort of very

142
01:07:24.570 --> 01:07:37.830
Duncan Watts: Much in the intersection of CISE and SBE and really thrilled that this initiative is coming together. In this panel, which is focused on the first of our thematic

143
01:07:38.370 --> 01:07:46.350
Duncan Watts: problem areas, improving trust in the information ecosystem. We have a diverse group of speakers.

144
01:07:47.040 --> 01:07:56.580
Duncan Watts: Claudia Deane from Pew Research Center, Kate Starbird Starbird from the University of Washington, David Lazer from Northeastern University, and Chris Wiggins from Columbia University.

145
01:07:57.120 --> 01:08:13.590
Duncan Watts: I am joined also by by my colleague, Nadya Bliss, who is co-moderating this session with me. She is the executive director of the Global Security Initiative at Arizona State University where she also is a professor in the School of

146
01:08:14.760 --> 01:08:32.490
Duncan Watts: Computing informatics and decision systems engineering. So the, the way the session will run each of the speakers will will present their own slides, I believe we have backup copies in case that doesn't work, and

147
01:08:33.660 --> 01:08:41.730
Duncan Watts: Each speaker will speak for 12 minutes and I believe that the twelve minute will be a warning time so that you have a signal to wrap up.

148
01:08:42.780 --> 01:09:01.680
Duncan Watts: Will have hopefully a few minutes for Q and A that will be focused on the presentation itself. And then, with any luck, we will get through all four presentations in one hour, and we'll have half an hour for general discussion as Beth mentioned just now.

149
01:09:02.700 --> 01:09:27.120
Duncan Watts: Nadya and I will be monitoring the the chat window where you can feel free to post questions and we will do our best to to select representative themes that we can then get the panel discussing. So with that, it's my pleasure to hand things over to our first speaker Claudia Deane.

150
01:09:29.940 --> 01:09:30.960
Duncan Watts: Claudia. Are you there?

151
01:09:32.130 --> 01:09:34.740
Claudia Deane: I'm here. Can you hear me.

152
01:09:36.660 --> 01:09:37.290
Duncan Watts: I can

153
01:09:40.950 --> 01:09:41.430
Duncan Watts: Okay.

154
01:09:42.240 --> 01:09:47.280
Claudia Deane: Yes, if you stop sharing. I will start sharing, everyone. It's nice to meet you. My name is Claudia Deane

155
01:09:47.880 --> 01:10:05.820
Claudia Deane: I am a regular of WebEx and not zoom. And so I do not have the tranquil background that Duncan has, which I'm very jealous of so you will be meeting today likely my colleagues, the 10th grader and seventh

grader who are studying down here and the nine year old who is now going to determine origin.

156
01:10:06.900 --> 01:10:09.930
Claudia Deane: So something to look forward! share

157
01:10:31.980 --> 01:10:34.320
Claudia Deane: Those of you not familiar, I noticed that

158
01:10:35.340 --> 01:10:43.470
Claudia Deane: Most of the folks on the panel are from universities, the academy. So I wanted to just tell you a little bit introduce

159
01:10:43.950 --> 01:10:53.670
Claudia Deane: My organization to give you a little bit of context for sort of the pluses and minuses that I'll lay out our entree into this world of cross disciplinary challenges.

160
01:10:54.420 --> 01:11:01.410
Claudia Deane: The Pew Research Center is a DC based institution. We are funded philanthropically. We're an independent subsidiary of the Pew Charitable Trust.

161
01:11:02.010 --> 01:11:10.680
Claudia Deane: And we have about 180 people, a lot of us are PhD social scientists primarily from disciplines like political science, sociology,

162
01:11:11.250 --> 01:11:16.080
Claudia Deane: demographers we prior to recently haven't had a lot of people from computational backgrounds.

163
01:11:16.530 --> 01:11:24.300
Claudia Deane: And we're trying to do rigorous social science and we are doing it without any sort of partisan or advocacy edge to it, which is

164
01:11:24.690 --> 01:11:37.740
Claudia Deane: both the big plus of our reputation and sometimes the limitation to the kinds of partnerships we can have as you hear about later. And we do make our data methods reports available to the public. We have traditionally been

165
01:11:39.150 --> 01:11:52.320
Claudia Deane: A more of a traditional social science shop we made our reputation doing public opinion research in general, we have a very

strong survey methodology shop. And so I think what you'll be hearing about is what it's like to sort of pivot that kind of shift.

166
01:11:55.440 --> 01:12:08.400
Claudia Deane: The one thing that I just want to flag that's a little bit different about us is that our work, our research is not channel necessarily through the academic journals where much of the academy puts their work.

167
01:12:09.480 --> 01:12:27.780
Claudia Deane: We are going instead somewhat direct to consumer we are trying to reach out to policymakers to the news media to the informed public. We have a lot of journalists on staff and so we are writing and sort of communicating in a bit of a different style than some of our colleagues.

168
01:12:29.790 --> 01:12:33.930
Duncan Watts: Claudia, sorry to interrupt you, can you go into presenter mode so we can see your slides a little bit better?

169
01:12:34.560 --> 01:12:35.280
Claudia Deane: Oh, yeah.

170
01:12:41.340 --> 01:12:46.470
Claudia Deane: I'm trying to figure out how to do that. I don't know why it's showing it this way.

171
01:12:48.090 --> 01:12:51.120
Claudia Deane: Sorry, the first slideshow. I knew this would happen.

172
01:12:52.470 --> 01:12:55.980
Claudia Deane: Does anybody know what I should hit on here? Give me a tip.

173
01:12:56.220 --> 01:12:57.930
Alondra Nelson: Click from beginning in the top left.

174
01:13:00.660 --> 01:13:01.470
Claudia Deane: Top left to

175
01:13:02.640 --> 01:13:06.630
Claudia Deane: Show taskbar display settings and slideshows and try again.

176
01:13:10.350 --> 01:13:18.720
Katie Siek - Indiana University - Informatics: I think you have to flip your screen. So we see your speaker view and you see your views. So I think you have to

177
01:13:19.470 --> 01:13:19.830
Change.

178
01:13:21.720 --> 01:13:22.890
Beth Mynatt: That's closer

179
01:13:24.330 --> 01:13:27.900
Claudia Deane: Alexis, are you trying to remote control it?

180
01:13:29.070 --> 01:13:31.050
Alexis Rodriguez: Yeah, I'm trying to get it set up for you. Give me one second.

181
01:13:31.410 --> 01:13:31.620
Thanks.

182
01:13:35.220 --> 01:13:37.080
Claudia Deane: I'll just keep chatting.

183
01:13:38.910 --> 01:13:46.290
Claudia Deane: I think our shop is here for two reasons today. One is that we have for the past 20 years been focused on understanding

184
01:13:46.650 --> 01:13:58.380
Claudia Deane: information and misinformation. There we go! Yes information and misinformation using the traditional tools of social science, including survey research. And we've done that across teams. If I can advance

185
01:13:59.400 --> 01:13:59.820
Claudia Deane: Or

186
01:14:01.470 --> 01:14:12.870
Claudia Deane: Here these are the research programs we have. And I've sort of highlight some of the ones in which we've done this work. But the second reason is that about five years ago, we decided to make that great leap.

187
01:14:13.860 --> 01:14:26.070
Claudia Deane: Into the waters of computational social science tempted by all the organic trace data that we saw being produced, and we created the data labs team of computational social scientists

188
01:14:26.400 --> 01:14:35.760
Claudia Deane: That are trying to sort of bridge the gap between the multiple methodologies that we're using, and over the five years we've grown to about a 10 person team. And I'll talk a little bit more about that.

189
01:14:37.680 --> 01:14:43.530
Claudia Deane: On that. So the research front on this information. Information we received some external funding.

190
01:14:43.740 --> 01:14:52.530
Claudia Deane: For a project we're calling American news pathways. We've done about where we will, over the course of the year. Do a six way panel survey based on a probability based panel.

191
01:14:53.280 --> 01:14:58.260
Claudia Deane: Trying to understand where people are getting their information about the election. Obviously, when the pandemic happened.

192
01:14:58.530 --> 01:15:11.850
Claudia Deane: We pivoted that to look also at people's reactions to the pandemic, the recession and, of course, whenever it will come back and as part of that we're using again what we know are sort of the

193
01:15:13.350 --> 01:15:23.730
Claudia Deane: somewhat problematic still traditional methods of asking people about misinformation and this would be just a sample finding of what we have been looking at trying to understand what parts of this story.

194
01:15:24.000 --> 01:15:40.470
Claudia Deane: Are resonating with different parts of the public. And as you can see here the storylines of the pandemic, some of which have more or less scientific evidence behind them are fairly widespread. This particular question you might notice that the bleach as a treatment for

195
01:15:41.640 --> 01:15:47.580
Claudia Deane: COVID-19 is fairly small. It was filled in before the President's remarks on a topic.

196
01:15:50.670 --> 01:15:51.630
Claudia Deane: We have

197
01:15:52.740 --> 01:15:57.900
Claudia Deane: also recognize that as our media environment has become so much more complex.

198
01:15:58.530 --> 01:16:03.360
Claudia Deane: It is really not possible for one human to remember all their interactions with information.

199
01:16:03.720 --> 01:16:15.390
Claudia Deane: And so our data labs team has been starting to join and partner with our other teams trying to track this down by looking at sources of information that are able to be analyzed via things like text analysis. So,

200
01:16:16.200 --> 01:16:22.440
Claudia Deane: on this page, you can see some of the projects that we have done in this space. We are trying to we have looked online.

201
01:16:23.340 --> 01:16:35.700
Claudia Deane: What people are hearing from the pulpit. We have looked at social media platforms by trying to connect them to survey to get a little bit more population base estimate of what's happening on there.

202
01:16:36.060 --> 01:16:42.090
Claudia Deane: And we are looking at the candidate social media platforms at the information people are getting from their members of Congress, which allows us to

203
01:16:42.480 --> 01:16:51.300
Claudia Deane: Sort of encompass the whole population. As per the pandemic topics that would lead to findings like this one that was released a little bit earlier this year

204
01:16:51.990 --> 01:17:01.380
Claudia Deane: Which could we just showed you the sort of time basis of when our representatives were beginning to communicate to the public about COVID-19 by

205
01:17:02.610 --> 01:17:03.000
Claudia Deane: Party.

206
01:17:06.870 --> 01:17:09.780
Claudia Deane: All right, now we get to the meat of it. And then this is my last slide.

207
01:17:11.160 --> 01:17:22.170
Claudia Deane: So those are what we are trying to do and some of the things that we've been able to do I think that part of what we're talking about today is the challenges that will run as a through line through our various perspectives on this.

208
01:17:22.680 --> 01:17:36.930
Claudia Deane: And this is a list of sort of the six of the challenges that we have faced. Our first question, of course, with people to source talent or re-train we, you know, survey methodologists are not known as sort of quantitatively.

209
01:17:38.070 --> 01:17:46.980
Claudia Deane: Challenged so they're a quite numerous group. But what we found when we went to start this kind of work that the idea of retraining our workforce was

210
01:17:47.490 --> 01:17:56.880
Claudia Deane: Not plausible for us in order to turn in a short time and we did hire folks from the field because it was five years ago, which in the field is

211
01:17:57.240 --> 01:18:01.110
Claudia Deane: A long time, at least from the social science view. I'm sure the computer, people are laughing at this but

212
01:18:01.470 --> 01:18:07.980
Claudia Deane: The humanities programs who had it at that point ramped up where we were seeing social scientists with computational training and so I

213
01:18:08.280 --> 01:18:16.470
Claudia Deane: Anecdotally would just note that a lot of the people we hired were people who are very adept social scientists who had just been interested in computers and computing and had

214
01:18:17.250 --> 01:18:26.700
Claudia Deane: Picked up some of these skills on the side on the side. We are still trying to work through some of the retraining challenges.

215

01:18:27.450 --> 01:18:39.870
Claudia Deane: They're not easy. The number of tools that you could potentially train on is large and they change much more quickly than some of the research techniques that we've been using in our more traditional work.

216
01:18:40.560 --> 01:18:47.460
Claudia Deane: The second thing was infrastructure. We obviously weren't set up to handle the kind of data or computational needs that we had

217
01:18:49.740 --> 01:18:50.460
Claudia Deane: The

218
01:18:52.380 --> 01:19:04.260
Claudia Deane: Jobs that we needed to hire for we did not have in our career architecture, we did not have we did not know how to describe them, I know this sounds silly, but there's a very practical challenge, right, you have to hire a DevOps engineer, nobody on your staff knows what that is

219
01:19:05.640 --> 01:19:12.840
Claudia Deane: So that was the, you know we we obviously now have filled out our career architecture a bit more, but that was a challenge for us at first.

