WEBVTT

1
00:00:00.520 --> 00:00:02.220
<v ->Good morning and welcome to the first session</v>

2
00:00:02.220 --> 00:00:05.120
of our four-part Data Science Career Seminar Series,

3
00:00:05.120 --> 00:00:07.500
bringing data science to addiction research.

4
00:00:07.500 --> 00:00:08.830
My name is Susan Wright.

5
00:00:08.830 --> 00:00:11.910
I'm from the Division of Neuroscience and Behavior or DNB.

6
00:00:11.910 --> 00:00:13.540
And I'm the program officer for big data

7
00:00:13.540 --> 00:00:16.200
and program for big data and computational science

8
00:00:16.200 --> 00:00:18.810
and leading our data science efforts here at NIHDA.

9
00:00:18.810 --> 00:00:21.260
Training and data science is a priority for NIHDA

10
00:00:22.096 --> 00:00:24.042
and it's supported by our new office

11
00:00:24.042 --> 00:00:25.790
of research training, diversity and disparities,

12
00:00:25.790 --> 00:00:27.730
or ORTDD.

13
00:00:27.730 --> 00:00:30.027
Organizers of the seminar series include myself,

14
00:00:30.027 --> 00:00:32.450
Dr. Roger Liddle, the Deputy Director

15
00:00:32.450 --> 00:00:34.830
of the Division of Neuroscience and Behavior,

16
00:00:34.830 --> 00:00:37.330
Dr. Bolson Compton (indistinct) Deputy Director

17
00:00:37.330 --> 00:00:39.260
and Acting Director of the Office of Research

18
00:00:39.260 --> 00:00:41.180
Training, Diversity, and Disparities,

19
00:00:41.180 --> 00:00:42.427
Dr. Albert Avella,

20
00:00:42.427 --> 00:00:44.470
Deputy Director of the Office of Research

21
00:00:44.470 --> 00:00:46.330
Training Diversity and Disparities,

22
00:00:46.330 --> 00:00:47.370
and the director of the office

23
00:00:47.370 --> 00:00:49.440
of disparities and health disparities,

24
00:00:49.440 --> 00:00:51.100
and Dr. Lindsay fund the research

25
00:00:51.100 --> 00:00:53.360

and career development program officer and the office

26
00:00:53.360 --> 00:00:56.070
of research, training, diversity and disparities.

27
00:00:56.070 --> 00:00:57.996
I want to thank Albert and Lindsay

28
00:00:57.996 --> 00:01:00.783
for their help with organizing the seminar series.

29
00:01:01.774 --> 00:01:02.790
And I also wanna thank the team who has been helping

30
00:01:02.790 --> 00:01:05.700
with the technical details, and that includes Lucia TaRita

31
00:01:05.700 --> 00:01:08.933
Susan Holbrook, Caitlin due Devor, and David Maza.

32
00:01:10.430 --> 00:01:13.050
This first session is in the format of a fireside chat

33
00:01:13.050 --> 00:01:15.300
and there will also be time for questions from the audience.

34
00:01:15.300 --> 00:01:17.840
Please use the chat box to submit your questions.

35
00:01:17.840 --> 00:01:19.860
And after the fireside chat concludes,

36
00:01:19.860 --> 00:01:23.320
we'll get to as many questions from the chat box as we can.

37
00:01:23.320 --> 00:01:25.308
We've organized the seminar series to include

38
00:01:25.308 --> 00:01:28.030
the feature exciting talks from both data science

39
00:01:28.030 --> 00:01:29.584
industry leaders and Ida funded scientists

40
00:01:29.584 --> 00:01:33.050
who were incorporating data science into their research.

41
00:01:33.050 --> 00:01:34.782
You'll hear more about the upcoming sessions later

42
00:01:34.782 --> 00:01:37.591
but by organizing the seminar series that way,

43
00:01:37.591 --> 00:01:39.800
we aim to generate some interest in discussion

44
00:01:39.800 --> 00:01:42.310
about how we can further efforts to bring data science

45
00:01:42.310 --> 00:01:43.480
to addiction research

46
00:01:43.480 --> 00:01:45.320
and hopefully will inspire a new generation

47
00:01:45.320 --> 00:01:47.714
of data scientists focused on addiction research.

48
00:01:47.714 --> 00:01:49.560
Throughout the series,

49
00:01:49.560 --> 00:01:50.450
we'll hear about a variety

50
00:01:50.450 --> 00:01:52.640

of interest in data science careers, and learn

51
00:01:52.640 --> 00:01:54.940
about the different paths the speakers took to get there,

52
00:01:54.940 --> 00:01:56.600
the skills needed, et cetera.

53
00:01:56.600 --> 00:01:58.147
I'm very excited to talk to you

54
00:01:58.147 --> 00:02:00.606
about this morning's fireside chat with DJ Patil.

55
00:02:00.606 --> 00:02:03.584
Before we get started, our NIDA director,

56
00:02:03.584 --> 00:02:07.050
Dr. Nora will make some opening remarks.

57
00:02:07.050 --> 00:02:08.817
Nora has been the director of NIDA

58
00:02:10.151 --> 00:02:11.960
for 18 years and is a stalwart advocate of data science.

59
00:02:11.960 --> 00:02:14.210
Please join me in welcoming her this morning.

60
00:02:15.470 --> 00:02:16.303
<v ->Susan, good morning.</v>

61
00:02:16.303 --> 00:02:17.240
Good morning, everyone.

62
00:02:17.240 --> 00:02:19.057
I wish we were at, I always say this,

63
00:02:19.057 --> 00:02:20.960
but really it would be so fantastic

64
00:02:20.960 --> 00:02:22.791
when we finally see each other.

65
00:02:22.791 --> 00:02:26.520
But I do want to thank the right for her leadership

66
00:02:26.520 --> 00:02:27.700
on the data science.

67
00:02:27.700 --> 00:02:31.070
And I think it is one of the most exciting areas

68
00:02:31.070 --> 00:02:32.278
that we've seen develop

69
00:02:32.278 --> 00:02:35.601
and actually starting to understand the enormous

70
00:02:35.601 --> 00:02:39.094
power that having data gives us for almost everything.

71
00:02:39.094 --> 00:02:44.094
And yet the challenge of taking advantage of it

72
00:02:46.010 --> 00:02:48.330
is to generate on the one hand,

73
00:02:48.330 --> 00:02:50.240
the databases that have the quality

74
00:02:50.240 --> 00:02:53.778
that will enable us to inquire into them on the one hand,

75
00:02:53.778 --> 00:02:55.214

but the other one,

76
00:02:55.214 --> 00:02:58.333
the scientist strain to take advantage of it.

77
00:02:59.260 --> 00:03:01.870
And these were upon these seminar series

78
00:03:01.870 --> 00:03:03.940
become so very important.

79
00:03:03.940 --> 00:03:06.731
It opens up a space for all sorts of dialogue,

80
00:03:06.731 --> 00:03:09.400
for understand the complexities

81
00:03:09.400 --> 00:03:11.688
and the challenges that we face,

82
00:03:11.688 --> 00:03:14.498
but also to get a view into a world

83
00:03:14.498 --> 00:03:17.673
that in many instances, we didn't even realize it existed.

84
00:03:17.673 --> 00:03:22.673
On my side as a scientist coming from brain imaging.

85
00:03:23.890 --> 00:03:27.370
What I can tell you from what data science has done

86
00:03:27.370 --> 00:03:30.420
it has basically revolutionized anything

87
00:03:30.420 --> 00:03:33.681
that we've ever thought would have been possible or happen.

88
00:03:33.681 --> 00:03:36.850
I need just the exploits on the one hand

89
00:03:36.850 --> 00:03:41.160
that the ability to get at complexity

90
00:03:41.160 --> 00:03:44.612
by obtaining sufficiently large data sets

91
00:03:44.612 --> 00:03:49.110
that allow to see associations and relationships

92
00:03:49.110 --> 00:03:51.473
that otherwise would be impossible to discern.

93
00:03:52.410 --> 00:03:54.910
I think that there are still major gaps

94
00:03:54.910 --> 00:03:57.040
that we need to build all of us.

95
00:03:57.040 --> 00:03:58.727
I mean, obviously from the science

96
00:03:58.727 --> 00:04:02.928
of analytics modeling mathematics prediction

97
00:04:02.928 --> 00:04:07.928
but from the big task of trying to understand

98
00:04:08.540 --> 00:04:11.720
how integrate diverse set of data.

99
00:04:11.720 --> 00:04:16.010
And that is an area that I think we are going to be seeing,

100
00:04:16.010 --> 00:04:18.860

may your advances in the next few years

101
00:04:18.860 --> 00:04:21.140
but we're still just at the beginning.

102
00:04:21.140 --> 00:04:23.163
It's extraordinary exciting,

103
00:04:24.144 --> 00:04:25.680
and I am glad to be part of this.

104
00:04:25.680 --> 00:04:29.080
And I'm glad that we have such a strong advocate

105
00:04:29.080 --> 00:04:30.050
in Dr. Wright,

106
00:04:30.050 --> 00:04:35.050
and I also want to thank Dr. Little Dr. Wilson and Dr. Abila

107
00:04:38.040 --> 00:04:40.670
for all of their efforts to make these seminars possible.

108
00:04:40.670 --> 00:04:45.209
So I look forward to today's talk and the chat

109
00:04:45.209 --> 00:04:47.310
and I think that it's wonderful

110
00:04:47.310 --> 00:04:49.597
to have Dr. Patil in this morning.

111
00:04:50.469 --> 00:04:53.323
So Dr. Wright, I return the phone back to you.

112
00:04:55.518 --> 00:04:56.630
Thank you, Nora.

113
00:04:56.630 --> 00:04:58.210
So now I'm very excited to introduce

114
00:04:58.210 --> 00:05:01.030
our special guest today, Dr. DJ Patel.

115
00:05:01.030 --> 00:05:03.880
Dr. Patel has held a variety of roles in academia,

116
00:05:03.880 --> 00:05:05.080
industry and government.

117
00:05:06.388 --> 00:05:07.740
He is CTO for devoted health

118
00:05:07.740 --> 00:05:09.500
and just recently announced that he is stepping back

119
00:05:09.500 --> 00:05:12.020
from operational duties and is joining the board.

120
00:05:12.020 --> 00:05:14.350
He's also a senior fellow at the Belfer center

121
00:05:14.350 --> 00:05:16.260
at the Harvard Kennedy school

122
00:05:16.260 --> 00:05:19.030
and an advisor to bedrock partners.

123
00:05:19.030 --> 00:05:22.110
Dr. Patel was appointed by president Obama

124
00:05:22.110 --> 00:05:24.630
to be the first us chief data scientist,

125
00:05:24.630 --> 00:05:26.880

where his efforts led to the establishment of nearly

126
00:05:26.880 --> 00:05:30.530
40 chief data officer roles across the federal government.

127
00:05:30.530 --> 00:05:32.805
He also established new healthcare programs

128
00:05:32.805 --> 00:05:34.671
including the precision medicine initiative

129
00:05:34.671 --> 00:05:38.300
and the cancer moonshot, new criminal justice reforms

130
00:05:38.300 --> 00:05:40.040
including the data-driven justice

131
00:05:40.040 --> 00:05:41.800
and police data initiatives

132
00:05:41.800 --> 00:05:44.840
that cover more than 94 million Americans,

133
00:05:44.840 --> 00:05:47.300
as well as leading the national data efforts.

134
00:05:47.300 --> 00:05:49.109
He has also been active in national security

135
00:05:49.109 --> 00:05:52.450
and for his efforts was awarded by secretary Carter,

136
00:05:52.450 --> 00:05:53.750
the department of defense medal

137
00:05:53.750 --> 00:05:55.610
for distinguished public service

138
00:05:55.610 --> 00:05:56.560
which is the highest honor

139
00:05:56.560 --> 00:05:58.246
the department on a civilian in industry.

140
00:05:58.246 --> 00:06:01.935
He led the product teams that relate IQ

141
00:06:01.935 --> 00:06:04.080
which was acquired by Salesforce.

142
00:06:04.080 --> 00:06:07.090
Was founding board member for crisis Textline

143
00:06:07.090 --> 00:06:09.380
which works to use new technologies to provide

144
00:06:09.380 --> 00:06:10.776
on demand mental and crisis support

145
00:06:10.776 --> 00:06:14.680
and was a member of venture firm, Ray lock partners.

146
00:06:14.680 --> 00:06:16.370
He was also chief scientist,

147
00:06:16.370 --> 00:06:18.860
chief security officer and head of analytics

148
00:06:18.860 --> 00:06:21.520
and data product teams at the LinkedIn corporation

149
00:06:21.520 --> 00:06:24.052
where he co coined the term data scientist.

150
00:06:24.052 --> 00:06:25.820

He has also held a number of roles

151
00:06:25.820 --> 00:06:27.883
at Skype, PayPal and eBay.

152
00:06:29.100 --> 00:06:31.890
As a member of faculty at the university of Maryland,

153
00:06:31.890 --> 00:06:35.320
his research focused on nonlinear dynamics and chaos theory.

154
00:06:35.320 --> 00:06:37.110
And he helped start a major research initiative

155
00:06:37.110 --> 00:06:39.330
on numerical weather prediction.

156
00:06:39.330 --> 00:06:42.040
As a AAA's science and technology policy fellow

157
00:06:42.040 --> 00:06:43.740
with the department of defense,

158
00:06:43.740 --> 00:06:45.559
Dr. Patel directed new efforts to leverage social

159
00:06:45.559 --> 00:06:48.810
network analysis and the melding of computational

160
00:06:48.810 --> 00:06:51.582
and social sciences to anticipate

161
00:06:51.582 --> 00:06:53.160
emerging threats to the U.S.

162
00:06:53.160 --> 00:06:55.610
He has also co-chaired a major review of U.S efforts

163
00:06:55.610 --> 00:06:58.057
to prevent bio weapons proliferation in central Asia,

164
00:06:58.057 --> 00:07:01.640
and co-founded the Iraqi virtual science library

165
00:07:01.640 --> 00:07:03.260
or the IBSL.

166
00:07:03.260 --> 00:07:07.260
In 2014, he was selected by the world economic forum

167
00:07:07.260 --> 00:07:09.570
as a global yet because a young global leader,

168
00:07:09.570 --> 00:07:12.070
as a member of the council for foreign relations.

169
00:07:12.070 --> 00:07:14.300
And according to him, he barely graduated from high school

170
00:07:14.300 --> 00:07:15.710
because of his math grades.

171
00:07:15.710 --> 00:07:17.580
Please join me in welcoming Dr. DJ Patel.

172
00:07:17.580 --> 00:07:20.180
And before we get started, Nora,

173
00:07:20.180 --> 00:07:22.480
would you like to ask Dr. Patel any questions?

174
00:07:26.160 --> 00:07:27.530
<v ->No, I'm fine.</v>

175
00:07:27.530 --> 00:07:28.670

Let's start.

176
00:07:28.670 --> 00:07:29.560
<v ->Okay.</v>

177
00:07:29.560 --> 00:07:32.510
I'm gonna turn it over to Wilson to ask the first question.

178
00:07:36.750 --> 00:07:38.240
<v ->Thanks, and good morning.</v>

179
00:07:38.240 --> 00:07:40.096
It's really a pleasure to have you here, Dr. Patel.

180
00:07:40.096 --> 00:07:42.867
And just to get us started,

181
00:07:42.867 --> 00:07:45.710
you mentioned that you didn't do very well

182
00:07:45.710 --> 00:07:48.980
in math during primary education.

183
00:07:48.980 --> 00:07:51.440
So what did you want to do when you were growing up

184
00:07:51.440 --> 00:07:53.519
and were there any clues in your childhood

185
00:07:53.519 --> 00:07:58.050
or your education that might make you might make us think

186
00:07:58.050 --> 00:08:01.080
that you were destined to be a data scientist?

187
00:08:01.080 --> 00:08:02.303
Yeah.

188
00:08:02.303 --> 00:08:04.143
Well, first everyone should just call me DJ, please.

189
00:08:04.143 --> 00:08:07.170
And apologies that we can't also be together

190
00:08:07.170 --> 00:08:09.240
because this would have been one of those great excuses

191
00:08:09.240 --> 00:08:13.290
to be back on campus at NIH.

192
00:08:13.290 --> 00:08:16.790
And I really miss spending time up there.

193
00:08:16.790 --> 00:08:21.150
When I was in federal service again, last time around

194
00:08:21.150 --> 00:08:24.180
I got to spend a lot of time in the agent.

195
00:08:24.180 --> 00:08:25.550
It was one of the great joys.

196
00:08:25.550 --> 00:08:28.860
So I just really miss being up there in person

197
00:08:28.860 --> 00:08:30.523
with all of you.

198
00:08:31.582 --> 00:08:36.582
I'm just a few miles from where I grew up,

199
00:08:37.272 --> 00:08:39.980
California, which is right down the street

200
00:08:39.980 --> 00:08:42.588

from where Apple was starting.