220
01:19:13.200 --> 01:19:23.730
Claudia Deane: The other challenge I didn't really foresee is the extent to which the programs that these folks would be using are open source programs and that a traditional IT department that has to be at the same university.

221
01:19:24.450 --> 01:19:30.840
Claudia Deane: Likes to have a lot more control about what programs are being helmed on people's machines. And that was an unexpected practical challenge.

222
01:19:32.400 --> 01:19:41.820
Claudia Deane: Number three, cross disciplinary teamwork. I probably don't need to tell you all that there are a lot of challenges here and they're not really just about norms and vocabulary for us.

223
01:19:42.630 --> 01:19:48.600
Claudia Deane: The pace of work was extremely different the amount of front end work that needs to happen on the data science program.

224

01:19:49.530 --> 01:19:55.860
Claudia Deane: Was not something we necessarily anticipated. They're also very different risk profiles.

225
01:19:56.610 --> 01:20:03.090
Claudia Deane: When you go out to design and field a survey unless something very serious happens during your field work, you're going to get data back

226
01:20:03.330 --> 01:20:15.630
Claudia Deane: When you go out and start a data science project, you don't know if you will get findings back, and trying to readjust an organization to be prepared to fail, even if you fail in a very interesting way, was not an easy challenge.

227
01:20:16.800 --> 01:20:19.530
Claudia Deane: Number for sourcing data. I think we all know this one.

228
01:20:19.950 --> 01:20:26.610
Claudia Deane: It seemed at first we were a swim and a possible ocean of data. And while we tried to capture it, or water just ran through our fingers about as quickly

229
01:20:26.910 --> 01:20:36.720
Claudia Deane: It's difficult to get good data, data that will work for research purposes it's difficult to get folks to who have the data in private hands to share that data with one

230
01:20:38.520 --> 01:20:48.930
Claudia Deane: Which sort of has related to five External partnerships. To the extent that data is in private hands, we've had lots of interesting discussions with

231
01:20:49.290 --> 01:20:55.110
Claudia Deane: Social media platforms and other places. And those discussions often would break down around the time that we

232
01:20:55.350 --> 01:21:03.090
Claudia Deane: Let them know that we would need to publish the results of the findings, whether or not those findings were necessarily flattering to whatever of the interest

233
01:21:03.780 --> 01:21:18.300
Claudia Deane: The corporate interest was, I don't mean to make this sound nefarious. I think that's perfectly within their rights, but it was

very, it has been frustrating for us. You can get a good, bad going with the research team at some of these places. And then, of course, when you get to legal and communications.

234
01:21:19.980 --> 01:21:36.060
Claudia Deane: Things get a little bit more challenging. Finally, I think the thing always on my mind is ethics. I often tell my teams, What we can do outpaces what we should do the norms on this are sort of behind

235
01:21:37.380 --> 01:21:50.910
Claudia Deane: Of course, where they are for more traditional research. Everyone is trying to keep up. We would like to be part of that effort to set those norms and we're just looking for ways to do that by being as transparent as possible.

236
01:21:52.230 --> 01:21:52.770
Claudia Deane: And

237
01:21:54.180 --> 01:22:02.490
Claudia Deane: So I end on my time which hopefully makes up for my slide sharing problems and I look forward to the session questions.

238
01:22:07.470 --> 01:22:14.460
Duncan Watts: Thanks Claudia. Yeah, I don't see any questions in the in the chat.

239
01:22:15.720 --> 01:22:18.210
Duncan Watts: At the moment, so why don't we

240
01:22:19.260 --> 01:22:21.330
Duncan Watts: Take advantage of your, of your

241
01:22:22.650 --> 01:22:30.060
Duncan Watts: Time with us and move straight along catch up a little bit and move right along to our second speaker Kate Starbird. Kate, are you there?

242
01:22:31.050 --> 01:22:32.070
Kate Starbird: I am, yes.

243
01:22:33.990 --> 01:22:34.680
Kate Starbird: Can you see me?

244

01:22:35.160 --> 01:22:39.720
Duncan Watts: I can see you. Claudia, can you turn off your presenter, and **Kate Starbird**.

245
01:22:53.940 --> 01:22:55.050
Claudia Deane: I believe I am off

246
01:22:57.600 --> 01:22:57.900
Duncan Watts: Right.

247
01:23:03.990 --> 01:23:07.050
Kate Starbird: Do you see a white slide with black text on it?

248
01:23:08.430 --> 01:23:11.250
Kate Starbird: It says building infrastructure to support research and development.

249
01:23:11.970 --> 01:23:12.420
Duncan Watts: Got it.

250
01:23:12.840 --> 01:23:18.870
Kate Starbird: Excellent. Thank you. All right, I'm Kate Starbird from the University of Washington and an Associate Professor in the Department of human centered

251
01:23:19.200 --> 01:23:24.900
Kate Starbird: design and engineering. I'm also a co founder for the new Center for an informed public at the University of Washington and

252
01:23:25.200 --> 01:23:32.610
Kate Starbird: Can talk a little bit about our work in the areas of misinformation, disinformation and talk about sort of our mission and infrastructure

253
01:23:33.120 --> 01:23:55.320
Kate Starbird: At the center a little bit as I end. And so a little research, a little infrastructure. And I want to start just by noting that we are in some strange times. And if you've been online in the last few weeks, you will have seen lots of interesting stories, misinformation rumors, conspiracy theories,

254
01:23:56.550 --> 01:24:01.200
Kate Starbird: And things like that. And I think a lot of us are experiencing this we're seeing this in our social media feeds.

255
01:24:01.620 --> 01:24:09.300
Kate Starbird: And it can seem a little bit overwhelming. I actually the picture over there of Bill Gates is a killer, I took that a mile from my house.

256
01:24:09.960 --> 01:24:18.300
Kate Starbird: Which is strange. I live in Seattle. And that's a very strange thing to see. So we can also see that these online, things are also manifesting online as well and

257
01:24:19.230 --> 01:24:27.720
Kate Starbird: The World Health Organization has talked about this from the very beginning about sort of the infodemic that would be around the pandemic of how we would be struggling with information.

258
01:24:28.440 --> 01:24:34.710
Kate Starbird: That may or may not be true. And it's going to be kind of intersecting with the impacts of the event. And certainly that's something we're experiencing.

259
01:24:35.700 --> 01:24:40.950
Kate Starbird: Well, we actually know that this isn't just about the internet. If we look back to the

260
01:24:41.370 --> 01:24:51.690
Kate Starbird: Research and psychology rumoring, disaster sociology, rumors are very common during crisis events. They are actually sort of a natural byproduct of the collective sense making process where we come together to make sense of

261
01:24:52.020 --> 01:25:00.090
Kate Starbird: Complex and incomplete information to try to resolve some of the anxiety and uncertainty that we experienced as a part of the

262
01:25:00.690 --> 01:25:09.570
Kate Starbird: Event and and also on top of this, it's important to kind of remember, and you probably see this with your friends as well, is that behavior during crisis events tends to be pro social

263
01:25:09.930 --> 01:25:13.650
Kate Starbird: And in that people are actually trying to do the right thing. And they're getting it wrong.

264

01:25:14.130 --> 01:25:24.720
Kate Starbird: So, people tend to share rumors and past things are wrong for altruistic reasons they're trying to inform their friends and neighbors and but we also see exploit exploitation happening as folks can converge.

265
01:25:25.500 --> 01:25:32.700
Kate Starbird: For, you know, to try to take advantage of it in different ways. With COVID-19 we're seeing lots of misinformation,

266
01:25:33.150 --> 01:25:46.440
Kate Starbird: Possibly because of this sort of persistent uncertainty. The scientific uncertainty is helping to stir the rumor mill. We also have these sort of global global impacts of the event and global participation and these kind of conditions seem to be producing

267
01:25:48.120 --> 01:25:51.810
Kate Starbird: This kind of collective vulnerability to spreading misinformation at this time.

268
01:25:52.290 --> 01:26:02.940
Kate Starbird: Early on, most of what we saw was primarily unintentional people just trying to figure out what was going on. But increasingly, we're seeing more conspiracy theories more politicized misinformation and disinformation.

269
01:26:03.600 --> 01:26:10.260
Kate Starbird: And I know most of you know this already, but I want to point out this very important distinction between misinformation, which is information that's false.

270
01:26:10.980 --> 01:26:16.350
Kate Starbird: But not necessarily intentionally so. In disinformation, which is information that's intentionally false

271
01:26:17.190 --> 01:26:24.990
Kate Starbird: It. Well, it's misleading or false and is intentionally spread for some purpose of financial, political or other objective.

272
01:26:25.800 --> 01:26:37.140
Kate Starbird: I think this is the super, this distinction is going to be even more interesting in this event because we're seeing covid-19 begin to intersect with election 2020. And I think that's just going to keep growing over time.

273

01:26:37.830 --> 01:26:43.620
Kate Starbird: As we go and I think also, there's this danger, there's this hope that I've had that this will be a coming together moment where people

274
01:26:43.890 --> 01:26:48.630
Kate Starbird: Will be able to, you know, kind of, we have a common thing that we're fighting, we can come together to struggle against it.

275
01:26:49.020 --> 01:26:59.460
Kate Starbird: But it also seems that this event is it's providing an opportunity for people. It's for some people to sort of stumbled down the rabbit hole. And so we've got these two different tensions and and and playing out in different ways.

276
01:27:01.020 --> 01:27:04.350
Kate Starbird: So stepping back. I want to talk about as these things converge, covid-19 and election 2020,

277
01:27:06.390 --> 01:27:17.370
Kate Starbird: What kind of the dangers are the threat that misinformation and disinformation pose to sort of democratic societies. One of those is a diminished trust and information and information systems.

278
01:27:18.930 --> 01:27:28.230
Kate Starbird: We see this perception that misinformation is everywhere, and we can't trust what we see. And so we begin to lose this trustworthy information systems we begin to lose, lose trust in our systems.

279
01:27:28.830 --> 01:27:40.470
Kate Starbird: We also have this situation where people that are going online to sort of do their own research can be drawn into these ecosystems of conspiracy theories, misinformation, purposeful disinformation.

280
01:27:40.800 --> 01:27:49.710
Kate Starbird: And that's something we've seen in the past is a graph of a past event that these same the same sites are active and working and connecting and sharing content related to the

281
01:27:50.580 --> 01:28:00.780
Kate Starbird: The pandemic as well. We're also seeing direct, direct attacks on media flow from disinformation campaigns through elected leaders and then out to the broader public.

282
01:28:01.470 --> 01:28:11.250
Kate Starbird: So professional journalism is being derided as fake news. And again, just contributing to distrust in the information system. We're also seeing increased division. Now some of that's our own

283
01:28:12.660 --> 01:28:21.210
Kate Starbird: You know, it's our own stuff. And some of that is is folks that want to exploit that for political gain. Both, both domestically and foreign land in foreign actors as well.

284
01:28:21.450 --> 01:28:27.810
Kate Starbird: This is a graph from 2016 about a politicized conversation that was happening in the US. It's a retweet network graph you

285
01:28:28.410 --> 01:28:37.590
Kate Starbird: See these over and over again people and you know kind of two sides not really talking to the other, but kind of shouting at each other and in this case the orange pieces are actually accounts from the

286
01:28:38.400 --> 01:28:46.350
Kate Starbird: rushes internet research agency who join that conversation became part of that conversation. And, you know, we kind of see them trying to aggregate those decisions.

287
01:28:46.890 --> 01:28:58.110
Kate Starbird: Divisions excuse me, but the problem is is that as we have that increased division and we lose common ground. We can't come together as a democratic society to sort of make those decisions that we need to govern ourselves.

288
01:28:58.770 --> 01:29:06.090
Kate Starbird: And we have this diminished trust in democratic institutions, government organizations, the public health officials that we need to trust in order to

289
01:29:06.810 --> 01:29:12.960
Kate Starbird: You know, take the right actions, the medical professionals we need to trust in order to have people take the vaccine when it becomes available.

290
01:29:13.380 --> 01:29:28.050
Kate Starbird: Things like election results. All of these things that make a healthy society, both physically and sort of just democratically are kind of in threat or vulnerable to sort of the effects of pervasive missing this information.