201
00:08:42.588 --> 00:08:45.880
And so I got to spend a lot of time playing our computers.

202
00:08:45.880 --> 00:08:47.950
And one of the things I remember was like,

203
00:08:47.950 --> 00:08:50.200
in those early days,

204
00:08:50.200 --> 00:08:52.680
some of the things that you could do with computers

205
00:08:52.680 --> 00:08:54.121
one of them was fractals.

206
00:08:54.121 --> 00:08:58.233
And there was these Mandel brought generator functions

207
00:08:58.233 --> 00:09:01.210
and they just create these beautiful pictures

208
00:09:01.210 --> 00:09:02.340
or you zoom in.

209
00:09:02.340 --> 00:09:04.340
And these ideas just fascinated me

210
00:09:04.340 --> 00:09:06.811
and I really spent a lot of time

211
00:09:06.811 --> 00:09:09.340
trying to understand these ideas.

212
00:09:09.340 --> 00:09:12.290
And at that time, chaos theory was becoming a thing

213
00:09:12.290 --> 00:09:14.270
but I just didn't have the math skills.

214
00:09:14.270 --> 00:09:17.760
And I just really was deeply excited

215
00:09:17.760 --> 00:09:22.460
by neuroscience and the brain and in these early ideas

216
00:09:22.460 --> 00:09:25.613
of plasticity and all of this stuff.

217
00:09:26.683 --> 00:09:29.500
But I just, I was growing up in an educational system

218
00:09:29.500 --> 00:09:33.520
where it just didn't work for me honestly.

219
00:09:33.520 --> 00:09:37.300
And so while I was super interested in coding

220
00:09:37.300 --> 00:09:38.690
and playing with data,

221
00:09:38.690 --> 00:09:41.879
and trying to just figure out things on the computer

222
00:09:41.879 --> 00:09:45.400
I just didn't have the backing or the training.

223
00:09:45.400 --> 00:09:48.868
And so I actually went to community college first.

224
00:09:48.868 --> 00:09:51.200
I didn't get into a traditional university.

225
00:09:51.200 --> 00:09:53.658

And so I went to one of our wonderful community colleges,

226
00:09:53.658 --> 00:09:56.030
DeAnza college which is on par

227
00:09:56.030 --> 00:09:59.723
with Montgomery community college, near NIH.

228
00:09:59.723 --> 00:10:02.580
And I took a calculus class

229
00:10:02.580 --> 00:10:06.470
because my girlfriend was taking a calculus class.

230
00:10:06.470 --> 00:10:09.830
And so I was in this thing and I was like, wow,

231
00:10:09.830 --> 00:10:13.130
I don't understand why don't always do things end with DX.

232
00:10:13.130 --> 00:10:18.130
And I just fell in love with the whole concept of calculus.

233
00:10:18.790 --> 00:10:20.358
And then from then on, it was just off to the races.

234
00:10:20.358 --> 00:10:24.403
But I think the real credit for me

235
00:10:24.403 --> 00:10:26.570
learning to love of math,

236
00:10:26.570 --> 00:10:28.583
rests with community college, honestly.

237
00:10:30.880 --> 00:10:33.240
<v ->That's an important, a bit of background.</v>

238
00:10:33.240 --> 00:10:35.250
Our goal is to inspire people

239
00:10:35.250 --> 00:10:37.209
in training to take tackle things that

240
00:10:37.209 --> 00:10:39.270
they may be interested in,

241
00:10:39.270 --> 00:10:41.730
but may not be be comfortable

242
00:10:41.730 --> 00:10:43.700
or be confident that they can figure out.

243
00:10:43.700 --> 00:10:45.330
And I love that idea

244
00:10:45.330 --> 00:10:47.550
that you found some local resources to really

245
00:10:47.550 --> 00:10:50.343
help you overcome that barrier.

246
00:10:51.830 --> 00:10:52.864
<v ->Yeah.</v>

247
00:10:52.864 --> 00:10:53.814
It's one of the things that I think

248
00:10:53.814 --> 00:10:58.680
if anything, you all have been on the cutting edge of that.

249
00:10:58.680 --> 00:11:00.623
And I think the rest of the countries,

250
00:11:01.508 --> 00:11:02.950

we had this a long time ago

251
00:11:02.950 --> 00:11:04.640
and we've kind of forgotten it

252
00:11:04.640 --> 00:11:07.060
which is what does it really look like to build

253
00:11:07.060 --> 00:11:09.761
the infrastructure at a community level,

254
00:11:09.761 --> 00:11:11.630
to support things.

255
00:11:11.630 --> 00:11:16.630
And what does that really mean to have high quality

256
00:11:16.750 --> 00:11:19.835
interactions, support layers,

257
00:11:19.835 --> 00:11:24.320
and all of that information also be distilled in a way

258
00:11:24.320 --> 00:11:28.840
that the sum of the parts gets to be put together

259
00:11:31.175 --> 00:11:32.121
in a way that benefits everybody.

260
00:11:32.121 --> 00:11:33.115
And how do you bring it together?

261
00:11:33.115 --> 00:11:35.050
How do you disseminate it, all of those different aspects?

262
00:11:35.050 --> 00:11:37.660
And that's something that we know intrinsically

263
00:11:37.660 --> 00:11:41.720
but I would argue over at least definitely

264
00:11:41.720 --> 00:11:45.340
over our federal budgets at our different aspects

265
00:11:45.340 --> 00:11:48.240
at the state level and the County and city levels

266
00:11:48.240 --> 00:11:50.020
we've really lost a lot of that.

267
00:11:50.020 --> 00:11:51.650
And one of the things I think

268
00:11:51.650 --> 00:11:53.800
which may be a little contrary is

269
00:11:53.800 --> 00:11:56.000
why we have the super power of the internet

270
00:11:56.000 --> 00:11:57.884
and a lot of these technologies,

271
00:11:57.884 --> 00:12:02.070
I'm actually more in favor of the local human connections,

272
00:12:02.070 --> 00:12:05.390
rather than just objects scaling of technologies

273
00:12:05.390 --> 00:12:07.570
and saying, look, we're just gonna build this into an app.

274
00:12:07.570 --> 00:12:10.413
I think the apps are really powerful and supportive

275
00:12:10.413 --> 00:12:12.560

but it's not a replacement

276
00:12:12.560 --> 00:12:16.020
for the infrastructure that really needs to happen

277
00:12:16.020 --> 00:12:17.270
at that local level.

278
00:12:17.270 --> 00:12:19.460
We should fundamentally be using technology

279
00:12:19.460 --> 00:12:20.720
to build superpowers,

280
00:12:20.720 --> 00:12:25.010
to augment the direct human interactions.

281
00:12:25.010 --> 00:12:27.565
And this is one of the reasons like even in healthcare,

282
00:12:27.565 --> 00:12:29.210
I'm not in favor of just

283
00:12:29.210 --> 00:12:31.310
some AI system replacing a physician.

284
00:12:31.310 --> 00:12:34.040
I don't think that's the value an AI system

285
00:12:34.040 --> 00:12:37.812
doesn't replace somebody holding a person's hand in crisis

286
00:12:37.812 --> 00:12:39.058
or being able to see

287
00:12:39.058 --> 00:12:42.070
or ask one additional question that gives an insight.

288
00:12:42.070 --> 00:12:44.720
Can it help spur a conversation?

289
00:12:44.720 --> 00:12:46.590
Can it help give us additional insights?

290
00:12:46.590 --> 00:12:49.053
Yes, but it's not a replacement tool.

291
00:12:51.160 --> 00:12:54.000
<v ->Certainly that resonates for me.</v>

292
00:12:54.000 --> 00:12:57.450
And I appreciate that philosophical approach of

293
00:12:57.450 --> 00:13:00.819
in some ways, reminding us that the technology is a support

294
00:13:00.819 --> 00:13:05.819
not a replacement for our brains and our human connections.

295
00:13:06.800 --> 00:13:07.633
Thank you.

296
00:13:08.520 --> 00:13:10.090
Roger, I think you're up.

297
00:13:10.090 --> 00:13:10.990
<v ->Yes, thank you.</v>

298
00:13:10.990 --> 00:13:14.090
Good morning, DJ, thank you for being with us.

299
00:13:14.090 --> 00:13:16.928
I'd like to ask you to expand a little bit on

300
00:13:16.928 --> 00:13:19.050

you've talked about community college

301
00:13:19.050 --> 00:13:20.850
and how important it was to you.

302
00:13:20.850 --> 00:13:23.185
And I thank you for bringing that up.

303
00:13:23.185 --> 00:13:26.370
Could you talk a little bit more about other points

304
00:13:26.370 --> 00:13:28.500
along your career path that were important

305
00:13:28.500 --> 00:13:31.070
and maybe if you were giving advice

306
00:13:31.070 --> 00:13:35.250
to someone coming along early in their career,

307
00:13:35.250 --> 00:13:37.973
maybe things to avoid or things to seek out.

308
00:13:39.190 --> 00:13:40.431
<v ->Yeah.</v>

309
00:13:40.431 --> 00:13:43.870
I think there's a few things that I see as commonalities

310
00:13:43.870 --> 00:13:45.570
and I think I was grateful,

311
00:13:45.570 --> 00:13:49.430
I did my doctorate at the university of Maryland

312
00:13:49.430 --> 00:13:53.690
and one of the most fantastic things there is

313
00:13:53.690 --> 00:13:55.180
it was so interdisciplinary.

314
00:13:55.180 --> 00:13:57.210
I was part of the math department

315
00:13:57.210 --> 00:13:59.310
but part of a program that was truly interdisciplinary.

316
00:13:59.310 --> 00:14:03.690
So, in our PhD program

317
00:14:03.690 --> 00:14:05.380
we have to have three qualifiers.

318
00:14:05.380 --> 00:14:09.499
And one of my qualifiers was in physical oceanography

319
00:14:09.499 --> 00:14:12.790
specifically El Nino and Southern oscillation.

320
00:14:12.790 --> 00:14:15.633
And you were forced to get out of your comfort zone.

321
00:14:16.561 --> 00:14:20.030
You're forced to get out into other disciplines.

322
00:14:20.030 --> 00:14:22.050
And because of that,

323
00:14:22.050 --> 00:14:24.580
I think one of the reasons I think

324
00:14:24.580 --> 00:14:26.489
that I really, it worked for me.

325
00:14:26.489 --> 00:14:31.489

in the traditional programs weren't filling this need,

326
00:14:33.589 --> 00:14:35.730
is what do you do for people of us

327
00:14:35.730 --> 00:14:37.880
that are in the intersection of things

328
00:14:37.880 --> 00:14:40.619
for those who remember many people came

329
00:14:40.619 --> 00:14:43.240
and started their career as much later with,

330
00:14:43.240 --> 00:14:46.360
that point of time around right

331
00:14:46.360 --> 00:14:48.960
when the human genome project was going

332
00:14:48.960 --> 00:14:50.200
there was this whole thing of like, well,

333
00:14:50.200 --> 00:14:52.153
what do you do if kind of sit in the intersection

334
00:14:52.153 --> 00:14:54.587
if you're working on some sequencing algorithm

335
00:14:54.587 --> 00:14:56.405
where you're computer scientist

336
00:14:56.405 --> 00:14:59.840
but you're not a geneticist,

337
00:14:59.840 --> 00:15:01.810
or you're sort of in this bullet

338
00:15:01.810 --> 00:15:04.510
but you're also both like, where do you fit?

339
00:15:04.510 --> 00:15:08.980
And how do you apply for funding sitting in these things?

340
00:15:08.980 --> 00:15:10.606
You can't get, if you're an academic,

341
00:15:10.606 --> 00:15:12.780
a true traditional academic,

342
00:15:12.780 --> 00:15:14.603
you can't get tenure in either department

343
00:15:14.603 --> 00:15:16.776
because they're like you don't fit you,

344
00:15:16.776 --> 00:15:19.050
you're not one of us.

345
00:15:19.050 --> 00:15:21.531
And then I think what we started to see is

346
00:15:21.531 --> 00:15:25.530
some of these more industry programs starting to move

347
00:15:25.530 --> 00:15:27.260
was people are like, well,

348
00:15:27.260 --> 00:15:29.080
what is this mean to be interdisciplinary?

349
00:15:29.080 --> 00:15:32.500
What does it mean to be multidisciplinary?

350
00:15:32.500 --> 00:15:35.532

And I think the power of data science is

351
00:15:35.532 --> 00:15:38.133
it's not supposed to fit.

352
00:15:39.140 --> 00:15:41.140
It's supposed to actually kind of cross,

353
00:15:42.290 --> 00:15:44.640
it gives flexibility.

354
00:15:44.640 --> 00:15:48.860
And maybe this is a very long winded way

355
00:15:48.860 --> 00:15:50.704
of answering your question,

356
00:15:50.704 --> 00:15:52.080
but maybe just kind of frame this as

357
00:15:52.080 --> 00:15:55.350
why did we come up with the term data scientist?

358
00:15:55.350 --> 00:15:57.960
And I think the story is actually helpful.

359
00:15:57.960 --> 00:15:59.450
So we weren't like

360
00:15:59.450 --> 00:16:02.140
we didn't have this like massive grand plan of like

361
00:16:02.140 --> 00:16:03.870
we're going to go name a new field.

362
00:16:03.870 --> 00:16:06.340
And then one day there's going to be degrees

363
00:16:06.340 --> 00:16:08.260
and, it's going to be amazing.

364
00:16:08.260 --> 00:16:09.640
And there'll be a chief data scientist.

365
00:16:09.640 --> 00:16:11.460
That was not the plan.

366
00:16:11.460 --> 00:16:16.460
The plan literally was how do we get HR off our backs?

367
00:16:16.650 --> 00:16:18.510
And so LinkedIn and both Facebook

368
00:16:18.510 --> 00:16:20.102
were heading towards the IPO

369
00:16:20.102 --> 00:16:23.090
Jeff Hammerbacher and I would use to get together

370
00:16:23.090 --> 00:16:25.252
and have lunch every so often and just compare notes.

371
00:16:25.252 --> 00:16:27.967
And some people may find that strange,

372
00:16:27.967 --> 00:16:30.770
like why LinkedIn and Facebook competitors,

373
00:16:30.770 --> 00:16:32.790
but we'd share technology.

374
00:16:32.790 --> 00:16:35.290
It's no different than like if we're researchers, right?

375
00:16:35.290 --> 00:16:39.078

We're born water Harbor floats all boats kind of argument.

376
00:16:39.078 --> 00:16:42.144
And as we're trading ideas, one of the things that came up

377
00:16:42.144 --> 00:16:44.660
is I said, hey, you know what?

378
00:16:44.660 --> 00:16:47.570
What are you starting to call yourself?

379
00:16:47.570 --> 00:16:48.783
Because we have this problem.

380
00:16:48.783 --> 00:16:51.969
Like HR came to us and said, look, you've got,

381
00:16:51.969 --> 00:16:55.019
all these titles, business analysts

382
00:16:55.019 --> 00:16:57.830
data scientists, statistician, and all of these.

383
00:16:57.830 --> 00:17:00.510
So like the engineers, they're engineers.

384
00:17:00.510 --> 00:17:03.405
Sales is sales, designers is designers,

385
00:17:03.405 --> 00:17:04.238
what are you all?

386
00:17:04.238 --> 00:17:06.100
And we're like, okay, well, no.

387
00:17:06.100 --> 00:17:07.250
And so we compared notes

388
00:17:07.250 --> 00:17:09.380
and literally we kind of went through

389
00:17:09.380 --> 00:17:12.670
all of them were like, huh, well, which ones,

390
00:17:12.670 --> 00:17:16.695
like analysts sounded too business oriented to wall street.

391
00:17:16.695 --> 00:17:19.600
Statistician was going to off the economists

392
00:17:19.600 --> 00:17:21.270
and economists are going to off the statistician.

393
00:17:21.270 --> 00:17:22.990
So we were like, we don't want that war.

394
00:17:22.990 --> 00:17:26.090
And then we just kind of went through

395
00:17:26.090 --> 00:17:28.260
and at the end, we were like, well

396
00:17:28.260 --> 00:17:31.239
data scientists seems to be the one that kind of gravitated

397
00:17:31.239 --> 00:17:33.640
but because we were LinkedIn,

398
00:17:33.640 --> 00:17:37.710
we actually posted all of our job applications,

399
00:17:37.710 --> 00:17:39.496
just for the LinkedIn corporation

400
00:17:39.496 --> 00:17:42.220

using the term data scienctist

401
00:17:42.220 --> 00:17:44.080
and all the others as well.

402
00:17:44.080 --> 00:17:47.196
And we saw which ones people would apply to and who we hired

403
00:17:47.196 --> 00:17:51.323
and almost came through the data science pipeline

404
00:17:51.323 --> 00:17:54.472
and which was a little surprising.

405
00:17:54.472 --> 00:17:56.297
'Cause we were like data scientists.