291
01:29:29.220 --> 01:29:33.420
Kate Starbird: So we're at the Center for the Center for an informed public are trying to

292
01:29:34.380 --> 01:29:48.270
Kate Starbird: Trying to address the problem of misinformation, disinformation strategic manipulation of online spaces and we set up the Center with funding from the Knight Foundation that we that came through about a year ago, this time center went up in

293
01:29:49.860 --> 01:29:59.100
Kate Starbird: December officially and we just started hiring and by the time we started hiring, the first days for some of our people, we were in COVID-19

294
01:30:00.000 --> 01:30:10.200
Kate Starbird: Shelter in place. And that's become really a focus of our research, but I want to talk about sort of like what we're trying to do and what our mission is. So we have about five years of funding.

295
01:30:10.830 --> 01:30:17.130
Kate Starbird: For half funding, we gotta, we gotta go get the other half from the Knight Foundation, which has been great. And this is a collaborative effort.

296
01:30:17.460 --> 01:30:25.650
Kate Starbird: Across the information schools at the University of Washington, the law school, human centered design and engineering. And then we aim to add affiliates. We already are from diverse fields.

297
01:30:25.860 --> 01:30:30.840
Kate Starbird: So this is, we think of these as really interdisciplinary problems that need interdisciplinary solutions.

298
01:30:31.500 --> 01:30:41.820
Kate Starbird: And we have four kind of different pillars of goals. We have a research pillar, we want to keep understanding these problems in and help to kind of address them in different ways. We have an education.

299
01:30:43.410 --> 01:30:45.780
Kate Starbird: Pillar where we're really trying to develop curriculum.

300
01:30:47.460 --> 01:30:56.190

Kate Starbird: In a university curriculum, but also kind of K through 12. And I actually think K through 99. This is a, this is a problem that crosses generations. We all need to understand it better.

301
01:30:56.550 --> 01:31:07.080
Kate Starbird: An outreach pillar and a tech policy pillar and we have this expertise combining from our PIs from from a couple different places. So, Emma Spiro and I have

302
01:31:08.310 --> 01:31:10.800
Kate Starbird: Experience with misinformation during crisis events.

303
01:31:12.330 --> 01:31:15.690
Kate Starbird: And gentlemen West comes from misinformation and science.

304
01:31:17.040 --> 01:31:29.040
Kate Starbird: I've been doing a lot of research on disinformation and we also have folks looking at information literacy and tech policy and we're kind of trying to put these all together to come up with solutions and to think about

305
01:31:31.320 --> 01:31:40.260
Kate Starbird: Different parts of this phenomenon. So over emerging research questions or the structure and dynamics of online misinformation sort of the same stuff we've been doing

306
01:31:41.010 --> 01:31:45.570
Kate Starbird: Trying to understand this intersection of conspiracy theorizing with the online spread of disinformation.

307
01:31:46.320 --> 01:31:56.040
Kate Starbird: We're also looking increasingly at science communication science credentialing and the spread of misinformation, especially in the context of covid-19, we've seen those intersections be really important

308
01:31:56.790 --> 01:32:01.680
Kate Starbird: And we're exploring the role of sort of intention attention dynamics, but we're thinking about that. It's like

309
01:32:01.980 --> 01:32:13.200
Kate Starbird: The importance is of the systems, the social networks, the influencers the status and how that shapes the spread of misinformation in this event and others. This is one example I want to

310

01:32:15.060 --> 01:32:25.560
Kate Starbird: Highlight something we've been studying. It's actually a an article that is was done by sort of an armchair epidemiologist, who turned out to be a political operative that was on medium, it ends up going

311
01:32:25.980 --> 01:32:31.290
Kate Starbird: It ends up going viral, and about the time that Trump that President Trump

312
01:32:31.800 --> 01:32:37.500
Kate Starbird: decides that he's going to open up by Easter informed by this and some other things that are going on and we went through and we looked

313
01:32:37.800 --> 01:32:45.600
Kate Starbird: At how the article went viral. And we were initially surprised that wasn't just like social media all stars or tech pros, but it actually was

314
01:32:45.870 --> 01:32:53.790
Kate Starbird: propagated by a bunch of cable news personalities from Fox News early on and their big audiences help this thing to spread. Eventually, they

315
01:32:54.060 --> 01:33:05.760
Kate Starbird: Deleted their tweets or sent corrections and things, but it was too late to actually spread pretty broadly and given the person who had spread it a lot of attention. We see this sort of politicization of science ended up

316
01:33:06.900 --> 01:33:09.660
Kate Starbird: Creating this sort of opportunity for the misinformation display.

317
01:33:10.890 --> 01:33:17.310
Kate Starbird: So, so a little bit. I want to talk about our methodological approach and then bring this around for the end so

318
01:33:17.550 --> 01:33:31.170
Kate Starbird: One of the things I think the reasons that we are in this space and and in contributing is that we have this this interdisciplinary methodological approach we have extensive experience studying social media during crisis events. We've bootstrapped

319

01:33:32.280 --> 01:33:42.210
Kate Starbird: That social socio technical infrastructure to conduct this kind of research. We mostly use publicly available data, but we're increasingly relying on access special access from platforms.

320
01:33:43.800 --> 01:33:58.230
Kate Starbird: And we've been adapting methods from Crisis Informatics to the study of misinformation we integrate visual quantitative, qualitative methods we really feel that, like, yes, the high level, big, big picture of us are important, but also to retain closeness to the data is really important.

321
01:33:59.280 --> 01:34:07.740
Kate Starbird: One of our long term objectives is we know we can't do all this research ourselves. And so we've been wondering how do we scale up the ability to do this kind of research for other researchers

322
01:34:08.730 --> 01:34:15.180
Kate Starbird: And researchers that have these questions, researchers and journalists as well. I want to talk about that, if we have time afterwards but

323
01:34:16.890 --> 01:34:27.480
Kate Starbird: They have the questions they have the the contextual expertise but they may not be able to ask their questions they might not be able to have the methodological support. So how do we help them have, you know, contribute to sort of

324
01:34:27.990 --> 01:34:37.530
Kate Starbird: Getting the data technical support for the data and methodological support for asking and answering their questions. And so we're trying to build and extend our socio technical infrastructure.

325
01:34:38.760 --> 01:34:49.200
Kate Starbird: Extend our systems make give access to our data, but we think it's more than just building a tool and opening up our data and publishing on a website somewhere, is that we really need to figure out

326
01:34:49.650 --> 01:34:57.600
Kate Starbird: solutions that are that involved collaboration between us and whether it's a journalist who's trying to move in real time to study these phenomena.

327
01:34:57.960 --> 01:35:06.840

Kate Starbird: Or it's a fellow researcher at the University of Washington or possibly more broadly. One example of what we've done is we've hired data engineer.

328
01:35:07.590 --> 01:35:16.320
Kate Starbird: And for years, we've had these PhD students coming in and doing the work. And we've realized that it's not good for the PhD student and it's not, it's probably not good for us. So as part of for the Center

329
01:35:16.530 --> 01:35:18.240
Kate Starbird: As part of the Center for and for public

330
01:35:18.480 --> 01:35:18.930
Regan: I've got

331
01:35:19.020 --> 01:35:28.380
Kate Starbird: Two more points. Hopefully that we can do that as part of the Center for an informed public we hired this person who is just enhanced our ability to react quickly.

332
01:35:29.130 --> 01:35:38.340
Kate Starbird: And and be able to studies phenomena at a speed and scale we weren't able to before, and I think it's the right thing to do, but it's a very different kind of funding model. And how do we

333
01:35:38.910 --> 01:35:44.220
Kate Starbird: Continue doing that doesn't fit in a in a research grant the NSF that we've written in the past. So how do we

334
01:35:44.820 --> 01:35:49.350
Kate Starbird: Keep that going. And then we have this challenge of balancing commitments to conduct our own research.

335
01:35:49.710 --> 01:35:57.990
Kate Starbird: And those supporting others research and how do we like maintain multiple simultaneous collaborations where we're scaling this up, letting other people come into the

336
01:35:58.320 --> 01:36:03.870
Kate Starbird: Space and at the same time we're still able to to make the contributions that we think are important to make

337
01:36:04.140 --> 01:36:08.250

Kate Starbird: And finally, how do we fund this effort going forward in times of financial uncertainty.

338
01:36:08.520 --> 01:36:13.710
Kate Starbird: We think we're really positioned. Well, both to make a contribution ourselves and to help the broader research community.

339
01:36:13.890 --> 01:36:26.670
Kate Starbird: But this is a very interesting, fun, you know, the way this funding works is not yet stable even for us. We're still trying to figure out how to make that work. Um. Anyways, thank you. Thanks for the stage and I appreciate the opportunity to talk here.

340
01:36:29.970 --> 01:36:30.360
Duncan Watts: Thank you Kate. Okay.

341
01:36:32.160 --> 01:36:35.970
Duncan Watts: Next speaker is. Oh wait, we

342
01:36:37.650 --> 01:36:42.180
Duncan Watts: Have any time for specific questions while David gets set up.

343
01:36:45.330 --> 01:37:05.010
Nadya Bliss: Yeah, there's some. Sorry, I can, I can pull one. Hi Kate. Um, so one of the questions that came in was, how do you measure erosion of trust. So you've mentioned there's erosion of trust in government systems and science institutions, maybe you can address that one.

344
01:37:05.880 --> 01:37:11.430
Kate Starbird: Yeah I we've actually relied on others for the kinds of survey work that you would use to measure.

345
01:37:11.880 --> 01:37:18.870
Kate Starbird: erosion of trust over time. It's not something that I think we can we can grab out of social media data we can grab like

346
01:37:19.290 --> 01:37:25.320
Kate Starbird: You know, direct attacks or how people talk about what are these these institutions to trust it, but social media isn't a great

347
01:37:26.220 --> 01:37:42.000
Kate Starbird: Social media data that we use as our primary source isn't great for that kind of analysis, but we do rely on other researchers

who've done research on trust and they those numbers move around. We've seen some some trusted media go a little bit backup.

348
01:37:43.170 --> 01:37:50.370
Kate Starbird: But the general trends have been down for you know for a while now. So we've been relying on existing research.

349
01:37:53.430 --> 01:37:56.430
Duncan Watts: Great, thank you Kate. Hi, David. Looks like you're ready to go.

350
01:37:57.150 --> 01:38:01.920
David Lazer: Great. Let's see this looking proper on the on the screen.

351
01:38:03.120 --> 01:38:05.010
Duncan Watts: It looks like you need to go into presenter mode.

352
01:38:05.310 --> 01:38:08.310
David Lazer: Okay, wait, sorry.

353
01:38:11.130 --> 01:38:19.620
David Lazer: It's okay, there we go. Sorry. It was just not doing anything for a moment. Alright, great. Well, good. Good to be here.

354
01:38:20.820 --> 01:38:29.670
David Lazer: I, I think this is a terrific event and I really appreciate Beth and Duncan inviting me to speak.

355
01:38:30.930 --> 01:38:45.180
David Lazer: I i'm I'm gonna mainly I think talk about data infrastructures and the kinds of things that we need to invest in, I think, a Shared. Shared infrastructures to to work together in these spaces.

356
01:38:46.470 --> 01:38:47.310
David Lazer: And

357
01:38:49.080 --> 01:38:54.180
David Lazer: And hopefully this will almost view this as a to do list for us collectively

358
01:38:55.770 --> 01:39:01.020

David Lazer: What are the data we need to study the information ecosystem. I think there are

359
01:39:02.610 --> 01:39:10.470
David Lazer: Six kinds of data that I would identify and this isn't meant to be exclusive but but these are I think six times that we really

360
01:39:10.980 --> 01:39:20.550
David Lazer: Need to think about. The first is static content, the stuff that just sort of sits there not changing quickly, dynamic content that produced on the fly,

361
01:39:21.210 --> 01:39:27.540
David Lazer: Exposure to understand what people have actually seen, attention in terms of understanding what has at least made it

362
01:39:28.410 --> 01:39:41.190
David Lazer: Briefly into the brain cognition in terms of understanding what has changing belief. And then finally, what's changing behavior. And I want to walk through each of these, and I hope that it's it's sort of rich for discussion.