406
00:17:56.297 --> 00:17:58.250
Isn't that like repetitive, isn't that like,

407
00:17:58.250 --> 00:18:01.630
why wouldn't we want to call ourselves that?

408
00:18:01.630 --> 00:18:04.040
But we literally data science our way into it.

409
00:18:04.040 --> 00:18:07.640
And the reason I think it's taken off is honestly

410
00:18:07.640 --> 00:18:11.030
because I don't think people know what the hell it means.

411
00:18:11.030 --> 00:18:13.910
And think about that for a second is

412
00:18:13.910 --> 00:18:16.950
if you go into a meeting, especially in industry

413
00:18:16.950 --> 00:18:19.830
if you go into a meeting and somebody is like a big sort

414
00:18:19.830 --> 00:18:21.840
of proverbial cable and it runs there

415
00:18:21.840 --> 00:18:24.230
and someone says, so which group are you with?

416
00:18:24.230 --> 00:18:26.492
And you're like, they're like,

417
00:18:26.492 --> 00:18:28.260
I'm with business analysis or something.

418
00:18:28.260 --> 00:18:30.510
And then I'm like, why are you in this space?

419
00:18:31.540 --> 00:18:34.270
Or if you're a statistician, like, why are you here?

420
00:18:34.270 --> 00:18:37.170
This is the meeting for the real meeting.

421
00:18:37.170 --> 00:18:38.300
But if you're the data scientists,

422
00:18:38.300 --> 00:18:39.590
everyone was like oh good,

423
00:18:39.590 --> 00:18:42.130
we need some smart people around this room.

424
00:18:42.130 --> 00:18:45.570
And it's because it has this ambiguity

425
00:18:45.570 --> 00:18:48.054

and because it allows it to touch all these things

426
00:18:48.054 --> 00:18:48.887
that allows you to be in meetings

427
00:18:48.887 --> 00:18:52.338
where you're able to get context.

428
00:18:52.338 --> 00:18:57.338
And when you have that context, you now can use your skills.

429
00:18:58.270 --> 00:19:02.430
So, how does this come back to your question?

430
00:19:02.430 --> 00:19:05.893
Is the training I think that is most essential,

431
00:19:05.893 --> 00:19:08.510
fundamentally is a couple of things.

432
00:19:08.510 --> 00:19:11.603
One it's passion and curiosity.

433
00:19:12.480 --> 00:19:14.490
It's actually, honestly, it's not data.

434
00:19:14.490 --> 00:19:16.640
It's kind of like one on one for all of us

435
00:19:16.640 --> 00:19:18.270
and like how we got into our fields,

436
00:19:18.270 --> 00:19:20.050
passion and curiosity.

437
00:19:20.050 --> 00:19:22.500
Then, it's like, how do you exercise that passion?

438
00:19:22.500 --> 00:19:26.190
And your curiosity is through the use of computational

439
00:19:27.978 --> 00:19:29.007
and technical means some of which is statistics

440
00:19:29.007 --> 00:19:31.720
some of which is computational,

441
00:19:31.720 --> 00:19:33.218
but a lot of it honestly,

442
00:19:33.218 --> 00:19:36.630
is being able to be clever,

443
00:19:36.630 --> 00:19:40.670
to go find a data set or a way to use a dataset.

444
00:19:40.670 --> 00:19:43.100
So that, or bring multiple data sets together

445
00:19:43.100 --> 00:19:45.470
in a way that somebody, huh,

446
00:19:45.470 --> 00:19:48.220
I didn't think that we could actually do that.

447
00:19:48.220 --> 00:19:49.900
That's weird.

448
00:19:49.900 --> 00:19:51.480
Oh, what if we brought this data in?

449
00:19:51.480 --> 00:19:54.257
Or what if we go and found a way to get this other insight?

450
00:19:54.257 --> 00:19:56.250

Because otherwise we're always looking

451
00:19:56.250 --> 00:19:58.820
for something pristine and it's data.

452
00:19:58.820 --> 00:20:00.254
And by the time we collect it,

453
00:20:00.254 --> 00:20:04.140
it's too expensive and time has taken too long.

454
00:20:04.140 --> 00:20:07.299
And so those pieces, like where can you get those skills?

455
00:20:07.299 --> 00:20:10.890
Well, do you have to have a PhD?

456
00:20:10.890 --> 00:20:12.195
I don't think so personally,

457
00:20:12.195 --> 00:20:15.197
I've met lots of people who don't have PhDs

458
00:20:15.197 --> 00:20:18.866
or even college degrees who are phenomenal at data

459
00:20:18.866 --> 00:20:21.870
but what they have found in their ability to do,

460
00:20:21.870 --> 00:20:25.460
is they found different projects

461
00:20:25.460 --> 00:20:28.940
different mentors from lots of different areas

462
00:20:28.940 --> 00:20:31.470
where they're able to kind of cut across

463
00:20:31.470 --> 00:20:35.950
and be this crucible for ideas or this instigator

464
00:20:35.950 --> 00:20:38.880
for conversations on the team that allow us to

465
00:20:38.880 --> 00:20:42.550
kind of go, well, what is it if we did this?

466
00:20:42.550 --> 00:20:44.610
How bout if we brought this together

467
00:20:44.610 --> 00:20:47.390
or, well, what if we just use this computational?

468
00:20:47.390 --> 00:20:52.110
Or what if we use these new graphic processing units,

469
00:20:54.650 --> 00:20:58.840
the GPU's to do this instead of something else.

470
00:20:58.840 --> 00:21:00.930
That's the training.

471
00:21:00.930 --> 00:21:04.397
That's the place where you get that.

472
00:21:04.397 --> 00:21:07.670
And so I think we're starting to get formalized training,

473
00:21:07.670 --> 00:21:10.670
but I don't think you really,

474
00:21:10.670 --> 00:21:12.770
even in the formalized training where I tell people is like

475
00:21:12.770 --> 00:21:15.500

go get multiple hats.

476
00:21:15.500 --> 00:21:16.660
Get the multiple hats

477
00:21:16.660 --> 00:21:19.063
and then you will be hyper effective.

478
00:21:22.050 --> 00:21:23.633
<v ->Thank you very much.</v>

479
00:21:23.633 --> 00:21:27.430
What you just said actually encapsulates

480
00:21:27.430 --> 00:21:30.320
one of the reasons we're having the series of talks

481
00:21:30.320 --> 00:21:33.960
is because we recognize that maybe our much

482
00:21:33.960 --> 00:21:38.960
of our future of data science workforce and neurobiology

483
00:21:39.013 --> 00:21:41.390
right now is sitting in a place

484
00:21:41.390 --> 00:21:43.260
where they don't recognize themselves

485
00:21:43.260 --> 00:21:45.800
as being able to do that kind of work.

486
00:21:45.800 --> 00:21:50.380
And so we truly do see the role of the data scientists

487
00:21:50.380 --> 00:21:53.425
in addiction research as someone who is transdisciplinary

488
00:21:53.425 --> 00:21:56.260
and brings multiple perspectives

489
00:21:56.260 --> 00:22:01.260
and non-traditional research skills to these problems.

490
00:22:01.500 --> 00:22:03.500
So thank you for those comments.

491
00:22:03.500 --> 00:22:06.710
<v ->Roger if I have one ask of every data scientist,</v>

492
00:22:06.710 --> 00:22:10.580
And this is definitely true at the actually every level

493
00:22:10.580 --> 00:22:12.570
very true at the federal level

494
00:22:12.570 --> 00:22:15.230
was get out of our proverbial offices,

495
00:22:15.230 --> 00:22:17.883
when we can meet again in person and do things,

496
00:22:19.286 --> 00:22:20.119
get out of your office

497
00:22:20.119 --> 00:22:22.870
and go have coffee or lunch with other people.

498
00:22:22.870 --> 00:22:24.682
One of the ones that we used to set up,

499
00:22:24.682 --> 00:22:26.840
and there's no reason

500
00:22:26.840 --> 00:22:29.000

the chief data sciences office ran this,

501
00:22:29.000 --> 00:22:30.850
but it doesn't have to be that way

502
00:22:30.850 --> 00:22:35.270
is every Tuesday there used to be a meeting at noon

503
00:22:35.270 --> 00:22:38.933
where anybody could dial in or come down to GSA

504
00:22:40.096 --> 00:22:42.529
and hang out with other data scientists

505
00:22:42.529 --> 00:22:45.260
and share what they're doing.

506
00:22:45.260 --> 00:22:49.820
And if you can just hang out and share and do stuff,

507
00:22:49.820 --> 00:22:51.090
we used to have a rule

508
00:22:51.090 --> 00:22:52.900
and the teams that work together

509
00:22:52.900 --> 00:22:54.600
you can't be more than four feet.

510
00:22:54.600 --> 00:22:56.988
Your desk had to be four feet from somebody else's.

511
00:22:56.988 --> 00:23:00.600
And the reason is we're just trying to collaborate

512
00:23:00.600 --> 00:23:02.870
and pass information back and forth.

513
00:23:02.870 --> 00:23:05.313
It's some of the most amazing ideas

514
00:23:05.313 --> 00:23:09.160
that were off the wall harebrained,

515
00:23:09.160 --> 00:23:11.270
kind of came from those processes

516
00:23:11.270 --> 00:23:12.800
where you're just like, wow,

517
00:23:12.800 --> 00:23:14.403
I never thought about using that.

518
00:23:14.403 --> 00:23:16.188
Well, what if we did that?

519
00:23:16.188 --> 00:23:17.687
What about those kinds of ideas?

520
00:23:17.687 --> 00:23:20.330
And I think we've all experienced a version of that

521
00:23:20.330 --> 00:23:22.190
especially those that have gone to grad school.

522
00:23:22.190 --> 00:23:24.850
And we're kind of in like the crappy, like labs

523
00:23:24.850 --> 00:23:28.350
or apartments where you're desperate kind of forcing that.

524
00:23:29.222 --> 00:23:30.452
And you were just like, you were just,

525
00:23:30.452 --> 00:23:31.440

you learn to rapidly through,

526
00:23:31.440 --> 00:23:33.440
to kind of looking over somebody's shoulder.

527
00:23:33.440 --> 00:23:35.015
Somebody asked me a question.

528
00:23:35.015 --> 00:23:39.090
The place and one of the discerning things

529
00:23:39.090 --> 00:23:40.958
that I can almost always

530
00:23:40.958 --> 00:23:42.365
tell that goes wrong.

531
00:23:42.365 --> 00:23:46.390
There's two things that I would tell at least

532
00:23:46.390 --> 00:23:50.190
with interviewing people or watching people kind

533
00:23:50.190 --> 00:23:53.580
of get slotted into roles is predictors.

534
00:23:53.580 --> 00:23:56.920
One is if I'm interviewing you,

535
00:23:56.920 --> 00:23:58.699
like easiest question I asked is,

536
00:23:58.699 --> 00:24:00.990
great, here's a dataset.

537
00:24:00.990 --> 00:24:03.310
What would you be interested in?

538
00:24:03.310 --> 00:24:05.790
Like if it was, this was LinkedIn, I'd say, hey, great.

539
00:24:05.790 --> 00:24:07.790
Here's all of LinkedIn's data, you have it.

540
00:24:07.790 --> 00:24:10.670
What's the first question that you're interested in.

541
00:24:10.670 --> 00:24:12.430
And then what's the second question?

542
00:24:12.430 --> 00:24:15.834
And usually what people who are really have nailed this,

543
00:24:15.834 --> 00:24:20.070
are like, in the 90 seconds

544
00:24:20.070 --> 00:24:23.100
they have come up with like 15 hypothesis.

545
00:24:23.100 --> 00:24:24.650
They would like to test.

546
00:24:24.650 --> 00:24:26.400
They're just like, oh, I'd have this.

547
00:24:26.400 --> 00:24:30.460
If you had like, suppose you had FRMI data

548
00:24:30.460 --> 00:24:34.537
of every scan over the past 20 years,

549
00:24:41.810 --> 00:24:44.449
what would you do with that data?

550
00:24:44.449 --> 00:24:49.080

And if you're kind of like takes you more

551
00:24:50.110 --> 00:24:52.778
than 10 seconds, like, what would it take

552
00:24:52.778 --> 00:24:54.960
for you to come up with like faster answers

553
00:24:54.960 --> 00:24:57.150
all those things, because you want all that.

554
00:24:57.150 --> 00:24:59.267
The second is ask people,

555
00:24:59.267 --> 00:25:01.773
how do you go about,

556
00:25:01.773 --> 00:25:04.497
what's your model for actually working on a problem?

557
00:25:04.497 --> 00:25:09.497
And if your first inclination isn't to ask somebody else

558
00:25:11.423 --> 00:25:14.470
then that that's the reframing.

559
00:25:14.470 --> 00:25:16.250
It needs to be asked somebody else

560
00:25:16.250 --> 00:25:18.644
because everybody can do deep research.

561
00:25:18.644 --> 00:25:21.210
And I don't want to say to anybody out there,

562
00:25:21.210 --> 00:25:22.953
like, don't do deep research.

563
00:25:24.172 --> 00:25:26.050
Don't be like doing the deep skull

564
00:25:26.050 --> 00:25:28.320
or the work or investigation.

565
00:25:28.320 --> 00:25:32.683
But what I really want to emphasize is efficiency of work.

566
00:25:33.570 --> 00:25:36.673
And so a lot of times, instead of just going in,

567
00:25:36.673 --> 00:25:39.410
sitting in the library and doing the stuff,

568
00:25:39.410 --> 00:25:42.500
first ask around if anybody else has done this,

569
00:25:42.500 --> 00:25:44.000
if anybody else has ideas

570
00:25:44.000 --> 00:25:48.730
and then use that as your mechanism into the project

571
00:25:48.730 --> 00:25:50.580
so that you go, you're scalable

572
00:25:50.580 --> 00:25:52.180
you're more efficient at getting

573
00:25:53.189 --> 00:25:55.108
to the problem because otherwise there's too many dead ends.

574
00:25:55.108 --> 00:25:58.760
And the computationally that chews up way too

575
00:25:58.760 --> 00:26:01.253

much dollars in time, as you lay code down.

576
00:26:03.710 --> 00:26:05.250
<v ->Thank you.</v>

577
00:26:05.250 --> 00:26:07.660
Albert, I believe you're up next.

578
00:26:07.660 --> 00:26:09.455
<v ->Hi DJ, thanks for that.</v>

579
00:26:09.455 --> 00:26:13.230
So I have a sort of a two-parter question.

580
00:26:13.230 --> 00:26:14.810
You kind of address the first one

581
00:26:14.810 --> 00:26:16.980
but we still want to kind of ask it.

582
00:26:16.980 --> 00:26:20.820
So what attracted you to data science

583
00:26:20.820 --> 00:26:24.760
what attracted you to chief of data science,

584
00:26:24.760 --> 00:26:27.600
but also, can you unpack a little bit, you touched upon this

585
00:26:27.600 --> 00:26:30.430
like data science that were just sort of getting trendy

586
00:26:30.430 --> 00:26:33.020
but a lot of young people who are watching you now,

587
00:26:33.020 --> 00:26:34.714
or may watch this video later

588
00:26:34.714 --> 00:26:36.971
after they wake up on the West

589
00:26:36.971 --> 00:26:39.740
they don't really know what data science is.

590
00:26:39.740 --> 00:26:42.665
We are talking about it like everybody understands.

591
00:26:42.665 --> 00:26:44.380
So can you talk from a little bit more

592
00:26:44.380 --> 00:26:45.420
of a practical standpoint?

593
00:26:45.420 --> 00:26:46.990
So what attracted you into it,

594
00:26:46.990 --> 00:26:48.260
unpack a little bit more

595
00:26:48.260 --> 00:26:50.100
of what that means to analyze data,

596
00:26:50.100 --> 00:26:52.813
what does that mean to be a data scientist?

597
00:26:54.090 --> 00:26:55.060
<v ->Yeah, I think,</v>

598
00:26:55.060 --> 00:27:00.060
the first is what is the idea of data science?

599
00:27:00.813 --> 00:27:03.740
I think it is fundamentally,

600
00:27:03.740 --> 00:27:08.250

how do we use data to gain insights

601
00:27:08.250 --> 00:27:11.920
or how do we use data in a way to build something?

602
00:27:11.920 --> 00:27:14.033
So one of the things that I think is there is,

603
00:27:14.033 --> 00:27:19.013
that is the most interesting aspect of this nowadays is

604
00:27:19.013 --> 00:27:22.900
that we're using data actually to make products.

605
00:27:22.900 --> 00:27:25.310
So an easy one to think of is like,

606
00:27:25.310 --> 00:27:27.380
okay, if you go to a social media site

607
00:27:27.380 --> 00:27:28.890
like LinkedIn or something else

608
00:27:28.890 --> 00:27:31.780
and you kind of build out your profile

609
00:27:31.780 --> 00:27:35.160
it can tell you like, hey, here's interesting jobs,

610
00:27:35.160 --> 00:27:38.050
those job recommenders, that's a data product.