363
01:39:42.960 --> 01:39:51.630
David Lazer: static content. I mean, I use this apricot kernel example as a motivating one, it's, it's the classic example of

364
01:39:53.130 --> 01:39:58.770
David Lazer: mis- health misinformation out there. There's, there's some sort of parties

365
01:40:00.270 --> 01:40:06.870
David Lazer: Who are energetically pushing the idea that apricot kernels will cure cancer which it doesn't. By the way, so

366
01:40:08.490 --> 01:40:22.710
David Lazer: bad advice there but you can go to web- Web MD and their static content that doesn't change very often, and you can get information, it's not terribly well written, but reasonably accurate information.

367
01:40:24.330 --> 01:40:28.980
David Lazer: If not, in you know about the use of apricot kernels for curing cancer.

368

01:40:30.570 --> 01:40:30.960
David Lazer: If

369
01:40:32.010 --> 01:40:42.210
David Lazer: If we interestingly, then there's also other static content that's a little bit more dynamic. If we look at user ratings and this again highlights the how precarious the information ecosystem can be

370
01:40:42.720 --> 01:40:54.600
David Lazer: Where we see that user ratings for apricot kernels are quite high on Web MD very, you know, a legit website where it gets a 4.7

371
01:40:55.440 --> 01:41:00.210
David Lazer: On effective to scale for curing cancer and so on with all these testimonials

372
01:41:00.930 --> 01:41:18.690
David Lazer: For for curing cancer. And so, you know, this gets at the notion of what are the robustness of our information ecosystem. But again, this is also getting a sense of what the kinds of data, you would want to get you'd want to know you would want to get data on the static data on what what

373
01:41:20.040 --> 01:41:27.690
David Lazer: What web pages, providing but also then this this commentary, which is more invadable

374
01:41:29.280 --> 01:41:29.820
David Lazer: Below.

375
01:41:31.200 --> 01:41:39.690
David Lazer: Then there's also dynamic content-- dynamic content is literally dynamic that's for content that's produced on the fly.

376
01:41:40.620 --> 01:41:53.790
David Lazer: Typically through interaction with some technical platforms where every time you do a Google search every time you go on Facebook or Twitter those platforms are in that moment, curating something for you to see.

377
01:41:55.050 --> 01:41:55.590
David Lazer: And

378
01:41:56.790 --> 01:42:14.550

David Lazer: And of course that curation process is changing day by day, probably not dramatically, day by day. But what Google will show you for something today for, let's say, apricot seeds and cancer is going to be quite different from a year ago or even a few months ago.

379
01:42:15.690 --> 01:42:22.290
David Lazer: So to give you an example. If you did, search for cancer apricot seeds, this is what I got a few months ago and

380
01:42:23.520 --> 01:42:25.830
David Lazer: And you can see there's an algorithmic component that's

381
01:42:27.210 --> 01:42:38.640
David Lazer: Pulling information out from a particular website that promotes that at the top, Medical News Today, and you can see that that curatorial process matters.

382
01:42:39.330 --> 01:42:46.950
David Lazer: Because even though this article talks about, you know, the one of the sentences there is "apricots seeds can kill you because they have

383
01:42:47.790 --> 01:42:51.870
David Lazer: Because of what's in them" that that

384
01:42:52.440 --> 01:43:05.220
David Lazer: It begins with this weirdly positive few sentences apricot kernels may have some health benefits, blah, blah, blah. And suggest they make your cancer and then it later says in a part not excerpted that there's no science to actually

385
01:43:05.730 --> 01:43:26.130
David Lazer: Actually says this, and and people have died because they've overdosed on on apricot seeds so you can see how that curatorial process matters quite a lot because this may be all that a lot of people see and and then as you go below on the saw on the on the site. You do.

386
01:43:27.300 --> 01:43:27.870
David Lazer: You know,

387
01:43:29.640 --> 01:43:42.030
David Lazer: You see, again, even something from Dana Farber that begins with let's re- let's restate the theories and then and then later has more negative things

388
01:43:42.420 --> 01:43:49.410
David Lazer: And as you go down on the website, you also see Amazon advertising based on the keywords apricots seeds cancer.

389
01:43:50.340 --> 01:43:58.560
David Lazer: And you see a 4.7 rating on Amazon. Because again, those ratings have been invaded. Now, you know, Amazon isn't

390
01:43:59.430 --> 01:44:08.040
David Lazer: It doesn't have some person they're saying, well, what suckers can we identify that will buy these apricots seeds for cancer, but there

391
01:44:08.580 --> 01:44:23.100
David Lazer: Again, they have an algorithmic buying ad buying system that that is identifying the people who do these kinds of searches will buy apricots seeds and then of course they're promoting this this

392
01:44:24.180 --> 01:44:32.970
David Lazer: The high rating the apricots seeds get to as well. And then if you go to the bottom, you see these searches proposed searches apricots seeds cancer testimonials and so on.

393
01:44:33.420 --> 01:44:44.580
David Lazer: And of course, if you click on those, you get lots of YouTube videos telling you how much good apricot seeds are for killing a cancer. So all that is to say,

394
01:44:45.000 --> 01:44:58.920
David Lazer: dynamic content matters because it offers the pathways into static content which may be influential on attitudes and behaviors in sometimes in quite a catastrophic ways

395
01:45:00.420 --> 01:45:21.780
David Lazer: Exposure: you can show people stuff, but that doesn't mean they're exposed to stuff. And so, so we need to understand what what people are least seeing a con briefly cognitively processing, and this is something from my from my newsfeed from again sometime ago on

396
01:45:23.250 --> 01:45:41.880
David Lazer: On Twitter and and so we could see you know that when it's prioritizing but then I don't necessarily read everything I may just read the top thing on the list. And so we we need to understand what people actually exposed to and seeing and. And again, this is really hard.

397
01:45:43.320 --> 01:45:51.600
David Lazer: You know, just throwing in something, you know, just an example of sort of the classic approach or one of the approaches that one might use is

398
01:45:51.900 --> 01:46:01.650
David Lazer: We did a survey on covert news consumption is also by the way gets at a question that was in the in the in the chats around generational effects.

399
01:46:02.160 --> 01:46:13.410
David Lazer: And this this is this is useful. We did a survey. We asked where people gotten information on COVID in the last 24 hours and we can see, you know, and this is consistent with lots of other stuff local TV matters a lot.

400
01:46:14.460 --> 01:46:24.150
David Lazer: A lot TV actually matters quite a lot. Social Media Matters, although somewhat less. There are big generational differences, older people rely a lot on local television younger people rely

401
01:46:24.420 --> 01:46:34.680
David Lazer: Actually more on social ties and social media and so on. But there two points here. This is, I think, state of the art survey approach, but it's also limited because it's it's actually

402
01:46:36.540 --> 01:46:56.250
David Lazer: It's, it's, we can do a lot better and actually observing what people are watching what people are seeing and so and so I think this is good, but I also think that we need to build infrastructure for alternatives, so that we also know you know what people are seeing on Twitter, for example.

403
01:46:57.780 --> 01:47:03.360
David Lazer: And and what and and and we just don't really have that collectively

404
01:47:05.730 --> 01:47:16.080
David Lazer: Now for and this is also from some current research, what we can do. These are the top 10 we built this panel of users on Twitter, one and a half billion people.

405
01:47:16.830 --> 01:47:25.440

David Lazer: This is based on science paper we had last year on consumption of fake news. And we can see these are the top 10 URLs that have been shared in the last several months about

406
01:47:25.770 --> 01:47:36.960
David Lazer: COVID-19 and then you can you can get into very granular here's the story in the Washington Post, which was the most shared story and how much it was shared and what age groups.

407
01:47:37.440 --> 01:47:41.910
David Lazer: Big generational difference again. Again, and so on. And we can look at

408
01:47:42.420 --> 01:47:57.390
David Lazer: partisanship and so on. But this, this gets us down to like, here's the precise content this person with these attributes consumed in this moment. And I think this gives you a flavor of the much more detailed granularity that is possible.

409
01:47:58.530 --> 01:47:59.040
David Lazer: Today,

410
01:48:01.050 --> 01:48:12.660
David Lazer: next category cognition. What do you remember, how do you remember, what do you trust? So when I do those searches on apricots curing cancer, do I remember that?

411
01:48:14.010 --> 01:48:17.160
David Lazer: You know, I forgot a lot of stuff nowadays. So, you know,

412
01:48:18.240 --> 01:48:25.830
David Lazer: does that or do I, how do I remember that. I remember the particular propositions to do. Do I remember things like a scorecard? So there are

413
01:48:26.070 --> 01:48:31.140
David Lazer: Various ways of thinking about how we think and cognition, in some sense, we should view the Internet.

414
01:48:31.470 --> 01:48:40.140
David Lazer: As this vast global cognitive science experiment where it's controlling in various ways, you know, I can take my phone out and it's controlling

415

01:48:40.440 --> 01:48:49.500
David Lazer: The flow of information I'm seeing. And then the question is, like, what's happening in here in my head and then how does that eventually manifest

416
01:48:49.890 --> 01:48:56.880
David Lazer: In in behavior. So for thinking of an anti vaxxers, do we think that, you know, it doesn't affect whether

417
01:48:57.630 --> 01:49:09.330
David Lazer: Let's say I have my child vaccinated and you can anticipate, you know, sometime next winter if we're lucky, we'll be having new antivax battles about whether to to get vaccines for COVID-19

418
01:49:10.680 --> 01:49:22.020
David Lazer: So, you know, so really what we want in terms of a scientific machine is collecting data on all of these things, what, how are people behaving where they actually attending to

419
01:49:22.860 --> 01:49:32.970
David Lazer: What are they, what are they seeing and. And the other thing I want to mention briefly when we're thinking about trusting the ecosystem is

420
01:49:35.610 --> 01:49:43.080
David Lazer: And I'll close in just one and a half minutes is is a crowd the crowd versus elites and I think there's a lot of focus on the role of the crowd.

421
01:49:44.160 --> 01:49:50.310
David Lazer: And unwashed masses sharing this information by also think we, you know, probably still more important. Our leads. This is

422
01:49:50.700 --> 01:50:02.730
David Lazer: A short article by Jane Brody, and she had a whole series of articles and she sort of the Dean of science reporting in in the New York Times for the last 30 years and she had a series of articles

423
01:50:03.780 --> 01:50:07.440
David Lazer: About how harmless opioids were really

424
01:50:08.130 --> 01:50:17.130
David Lazer: And there's a quote here from a particular doctor. She's quoted many years, saying, think, you know, talk about thinking that

there, there'll be a drink dangerous drugs that will do bad things to them.

425
01:50:17.580 --> 01:50:31.680
David Lazer: It was known when this article is written in fact opioids will do bad things to them to people. And I think that this was a kind of misinformation, but it was through one of the most respected science reporters in the New York Times, and so we need to think about the different kinds of

426
01:50:32.910 --> 01:50:38.940
David Lazer: Behaviors and there was a lovely piece in Columbia Journalism Review talking about how the news coverage really affected

427
01:50:39.450 --> 01:50:51.750
David Lazer: The behaviors of primary care physicians. So, oh, I want to conclude with this slide, I have my grades for how we are collectively doing on each of these in terms of creating an infrastructure. And generally, I don't think we're doing well.

428
01:50:52.500 --> 01:50:57.330
David Lazer: I think we're doing something on static content, we're getting some bits on exposure.

429
01:50:58.170 --> 01:51:08.370
David Lazer: Some elements of research on attention, but when we think about dynamic content, yhere's really no ongoing infrastructure cognition behavior, maybe you know there are bits of research, but no ongoing

430
01:51:09.150 --> 01:51:20.580
David Lazer: Infrastructure for understanding how these things all operate together. So for me this is sort of the, the to do list for us as a research community at the intersection of computer science and the social sciences.

431
01:51:25.020 --> 01:51:32.520
Duncan Watts: Thank you, David. It's now 1148 so let's go straight to Chris Wiggins and

432
01:51:33.810 --> 01:51:34.950
David Lazer: Me stop sharing

433
01:51:36.120 --> 01:51:37.860
Duncan Watts: And Chris does not have slides. Correct.

434
01:51:38.160 --> 01:51:38.610
Correct.

435
01:51:40.740 --> 01:51:41.280
Christopher Wiggins: I can start

436
01:51:42.510 --> 01:51:42.750
Duncan Watts: Yep.

437
01:51:43.110 --> 01:51:52.530
Christopher Wiggins: Thank you, Duncan. So following the lead of my many of my humanist colleagues, I'm using prepared remarks, rather than slides also like Duncan

438
01:51:53.220 --> 01:52:00.840
Christopher Wiggins: I'll use a virtual background in part to emphasize that I'm today I'm going to be wearing my hat as a Columbia professor and not as a

439
01:52:01.320 --> 01:52:06.060
Christopher Wiggins: Chief data scientist at the New York Times, so I invoke that because the organizers encouraged me to

440
01:52:06.720 --> 01:52:13.680
Christopher Wiggins: Among other things, address the role of the New York Times, as well as what funding agencies could be doing so.

441
01:52:14.310 --> 01:52:23.760
Christopher Wiggins: I'll try to give two parts around both of those most on the first time. So just to elaborate on the context for I started a sabbatical at the New York Times in 2013

442
01:52:24.150 --> 01:52:34.110
Christopher Wiggins: To start up a data science group on the on the business side that is looking at the New York Times data rather than reporting about data journalism. And so some of my comments are really

443
01:52:35.400 --> 01:52:42.930
Christopher Wiggins: Ethnographic meaning, I can, I can speak to you as somebody who's been embedded in the New York Times for seven years, but I do not speak for the New York Times, and of course

444
01:52:43.380 --> 01:52:54.360

Christopher Wiggins: My background, I should say is, is in machine learning applied to biology. It's not a journalism or media. So I'll make some comments that respond directly to some of the things Kate said and Claudia. And also, David.

445
01:52:56.370 --> 01:53:00.900
Christopher Wiggins: First is some comments about the role of an institution in information so

446
01:53:02.340 --> 01:53:10.560
Christopher Wiggins: Again, I was encouraged to just speak a little bit about the role of journalism institution like the New York Times and disinformation. One thing I would say is that it's very easy. It

447
01:53:11.340 --> 01:53:18.870
Christopher Wiggins: Coming from academia is to anthropomorphize a company or to anthropomorphize a different group of people, as though the New York Times does something or other.

448
01:53:19.110 --> 01:53:23.220
Christopher Wiggins: Or journalists do something or other. And of course we have to remember that journalists are individuals and they have their own

449
01:53:23.790 --> 01:53:33.030
Christopher Wiggins: Motivations when we think about why journalists report in the way they do. And David and I have talked before about that opioid example as an example of one particular journalist in her report.

450
01:53:34.740 --> 01:53:44.310
Christopher Wiggins: There is much in common with scientists. Scientists, you know, are interested in truth, researchers are interested in truth, journalists are interested in truth, but they're also interested in truth on a deadline.

451
01:53:45.090 --> 01:53:53.940
Christopher Wiggins: They share with scientists, a form of peer review that peer review means that they want to earn the respect of people they respect that also means they don't like getting scooped

452
01:53:54.390 --> 01:54:03.810
Christopher Wiggins: And that is relevant, I think, to misinformation in particular misinformation is very dynamic a new medium post appears to invoke Kate’s example

453

01:54:04.680 --> 01:54:14.670
Christopher Wiggins: Or a new scientific results appears and there's a rush to get there first, not, not only because of prestige and the respect of your peers, which has always been true.

454
01:54:15.120 --> 01:54:23.670
Christopher Wiggins: But in the digital age, that also means many more views. So when you are the first to report on a news story and you get ahead of your peers in terms of Google

455
01:54:24.030 --> 01:54:29.610
Christopher Wiggins: search ranking algorithms, you tend to stay there because you're getting more clicks on that story than your peers. And so when

456
01:54:29.820 --> 01:54:40.440
Christopher Wiggins: When Google represents a topic they will represent the thing that gets published first. So there's a lot of there's extra digital incentive to what was already a very dangerous dynamic which disinformation

457
01:54:41.070 --> 01:54:49.200
Christopher Wiggins: Amplifies so within thinking about the New York Times, one thing I would say is about the role of people, you should remember what are the motivations of the journalists.

458
01:54:50.220 --> 01:54:57.060
Christopher Wiggins: Another thing to think about in media companies is the role of technology. So in technology, they're both reporting on

459
01:54:57.450 --> 01:55:05.040
Christopher Wiggins: Technology and they're reporting about technology and they are using technology. So we should remember that journalism, I would say, is a craft.

460
01:55:05.430 --> 01:55:13.890
Christopher Wiggins: It has largely developed for centuries, without being a particularly technology graft. And so many of the journalists, you know, their technique is shoe leather work.

461
01:55:14.580 --> 01:55:22.560
Christopher Wiggins: Meaning recording where you go talk to people. That is also true of data scientists, you know it's it's very important to understand the biases in any data set.

462

01:55:23.040 --> 01:55:33.450
Christopher Wiggins: That's also true of journalists as well and we should keep that in mind as we read critically works reported by journalists and as we partner with journalists to try to make sure that the truth gets out there.

463
01:55:34.980 --> 01:55:44.130
Christopher Wiggins: The role of technology therefore means journalists are often critical of technology platforms. Right. That's a, that's a very different community from the community of journalists.

464
01:55:44.490 --> 01:55:50.100
Christopher Wiggins: It also meant that journalists are often sometimes skeptical of research that we in the research community do

465
01:55:51.090 --> 01:55:57.930
Christopher Wiggins: Do using a lot of machine learning or technology otherwise that I would also say is not entirely different from scientists

466
01:55:58.800 --> 01:56:04.500
Christopher Wiggins: I can't help but think about the physics, quote, if your experiment needs a statistician, you've ought to have done a better experiment.

467
01:56:05.070 --> 01:56:16.080
Christopher Wiggins: So, you know, many, many scientists also come with a skepticism of data, which I think complicates those of us who who do research that we want to influence the journalistic narrative.

468
01:56:17.850 --> 01:56:22.740
Christopher Wiggins: But there's one thing I'd say about a news institution like the New York Times and how it relates to society.

469
01:56:23.490 --> 01:56:31.710
Christopher Wiggins: Is the role of trusted institutions, which I think is, is really extremely dynamic we we often think that, you know, trust in

470
01:56:32.040 --> 01:56:38.940
Christopher Wiggins: Institutions has precipitously dropped recently. I think one of the things that the Pew Foundation has done valuable research on is to chart out

471
01:56:39.210 --> 01:56:49.860

Christopher Wiggins: Trust in public institutions as a function of time, which has been steadily decreasing for decades. So we feel like it's a recent phenomenon, but in fact trust in institutions has, in fact, I've been dropping for a long time.

472
01:56:51.390 --> 01:56:57.600
Christopher Wiggins: What is the role of public? What are the motivations of a public institution today, if I can go back to

473
01:56:58.530 --> 01:57:05.790
Christopher Wiggins: Against my advice anthropomorphizing it's fun to look back on an ancient propaganda model. So from from 30 years ago, Chomsky

474
01:57:06.180 --> 01:57:14.370
Christopher Wiggins: Put forward this propaganda model of how media functions of propaganda that media companies are large corporations and they're interested in maximizing profit.

475
01:57:14.790 --> 01:57:19.740
Christopher Wiggins: And sometimes true sometimes not true, for example, Jeff Bezos is on a ship, currently, the Washington Post.

476
01:57:20.580 --> 01:57:30.990
Christopher Wiggins: They run by the advertising model that was a second filter Chomsky pointed out 30 years ago. That's definitely a challenge, you know, companies like the New York Times have turned aggressively towards subscription and away from advertising.

477
01:57:31.800 --> 01:57:40.530
Christopher Wiggins: The reporting is biased by access to sources. This I think is complicated when the people in power are not politicians with whom you cultivate relationships for decades.

478
01:57:40.860 --> 01:57:46.770
Christopher Wiggins: But instead information platform companies and I'll come back to the relationship between journalist and information platform companies in a second.

479
01:57:47.610 --> 01:57:55.410
Christopher Wiggins: And the fourth pillar is the role of people who try to influence things that certainly still true. In fact, more true than ever since there's so much money

480
01:57:56.070 --> 01:58:03.000

Christopher Wiggins: Involved in trying to influence journalism. And finally having a common enemy, which in 1988 when the propaganda model was proposed, it was communism.

481
01:58:03.750 --> 01:58:13.470
Christopher Wiggins: Now there's not clearly a communism going to be for all journalistic properties. So that's one that's one type of analysis of journalism that I think it would be useful for us to update

482
01:58:15.360 --> 01:58:25.200
Christopher Wiggins: Another aspect for us, which I do think gets it. There's a little bit of this in Kate's work is the role of adversaries. So in addition to this info, we have this info journalists do get

483
01:58:26.010 --> 01:58:35.640
Christopher Wiggins: Not only aggressively PR'd to invoke David's example with opioids, but journalists are subject to disinformation that's planted. So people certainly misrepresent things

484
01:58:36.120 --> 01:58:44.460
Christopher Wiggins: And journalists have an ongoing adversarial information relationship with anything that trends as Ren√©e DiResta heavily says, if it trends, it's true

485
01:58:44.790 --> 01:58:55.350
Christopher Wiggins: So if you report on something that's trending but it's misinformation you amplify it. If you don't report on something that's trending and is misinformation then you have the risk of sparking a counter conspiracy in which people look at your

486
01:58:55.890 --> 01:58:59.580
Christopher Wiggins: Company and say, You are not reporting on the truth. It's being suppressed.

487
01:59:01.830 --> 01:59:10.080
Christopher Wiggins: What is the intersection between journalism and research and civic interest. Well, one of the most important intersections is the role of leakers and whistleblowers.

488
01:59:11.400 --> 01:59:23.940
Christopher Wiggins: Great changes in society on all the two minutes too numerous to mention have been one thanks to whistleblowers who eventually went to the press. This is particularly true when we have our realities mediated by opaque

489

01:59:24.660 --> 01:59:33.360
Christopher Wiggins: Platform companies, right, we have very little transparency about the algorithms that design choices. We actually reali- we all citizens rely on whistleblowers quite a bit.

490
01:59:34.380 --> 01:59:41.370
Christopher Wiggins: One thing that supports whistleblowers from coming to the presses when tech companies successfully convinced their employees that the press

491
01:59:41.700 --> 01:59:49.620
Christopher Wiggins: Are vultures and are not to be trusted. And actually, it's actually successful move in the interest of maintaining opacity. If you convince your employees that the press is out to get you

492
01:59:50.730 --> 02:00:00.390
Christopher Wiggins: Which I don't know how we're going to fix directly, but it's something, those of us who've tried to maintain good relations with employees at tech platform companies in some ways are doing one small part

493
02:00:01.470 --> 02:00:07.410
Christopher Wiggins: The disinformation and the adversary is special in the time of COVID if I can amplify one point of Kate's

494
02:00:08.610 --> 02:00:17.430
Christopher Wiggins: Not only because the disinformation is scientific so it's not necessarily a dispute over facts of who was in a particular garage leaking a particular piece of political information.

495
02:00:18.150 --> 02:00:19.650
Christopher Wiggins: It's over a scientific facts.

496
02:00:20.010 --> 02:00:31.740
Christopher Wiggins: However, in the case of COVID, the scientific facts themselves are still not worked out. So unlike climate change where there's overwhelming scientific consensus and there's still, you know, tremendous money to be spent influencing people to turn away from science, COVID is very

497
02:00:32.820 --> 02:00:41.250
Christopher Wiggins: Beatrice dynamic time in which the scientists themselves still are working out the fact of the matter. So there's plenty of room for disinformants to get in there and win the day

498

02:00:43.860 --> 02:00:49.440
Christopher Wiggins: Okay, I wanted to move from some ethnographic comments about New York Times in journalism, things like that.

499
02:00:49.770 --> 02:00:56.220
Christopher Wiggins: And then pick up a something that David said about what data we want and connect it to something Claudia said around ethics.

500
02:00:56.520 --> 02:01:01.770
Christopher Wiggins: So David said, it's really hard to get the data you want it, you know, in part you can get partial data for example from surveys

501
02:01:02.100 --> 02:01:07.590
Christopher Wiggins: But as a scientist, we're often interested in doing direct experiments. So how can we go about doing direct experiments?

502
02:01:08.520 --> 02:01:16.260
Christopher Wiggins: It's not easy on information platform companies and, in particular, it's not information is not easy and what I'll call multiplayer mode. So to illustrate the point

503
02:01:16.710 --> 02:01:22.620
Christopher Wiggins: David showed an example of an experiment where you run a search query and then you report the results of that search query.

504
02:01:23.370 --> 02:01:27.420
Christopher Wiggins: There's a bunch of work in that field under you know auditing algorithms.