611
00:27:38.050 --> 00:27:40.410
If you go look on an e-commerce site and says

612
00:27:40.410 --> 00:27:42.169
people who bought this also bought this

613
00:27:42.169 --> 00:27:43.568
that's a data product,

614
00:27:43.568 --> 00:27:48.430
but on another data product is a self-driving car.

615
00:27:48.430 --> 00:27:50.270
Autonomous vehicles are a data product

616
00:27:50.270 --> 00:27:53.530
because it takes in all of this information

617
00:27:53.530 --> 00:27:56.632
it processes it and it makes decisions.

618
00:27:56.632 --> 00:28:00.210
But also those are data products that could help

619
00:28:00.210 --> 00:28:02.372
somebody else make a decision.

620
00:28:02.372 --> 00:28:04.640
I think of a data product,

621
00:28:04.640 --> 00:28:06.650
if you're looking at something of a brain scan

622
00:28:06.650 --> 00:28:09.500
or something and it highlights for interesting areas

623
00:28:09.500 --> 00:28:12.700
for you to look at from a diagnostic perspective

624
00:28:12.700 --> 00:28:15.500
or something else, that's a data product.

625
00:28:15.500 --> 00:28:20.500

How does that actual, the tangible experience

626
00:28:21.664 --> 00:28:24.770
of looking at something,

627
00:28:24.770 --> 00:28:25.920
being able to do something

628
00:28:25.920 --> 00:28:27.760
and how do you make a decision out it?

629
00:28:27.760 --> 00:28:29.940
How does it team change a course direction?

630
00:28:29.940 --> 00:28:32.010
One of the data products that we see every day

631
00:28:32.010 --> 00:28:34.730
that we often don't give a lot of credence to

632
00:28:34.730 --> 00:28:38.660
or a thought into is our weather forecast.

633
00:28:38.660 --> 00:28:40.080
And this is one of the important things,

634
00:28:40.080 --> 00:28:41.320
is these data products.

635
00:28:41.320 --> 00:28:45.600
Oftentimes don't have to give you a lot of data back.

636
00:28:45.600 --> 00:28:47.530
They can give you just a little icon

637
00:28:47.530 --> 00:28:49.610
like today it's like, hey, there's a sun

638
00:28:49.610 --> 00:28:51.910
and it's really nice and happy.

639
00:28:51.910 --> 00:28:53.530
And it's like, oh wow, okay.

640
00:28:53.530 --> 00:28:55.240
That's like, I don't need to take the umbrella.

641
00:28:55.240 --> 00:28:56.840
Or there's like a dark and stormy cloud

642
00:28:56.840 --> 00:28:59.020
that you don't need the detailed data.

643
00:28:59.020 --> 00:29:01.040
Now, if you're a pilot maybe you need a data,

644
00:29:01.040 --> 00:29:02.220
different data product

645
00:29:02.220 --> 00:29:04.951
because you need different levels of sophistication.

646
00:29:04.951 --> 00:29:08.590
And so, who builds that?

647
00:29:08.590 --> 00:29:10.733
Who makes that come alive?

648
00:29:11.860 --> 00:29:13.850
That's the data scientist.

649
00:29:13.850 --> 00:29:16.436
The data scientist is a person who actually,

650
00:29:16.436 --> 00:29:17.965

is the person who says, hey,

651
00:29:17.965 --> 00:29:22.965
what if we brought these things together and kind of dizzy

652
00:29:23.820 --> 00:29:25.530
and maybe this is helpful to tell the story

653
00:29:25.530 --> 00:29:26.500
of people you may know,

654
00:29:26.500 --> 00:29:27.520
this is this idea I think

655
00:29:27.520 --> 00:29:29.020
that we've all seen on social media things.

656
00:29:29.020 --> 00:29:30.242
This kind of like,

657
00:29:30.242 --> 00:29:32.835
you go on a thing and you sign up

658
00:29:32.835 --> 00:29:34.280
and it's like, hey, did you know

659
00:29:34.280 --> 00:29:36.242
these are the other people that are actually here?

660
00:29:36.242 --> 00:29:40.368
So the story of that actually is,

661
00:29:40.368 --> 00:29:45.368
another person who grew up in Bethesda, Jonathan Goldman.

662
00:29:45.508 --> 00:29:48.500
And Jonathan had this idea at,

663
00:29:48.500 --> 00:29:50.120
he's a theoretical physicist.

664
00:29:50.120 --> 00:29:53.763
And he was data scientist at LinkedIn, early person there.

665
00:29:53.763 --> 00:29:56.690
And, he went around and he said, hey,

666
00:29:56.690 --> 00:29:58.453
when people sign up for LinkedIn

667
00:29:58.453 --> 00:30:02.501
how do people supposed to know who to connect to?

668
00:30:02.501 --> 00:30:03.334
And we're going to said,

669
00:30:03.334 --> 00:30:04.990
they'll just upload their address book.

670
00:30:04.990 --> 00:30:06.647
And then we'll tell them, it's that's,

671
00:30:06.647 --> 00:30:09.382
if you just show up to the site for the first time,

672
00:30:09.382 --> 00:30:11.810
you really want to give it, give up your whole address book.

673
00:30:11.810 --> 00:30:14.270
That's a lot to ask for.

674
00:30:14.270 --> 00:30:17.020
And the product teams said, okay, Jonathan

675
00:30:17.020 --> 00:30:20.330

you and your harebrained ideas, whatever.

676
00:30:20.330 --> 00:30:24.230
Like, and so Jonathan, like being a good data scientist

677
00:30:24.230 --> 00:30:27.101
he didn't just kind of say, okay, what he did is he went

678
00:30:27.101 --> 00:30:32.101
and actually got access to all the LinkedIn information.

679
00:30:32.470 --> 00:30:34.670
And then he built a bunch of ads

680
00:30:34.670 --> 00:30:36.940
like literal ads that would go on

681
00:30:36.940 --> 00:30:40.350
your page when you first logged in.

682
00:30:40.350 --> 00:30:43.742
And they just used a set of very simple heuristic.

683
00:30:43.742 --> 00:30:46.970
You triangle closing.

684
00:30:46.970 --> 00:30:50.388
If I know Susan and Albert you know Susan,

685
00:30:50.388 --> 00:30:52.700
maybe there's a chance we know each other.

686
00:30:52.700 --> 00:30:54.440
It's like triangle closing,

687
00:30:54.440 --> 00:30:56.445
then oh, did you overlapping jobs?

688
00:30:56.445 --> 00:30:58.170
Did you overlap in schools?

689
00:30:58.170 --> 00:30:59.967
Where do you get those questions from?

690
00:30:59.967 --> 00:31:02.670
It's the same questions we ask

691
00:31:02.670 --> 00:31:03.503
if we were running into each other at a barbecue.

692
00:31:03.503 --> 00:31:07.270
So how did you get here?

693
00:31:07.270 --> 00:31:08.450
Who do you know here?

694
00:31:08.450 --> 00:31:09.490
Like what, like,

695
00:31:09.490 --> 00:31:12.610
these are the things we're just sort of curistics,

696
00:31:12.610 --> 00:31:14.010
no AI, nothing fancy.

697
00:31:14.010 --> 00:31:15.583
And then he put them up,

698
00:31:16.770 --> 00:31:17.780
he just kind of gave these as a static ad,

699
00:31:17.780 --> 00:31:19.410
nothing sophisticated.

700
00:31:19.410 --> 00:31:21.660

And then the next morning, everyone was like,

701
00:31:21.660 --> 00:31:23.490
what is going on the site?

702
00:31:23.490 --> 00:31:27.070
Like we've never seen this kind of engagement or activity.

703
00:31:27.070 --> 00:31:29.680
And people were like this is crazy

704
00:31:29.680 --> 00:31:32.323
and well, like what happened?

705
00:31:33.970 --> 00:31:35.494
And then finally I was like, did Jonathan do something?

706
00:31:35.494 --> 00:31:36.590
Jonathan's like, I just put this experiment out there.

707
00:31:36.590 --> 00:31:38.000
And everyone's like, this is amazing.

708
00:31:38.000 --> 00:31:40.482
And what happened is after that

709
00:31:40.482 --> 00:31:45.310
is the entire social networking landscape.

710
00:31:45.310 --> 00:31:47.610
All the companies realize like, oh my gosh,

711
00:31:47.610 --> 00:31:48.890
this is the answer.

712
00:31:48.890 --> 00:31:49.830
This is the thing.

713
00:31:49.830 --> 00:31:52.144
And so every company then raised to actually implement this.

714
00:31:52.144 --> 00:31:55.220
And then of course they became very sophisticated algorithms

715
00:31:55.220 --> 00:31:56.070
and other things.

716
00:31:58.655 --> 00:32:00.680
But the point I'm trying to make is one data scientist

717
00:32:01.628 --> 00:32:02.840
with them, an idea changed the trajectory

718
00:32:02.840 --> 00:32:04.850
of the whole thing industry.

719
00:32:04.850 --> 00:32:07.510
And we've seen that time and time again.

720
00:32:07.510 --> 00:32:11.090
And, that person with the eScience now

721
00:32:11.090 --> 00:32:12.450
there's a similar version

722
00:32:12.450 --> 00:32:14.250
of the data scientists who comes in.

723
00:32:14.250 --> 00:32:17.630
So with that remarkable insight that says, you know

724
00:32:17.630 --> 00:32:19.750
something like akin to, you know

725
00:32:19.750 --> 00:32:23.240

it's sticking on the social media kind of subject

726
00:32:23.240 --> 00:32:27.680
is this kind of idea of, if you get a number

727
00:32:27.680 --> 00:32:31.421
of X number of people to connect the probability

728
00:32:31.421 --> 00:32:35.890
of sticking around, goes up dramatically.

729
00:32:35.890 --> 00:32:38.560
Versions of this as on Netflix is,

730
00:32:38.560 --> 00:32:42.870
if a person checks like five movies

731
00:32:42.870 --> 00:32:44.210
or shows that they're interested

732
00:32:44.210 --> 00:32:47.750
in watching the retention is going to increase.

733
00:32:47.750 --> 00:32:52.750
And those like insights have been extraordinarily powerful.

734
00:32:52.870 --> 00:32:54.590
And I also, in many cases

735
00:32:54.590 --> 00:32:56.490
I do need to highlight many of those

736
00:32:58.804 --> 00:32:59.854
who've been abused also.

737
00:32:59.854 --> 00:33:00.862
And that is one

738
00:33:00.862 --> 00:33:01.695
of the reckonings that we are having to do

739
00:33:01.695 --> 00:33:02.528
on social media right now.

740
00:33:03.764 --> 00:33:04.597
There's a flip version of this at

741
00:33:04.597 --> 00:33:06.340
for crisis Textline that we've seen

742
00:33:06.340 --> 00:33:08.517
that our data scientists figured out, which is if you're

743
00:33:08.517 --> 00:33:10.175
cause we do a lot of natural language processing

744
00:33:10.175 --> 00:33:14.420
on top of the text messaging and the Corpus

745
00:33:14.420 --> 00:33:17.550
what they realize that if you're in your first set

746
00:33:17.550 --> 00:33:21.100
of conversations with a person and you're texting back

747
00:33:21.100 --> 00:33:25.169
with a counselor and you use the word heard any

748
00:33:25.169 --> 00:33:28.390
or any pharmaceutical use that any pharmaceutical or Tylenol

749
00:33:28.390 --> 00:33:30.090
or thing you were eight times more

750
00:33:32.988 --> 00:33:35.130

likely to have actually taken the pharmaceutical already.

751
00:33:35.130 --> 00:33:36.243
You have already ODI.

752
00:33:37.480 --> 00:33:40.753
If you say, though, it might be actually

753
00:33:40.753 --> 00:33:42.270
that one might be the 16 times.

754
00:33:42.270 --> 00:33:43.350
I can't remember exactly.

755
00:33:43.350 --> 00:33:45.500
But then if you use the word, if you say

756
00:33:45.500 --> 00:33:49.600
that you have a gun or you're in the military, we know

757
00:33:49.600 --> 00:33:51.680
that that you're X more time is more likely to

758
00:33:51.680 --> 00:33:54.700
actually have that gun locked and ready next to you.

759
00:33:54.700 --> 00:33:56.000
So what do we do with that?

760
00:33:56.000 --> 00:33:57.590
We go, Oh, wow.

761
00:33:57.590 --> 00:34:01.040
Now, well, we have to change the way we prioritize the queue

762
00:34:01.040 --> 00:34:02.190
so that we say, all right

763
00:34:02.190 --> 00:34:05.860
we need to escalate this right now in take different action.

764
00:34:05.860 --> 00:34:06.870
And we highlight that

765
00:34:06.870 --> 00:34:09.688
to the counselor saying this is higher risk.

766
00:34:09.688 --> 00:34:14.570
And so that data center scientist who's come

767
00:34:14.570 --> 00:34:18.180
up with that novel insight has changed the paradigm

768
00:34:18.180 --> 00:34:19.860
of those things.

769
00:34:19.860 --> 00:34:24.760
What does it take to get there is playing,

770
00:34:24.760 --> 00:34:27.440
a lot of play with the data

771
00:34:27.440 --> 00:34:29.570
and kind of going, what if, what about the, yes.

772
00:34:29.570 --> 00:34:30.403
What about this?

773
00:34:32.155 --> 00:34:35.602
But when, when you have that nugget

774
00:34:35.602 --> 00:34:37.610
all of a sudden, it's like, Ooh, that's really interesting.

775
00:34:37.610 --> 00:34:40.410

And we know this all the time in science

776
00:34:40.410 --> 00:34:42.972
because we see a clever experiment.

777
00:34:42.972 --> 00:34:43.805
I'm like, dang, that was clever.

778
00:34:43.805 --> 00:34:46.970
What a way to get an insight on it?

779
00:34:46.970 --> 00:34:51.350
You know, the one that always strikes me is

780
00:34:52.658 --> 00:34:56.910
the early ones on plasticity

781
00:34:56.910 --> 00:35:01.910
where the experiments were done on Phantom limb phenomena

782
00:35:02.170 --> 00:35:05.210
with motorcyclists who had an accident

783
00:35:05.210 --> 00:35:07.710
and then had they still had their arm

784
00:35:07.710 --> 00:35:09.915
but the nerves had been ripped

785
00:35:09.915 --> 00:35:11.540
out effectively from the spinal column.

786
00:35:11.540 --> 00:35:15.250
And then they still have Phantom limb, but they, they that

787
00:35:16.486 --> 00:35:17.319
and how that mapping like, you know, itching

788
00:35:17.319 --> 00:35:19.950
on their forehead would alleviate the Phantom lineage.

789
00:35:19.950 --> 00:35:22.060
Like these kinds of things like that.

790
00:35:22.060 --> 00:35:24.950
Those kinds of like insights is what we're trying to go for.

791
00:35:24.950 --> 00:35:27.760
And so the way I try to emphasize this a lot

792
00:35:27.760 --> 00:35:32.350
of times is an early work is saying the phrase, clever

793
00:35:32.350 --> 00:35:34.383
be smart, nine times out of 10.

794
00:35:35.390 --> 00:35:38.470
Now smart is what want to apply when we're trying to

795
00:35:38.470 --> 00:35:42.520
scale something because of that first initial

796
00:35:42.520 --> 00:35:45.530
like let's start figuring out what the problem is great.

797
00:35:45.530 --> 00:35:47.588
Like the people who have an algorithm that's clever

798
00:35:47.588 --> 00:35:50.920
is that going to scale no way.

799
00:35:50.920 --> 00:35:53.950
Now we need to bring the heavy machinery in place.

800
00:35:53.950 --> 00:35:55.010

And depending on what kind

801
00:35:55.010 --> 00:35:57.120
of data scientists you want to be

802
00:35:57.120 --> 00:35:59.990
maybe you're a little bit more on the insight side.

803
00:35:59.990 --> 00:36:01.625
Maybe you're a little bit more on the

804
00:36:01.625 --> 00:36:04.370
the data scientists that wants to build something.

805
00:36:04.370 --> 00:36:05.820
Maybe you kind of flow in between.

806
00:36:05.820 --> 00:36:07.820
And now we're starting to see this emergence

807
00:36:07.820 --> 00:36:10.880
of people who are much more data engineering.

808
00:36:10.880 --> 00:36:12.500
They move data around.

809
00:36:12.500 --> 00:36:15.460
They build the infrastructure to help things happen.

810
00:36:15.460 --> 00:36:17.630
Like these event based processing cues

811
00:36:17.630 --> 00:36:18.790
the other different pieces.

812
00:36:18.790 --> 00:36:21.280
All of that I start thinking is starting to

813
00:36:21.280 --> 00:36:24.020
become the place where things are actually moving

814
00:36:24.020 --> 00:36:27.720
to in this more specialization and the data science programs

815
00:36:27.720 --> 00:36:30.653
especially at the undergraduate level,

816
00:36:32.252 --> 00:36:34.252
what they're trying to do is foster this

817
00:36:35.244 --> 00:36:37.640
multidisciplinary thing that sort of allows us to

818
00:36:37.640 --> 00:36:41.053
kind of go back and forth between kind of places.

819
00:36:44.360 --> 00:36:45.193
<v ->Thanks so much.</v>

820
00:36:45.193 --> 00:36:46.720
It sounds like it it's wide open.

821
00:36:46.720 --> 00:36:49.180
I mean, there's a lot that these, that, you know

822
00:36:49.180 --> 00:36:51.260
new people, young folks could do with it.

823
00:36:51.260 --> 00:36:53.690
So that's really exciting.

824
00:36:53.690 --> 00:36:54.523
I'm going to turn it over to Lindsay.

825
00:36:56.135 --> 00:36:57.447

Awesome, thank you.

826
00:36:57.447 --> 00:36:58.280
And thank you for being here.

827
00:36:58.280 --> 00:36:59.340
I wanted to switch gears a little bit

828
00:36:59.340 --> 00:37:02.930
and ask about your job as the U.S chief of data science.

829
00:37:02.930 --> 00:37:04.080
How did you land that job

830
00:37:04.080 --> 00:37:05.597
and kind of your experiences leading up to that?

831
00:37:05.597 --> 00:37:09.910
<v ->Yeah, so the story of that was,</v>

832
00:37:09.910 --> 00:37:13.390
I think which is historical, especially

833
00:37:13.390 --> 00:37:16.410
during not only the Obama presidency, but as we've seen

834
00:37:16.410 --> 00:37:21.410
in society is president Obama had really seen, I think

835
00:37:21.600 --> 00:37:22.790
through his campaigns.

836
00:37:22.790 --> 00:37:27.530
And then through research that was taking place

837
00:37:27.530 --> 00:37:30.340
and everything was the world is becoming to

838
00:37:30.340 --> 00:37:32.650
become increasingly reliant on data.

839
00:37:32.650 --> 00:37:34.720
You know, in his campaign, he was using data

840
00:37:34.720 --> 00:37:37.730
in very novel ways, but then you were starting to watch.

841
00:37:37.730 --> 00:37:41.000
Society is not only social media companies were coming up

842
00:37:41.000 --> 00:37:43.070
but remember this was a time period

843
00:37:43.070 --> 00:37:45.590
when maps were being replaced.

844
00:37:45.590 --> 00:37:47.390
Like we used to carry stacks of map.

845
00:37:47.390 --> 00:37:50.410
And then it was like, Oh no, what we had was MapQuest

846
00:37:50.410 --> 00:37:52.640
and other things where we then print out the map

847
00:37:52.640 --> 00:37:55.016
just directions for MapQuest and those things.

848
00:37:55.016 --> 00:37:58.150
And then it was being replaced by the mobile phone.

849
00:37:58.150 --> 00:38:01.140
You know, then you had this device and the maps were

850
00:38:01.140 --> 00:38:03.010

on there the same time.

851
00:38:03.010 --> 00:38:06.930
You were starting to see like companies like Uber and left

852
00:38:06.930 --> 00:38:08.900
and people being able to access things.

853
00:38:08.900 --> 00:38:12.530
So there was a shift fundamentally happening simultaneously

854
00:38:12.530 --> 00:38:15.920
in the creation of the U S CTO role, the role

855
00:38:15.920 --> 00:38:18.000
of the first two chief data scientist.

856
00:38:18.000 --> 00:38:20.630
First I need to Chopra and then taught part

857
00:38:20.630 --> 00:38:24.060
the fundamental focus was opening data and specifically

858
00:38:24.060 --> 00:38:25.620
for healthcare.

859
00:38:25.620 --> 00:38:28.100
And then as we saw that was happening with healthcare

860
00:38:28.100 --> 00:38:30.700
the question was like, well, what about research?

861
00:38:30.700 --> 00:38:33.220
What about all these other aspects?

862
00:38:33.220 --> 00:38:34.984
And then as Megan Smith, the next

863
00:38:34.984 --> 00:38:37.482
Chief technology officer, the next U.S CTO

864
00:38:37.482 --> 00:38:42.482
and her focus was her direction from

865
00:38:43.310 --> 00:38:44.870
the president was really

866
00:38:44.870 --> 00:38:47.812
about how do we actually make sure

867
00:38:47.812 --> 00:38:49.948
that we're using technology

868
00:38:49.948 --> 00:38:53.787
in a much more broad way education

869
00:38:53.787 --> 00:38:56.550
making sure we scale these systems

870
00:38:56.550 --> 00:38:59.250
inside government procurement, all these other things.

871
00:39:02.453 --> 00:39:04.490
But we couldn't lose out on the data science aspect

872
00:39:04.490 --> 00:39:06.670
who is going to make sure that that portfolio,

873
00:39:06.670 --> 00:39:08.500
that focus was there.

874
00:39:08.500 --> 00:39:11.070
And so president Obama said like

875
00:39:11.070 --> 00:39:14.291

there has to be somebody who's going to advise the president

876
00:39:14.291 --> 00:39:17.270
on how to make sure to use data responsible.

877
00:39:17.270 --> 00:39:20.982
And so the mission statement of the chief data scientist is

878
00:39:20.982 --> 00:39:25.450
has specifically the words responsible

879
00:39:25.450 --> 00:39:28.660
in there to ensure that there's responsibility

880
00:39:29.522 --> 00:39:32.700
of data and making sure it benefits all Americans

881
00:39:32.700 --> 00:39:34.150
and it's to actually the full

882
00:39:34.150 --> 00:39:36.680
of it is to unleash the power of data, to benefit

883
00:39:36.680 --> 00:39:38.380
to responsibly unleash the power of data

884
00:39:38.380 --> 00:39:40.490
to benefit all Americans.

885
00:39:40.490 --> 00:39:43.280
And so in that framing of it,

886
00:39:43.280 --> 00:39:44.113
they had asked me to say, hey

887
00:39:44.113 --> 00:39:48.240
would I come in and take on this role?

888
00:39:48.240 --> 00:39:51.150
And I actually spent my first time portion

889
00:39:51.150 --> 00:39:53.930
of coming into that role at NIH.

890
00:39:53.930 --> 00:39:55.840
I actually spent my first several weeks

891
00:39:55.840 --> 00:39:59.080
at NIH working with Francis

892
00:39:59.080 --> 00:40:01.180
and the team to really start thinking

893
00:40:01.180 --> 00:40:04.190
about how do we get precision medicine off the ground

894
00:40:04.190 --> 00:40:06.307
and then moved over fully to the white house to

895
00:40:06.307 --> 00:40:09.220
get actually to kind of say, okay, now that we've got

896
00:40:09.220 --> 00:40:12.030
NIH has kind of the core mechanism for this.

897
00:40:12.030 --> 00:40:16.420
How do we expand this across the federal sector?

898
00:40:16.420 --> 00:40:20.830
And so that's really kind of was the Genesis.

899
00:40:20.830 --> 00:40:23.460
And then similarly, a number of people may have seen

900
00:40:23.460 --> 00:40:25.480

on social media, we talked about,

901
00:40:25.480 --> 00:40:28.380
'cause we weren't trying to get a lot of attention for it

902
00:40:28.380 --> 00:40:33.380
but I was CTO and senior staff for the buying

903
00:40:33.550 --> 00:40:35.140
the Harris transition.

904
00:40:35.140 --> 00:40:37.880
And one of the things that you'll see in the a hundred day

905
00:40:37.880 --> 00:40:39.223
and 200 day plans

906
00:40:39.223 --> 00:40:43.511
is that heavy emphasis of how do we use technology

907
00:40:43.511 --> 00:40:45.470
through everything that includes

908
00:40:45.470 --> 00:40:49.580
the it monitorization efforts that have been put

909
00:40:49.580 --> 00:40:51.600
into the American recovery plan

910
00:40:51.600 --> 00:40:55.240
but also ensuring that we have surveillance technologies

911
00:40:55.240 --> 00:40:57.899
for genetic genomic testing,

912
00:40:57.899 --> 00:41:00.530
but also to continue to double down

913
00:41:00.530 --> 00:41:04.590
on our R and D portfolios, particularly through NIH.

914
00:41:04.590 --> 00:41:07.351
And now Eric Lander is going to be taking this on as

915
00:41:07.351 --> 00:41:11.169
as the head of OSTP, but you're going to see data

916
00:41:11.169 --> 00:41:15.707
and technology through web through all of this.

917
00:41:15.707 --> 00:41:17.980
And the how important is it Lindsey,

918
00:41:17.980 --> 00:41:20.830
is remember when we announced precision medicine

919
00:41:20.830 --> 00:41:23.270
initially the president Obama announced it,

920
00:41:23.270 --> 00:41:27.031
there was something like $20 million associated with it.

921
00:41:27.031 --> 00:41:29.936
And China followed up with $9 billion

922
00:41:29.936 --> 00:41:34.580
as their funding component.

923
00:41:34.580 --> 00:41:37.360
Now, thanks to 21st century cures

924
00:41:37.360 --> 00:41:41.140
and the work many of which was the heavy lifting was done

925
00:41:41.140 --> 00:41:45.180

by NIH and HHS efforts to get it through Congress,

926
00:41:45.180 --> 00:41:46.986
that moved the funding up to a billion.

927
00:41:46.986 --> 00:41:51.235
But we know that that the investment that is taking place

928
00:41:51.235 --> 00:41:53.990
and the reason why so many other countries

929
00:41:53.990 --> 00:41:55.197
want to go after this

930
00:41:55.197 --> 00:41:57.990
fundamentally is because they see

931
00:41:57.990 --> 00:42:01.790
the next great set of innovation.

932
00:42:01.790 --> 00:42:03.590
The next set of things that are coming,

933
00:42:03.590 --> 00:42:06.792
is the melding of large scale data sets.

934
00:42:06.792 --> 00:42:11.792
And also the physical and the wet labs.

935
00:42:12.530 --> 00:42:17.530
That combination is where the puck is going to be.

936
00:42:17.820 --> 00:42:20.170
Like, we're seeing it with the amount of genomics,

937
00:42:20.170 --> 00:42:21.950
we're seeing it with the vaccines

938
00:42:21.950 --> 00:42:24.420
and everything that we're learning with COVID,

939
00:42:24.420 --> 00:42:25.888
we're seeing it everywhere.

940
00:42:25.888 --> 00:42:28.859
And so the next great set of innovation is truly

941
00:42:28.859 --> 00:42:33.310
going to be this intersection of these places.

942
00:42:33.310 --> 00:42:36.080
And that means that we got to have this

943
00:42:36.080 --> 00:42:38.687
better hybridization of technology,

944
00:42:38.687 --> 00:42:41.750
the wet labs, the physical labs,

945
00:42:41.750 --> 00:42:46.390
because we have a lot of small blockers

946
00:42:46.390 --> 00:42:49.644
that are preventing us from making this a lot easier

947
00:42:49.644 --> 00:42:53.090
for everybody right now is too expensive,

948
00:42:53.090 --> 00:42:56.930
it is too hard for people to deploy technology

949
00:42:57.890 --> 00:42:59.390
on these problems.

950
00:43:03.934 --> 00:43:04.957

And it's hard to get data,

951
00:43:04.957 --> 00:43:05.975
it's still way too hard to get data

952
00:43:05.975 --> 00:43:07.060
in that way from electronic medical records,

953
00:43:07.060 --> 00:43:08.720
to all the other places

954
00:43:09.825 --> 00:43:11.570
that is still not an acceptable place.

955
00:43:11.570 --> 00:43:13.450
And we're seeing the aspects right now

956
00:43:13.450 --> 00:43:18.308
with clinical trials.gov to COVID

957
00:43:18.308 --> 00:43:21.960
to so many other diseases where genomic diseases,

958
00:43:21.960 --> 00:43:23.320
other areas.

959
00:43:23.320 --> 00:43:24.360
And so it's one of the areas

960
00:43:24.360 --> 00:43:29.360
that I'm most bullish on and excited about, honestly.

961
00:43:31.010 --> 00:43:31.843
<v ->Awesome, thank you.</v>

962
00:43:31.843 --> 00:43:33.780
It's good to hear like the beginning of

963
00:43:33.780 --> 00:43:35.380
this big investment data science

964
00:43:37.016 --> 00:43:37.849
and where it's moving hopefully soon.

965
00:43:37.849 --> 00:43:38.943
Susan, back to you.

966
00:43:41.470 --> 00:43:42.350
<v ->Thanks, Lindsay.</v>

967
00:43:42.350 --> 00:43:45.550
This has been great to hear about your experiences, DJ.

968
00:43:45.550 --> 00:43:47.470
So I'm curious, can you tell us a little bit about

969
00:43:47.470 --> 00:43:49.700
what some of your biggest challenges you faced

970
00:43:49.700 --> 00:43:51.690
as the U.S chief of data science?

971
00:43:51.690 --> 00:43:53.050
I know you've talked about it a little bit,

972
00:43:53.050 --> 00:43:54.700
but I'd love for you to expand on it some more

973
00:43:54.700 --> 00:43:56.260
and possibly even some of your challenges

974
00:43:56.260 --> 00:43:57.853
from other roles as well.

975
00:43:59.300 --> 00:44:00.133

Yeah.

976
00:44:00.133 --> 00:44:02.035
So the biggest challenge honestly

977
00:44:02.035 --> 00:44:06.066
is probably the one that I think many

978
00:44:06.066 --> 00:44:10.743
of us all face is why does anybody take us seriously?

979
00:44:11.810 --> 00:44:12.900
Like, think about it.

980
00:44:12.900 --> 00:44:14.910
Like we kind of have these ideas

981
00:44:14.910 --> 00:44:16.296
and you have to go fight for funding

982
00:44:16.296 --> 00:44:17.570
and you have to go fight

983
00:44:18.810 --> 00:44:19.740
for somebody to take your ideas seriously.

984
00:44:19.740 --> 00:44:23.021
People it's very easy to double down on the traditional,

985
00:44:23.021 --> 00:44:26.593
but it's very hard to get someone

986
00:44:26.593 --> 00:44:29.270
to take the radical next step.

987
00:44:29.270 --> 00:44:33.110
And I have to applaud Francis Collins

988
00:44:33.110 --> 00:44:36.310
for really starting to push NIH

989
00:44:36.310 --> 00:44:37.990
to start thinking about data science

990
00:44:37.990 --> 00:44:40.890
and being very amenable to kind of saying,

991
00:44:40.890 --> 00:44:43.020
hey, what's the next generation

992
00:44:43.020 --> 00:44:46.023
of next iteration of this look like.

993
00:44:47.923 --> 00:44:52.630
And part of that the challenge was how can data help?

994
00:44:52.630 --> 00:44:54.310
And I'll give you a very specific example

995
00:44:54.310 --> 00:44:55.920
criminal justice reform.

996
00:44:55.920 --> 00:44:58.939
And so, like criminal justice reform,

997
00:44:58.939 --> 00:45:03.150
you have really great people who are experts on this

998
00:45:03.150 --> 00:45:06.850
the public defenders, community organizers

999
00:45:06.850 --> 00:45:09.110
other activists alert other groups.

1000
00:45:09.110 --> 00:45:11.300

And so there was these meetings

1001
00:45:11.300 --> 00:45:13.378
at the white house, and people would

1002
00:45:13.378 --> 00:45:18.378
they'd let us in because I was the chief data scientist

1003
00:45:18.640 --> 00:45:21.540
but they were like, why are you here?

1004
00:45:21.540 --> 00:45:23.370
Don't you have something better to do,

1005
00:45:23.370 --> 00:45:25.820
like this is not the table for nerds.

1006
00:45:25.820 --> 00:45:28.211
And this is kind of like that classic data science thing.

1007
00:45:28.211 --> 00:45:30.670
It was like, well, they're like, okay, well,

1008
00:45:30.670 --> 00:45:33.182
we could use the smart person here I guess,

1009
00:45:33.182 --> 00:45:36.547
but this is not the table you're really supposed to be at.

1010
00:45:36.547 --> 00:45:38.800
And think about this for the same reason.

1011
00:45:38.800 --> 00:45:39.900
How many times did that happen?

1012
00:45:39.900 --> 00:45:41.780
Like, one of you has been there and they were like,

1013
00:45:41.780 --> 00:45:43.250
why are you in this conference?

1014
00:45:43.250 --> 00:45:45.020
Why are you in this meeting?

1015
00:45:45.020 --> 00:45:46.610
Like why are you here?

1016
00:45:46.610 --> 00:45:48.170
And it's that kind of thing.

1017
00:45:48.170 --> 00:45:50.630
Now, what happened was interestingly,

1018
00:45:50.630 --> 00:45:53.259
there was a blue ribbon task force that was created,

1019
00:45:53.259 --> 00:45:58.259
and the results of that task force came back.