505
02:01:28.140 --> 02:01:36.420
Christopher Wiggins: Plenty of great work by, for example, Latanya Sweeney, looking at search results for how name for her name algorithms of oppression in some ways motivated by

506
02:01:36.870 --> 02:01:42.990
Christopher Wiggins: Different search results that you get with slight variations. Excellent form of experimental work you can do in single player mode.

507
02:01:43.560 --> 02:01:49.680
Christopher Wiggins: Much more difficult to do any choice of research on an information platform company that works in network mode. In order to do so,

508
02:01:50.280 --> 02:01:57.930
Christopher Wiggins: You need to actually intervene and see what affects our other people sometimes, life is good and you get either an instrumental variable or a natural experiment. That's true.

509
02:01:58.620 --> 02:02:07.320
Christopher Wiggins: But ever since 1925 when Fisher said we should be doing randomized controlled trial. There's been a lot of interest in how we could do some sort of experimental trial to learn causal impact.

510
02:02:07.980 --> 02:02:19.110
Christopher Wiggins: The problem we're looking at here, in some extent, is how are the subjective design choices made by information platform companies impacting our understanding of reality, it's very difficult for us to do that.

511
02:02:19.530 --> 02:02:26.430
Christopher Wiggins: Not only is it difficult for us to do that. But even for companies that actually do produce an API we could do some experiments, we are very--

512
02:02:28.590 --> 02:02:30.570
Christopher Wiggins: We need to have a conversation about ethics, so

513
02:02:31.710 --> 02:02:43.500
Christopher Wiggins: I'll say something slightly provocative. This is closed meeting. So it's a good time to say something provocative although it's an august group so maybe I shouldn't say anything too provocative. So if I can thread the needle, a little bit, um, you know, the 1970s

514
02:02:44.520 --> 02:02:53.220
Christopher Wiggins: As a result of a whistleblower, there's a lot of concerted work commissioned by the federal government to try to think about what it means to do ethical experiments in the context of human subject research.

515
02:02:54.750 --> 02:03:01.350
Christopher Wiggins: For many of us, we only encounter that in a sort of a checklist form by interacting with our IRB, but a point made by

516
02:03:01.860 --> 02:03:08.520
Christopher Wiggins: Duncan's student, Matt Salganik, in his book on computational social science is that a lot of work was done in the 1970s to think through

517
02:03:09.360 --> 02:03:18.180
Christopher Wiggins: One of the principles of ethics that then led to the IRB process and those principles are quite broad. And then, just as we use the United States Constitution stills centuries later.

518
02:03:18.600 --> 02:03:32.370
Christopher Wiggins: I think it might be useful to revisit what those principles look like for a digital age and how we can go about doing ethical experiments that help us all understand the causal impact of subjective design decisions made by information platform companies.

519
02:03:33.420 --> 02:03:40.890
Christopher Wiggins: Heretofore there's been a lot of critical take on that and we all would like not to do anything unethical and when experiments have been done.

520
02:03:41.850 --> 02:03:47.640
Christopher Wiggins: The experiments that have really rise risen to public attention, have been ones that are you should use universally

521
02:03:48.630 --> 02:03:58.380
Christopher Wiggins: Claimed, for example, the Alabama experiment in politics was reported on widely by Scott Shane at the New York Times, among others, as an example of somebody putting up fake

522
02:03:58.830 --> 02:04:07.680
Christopher Wiggins: Accounts on Facebook in order to investigate their, their influence, but it was universally derided. So a provocative idea is, it might be of interest for a group like

523
02:04:09.300 --> 02:04:16.500
Christopher Wiggins: NSF or somebody else to ask, is there something like the Belmont Principles or is there a reaffirmation of the Belmont principles

524
02:04:16.770 --> 02:04:21.420
Christopher Wiggins: That allows researchers who are interested in getting the data they want to invoke this title David's title.

525
02:04:22.050 --> 02:04:31.770
Christopher Wiggins: In a way that nonetheless comports with our communal understanding of what it means to do so ethically to invoke a phrase by Celia and with that the door bell has rung so close. Thank you.

526

02:04:35.460 --> 02:04:36.000
Duncan Watts: Great, thanks.

527
02:04:38.160 --> 02:04:44.940
Duncan Watts: Yeah, sorry. Can I just get all the speakers to turn off their videos.

528
02:04:46.440 --> 02:04:53.190
Duncan Watts: And then we will have Nadya and I will will build some kind of general questions.

529
02:04:54.360 --> 02:05:06.540
Duncan Watts: That sort of try to sum up the spirit of the of the questions we have seen in some of our own thoughts to get the the panelists discussing so Nadya, you want to lead off?

530
02:05:07.080 --> 02:05:14.550
Nadya Bliss: Yeah, sure. And just a reminder, if all the panelists could turn on their video cameras and be ready to

531
02:05:15.600 --> 02:05:27.960
Nadya Bliss: unmute be ready to go for the conversation. So thank you so much, everybody. And good morning still good morning over here in Arizona, and it's it's great to see everyone, everyone here.

532
02:05:28.650 --> 02:05:38.430
Nadya Bliss: So there were a lot of questions during Kate's talks specifically about the goals of her center. And essentially, how do we combat

533
02:05:38.790 --> 02:05:45.750
Nadya Bliss: disinformation so I would like to, well, starting with Kate, have this be a more general question-- it seems that there are

534
02:05:46.200 --> 02:05:53.130
Nadya Bliss: Issues essentially of resiliency-- individual resiliency to disinformation, societal resiliency to disinformation,

535
02:05:53.430 --> 02:06:02.190
Nadya Bliss: And system resiliency to disinformation and misinformation is in there as well. So how do we forward at each of move forward on each of those individual levels?

536
02:06:02.520 --> 02:06:14.430

Nadya Bliss: How do we cross between levels and what are the research challenges? And I know that this is a loaded question. So, pick, pick an element that you think is most relevant from your work to address.

537
02:06:16.710 --> 02:06:18.510
Nadya Bliss: And Kate, you can start

538
02:06:18.840 --> 02:06:20.520
Kate Starbird: Yeah, I guess we've been thinking about

539
02:06:21.300 --> 02:06:25.440
Duncan Watts: Everyone should turn their video is on not off so

540
02:06:26.700 --> 02:06:27.510
Duncan Watts: Can you please

541
02:06:28.050 --> 02:06:29.100
Nadya Bliss: Oh, she's on it.

542
02:06:30.060 --> 02:06:30.840
Duncan Watts: Yeah yeah

543
02:06:32.430 --> 02:06:43.470
Kate Starbird: Um, so we think about that we just changed it a little bit, but I think I think you're absolutely right, we've been thinking about sort of three different directions. If we're going to be more resilient, there's, you know, sort of a

544
02:06:44.190 --> 02:06:49.770
Kate Starbird: Collective ed- education and it doesn't mean like in a classroom, necessarily, but that's about us becoming more

545
02:06:51.480 --> 02:07:01.740
Kate Starbird: Aware and responsible and resilient information participants and that's that's one, there's also that the platforms could work differently.

546
02:07:02.040 --> 02:07:06.390
Kate Starbird: And so there's a there's a technology component that the platforms could be designed in different ways.

547
02:07:06.720 --> 02:07:17.160

Kate Starbird: And and related, those aren't totally separate but there's two different kinds of directions. And the third one is that policies could be different, not just platform policy but government policy, we could actually

548
02:07:17.640 --> 02:07:20.340
Kate Starbird: Collectively come together and decide that things should work in a different way.

549
02:07:20.640 --> 02:07:27.030
Kate Starbird: And so we've been thinking about those three different directions which actually map a little bit to our center into three different what we have four pillars, but

550
02:07:27.240 --> 02:07:41.790
Kate Starbird: They kind of map into those three different kinds of solutions spaces. I don't necessarily think that's it, all of it, but that's kind of how we've been thinking about it and trying it. And we literally have, you know, different researchers in charge of in it or trying to kind of

551
02:07:43.080 --> 02:07:48.990
Kate Starbird: Lead research and and interventions in in each of those different directions.

552
02:07:51.720 --> 02:07:55.440
Nadya Bliss: Thank you. I have Chris on my screen. So Chris, you want to go next.

553
02:07:55.800 --> 02:08:05.670
Christopher Wiggins: Sure. Um, I guess, amplifying- so I can't help but think about this class I've been teaching. So I've been teaching this class that I co-teach with a historian about the history and ethics of data.

554
02:08:06.120 --> 02:08:20.520
Christopher Wiggins: And when we get to the very last two lectures, it's sort of problems and solutions. So, in the solutions lecture, the analysis we try to use is about this unstable game of three powers: people power, state power, and corporate power so

555
02:08:23.070 --> 02:08:31.560
Christopher Wiggins: I do think there's room for some people power. We often look to corporate governance as as though that's the only way to think about forces that shape companies, but

556
02:08:32.820 --> 02:08:46.890
Christopher Wiggins: Well one point, governance is not just the federal government there's also state and local and international government, but I do think there's a role for people power in the sense of the tech employees and their power around walkouts and leaking.

557
02:08:48.030 --> 02:08:55.620
Christopher Wiggins: And other disruptive functions. There's a great law review article about this called private ordering in the world of tech employees, which I encourage you all to check out

558
02:08:56.790 --> 02:09:02.280
Christopher Wiggins: That's on the short time scale and on the long time scale and I still- where my head is a professor- I still believe in education so

559
02:09:02.580 --> 02:09:12.900
Christopher Wiggins: I do think one thing that we can all do, and those of us who were at the research-education intersection is to think hard about what we're teaching students, which is in part why did we develop this new class.

560
02:09:13.530 --> 02:09:16.950
Christopher Wiggins: Which is a class, both for as I sometimes say the techies and the fuzzies.

561
02:09:17.490 --> 02:09:25.920
Christopher Wiggins: To teach a little bit about the functional capabilities of data science as well as the ethical import and this and the social technical realities.

562
02:09:26.400 --> 02:09:37.290
Christopher Wiggins: These students we're teaching are going to become the product managers and the senators of the future. So my long bet is to be bullish on education, provided that

563
02:09:38.550 --> 02:09:52.200
Christopher Wiggins: Educational reformers reform. In short term though, I do think that we should remember that the governance doesn't only come from federal government. Right. There's, there's, there's many powers that constrain technology companies, I hope that speaks to your question, Nadya

564
02:09:54.540 --> 02:09:55.170
Nadya Bliss: Thank you.

565
02:09:57.630 --> 02:09:58.200
Nadya Bliss: David

566
02:09:59.190 --> 02:10:05.010
David Lazer: Oh, sure. I think I think Kate and Chris covered it nicely, I guess.

567
02:10:06.330 --> 02:10:17.130
David Lazer: It's like reformulation would just be to say that, you know, we need to think about the machinery, by which individuals are susceptible

568
02:10:17.580 --> 02:10:22.950
David Lazer: To miss information. Because if people aren't susceptible wouldn't really matter how it spread. And then we have to think about the

569
02:10:23.400 --> 02:10:38.940
David Lazer: Machinery of spreading, which is in part a batter of individual choice. Do I tell other people do I try to persuade other people, but it's also a structural thing about how the platforms and other mass disseminators of information

570
02:10:40.320 --> 02:10:46.770
David Lazer: Either can be manipulated and amplify misinformation or could dampen it. And I'd say,

571
02:10:48.240 --> 02:11:02.010
David Lazer: I'd say, really, those are those are the key, the key ingredients. When we think about a robust system if the if and and you put those all together and you have a system that's either pretty robust or pretty fragile so

572
02:11:05.820 --> 02:11:06.540
David Lazer: I'll keep it there.

573
02:11:08.880 --> 02:11:18.450
Claudia Deane: Because, just to add that, you know, there is some optimistic news here, which is that we know the American public is very concerned about misinformation. So it's not like it's a

574
02:11:18.570 --> 02:11:24.600
Claudia Deane: stealth, you know a stealth attack going on the challenge. You know, they even think they see it, the challenges that we

575
02:11:25.530 --> 02:11:35.760
Claudia Deane: put our trust you know we are in such a polarized partisan environment that where people go, you know, who people trust their information has become increasingly polarized and so

576
02:11:36.390 --> 02:11:44.250
Claudia Deane: They just seem this information is coming from whatever the opposite side is that are not as resistant to it when it comes from their own perceived side.