1020
00:45:58.986 --> 00:46:02.510
And then that we're sitting around the table,

1021
00:46:02.510 --> 00:46:05.158
everyone's like, huh, let's look at these assessments,

1022
00:46:05.158 --> 00:46:09.650
body cameras, data sets on use of force.

1023
00:46:09.650 --> 00:46:11.939
And they kind of kept going around and they're like

1024
00:46:11.939 --> 00:46:16.939
aren't all these technical guests.

1025
00:46:17.590 --> 00:46:22.200

Well, I've been at these meetings and suddenly

1026
00:46:22.200 --> 00:46:23.450
people realize like, wait a second.

1027
00:46:23.450 --> 00:46:26.690
Oh, maybe you do have something that valued add.

1028
00:46:26.690 --> 00:46:28.064
And one of the things that happened was

1029
00:46:28.064 --> 00:46:32.440
we got all these police chiefs and technologists

1030
00:46:32.440 --> 00:46:35.190
and civic, activists, and political organizers

1031
00:46:35.190 --> 00:46:37.480
and all these groups together at the white house.

1032
00:46:37.480 --> 00:46:39.160
And we said, we're going to put them all

1033
00:46:39.160 --> 00:46:41.110
into a room and lock the door

1034
00:46:41.110 --> 00:46:45.070
and see what happens and see what they could come up with.

1035
00:46:45.070 --> 00:46:46.330
And some of the stuff that came up with

1036
00:46:46.330 --> 00:46:49.237
they were like, actually, nobody has any data.

1037
00:46:49.237 --> 00:46:52.570
Well, could we actually create these insights?

1038
00:46:52.570 --> 00:46:54.534
So they all got together and said

1039
00:46:54.534 --> 00:46:56.543
what if we committed to opening up

1040
00:46:56.543 --> 00:46:58.261
these hundred and one different datasets together

1041
00:46:58.261 --> 00:47:00.233
and started to see what happened?

1042
00:47:03.480 --> 00:47:04.950
And because nobody could say like, hey, how is,

1043
00:47:04.950 --> 00:47:09.030
PG county's police force compared to Montgomery counties?

1044
00:47:09.030 --> 00:47:11.840
And how does it compare to LA counties

1045
00:47:13.243 --> 00:47:16.080
or somewhere in Juneau, Alaska, like nobody knew.

1046
00:47:16.080 --> 00:47:17.610
And there were some really fascinating

1047
00:47:17.610 --> 00:47:18.580
data sets that came out.

1048
00:47:18.580 --> 00:47:20.469
For example, one that came out was

1049
00:47:20.469 --> 00:47:24.830
if you look at stop rates

1050
00:47:24.830 --> 00:47:25.950

and they had it by race,

1051
00:47:25.950 --> 00:47:28.640
and they had them over like the last 20 years,

1052
00:47:28.640 --> 00:47:32.360
that stop rates, actually they're pretty consistent

1053
00:47:32.360 --> 00:47:35.610
black and white stop rates

1054
00:47:35.610 --> 00:47:38.350
at least the way they recorded the data.

1055
00:47:38.350 --> 00:47:41.792
Pretty good, pretty actually pretty decent split.

1056
00:47:41.792 --> 00:47:44.547
But then if you look at the search rate data

1057
00:47:44.547 --> 00:47:47.969
and it's black search rates versus white search rates

1058
00:47:47.969 --> 00:47:51.601
on the conditional probability, and it's crazy,

1059
00:47:51.601 --> 00:47:54.810
and people had never just even put that thing.

1060
00:47:54.810 --> 00:47:56.007
So then the question was like,

1061
00:47:56.007 --> 00:48:01.007
who are the officers that contribute to this?

1062
00:48:01.260 --> 00:48:03.740
Is it the officers as a proxy for is it in one area,

1063
00:48:03.740 --> 00:48:08.740
or as another, is it certain beats what's going on here?

1064
00:48:09.380 --> 00:48:11.075
And then they sort of started to look at

1065
00:48:11.075 --> 00:48:13.230
and this is sort of the natural questions.

1066
00:48:13.230 --> 00:48:14.899
Like we all ask the scientist is like,

1067
00:48:14.899 --> 00:48:15.911
what led to this?

1068
00:48:15.911 --> 00:48:18.160
What's behind this data?

1069
00:48:18.160 --> 00:48:19.650
And then they found that in the case

1070
00:48:19.650 --> 00:48:21.231
like they saw like, oh my gosh

1071
00:48:21.231 --> 00:48:25.340
it's like three officers responsible for all of this.

1072
00:48:25.340 --> 00:48:27.460
And then they looked at and pulled on that thread

1073
00:48:27.460 --> 00:48:29.094
and they saw, hey, these officers, it's actually not

1074
00:48:29.094 --> 00:48:33.000
because of where they are, where they work.

1075
00:48:33.000 --> 00:48:36.070

You actually had in this case, a couple bad apples.

1076
00:48:36.070 --> 00:48:38.160
And then the chiefs were like

1077
00:48:38.160 --> 00:48:40.633
how come I don't have this data?

1078
00:48:40.633 --> 00:48:43.140
And you're like who's the person

1079
00:48:43.140 --> 00:48:44.940
that actually puts this together?

1080
00:48:44.940 --> 00:48:46.840
Who puts the dashboard together?

1081
00:48:46.840 --> 00:48:49.670
That's the data scientist.

1082
00:48:49.670 --> 00:48:51.550
Now equally as much,

1083
00:48:51.550 --> 00:48:56.550
there's a police officers who have really hard jobs.

1084
00:48:56.930 --> 00:48:59.263
And so one of the ones that we found in this case

1085
00:48:59.263 --> 00:49:01.970
was we put a team of data scientists

1086
00:49:01.970 --> 00:49:05.380
in to this police department down in the South.

1087
00:49:05.380 --> 00:49:07.630
And they were looking at excessive use of force.

1088
00:49:07.630 --> 00:49:11.437
And this is a very locked down dataset

1089
00:49:13.630 --> 00:49:14.820
because it had so much information.

1090
00:49:14.820 --> 00:49:16.510
It wasn't a public dataset.

1091
00:49:16.510 --> 00:49:18.510
And they started looking at use of force.

1092
00:49:18.510 --> 00:49:22.195
And as they started to do the kind of the classic,

1093
00:49:22.195 --> 00:49:25.700
like what are the feature sets that

1094
00:49:25.700 --> 00:49:30.150
that provide the lift in the cases of excessive use force,

1095
00:49:30.150 --> 00:49:31.890
there was sort of the first set of features

1096
00:49:31.890 --> 00:49:35.351
that were all very obvious, excessive,

1097
00:49:35.351 --> 00:49:39.840
lots of accidents, lots of complaints, lots of easy signal

1098
00:49:39.840 --> 00:49:41.470
but then you kind of go down a little bit further

1099
00:49:41.470 --> 00:49:43.927
a feature set in two very interesting ones pop out

1100
00:49:43.927 --> 00:49:46.810

as features.

1101
00:49:46.810 --> 00:49:50.240
One is that you reported you've been participant

1102
00:49:50.240 --> 00:49:52.710
in multiple suicide calls

1103
00:49:52.710 --> 00:49:56.930
or you've responded to a domestic violence

1104
00:49:56.930 --> 00:49:59.010
where children are present.

1105
00:49:59.010 --> 00:50:01.010
And so they've started to kind of, the data scientists were

1106
00:50:01.010 --> 00:50:02.300
like, what's going on?

1107
00:50:02.300 --> 00:50:03.920
But in this case, one of the most

1108
00:50:03.920 --> 00:50:05.750
and this is one of the most important things is,

1109
00:50:05.750 --> 00:50:09.414
the data scientists were embedded with the officers.

1110
00:50:09.414 --> 00:50:12.230
Like they were doing ride alongs,

1111
00:50:12.230 --> 00:50:13.063
they were doing everything.

1112
00:50:13.063 --> 00:50:14.590
They were like really embedded.

1113
00:50:14.590 --> 00:50:15.440
So they had context.

1114
00:50:15.440 --> 00:50:17.151
And so right away, they knew what would have happened

1115
00:50:17.151 --> 00:50:21.160
in both those cases of response to suicide

1116
00:50:21.160 --> 00:50:23.070
and domestic violence, where children presents,

1117
00:50:23.070 --> 00:50:26.472
they're both highly emotionally charged events,

1118
00:50:26.472 --> 00:50:31.472
very physically messy, dangerous, and just tough situations.

1119
00:50:34.080 --> 00:50:36.240
And so then if you're not the officer

1120
00:50:36.240 --> 00:50:37.790
that's just filing everything

1121
00:50:37.790 --> 00:50:39.591
but maybe you're the one that responded,

1122
00:50:39.591 --> 00:50:42.453
dispatch tells you to get back out on the street.

1123
00:50:43.470 --> 00:50:46.311
And now you pull somebody over and they're flippant with you

1124
00:50:46.311 --> 00:50:51.311
and you haven't had that time to really decompress.

1125
00:50:51.426 --> 00:50:56.010

So it's amazing that we can design software

1126
00:50:56.010 --> 00:51:01.010
to get you a package from e-commerce in two days,

1127
00:51:01.300 --> 00:51:04.890
we can get you a car in five minutes,

1128
00:51:04.890 --> 00:51:09.052
we can get you a pizza or food delivered instantly

1129
00:51:09.052 --> 00:51:12.190
but we can't have a dispatch system that takes us

1130
00:51:12.190 --> 00:51:16.720
into account that it can think about these things.

1131
00:51:16.720 --> 00:51:19.206
So what they did is they started to build

1132
00:51:19.206 --> 00:51:21.900
that smarter dispatch system.

1133
00:51:21.900 --> 00:51:24.760
And they've been trying it out and starting to deploy it.

1134
00:51:24.760 --> 00:51:26.760
Who is that person?

1135
00:51:26.760 --> 00:51:29.460
That's the data scientist that's doing that.

1136
00:51:29.460 --> 00:51:32.890
And so I kind of belabor this point through a couple

1137
00:51:33.929 --> 00:51:36.029
of these stories to really emphasize that,

1138
00:51:37.462 --> 00:51:39.430
that the role of being in the room,

1139
00:51:39.430 --> 00:51:41.650
being at the table when you're

1140
00:51:41.650 --> 00:51:44.610
in there is one of the most important things.

1141
00:51:44.610 --> 00:51:47.160
And one of the big challenge was

1142
00:51:48.554 --> 00:51:52.787
how do we create better pathways

1143
00:51:52.787 --> 00:51:56.450
for more people to be at those tables?

1144
00:51:56.450 --> 00:52:01.450
How do we get our data scientist, friends, our colleagues

1145
00:52:02.822 --> 00:52:05.110
how do we get one more person along on the journey with us?

1146
00:52:05.110 --> 00:52:06.197
And it was to break out of our classic bureaucracy

1147
00:52:06.197 --> 00:52:11.197
and say look, add another person here,

1148
00:52:13.536 --> 00:52:15.470
let somebody else help.

1149
00:52:15.470 --> 00:52:18.360
And that was kind of the singular goal

1150
00:52:19.292 --> 00:52:22.650

of mine was to get us to be horizontals,

1151
00:52:22.650 --> 00:52:26.390
to be accepted whether it's research and development

1152
00:52:26.390 --> 00:52:30.890
whether it's a census, whether it's national security

1153
00:52:30.890 --> 00:52:33.603
all these different aspects benefit from us

1154
00:52:33.603 --> 00:52:36.513
just being able to collaborate more effectively together.

1155
00:52:39.960 --> 00:52:44.020
<v ->DJ, that's an amazing story about criminal justice reform</v>

1156
00:52:44.020 --> 00:52:46.750
coming from data science and your sitting in

1157
00:52:46.750 --> 00:52:50.330
on a meeting somewhat unexpectedly.

1158
00:52:50.330 --> 00:52:51.163
I also saw, I have another question for you,

1159
00:52:51.163 --> 00:52:56.117
but I am struck by the importance of

1160
00:52:58.364 --> 00:52:59.197
sort of mixed modes of data.

1161
00:52:59.197 --> 00:53:01.020
In other words, what you described was the importance

1162
00:53:01.020 --> 00:53:04.220
of a qualitative approach as well as quantitative.

1163
00:53:04.220 --> 00:53:07.160
Cause you said that it was the ride alongs that provided an

1164
00:53:07.160 --> 00:53:09.743
essential ingredient for interpreting,

1165
00:53:09.743 --> 00:53:12.900
what otherwise you might've looked at, why are suicides

1166
00:53:12.900 --> 00:53:14.420
and domestic violence associated.

1167
00:53:14.420 --> 00:53:16.570
You still might've been able to design a system to adapt

1168
00:53:16.570 --> 00:53:19.360
to those, but it wouldn't have generalized

1169
00:53:19.360 --> 00:53:21.780
to other emotionally charged events too

1170
00:53:21.780 --> 00:53:25.060
that could help the dispatch a very interesting set

1171
00:53:25.060 --> 00:53:26.250
of implications.

1172
00:53:26.250 --> 00:53:30.150
What are some other memorable moments

1173
00:53:30.150 --> 00:53:33.877
from you being a chief data scientist for the U.S?

1174
00:53:36.863 --> 00:53:39.020
<v ->The some of the ones that were</v>

1175
00:53:39.020 --> 00:53:42.250

the greatest joy was actually being able to hang

1176
00:53:42.250 --> 00:53:46.120
out with amazing colleagues across the federal government

1177
00:53:46.120 --> 00:53:50.910
across the country and see what people were doing.

1178
00:53:50.910 --> 00:53:53.057
And many times was to just help kind of say,

1179
00:53:53.057 --> 00:53:57.740
hey, if you and you work together

1180
00:53:57.740 --> 00:54:00.500
this is going to be a heck of a lot more interesting

1181
00:54:00.500 --> 00:54:02.620
you know, cancer moonshot was a great example

1182
00:54:02.620 --> 00:54:04.173
of this is,

1183
00:54:07.523 --> 00:54:11.140
when then by vice president Biden was mentioning,

1184
00:54:11.140 --> 00:54:13.160
like he wants to go after this.

1185
00:54:13.160 --> 00:54:15.610
Well, what really, what kind of,

1186
00:54:15.610 --> 00:54:17.383
where's the puck gonna kind of go?

1187
00:54:19.733 --> 00:54:20.566
It was like, well, why

1188
00:54:20.566 --> 00:54:21.960
aren't we using technology and data more effectively here?

1189
00:54:21.960 --> 00:54:23.310
And then there was this kind of question

1190
00:54:23.310 --> 00:54:25.120
of what would it take to massively accelerate this?

1191
00:54:25.120 --> 00:54:26.110
And we said, well

1192
00:54:26.110 --> 00:54:28.743
we haven't actually brought these efforts together.

1193
00:54:30.044 --> 00:54:32.190
And I think this is one of our challenges is,

1194
00:54:32.190 --> 00:54:36.200
NIH does what it does phenomenally

1195
00:54:36.200 --> 00:54:40.910
but NIH is not a computational superpower in the same way

1196
00:54:40.910 --> 00:54:44.030
like DOE has historically been right

1197
00:54:46.139 --> 00:54:48.447
because that's where the supercomputers have been built.

1198
00:54:48.447 --> 00:54:49.280
That's where we federally allocate dollars.

1199
00:54:49.280 --> 00:54:52.760
NSF does a whole lot on algorithms.

1200
00:54:52.760 --> 00:54:53.940

That's where that's been.

1201
00:54:53.940 --> 00:54:55.150
So what would it look like

1202
00:54:55.150 --> 00:54:58.830
if we started to blend that more efficiently together?

1203
00:54:58.830 --> 00:55:01.586
It's not like NIH can't work on algorithms

1204
00:55:01.586 --> 00:55:05.416
but what if we started to make this more hybridized?

1205
00:55:05.416 --> 00:55:07.350
What if we brought in the best

1206
00:55:07.350 --> 00:55:11.171
from NASA and different technologies to do things

1207
00:55:11.171 --> 00:55:15.390
and what I think we found in all of these efforts

1208
00:55:15.390 --> 00:55:19.145
these big ticket projects, especially with president Obama

1209
00:55:19.145 --> 00:55:22.920
was you need to really start

1210
00:55:22.920 --> 00:55:26.368
to bring a more multidisciplinary approach.

1211
00:55:26.368 --> 00:55:27.760
And by the way,

1212
00:55:27.760 --> 00:55:30.073
this was definitely true for me during, when I served

1213
00:55:30.073 --> 00:55:34.620
in the Bush administration as well around post nine 11 was

1214
00:55:34.620 --> 00:55:37.420
as we're trying to figure out many of these ideas

1215
00:55:37.420 --> 00:55:40.940
around national security, it wasn't like

1216
00:55:40.940 --> 00:55:45.400
these ideas existed in traditional NASA security apparatus.