577
02:11:45.840 --> 02:11:57.390
David Lazer: The only thing I would add to that, I think both Pew's surveys and our surveys highlight how actually scientists and doctors and medical professional are still very highly trusted

578
02:11:58.320 --> 02:12:08.580
David Lazer: On both sides of the aisle bit more trusted on the left and on the right, but so there's a little bit polarization there, but there is general agreement. Now we'll see how that holds up

579
02:12:09.720 --> 02:12:18.720
David Lazer: Over the next few months. I suspect that unfortunately we're going to see some erosion of that consensus, but little hope.

580
02:12:24.360 --> 02:12:34.020
Duncan Watts: All right. I mean, I'd like to touch on a theme that came up in different ways in all four of the presentations, listening to you,

581
02:12:34.830 --> 02:12:48.780
Duncan Watts: It's very clear that, you know, we want to wrap our arms around this problem of, you know, you know. How is information being produced? How is it being consumed? How is it being absorbed? How is that ultimately

582
02:12:49.830 --> 02:12:57.060
Duncan Watts: Impacting the kinds of things that we care about like public trust and institutions political polarization functioning of democracy, etc.

583
02:12:57.780 --> 02:13:12.000
Duncan Watts: That the sort of meta problem that we have to solve is that the data that we would need to to to even sort of describe what is going on is extremely distributed across many different

584
02:13:12.810 --> 02:13:34.260
Duncan Watts: Sources, it's, it's very heterogeneous in nature, you know, some of its survey data, some of its social media data, some of its traditional you know online data or TV data. So it comes in many different forms is very noisy, it has all kinds of problems with so like with bias and and selection.

585
02:13:35.880 --> 02:13:49.920
Duncan Watts: So, and nobody currently has the ability to sort of, you know, query that database, right. So it's sort of the world has not been set up for us to study the kinds of problems. And so we end up with

586
02:13:50.310 --> 02:13:56.850
Duncan Watts: With something like, you know, read, read the Kipling's poem about the blind men and the elephant where everybody's sort of got their own

587
02:13:57.540 --> 02:14:07.770
Duncan Watts: Their own data set that they've collected themselves and they're they're drawing inferences from that that they're trying to generalize and everybody's coming up with different generalizations with with the state of the world is

588
02:14:08.310 --> 02:14:21.420
Duncan Watts: So in terms of, you know, going back to the sort of objectives of this meeting, you know, what are some things that that we, that the NSF could do to, to, to, you know, to help with the

589
02:14:22.140 --> 02:14:32.580
Duncan Watts: The, the, the, just the just this the straightforward or conceptually straightforward act of just measuring the things that we care about?

590
02:14:35.610 --> 02:14:45.720
Christopher Wiggins: Maybe I can celebrate one of Kate's answers, which is better tooling, which we provide both the academics and to journalists. Sometimes journalists rush in and where academics are reticent to tread.

591
02:14:46.860 --> 02:14:56.190
Christopher Wiggins: You know, journalists have some tooling available to them mostly vendors. But the more we create tooling that allows keep investigation, the more that would open up one new

592
02:14:57.870 --> 02:15:06.810

Christopher Wiggins: Avenue of research. There's certainly survey data, which I think could be expanded modernized expanded digitalized that's another possible avenue.

593
02:15:11.430 --> 02:15:13.590
Kate Starbird: We've been thinking about this a lot. I mean, even

594
02:15:16.170 --> 02:15:20.880
Kate Starbird: Even for folks that have the ability to collect data and they have the ability to analyze it, which is

595
02:15:21.900 --> 02:15:22.590
Kate Starbird: Just a few

596
02:15:24.150 --> 02:15:32.400
Kate Starbird: We don't have all the data. Right. So there's both like this data sharing of like how do we get these different perspectives. But how do we bring those perspectives together, either for a single sort of

597
02:15:32.850 --> 02:15:46.380
Kate Starbird: Misinformation narrative that's online. So let me focus it just on that. That thing is moving across so many different platforms and no researcher has has is following it across all of them. So how do we, you know, how do we bring together resources.

598
02:15:47.460 --> 02:15:51.120
Kate Starbird: Both sort of academic resources, there's also these sort of

599
02:15:52.350 --> 02:16:00.570
Kate Starbird: Other groups like Jason Baumgartner's group, I forget what it's called, that are collecting all this data and making it available to researchers, we're not really sure if we could use it. We don't know about the terms of service.

600
02:16:01.650 --> 02:16:07.350
Kate Starbird: But trying to bring together these different kinds of things and and i don't necessarily think it's always just

601
02:16:08.100 --> 02:16:14.610
Kate Starbird: The tool and the data, but I really do think it's it's collaborations between the people that know different sets and sort of a

602
02:16:15.330 --> 02:16:22.170

Kate Starbird: Different sets, different methods have contextual understanding have the questions. And how do you bring those things together, especially in the cases of journalists.

603
02:16:22.560 --> 02:16:27.570
Kate Starbird: Who you know have the ability to go look at the things they've been staring at the data really close, they need

604
02:16:28.230 --> 02:16:35.520
Kate Starbird: They need assistance in in looking at at the at the data at scale in a different way and combining that with their kind of integrated perspective.

605
02:16:36.180 --> 02:16:45.390
Kate Starbird: But I don't think necessarily we talked to a few in its research we're doing it just giving them a tool is going to solve the problem. It's often about how did they set up a collaboration where they have access to kind of

606
02:16:46.380 --> 02:16:53.670
Kate Starbird: Some something to help them ask those questions that are hard to ask without some really strong methodological experience that many of them don't have

607
02:16:57.000 --> 02:16:57.480
Claudia Deane: I think

608
02:16:58.770 --> 02:17:04.380
Claudia Deane: I would point to Bob Groves who I don't know if he's on this call from Georgetown, but

609
02:17:05.490 --> 02:17:15.540
Claudia Deane: Is on our board and has been constantly encouraging us to do one thing, which is to find out how these new methodology, the new data collection efforts. How do they relate to our traditional

610
02:17:16.050 --> 02:17:23.310
Claudia Deane: Sort of- so take for example survey efforts, we know we're going to get different results. But we anticipate, they're probably correlated in some way and that would let you

611
02:17:23.520 --> 02:17:32.370
Claudia Deane: If you could figure out that pattern, you could model them a little bit back and forth. And I don't know David on the panel, if you want to talk a little bit about how we've been trying to

612
02:17:33.420 --> 02:17:40.320
Claudia Deane: Sort of square the two ways that we've been looking at trying to get a representative population of Twitter users.

613
02:17:42.360 --> 02:17:43.380
Claudia Deane: You in Northeastern, yeah

614
02:17:43.830 --> 02:17:53.010
David Lazer: I have this wonderful project with Pew where where we've been working on these two approaches to developing Twitter panels

615
02:17:53.820 --> 02:17:59.520
David Lazer: that we worked on matching to voter data, administrative data, and we have a very large sample.

616
02:18:00.120 --> 02:18:08.580
David Lazer: Whereas, whereas Pew is built a smaller panel based on address based sampling where you then ask people for their handles.

617
02:18:08.970 --> 02:18:22.740
David Lazer: And and we look at where we had resonant results and where they deviated in so we can see the kinds of biases that can creep in both kinds of methods and then hopefully adjust accordingly. So, for example,

618
02:18:23.880 --> 02:18:35.040
David Lazer: Matching with administrative data to take a simple example we we we match too many women relative to men because women are more likely to have statistically unusual names.

619
02:18:36.270 --> 02:18:44.940
David Lazer: And and and so so we end up looking like we have too many women in our sample, but on the other hand,

620
02:18:45.660 --> 02:18:53.220
David Lazer: What Pew can do is they can see who turned them down. So yes, I use Twitter, but I don't give you, I'm not going to give you my Twitter handle

621
02:18:53.520 --> 02:18:58.140
David Lazer: And unsurprisingly women tend to be underrepresented in that approach. And so we can

622

02:18:58.710 --> 02:19:13.290
David Lazer: link that to survey data or link that to administrative data and we get a sense of the relative weaknesses. But I do think that it is it is really not about just using the data that happened to fall off the truck, but it is about the data.

623
02:19:14.340 --> 02:19:18.720
David Lazer: That we want to construct and build and evaluating the kinds of biases

624
02:19:19.170 --> 02:19:28.620
David Lazer: That are intrinsic to the methodology. In this case, we have two different methods for getting to the same place, and they have somewhat different biases. Another bias in our method is, it turns out that

625
02:19:29.160 --> 02:19:40.500
David Lazer: We under match Asian American names because I suspect that there's just less good data in terms of of of shortened versions of names

626
02:19:41.250 --> 02:19:45.990
David Lazer: That for for more Anglo names as compared to Asian American names and thus we lose a lot of

627
02:19:46.350 --> 02:19:57.780
David Lazer: Representations of individuals that as they do on social media. So all of this is to get to. And this is also a lot of the questions I've seen in the group chat is how do we really come create this fusion of

628
02:19:58.410 --> 02:20:11.160
David Lazer: Of methodologies around computer science and social science and think about the kinds of biases that get entered in. And how do we characterize the population and how do we how do we bring some of those sensibilities to these large scale digital trees data.

629
02:20:14.850 --> 02:20:19.050
Nadya Bliss: Right. So, another really interesting question that came in was

630
02:20:20.040 --> 02:20:31.800
Nadya Bliss: Essentially, can we use parallels from more traditional computer security in treating information as essentially a new programming language. And what are some of the techniques

631

02:20:32.190 --> 02:20:39.480
Nadya Bliss: That we can bring from that community into the information trust communities. So essentially,

632
02:20:39.780 --> 02:20:54.630
Nadya Bliss: Right now, a lot of the machine learning algorithms are being used to manipulate which information is being served in which information targets which populations. So can we think about it from an information security perspective how cybersecurity professionals do

633
02:20:56.730 --> 02:21:02.760
Christopher Wiggins: One thing that's great from Info Sec is just categorizing different types of the tracks types of attacks and the

634
02:21:03.030 --> 02:21:12.990
Christopher Wiggins: Severity of attacks. So not that, not that it's universally used within info sack or still considered state of the art but like stride and dread, a sort of taxonomy of what is the type of attack?

635
02:21:13.440 --> 02:21:16.290
Christopher Wiggins: How bad is the attack? Is it something that's repeatable, for example,

636
02:21:16.890 --> 02:21:24.660
Christopher Wiggins: is an escalation of privilege or something else. I think you know taxonomy is like they can at least get started just talking on the same page, right, this is not going to happen in journalism.

637
02:21:24.990 --> 02:21:31.320
Christopher Wiggins: Careful taxonomy of disinformation and attacks in their, in their the threat vector and the severity is not going to happen in journalism.

638
02:21:31.650 --> 02:21:40.680
Christopher Wiggins: So getting to consensus there which is taking a long time and the Info Sec community. I think will will help us at least having you're using using the right terms and speaking the same language.

639
02:21:41.910 --> 02:21:51.840
Nadya Bliss: Though I would say in the computer security community, we're still struggling with metrics. That's one of the biggest unsolved challenges. Any other comments on that one?

640

02:21:57.570 --> 02:22:05.880
Duncan Watts: I'd just like to come back to a couple of comments from the, from the chat thread and there's been a number of comments, sort of in the general area of

641
02:22:07.320 --> 02:22:08.190
Duncan Watts: You know of

642
02:22:09.240 --> 02:22:17.340
Duncan Watts: Combining the, the kind of traditional emphasis that the social sciences have placed on

643
02:22:18.720 --> 02:22:20.040
Duncan Watts: On you know

644
02:22:21.510 --> 02:22:30.570
Duncan Watts: On sampling methodology and and worrying about selection bias in data as well as

645
02:22:31.950 --> 02:22:35.850
Duncan Watts: You know, careful thinking about causal mechanisms.

646
02:22:36.900 --> 02:22:44.100
Duncan Watts: With the, you know, with, you know, more sort of computational methods and and you know

647
02:22:44.790 --> 02:22:45.630
Duncan Watts: How is that

648
02:22:46.890 --> 02:22:47.640
Duncan Watts: How is that

649
02:22:49.650 --> 02:23:07.770
Duncan Watts: High hybrid approach being currently manifested in computational social science and and how could it be done better? What we need to do to to to do a better job of bringing together computational thinking and more sort of traditional social scientific thinking

650
02:23:09.090 --> 02:23:10.020
Duncan Watts: Around these problems?

651
02:23:13.470 --> 02:23:18.450

Christopher Wiggins: cross disciplinary respect and training. We see sort of a long term investment, a long term investment is

652
02:23:18.900 --> 02:23:26.760
Christopher Wiggins: Respect for social scientists among people who teach computer science, respect for computer science among those who keep social science and then training people in the two

653
02:23:27.240 --> 02:23:34.800
Christopher Wiggins: And then you plant the seed and you wait two decades, so it takes a while. Um, the other of course its funding mechanisms to get those people together, which is one way of kickstarting

654
02:23:35.790 --> 02:23:44.250
Duncan Watts: So we will probably talk a lot about the first of those things in our, in our third session today, which is about the technical workforce.

655
02:23:45.330 --> 02:23:48.180
Duncan Watts: But it may be if if

656
02:23:49.770 --> 02:23:57.060
Duncan Watts: David, Kate, Claudia had thoughts about you know this, you know, the second point there. What can we do sort of in the short run to

657
02:23:58.620 --> 02:24:04.410
Duncan Watts: To to build teams or or whatever we need to work on these problems.

658
02:24:06.210 --> 02:24:18.810
David Lazer: I'll throw out a few thoughts and I do think that this is it is really at this fusion, that is sort of a long term scientific hope. I think that, and I actually think things are much better now than they were,

659
02:24:19.890 --> 02:24:28.920
David Lazer: You know, just a matter of years ago, I think, actually, you know, but. But part of what needs to be done in the short run is

660
02:24:30.210 --> 02:24:40.800
David Lazer: You know, training, I think that I think that there are actually some very on the social science side as well as in computer science side, there are some very core set of skills and ideas

661

02:24:41.160 --> 02:24:47.220
David Lazer: That would be great if each community could adapt, right, like a lot of what I'm seeing in comments is like, Well, how do we think about

662
02:24:47.490 --> 02:24:52.770
David Lazer: Bias in sampling and bias, not just individuals, but we should also be thinking about bias and sampling

663
02:24:53.220 --> 02:24:59.400
David Lazer: Of behaviors. When we look at Twitter, but not Facebook or what people are saying it's like we're getting this very weird slice of people and I think

664
02:24:59.700 --> 02:25:03.690
David Lazer: Actually social science collectively has thought a lot about these kinds of

665
02:25:04.260 --> 02:25:14.940
David Lazer: Biases and that you know that it's amazing. I think how far for people would study social media on CSI and how far just so some basic ideas and research design would go

666
02:25:15.420 --> 02:25:28.140
David Lazer: And then similarly, I also think on the social science side getting some basic data skills is not, you know, to deal with digital trace data is not is not a huge reach. And so obviously that this all works better for

667
02:25:28.950 --> 02:25:32.880
David Lazer: More junior people, which is, I think, where we see a lot of the innovation, but I think that

668
02:25:33.180 --> 02:25:43.020
David Lazer: In the short run, they're sort of a matter of a little knowledge actually could help a lot in terms of building these bridges and then building people in long term building programs that

669
02:25:43.470 --> 02:25:46.260
David Lazer: Create a more thorough integration is important.

670
02:25:46.680 --> 02:25:57.000
David Lazer: And then I think we're always going to need multi disciplinary teams and we need to think about funding mechanisms. I think

there's no more effective way to get a bunch of computer scientists and social scientists in a room together than to say

671
02:25:57.390 --> 02:26:12.900
David Lazer: You know you- this is important problem for you to solve, and here's a bunch of money that will that you need to use together. And I think that a lot, you know that those if done properly, that that would be very effective, combined with the other steps so

672
02:26:15.300 --> 02:26:15.990
Duncan Watts: We've had a

673
02:26:17.460 --> 02:26:24.510
Duncan Watts: Mindful of time here. We've had quite a few questions about about ethics and

674
02:26:26.070 --> 02:26:30.810
Duncan Watts: You know, sort of bleeding into our next session. Also, how

675
02:26:31.890 --> 02:26:53.010
Duncan Watts: How that thinking about ethics intersects with with historical disparities and the the differential treatment of different groups in the population by by, you know, scientific institutions, historically, and how that may relate to, to, to quite legitimate trust issues.

676
02:26:54.750 --> 02:27:01.530
Duncan Watts: But I think one, you know, this is a problem that I have like worried about a lot for years is that is that, you know,

677
02:27:02.610 --> 02:27:14.010
Duncan Watts: That you know at the at the data generation level, you know, most if not all of this data has been generated by individual people. And so that is something that

678
02:27:14.340 --> 02:27:31.320
Duncan Watts: You know, I think, has has attracted a lot of increasing attention in the last few years and and what are the rights that individuals have to their data and privacy into being left alone and all of the kinds of principles that come from the Belmont report.

679
02:27:33.150 --> 02:27:36.900
Duncan Watts: But as we are learning in particularly in this

680

02:27:38.220 --> 02:27:38.910
Duncan Watts: Pandemic

681
02:27:39.930 --> 02:27:50.730
Duncan Watts: At the at the mass aggregate collective level, this data is also an incredibly valuable public good that it is something that transcends

682
02:27:51.690 --> 02:28:02.580
Duncan Watts: The, the, you know the the particular individuals that have produced it and it gives us insights into movements of whole societies and with potentially quite profound

683
02:28:03.120 --> 02:28:11.040
Duncan Watts: Public health impact. And I think something like that is true for lots of these problems that we're talking about. And, and I'm not really aware of, you know, sort of,

684
02:28:12.150 --> 02:28:29.370
Duncan Watts: You know, many public conversations that are really trying to address this. The core of this issue of how do we sort of balance the private rights of individuals with the public value of the data that that that exists at an aggregate level. So just

685
02:28:30.570 --> 02:28:40.230
Christopher Wiggins: COVID trackers is one is one modern- er, contemporary narrative. So the narrative a narrative around COVID trackers, and this is very different in different countries is exactly that beneficence versus

686
02:28:41.040 --> 02:28:48.300
Christopher Wiggins: Rights, you know, like the rest of the right to consent to my information versus the fact that I could be def- flattening the curve by sharing my information.

687
02:28:49.710 --> 02:28:53.550
Kate Starbird: And we can look at different cultural contexts where that

688
02:28:55.260 --> 02:29:07.590
Kate Starbird: Have different arrangements with that tension, like China has a very different arrangement of the attention of like what data is available and what data you know becomes part of the collective good versus maybe some of the ways we see it here. So there's some really interesting value.

689

02:29:09.360 --> 02:29:10.710
Kate Starbird: tensions with that.

690
02:29:12.120 --> 02:29:21.240
Kate Starbird: With with these kinds of trade offs and clearly COVID-19 gives us an example of where it's, you know, the public good.

691
02:29:22.080 --> 02:29:30.300
Kate Starbird: To have more of this data, but it's still, you know, in the long term, loss of privacy is still something we should be concerned about. And we should be having those conversations. I think they're really important

692
02:29:32.220 --> 02:29:45.480
David Lazer: And if I could jump in and this is a somewhat different angle. But I also think it has to inform our the core research questions. I think, you know, sometimes associated technical systems are intrinsically

693
02:29:46.950 --> 02:29:49.110
David Lazer: Majoritarian in terms of

694
02:29:50.310 --> 02:29:52.830
David Lazer: What gets shared and so on and

695
02:29:53.850 --> 02:29:56.160
David Lazer: And that there needs to be

696
02:29:57.690 --> 02:30:03.060
David Lazer: There is a need for hardcore researchers to be socio technical

697
02:30:04.140 --> 02:30:15.840
David Lazer: Critics and this is a particular, this is a different angle, and our ethical conundrums because I think we need to, on the one hand, often work with industry and, on the other hand, be critics of industry and I think that there are a lot of ways

698
02:30:17.100 --> 02:30:28.020
David Lazer: that that we need--we need to find that, you know, inner passion for thinking about social justice when we're looking at these data, but then

699
02:30:28.680 --> 02:30:41.520

David Lazer: But then, you know, we have to, let's say, go talk to Facebook and say, Well, can we, under what terms. Can we look at these data or or try to negotiate, you know, enhanced access to Twitter data and so on and so I think we are collectively

700
02:30:42.270 --> 02:30:48.510
David Lazer: In a bit of a pickle in terms of what what our role needs to be in terms of making our society better

701
02:30:48.750 --> 02:30:55.650
David Lazer: And practically what we need to do to negotiate access to these data. And that's not that about the privacy stuff. So it's a different angle and the

702
02:30:55.860 --> 02:31:03.210
David Lazer: Ethical stuff, which I think is also really important. But I also want to think about that. What we are doing is good, and it has massive potential to improve things

703
02:31:03.480 --> 02:31:15.300
David Lazer: And we need to be motivated by that and yet we also have to somehow negotiate these other things, and I wish I had the answer for that as well. But I'm just raising the conundrum.

704
02:31:18.990 --> 02:31:20.910
Kate Starbird: Want to follow on that, that conundrum on the

705
02:31:21.360 --> 02:31:31.650
Kate Starbird: The sort of data divide issue and I kind of hinted that into my slides is that we're we're feeling these benefits of special access as we get invited into spaces that everyone doesn't have this is true of researchers and this is

706
02:31:32.100 --> 02:31:39.330
Kate Starbird: True of journalists and so really difficult situation because we get in a situation where we need this information from certain platforms that are not

707
02:31:39.660 --> 02:31:47.340
Kate Starbird: Not wholly public in order to ask important questions at the same time we're so dependent on them for that data that there's this kind of

708
02:31:47.880 --> 02:31:58.800

Kate Starbird: Tension in even how we report on things about, you know, making sure we meet we maintain that data. But what does that mean, it means, you know, it means that we do have these conflicts and inherit kind of conflict.

709
02:31:59.820 --> 02:32:13.230
Kate Starbird: And it's something that we really need to talk about as a larger community about how do we make sure that that that access is equitable that other people have access as well, including up and coming journalists and so, you know, becoming researchers and

710
02:32:14.640 --> 02:32:20.250
Kate Starbird: And how do we make sure that we don't feel conflicted and how we report on that data, due to the access that we have

711
02:32:21.210 --> 02:32:26.310
Christopher Wiggins: You can imagine expanded version of broader impact that's a funding criterion that involves

712
02:32:27.090 --> 02:32:32.790
Christopher Wiggins: Deep ethnographic work with effective Community or evidence that you've co created the research with somebody from a community.

713
02:32:33.600 --> 02:32:45.480
Christopher Wiggins: You know who you're not just going to represent as number 60 Virginia Eubanks focus, great example starting up and going and interviewing people and then making really making visceral to the reader who may not come from that community idea that what they're doing to us first because

714
02:32:47.460 --> 02:32:53.400
Christopher Wiggins: There's a lot to be done by scientists for whom that's not part of their, their usual MO in an experimental design.

715
02:32:56.490 --> 02:33:01.050
Nadya Bliss: I think also mean in general, even in this discussion, the panelists are essentially sending out

716
02:33:01.560 --> 02:33:11.610
Nadya Bliss: A pretty deep research question as to what are the benefits of having the data versus the privacy implications and a lot of the privacy and ethics implications.

717
02:33:11.910 --> 02:33:24.780

Nadya Bliss: That's the social science aspect of it while the benefits. What's discoverable is almost like an information theoretic machine learning type of question. So I can certainly see that as essentially being one of the

718
02:33:26.070 --> 02:33:38.550
Nadya Bliss: pillars that could use a lot more research in a longer term research agenda that brings that to just the disciplines together, not two, there's a lot more here than two. We're at 930 Duncan, should we wrap up?

719
02:33:39.960 --> 02:33:40.680
Nadya Bliss: 1230

720
02:33:41.730 --> 02:33:44.280
Duncan Watts: At night around the world. And this again.

721
02:33:45.090 --> 02:33:55.560
Duncan Watts: So yes, thank you all. Thank you to all of our speakers and thank you for the questions. Unfortunately, we didn't get to all of them, but I think we got some good ones.

722
02:33:56.940 --> 02:34:10.590
Duncan Watts: And now we'll be taking a 30 minute break and we'll be back here at 1pm Eastern Daylight Time for a second session, but did you want to add anything

723
02:34:10.890 --> 02:34:22.920
Beth Mynatt: No, just thanks again terrific work by our panelists and just to let everyone know we are deeply appreciating the conversation discussions going on in the channels and we're capturing all of that as well. So

724
02:34:23.520 --> 02:34:33.540
Beth Mynatt: Great work. Everyone please take advantage of your 30 minutes and stretch recover and please come back to you at 1pm Eastern Time.