1217
00:55:45.400 --> 00:55:49.970
A lot of these ideas came from other places, other ideas

1218
00:55:49.970 --> 00:55:51.030
other disciplines.

1219
00:55:51.030 --> 00:55:53.080
And we said, what if we brought these in

1220
00:55:53.080 --> 00:55:56.100
and use these techniques, would that be more efficient?

1221
00:55:56.100 --> 00:55:57.990
Would that be more effective?

1222
00:55:57.990 --> 00:56:01.429
And the answer is massively so massively.

1223
00:56:01.429 --> 00:56:05.280
And so it was kind of being in there, maybe, you know

1224
00:56:05.280 --> 00:56:07.180
like what some of the other ones that I think

1225
00:56:07.180 --> 00:56:09.794

were very special moments for me personally were,

1226
00:56:09.794 --> 00:56:14.794
which may seem a little strange from a data scientist

1227
00:56:17.071 --> 00:56:18.739
was actually making sure we tell the stories

1228
00:56:18.739 --> 00:56:23.739
of real people and maybe using as our opening,

1229
00:56:25.460 --> 00:56:29.710
if you will to allow deeper research for work,

1230
00:56:29.710 --> 00:56:30.883
to take place.

1231
00:56:32.740 --> 00:56:35.323
And you know, some of the ones

1232
00:56:36.499 --> 00:56:38.680
that is there's a jealous Zoe Keating.

1233
00:56:38.680 --> 00:56:41.956
And some of you may know her for her work.

1234
00:56:41.956 --> 00:56:45.010
She is phenomenal cellos her, she

1235
00:56:45.010 --> 00:56:48.790
her husband died of a brain cancer.

1236
00:56:48.790 --> 00:56:49.840
Do you have all Soma?

1237
00:56:51.242 --> 00:56:52.760
And like so many of hers

1238
00:56:52.760 --> 00:56:56.550
she just couldn't get the electronic medical records.

1239
00:56:56.550 --> 00:56:58.810
And so she was trying to get care for him

1240
00:56:58.810 --> 00:57:01.580
and was going around from institution after institution

1241
00:57:01.580 --> 00:57:06.580
after institution, and was just getting the runaround

1242
00:57:08.779 --> 00:57:11.430
for ferry, very basic things.

1243
00:57:11.430 --> 00:57:14.407
Now she happens to be really good at data

1244
00:57:14.407 --> 00:57:19.160
and she's a really fast at typing.

1245
00:57:19.160 --> 00:57:22.270
So she used, she transcribed all of her notes.

1246
00:57:22.270 --> 00:57:25.673
And so she walked me through her experience.

1247
00:57:27.669 --> 00:57:31.393
Now it turns out her experience is exact same one

1248
00:57:32.250 --> 00:57:34.090
Joe Biden went through.

1249
00:57:34.090 --> 00:57:35.864
And I remember literally sitting

1250
00:57:35.864 --> 00:57:40.864

in the kitchen at the Vice-President's a residence one time.

1251
00:57:41.680 --> 00:57:46.680
And we were kind of talking about the cancer moonshot.

1252
00:57:47.300 --> 00:57:50.970
And he couldn't believe that other people

1253
00:57:50.970 --> 00:57:55.210
were facing the same problem that Bo his son was facing

1254
00:57:55.210 --> 00:57:58.368
as they were trying to get his medical records between DC

1255
00:57:58.368 --> 00:58:03.368
and Delaware to just get the same kind of treatment.

1256
00:58:04.964 --> 00:58:06.340
And so you had the vice-presidents then

1257
00:58:06.340 --> 00:58:09.190
vice-president son having the same issue

1258
00:58:09.190 --> 00:58:13.303
that Zoe Keating's husband was facing and so many others.

1259
00:58:15.762 --> 00:58:18.320
And that's really what then helped kick off part

1260
00:58:18.320 --> 00:58:22.880
of that discussion for how do we actually

1261
00:58:22.880 --> 00:58:27.230
massively start to use the electronic medical records

1262
00:58:27.230 --> 00:58:31.630
for efforts like precision medicine.

1263
00:58:31.630 --> 00:58:35.380
They had to be the foundation aspect of it, but

1264
00:58:35.380 --> 00:58:38.270
in one of the things I would just encourage all of

1265
00:58:42.991 --> 00:58:44.564
all of you to remember is that, and this is why

1266
00:58:44.564 --> 00:58:46.410
I say it all the time is the data points have names.

1267
00:58:46.410 --> 00:58:48.520
The D the data points have names

1268
00:58:48.520 --> 00:58:53.290
and we have to remember those names and tell their stories.

1269
00:58:53.290 --> 00:58:56.560
That's one of the most important things.

1270
00:58:56.560 --> 00:59:00.200
One of the really big ass kickings that I had

1271
00:59:01.528 --> 00:59:02.511
from president Obama

1272
00:59:02.511 --> 00:59:04.711
which is not an kicking that you really want

1273
00:59:08.376 --> 00:59:10.810
was during precision medicine, he said, hey, have you

1274
00:59:10.810 --> 00:59:13.840
it was very clear cut order of making sure

1275
00:59:13.840 --> 00:59:16.420

that people who are going to be impacted

1276
00:59:16.420 --> 00:59:19.703
by this program have a say in the design of the program.

1277
00:59:20.700 --> 00:59:23.140
And so we went out and we said, yes

1278
00:59:23.140 --> 00:59:27.530
we have the groups from cystic fibrosis, cystic fibrosis

1279
00:59:27.530 --> 00:59:29.840
the community, we have this group

1280
00:59:29.840 --> 00:59:32.452
we have this group, we have groups this group

1281
00:59:32.452 --> 00:59:35.060
and he kept saying, that's the group that's represented.

1282
00:59:35.060 --> 00:59:37.573
Do you have the actual people?

1283
00:59:38.800 --> 00:59:40.352
And we thought about it.

1284
00:59:40.352 --> 00:59:42.597
And they're like, no, sir, we need to do a better job here.

1285
00:59:42.597 --> 00:59:44.590
And so we went and started to get those people in.

1286
00:59:44.590 --> 00:59:47.190
One of the ones that we had is, is there a recall

1287
00:59:47.190 --> 00:59:51.150
we had a big science conference out in Pittsburgh

1288
00:59:51.150 --> 00:59:56.146
and Carnegie Mellon, and we kinda had a lot of groups there.

1289
00:59:56.146 --> 00:59:59.400
And one of the ones that we did is we had a round table

1290
00:59:59.400 --> 01:00:02.170
on precision medicine with the local community

1291
01:00:02.170 --> 01:00:05.110
Hispanic community, the black community, researchers

1292
01:00:05.110 --> 01:00:10.110
others there, and had a big long room and everyone

1293
01:00:10.400 --> 01:00:14.730
and Stephanie Davine, Francis Collins, myself

1294
01:00:14.730 --> 01:00:16.230
were there a number of others.

1295
01:00:17.984 --> 01:00:19.730
And we kind of were listening to everybody, go around

1296
01:00:19.730 --> 01:00:23.340
and we go around and then there's this kind of

1297
01:00:23.340 --> 01:00:25.050
everyone has spoken given us their input.

1298
01:00:25.050 --> 01:00:28.959
There's this elderly black woman in the corner.

1299
01:00:28.959 --> 01:00:31.720
I remember saying, ma'am we haven't heard

1300
01:00:31.720 --> 01:00:33.187

from you love your opinion.

1301
01:00:33.187 --> 01:00:34.523
She's like, do you really want my opinion?

1302
01:00:35.695 --> 01:00:37.440
What did he say to that?

1303
01:00:38.290 --> 01:00:40.750
Yes, ma'am really want to know.

1304
01:00:40.750 --> 01:00:43.910
And, oh my gosh, if she let us have it.

1305
01:00:43.910 --> 01:00:46.200
Wow, did she tear into us?

1306
01:00:46.200 --> 01:00:47.820
She's like, have you thought about this?

1307
01:00:47.820 --> 01:00:49.547
You haven't thought about this.

1308
01:00:49.547 --> 01:00:50.531
You haven't done this.

1309
01:00:50.531 --> 01:00:51.557
What about this?

1310
01:00:51.557 --> 01:00:52.524
What about this?

1311
01:00:52.524 --> 01:00:53.535
What about this for the community?

1312
01:00:53.535 --> 01:00:54.554
I remember like walking out

1313
01:00:54.554 --> 01:00:57.000
of that room with Francis and saying

1314
01:00:58.470 --> 01:01:00.570
she just crushed us.

1315
01:01:00.570 --> 01:01:02.883
I thought, can we hire her?

1316
01:01:03.920 --> 01:01:06.290
Like she just gave us a playbook.

1317
01:01:06.290 --> 01:01:08.210
She just literally gave us a playbook

1318
01:01:08.210 --> 01:01:11.913
of what we need to do properly to engage the community.

1319
01:01:12.930 --> 01:01:15.320
Otherwise we're not, they're not there's data points.

1320
01:01:15.320 --> 01:01:16.290
Aren't gonna be,

1321
01:01:16.290 --> 01:01:18.190
we're not gonna get those data points.

1322
01:01:21.825 --> 01:01:22.817
We're not going to have representation.

1323
01:01:22.817 --> 01:01:23.832
We are going to fail on the mission.

1324
01:01:23.832 --> 01:01:24.665
And I just emphasize, like

1325
01:01:25.500 --> 01:01:26.333

you have to get into the community.

1326
01:01:26.333 --> 01:01:27.680
You have to get out of your comfort zone

1327
01:01:27.680 --> 01:01:28.513
get out of the office,

1328
01:01:28.513 --> 01:01:30.323
get out of the lab and meet with those people.

1329
01:01:30.323 --> 01:01:32.444
Because that's where the insights

1330
01:01:32.444 --> 01:01:34.003
truly gonna come from.

1331
01:01:34.890 --> 01:01:36.030
Person's book that we read.

1332
01:01:36.030 --> 01:01:38.110
A lot of us read a tool one day.

1333
01:01:38.110 --> 01:01:40.630
If he's kind of thinking about what he really does

1334
01:01:40.630 --> 01:01:42.913
what he does best at the end of the day.

1335
01:01:43.856 --> 01:01:45.080
And I've seen, I told you this, especially

1336
01:01:45.080 --> 01:01:48.280
around COVID is you got to meet and talk to the people

1337
01:01:48.280 --> 01:01:49.520
and that's where you get the glimmer

1338
01:01:49.520 --> 01:01:51.783
of insight that you then can carry forward.

1339
01:01:54.670 --> 01:01:55.743
<v ->Thanks very much.</v>

1340
01:01:59.200 --> 01:02:00.570
Roger, I think internet,

1341
01:02:00.570 --> 01:02:01.720
<v ->Francis would be like</v>

1342
01:02:02.895 --> 01:02:03.728
you can't tell all these stories.

1343
01:02:05.150 --> 01:02:05.983
<v ->You haven't.</v>

1344
01:02:07.330 --> 01:02:08.163
Thank you.

1345
01:02:09.216 --> 01:02:10.520
<v ->Yes, thank you.</v>

1346
01:02:10.520 --> 01:02:15.140
So perhaps negative examples are as effective

1347
01:02:16.095 --> 01:02:16.928
as positive examples.

1348
01:02:16.928 --> 01:02:20.060
So what is some of the worst and best advice that

1349
01:02:20.060 --> 01:02:21.623
you've received along the way?

1350
01:02:22.490 --> 01:02:23.323

Yeah.

1351
01:02:26.360 --> 01:02:29.470
So one of the most powerful, I mean

1352
01:02:29.470 --> 01:02:31.949
one of the things I need to emphasize is I had

1353
01:02:31.949 --> 01:02:35.713
catastrophic failures.

1354
01:02:37.725 --> 01:02:38.558
I think some people don't know.

1355
01:02:38.558 --> 01:02:43.050
I had a $41 million failures company that failed

1356
01:02:43.050 --> 01:02:45.300
because we were building a photo sharing app.

1357
01:02:46.835 --> 01:02:48.090
I kid you not.

1358
01:02:48.090 --> 01:02:50.100
And people thought my career was dead.

1359
01:02:50.100 --> 01:02:52.090
People really thought my career was dead.

1360
01:02:52.090 --> 01:02:54.240
I remember people being like

1361
01:02:54.240 --> 01:02:55.073
with that kind of failure in Silicon Valley, you're done.

1362
01:02:55.073 --> 01:02:59.847
One of the most important things I lessons I learned

1363
01:03:01.423 --> 01:03:02.406
with that is,

1364
01:03:02.406 --> 01:03:04.690
if you are constantly finding a way to help somebody

1365
01:03:04.690 --> 01:03:08.680
and approach of constantly creating more than you take

1366
01:03:08.680 --> 01:03:10.810
people will help you out.

1367
01:03:10.810 --> 01:03:14.420
People will help you out massively in phenomenal ways.

1368
01:03:14.420 --> 01:03:15.640
They'll give you a second chance

1369
01:03:15.640 --> 01:03:18.883
a third chance to try to do more things.

1370
01:03:20.610 --> 01:03:21.623
And it's one of the reasons I emphasize collaboration is

1371
01:03:21.623 --> 01:03:25.389
because if I didn't have those collaborative experiences

1372
01:03:25.389 --> 01:03:30.389
I think my career would have would've tanked at that point.

1373
01:03:31.670 --> 01:03:33.520
I would have had a really tough time.

1374
01:03:35.330 --> 01:03:37.997
The other one that I emphasize is,

1375
01:03:37.997 --> 01:03:41.400

from a career planning perspective

1376
01:03:41.400 --> 01:03:43.253
what are the skills you need?

1377
01:03:44.300 --> 01:03:48.560
One of the ones I went to after it was early on in eBay

1378
01:03:48.560 --> 01:03:50.010
an early in my tenure at eBay.

1379
01:03:50.010 --> 01:03:54.820
And I had some pretty good wins

1380
01:03:54.820 --> 01:03:57.930
and I kind of went to this VC and I said,

1381
01:03:57.930 --> 01:04:01.540
hey, I'd like to be part of the early stage team.

1382
01:04:01.540 --> 01:04:03.750
And I would like to work on these things.

1383
01:04:03.750 --> 01:04:05.890
And he kind of was like

1384
01:04:05.890 --> 01:04:08.140
why would I ever give you that shot?

1385
01:04:08.140 --> 01:04:10.527
You're haven't had our proven track record

1386
01:04:10.527 --> 01:04:14.440
of building things.

1387
01:04:14.440 --> 01:04:17.460
You haven't had a proven track record or building teams.

1388
01:04:17.460 --> 01:04:20.680
You haven't had a proven track record of raising funding

1389
01:04:20.680 --> 01:04:22.100
all these different things.

1390
01:04:22.100 --> 01:04:24.428
And you know, one side you can take that as an kicking.

1391
01:04:24.428 --> 01:04:26.530
The other side, you can look

1392
01:04:26.530 --> 01:04:29.110
at it and be like, oh, that's the check boxes.

1393
01:04:29.110 --> 01:04:30.350
Those are all the checks.

1394
01:04:30.350 --> 01:04:32.603
That's the things I need to do, great.

1395
01:04:37.741 --> 01:04:39.972
Now I've got a roadmap and I can go do those things.

1396
01:04:39.972 --> 01:04:42.922
The other part I would say career failure that is there is,

1397
01:04:45.325 --> 01:04:46.158
one of the things that people don't

1398
01:04:46.158 --> 01:04:48.939
realize is like, oh, like I just went from,

1399
01:04:48.939 --> 01:04:53.230
being an academic to going to government industry.

1400
01:04:53.230 --> 01:04:54.673

Those were very hard.

1401
01:04:57.830 --> 01:04:58.830
I don't want to make it seem like I just jumped in.

1402
01:04:58.830 --> 01:05:00.580
And it was like easy.

1403
01:05:01.586 --> 01:05:03.617
It was excrutiatingly difficult.

1404
01:05:03.617 --> 01:05:06.320
The first from Maryland to the government

1405
01:05:06.320 --> 01:05:07.690
part of that was after 911.

1406
01:05:07.690 --> 01:05:08.890
But second of part

1407
01:05:10.309 --> 01:05:12.810
of the reason I went in was in the Bush administration

1408
01:05:12.810 --> 01:05:14.860
our budgets were just getting crushed

1409
01:05:14.860 --> 01:05:16.720
from a scientific perspective.

1410
01:05:16.720 --> 01:05:18.920
And so, because of that, I couldn't

1411
01:05:18.920 --> 01:05:21.170
I wasn't going to be able to fund the team

1412
01:05:21.170 --> 01:05:22.880
in a way that that needed to be funded.

1413
01:05:22.880 --> 01:05:26.802
So we're going to run out of classic research problems

1414
01:05:26.802 --> 01:05:31.670
being able to support them going into, from a government

1415
01:05:34.403 --> 01:05:35.425
then going back to university of Maryland

1416
01:05:35.425 --> 01:05:36.406
then deciding to leave for industry.

1417
01:05:36.406 --> 01:05:37.863
It took me six months to get my job,

1418
01:05:39.663 --> 01:05:40.613
to get my first job for someone to take a chance on me,

1419
01:05:42.886 --> 01:05:43.904
six months of pounding the pavement

1420
01:05:43.904 --> 01:05:44.879
and no one taking a shot.

1421
01:05:44.879 --> 01:05:47.420
And by the way, Sergei Brin, the founder of Google

1422
01:05:47.420 --> 01:05:49.700
his dad was on my thesis committee.

1423
01:05:49.700 --> 01:05:52.197
His mom was on research grants with me

1424
01:05:52.197 --> 01:05:54.420
and I still couldn't get a job at Google.

1425
01:05:54.420 --> 01:05:55.900

The way I got my job.

1426
01:05:55.900 --> 01:05:58.230
And this is definitely a story of privilege

1427
01:06:01.267 --> 01:06:03.130
is my mom happened to be at a cocktail party

1428
01:06:03.130 --> 01:06:06.470
with the president of Skype at the time rigid data.

1429
01:06:06.470 --> 01:06:09.610
And she harassed him so much to take a call with me

1430
01:06:09.610 --> 01:06:12.196
that I actually was able to get on a call with them.

1431
01:06:12.196 --> 01:06:15.010
And then he's like, oh, you seem smart.

1432
01:06:15.010 --> 01:06:17.880
Maybe we should say, you're going to interview at eBay.

1433
01:06:17.880 --> 01:06:19.850
That's how I ended up at eBay.

1434
01:06:19.850 --> 01:06:23.068
Later on, people were like, oh, you're actually have

1435
01:06:23.068 --> 01:06:27.820
you can benefit things here in Silicon Valley only

1436
01:06:27.820 --> 01:06:31.370
because somebody was able to take a chance on me.

1437
01:06:31.370 --> 01:06:35.050
And so what translate that into for all of us

1438
01:06:35.050 --> 01:06:37.310
because I had that opportunity to privilege is how

1439
01:06:37.310 --> 01:06:42.300
do we use our respective privileges to create an opportunity

1440
01:06:42.300 --> 01:06:44.130
for another person?

1441
01:06:44.130 --> 01:06:47.320
What would it look like if you gave just a little bit

1442
01:06:47.320 --> 01:06:50.310
of time to help somebody else in need?

1443
01:06:50.310 --> 01:06:52.040
What if you just took a chance on somebody?

1444
01:06:52.040 --> 01:06:54.450
Cause none of us got to where we are without somebody

1445
01:06:54.450 --> 01:06:56.099
taking your shot on me on us.

1446
01:06:56.099 --> 01:06:59.750
And I certainly didn't get to where I was without lots

1447
01:06:59.750 --> 01:07:02.250
of different people taking a shot on me.

1448
01:07:02.250 --> 01:07:06.040
And I think that was whether it

1449
01:07:07.092 --> 01:07:10.347
was this idea of having this rebooting LinkedIn

1450
01:07:10.347 --> 01:07:12.897

as a data company, whether it was that there, the idea

1451
01:07:12.897 --> 01:07:15.460
that there, the chief data scientist should work

1452
01:07:15.460 --> 01:07:17.545
on criminal justice issues, whether it should be

1453
01:07:17.545 --> 01:07:20.260
that there any of these things.

1454
01:07:20.260 --> 01:07:24.880
And so my ask is, let's just, let's push ourselves.

1455
01:07:24.880 --> 01:07:27.210
Let's challenge ourselves to open the door

1456
01:07:27.210 --> 01:07:29.800
for somebody that may not have a chance

1457
01:07:32.242 --> 01:07:33.253
that we may have had,

1458
01:07:33.253 --> 01:07:34.413
or the chance that we wish we had.

1459
01:07:36.757 --> 01:07:37.717
<v ->Thank you, DJ.</v>

1460
01:07:37.717 --> 01:07:38.745
This is Albert again.

1461
01:07:38.745 --> 01:07:40.920
I just want to follow up on your last point actually.

1462
01:07:40.920 --> 01:07:42.563
So from the other perspective,

1463
01:07:43.738 --> 01:07:44.760
I think what you said is so true

1464
01:07:44.760 --> 01:07:46.377
we have to give back, but let's say

1465
01:07:46.377 --> 01:07:50.010
there's those early career students, undergrads

1466
01:07:50.010 --> 01:07:51.850
high school students out there who are

1467
01:07:51.850 --> 01:07:53.410
from Silicon Valley or Bethesda

1468
01:07:53.410 --> 01:07:56.180
or the Montgomery corridor what can they do?

1469
01:07:56.180 --> 01:07:58.329
What advice can you give them?

1470
01:07:58.329 --> 01:08:00.550
We're always looking at ways to reach out

1471
01:08:00.550 --> 01:08:01.910
to the community and Dr. Volkov,

1472
01:08:01.910 --> 01:08:03.330
this is really high on her radar.

1473
01:08:03.330 --> 01:08:05.505
How do we reach out to the community to give them a chance?

1474
01:08:05.505 --> 01:08:07.533
What advice can you give them to get involved

1475
01:08:07.533 --> 01:08:09.663

with data science and the follow up question to that

1476
01:08:09.663 --> 01:08:11.600
because we're running short on time is

1477
01:08:14.045 --> 01:08:15.254
what should students do

1478
01:08:15.254 --> 01:08:17.180
during the pandemic to best themselves

1479
01:08:17.180 --> 01:08:19.130
to get more involved with data science.

1480
01:08:19.980 --> 01:08:23.090
<v ->Yeah, so thank you for asking this question.</v>

1481
01:08:23.090 --> 01:08:26.403
So the first is,

1482
01:08:27.720 --> 01:08:32.720
at the high school level, the collegiate level, I think

1483
01:08:32.730 --> 01:08:35.538
the models that the two models I often point to

1484
01:08:35.538 --> 01:08:40.538
are the Montgomery County,

1485
01:08:40.940 --> 01:08:44.230
the Montgomery Blair high school that

1486
01:08:44.230 --> 01:08:47.730
that the Fairfax County County schools

1487
01:08:47.730 --> 01:08:49.510
and some of the Silicon Valley schools

1488
01:08:49.510 --> 01:08:51.440
I think those are the high water marks

1489
01:08:51.440 --> 01:08:53.840
of our school systems.

1490
01:08:53.840 --> 01:08:56.670
Because the students it's very easy for us to reach

1491
01:08:56.670 --> 01:08:58.550
out and let one of them into our labs

1492
01:08:58.550 --> 01:09:01.700
or how internships and all of those different aspects.

1493
01:09:01.700 --> 01:09:04.460
I think the data science space is actually some

1494
01:09:04.460 --> 01:09:07.234
of the most amenable because you can just label data

1495
01:09:07.234 --> 01:09:10.540
in turn and just start to get used to stuff and play

1496
01:09:10.540 --> 01:09:12.290
with stuff where your coding skills

1497
01:09:14.387 --> 01:09:15.220
don't have to be up to par,

1498
01:09:15.220 --> 01:09:17.790
or you can plug things in wire stuff together in packages.

1499
01:09:17.790 --> 01:09:20.130
What I think is the challenge is how do we reach

1500
01:09:20.130 --> 01:09:22.120

beyond our local communities?

1501
01:09:22.120 --> 01:09:24.222
How do we reach to communities that

1502
01:09:24.222 --> 01:09:26.410
don't have that access that are

1503
01:09:29.674 --> 01:09:30.693
around the country and be able to do things.

1504
01:09:30.693 --> 01:09:31.711
And I think that we need to start figuring

1505
01:09:31.711 --> 01:09:32.640
out and be clever about new ways to do it.

1506
01:09:32.640 --> 01:09:35.420
Because as we found with the pandemic

1507
01:09:35.420 --> 01:09:38.220
is there are ways to do this a little bit more virtually.

1508
01:09:38.220 --> 01:09:41.390
There are ways that we can start to do those things

1509
01:09:41.390 --> 01:09:42.610
but we honestly need some

1510
01:09:42.610 --> 01:09:45.370
of the programs that bring these students into us.

1511
01:09:45.370 --> 01:09:47.470
And I think we've all had that moment where one

1512
01:09:47.470 --> 01:09:48.810
of us got to be in a lab

1513
01:09:50.406 --> 01:09:51.490
or around somebody that changed our trajectory.

1514
01:09:51.490 --> 01:09:53.420
You know, for me, it was

1515
01:09:53.420 --> 01:09:56.760
doing early PCR experiments in high school

1516
01:09:56.760 --> 01:09:59.836
because I talked much because my grades weren't good.

1517
01:09:59.836 --> 01:10:01.545
I wasn't supposed to be the kid that was allowed to go

1518
01:10:01.545 --> 01:10:03.329
to lab, but I talked my way

1519
01:10:03.329 --> 01:10:05.010
into the field trip and I got to do this PCR

1520
01:10:05.010 --> 01:10:06.993
like kind of lab experiment.

1521
01:10:08.106 --> 01:10:10.326
It changed my thinking on so many different levels.

1522
01:10:10.326 --> 01:10:13.299
But if, and the reason I kind of talked my way

1523
01:10:13.299 --> 01:10:14.520
there was one teacher was like, well

1524
01:10:14.520 --> 01:10:16.270
if this guy is this passionate

1525
01:10:16.270 --> 01:10:20.050

about doing this creates be damned.

1526
01:10:20.050 --> 01:10:23.314
Let's let him into the program anyway and see what he does.

1527
01:10:23.314 --> 01:10:25.014
And then that same physics teacher

1528
01:10:26.454 --> 01:10:27.287
because even though my grades didn't qualify

1529
01:10:27.287 --> 01:10:30.530
she gave me a ton of scientific equipment to take home.

1530
01:10:30.530 --> 01:10:32.680
She just said, return it at the end of the year.

1531
01:10:32.680 --> 01:10:34.850
And so I built a little mini lab where

1532
01:10:34.850 --> 01:10:38.170
I was reproducing editor photographic experiments

1533
01:10:39.610 --> 01:10:40.810
on my own at home.

1534
01:10:40.810 --> 01:10:42.220
And so what does it look like

1535
01:10:42.220 --> 01:10:45.430
to do a version of that for kids?

1536
01:10:45.430 --> 01:10:46.710
I don't know.

1537
01:10:46.710 --> 01:10:48.210
I don't honestly

1538
01:10:48.210 --> 01:10:50.060
but my challenge tell us is like

1539
01:10:50.060 --> 01:10:51.592
let's go figure it out together.

1540
01:10:51.592 --> 01:10:53.432
And I think if we do we're going to do that.

1541
01:10:53.432 --> 01:10:57.170
In terms of the pandemic,

1542
01:10:57.170 --> 01:10:59.997
some of the things that I think are really powerful is

1543
01:10:59.997 --> 01:11:02.000
because there's so much open data

1544
01:11:02.000 --> 01:11:03.810
you just get to play with it.

1545
01:11:03.810 --> 01:11:05.627
Like, what I've been telling students is like

1546
01:11:05.627 --> 01:11:08.070
now this is kind of shut down.

1547
01:11:08.070 --> 01:11:09.680
But like before it was like, Oh

1548
01:11:10.551 --> 01:11:11.713
there's a COVID tracking project.

1549
01:11:11.713 --> 01:11:15.830
Go join the COVID tracker project, like think about

1550
01:11:15.830 --> 01:11:20.450

because we have had such a collapse of our infrastructure

1551
01:11:20.450 --> 01:11:23.220
of our federal solid layer during COVID.

1552
01:11:23.220 --> 01:11:24.800
We have seen the rise

1553
01:11:26.476 --> 01:11:29.623
of the data scientists and citizen scientists jumping

1554
01:11:29.623 --> 01:11:31.173
in to COVID to fill those gaps.

1555
01:11:32.100 --> 01:11:34.126
COVID tracking project COVID act now

1556
01:11:34.126 --> 01:11:36.221
different modeling efforts, different ways

1557
01:11:36.221 --> 01:11:37.740
of playing some of them good

1558
01:11:37.740 --> 01:11:40.503
some of them bad, but creative.

1559
01:11:42.890 --> 01:11:45.157
And you can just get in there and start playing.

1560
01:11:45.157 --> 01:11:49.040
And that is is how I think stuff has

1561
01:11:49.040 --> 01:11:50.670
the models that were built

1562
01:11:50.670 --> 01:11:53.269
for the assessment of California were built

1563
01:11:53.269 --> 01:11:57.393
and put in open source into Amazon servers.

1564
01:11:58.290 --> 01:11:59.740
And so anybody could take that code

1565
01:11:59.740 --> 01:12:01.749
and build and play on it.

1566
01:12:01.749 --> 01:12:04.921
And I would love to see more of that,

1567
01:12:04.921 --> 01:12:07.301
as what's happening in us to think about, well

1568
01:12:07.301 --> 01:12:09.050
what could we open up?

1569
01:12:09.050 --> 01:12:10.230
What can we put out there?

1570
01:12:10.230 --> 01:12:11.740
But then we have to get the word out.

1571
01:12:11.740 --> 01:12:12.910
You have to get advocate.

1572
01:12:12.910 --> 01:12:16.040
It's not okay just to say, oh, it's a website.

1573
01:12:16.040 --> 01:12:19.240
You got to do a little bit more of saying

1574
01:12:19.240 --> 01:12:21.547
come work on this, or here's examples or ideas

1575
01:12:21.547 --> 01:12:23.803

and then it'll carry forward.

1576
01:12:27.530 --> 01:12:28.363
<v ->Thank you.</v>

1577
01:12:30.190 --> 01:12:31.360
<v ->Awesome.</v>

1578
01:12:31.360 --> 01:12:32.580
I know we're running short on time,

1579
01:12:32.580 --> 01:12:34.628
but I want to follow up with another question about COVID.

1580
01:12:34.628 --> 01:12:37.950
So looking back at the beginning

1581
01:12:37.950 --> 01:12:39.620
how could we have used data science

1582
01:12:39.620 --> 01:12:42.119
or AI to better predict the growth and then looking forward

1583
01:12:42.119 --> 01:12:44.790
what lessons can we take away from this?

1584
01:12:44.790 --> 01:12:47.470
And how can we apply data science looking forward

1585
01:12:47.470 --> 01:12:49.972
to whatever kind of coming your way next?

1586
01:12:49.972 --> 01:12:50.805
<v ->Yeah,</v>

1587
01:12:50.805 --> 01:12:51.638
thanks for asking.

1588
01:12:51.638 --> 01:12:53.652
So the first lesson is when COVID was starting,

1589
01:12:53.652 --> 01:12:54.974
the models were toy models.

1590
01:12:54.974 --> 01:12:57.593
Like all of us know that the Seir models

1591
01:12:57.593 --> 01:13:01.634
and all these other different types, but they were not okay,

1592
01:13:01.634 --> 01:13:04.011
or have the ability to really help policy makers.

1593
01:13:04.011 --> 01:13:09.011
So very quickly you have this model

1594
01:13:12.940 --> 01:13:16.410
and you'd be like, ah, look at this exponential growth.

1595
01:13:16.410 --> 01:13:18.661
But then you'd ask one more question

1596
01:13:18.661 --> 01:13:21.220
and then you couldn't answer them.

1597
01:13:21.220 --> 01:13:23.331
Many times the granularity of the models

1598
01:13:23.331 --> 01:13:27.500
did not support the type of questions we want.

1599
01:13:27.500 --> 01:13:30.870
So like, you could have this model for Connecticut

1600
01:13:30.870 --> 01:13:33.531

and that same model is being used for California

1601
01:13:33.531 --> 01:13:37.526
but given the geographic sites, no, there was no, like

1602
01:13:37.526 --> 01:13:38.660
what is transport look like?

1603
01:13:38.660 --> 01:13:41.700
It does not take into account how the different

1604
01:13:41.700 --> 01:13:43.450
community structures are there.

1605
01:13:43.450 --> 01:13:44.720
And so what did it look like

1606
01:13:44.720 --> 01:13:48.420
to take kind of a toy weather model and graduates?

1607
01:13:48.420 --> 01:13:50.160
And the national weather service is one of the reasons

1608
01:13:50.160 --> 01:13:53.710
why we call for a national Reese forecasting center.

1609
01:13:53.710 --> 01:13:55.160
Like at that level, some kids

1610
01:13:58.671 --> 01:14:00.372
we now need to get the infrastructure in place.

1611
01:14:00.372 --> 01:14:01.345
Some of the concrete things that we need to do

1612
01:14:01.345 --> 01:14:02.540
for COVID is we need to do the deep dive

1613
01:14:02.540 --> 01:14:04.784
post-mortem lessons learned that we saw

1614
01:14:04.784 --> 01:14:06.514
in what is the infrastructure we want to place.

1615
01:14:06.514 --> 01:14:08.091
We still need to get massive genomic surveillance in place

1616
01:14:08.091 --> 01:14:10.290
whether it's static testing, randomized testing,

1617
01:14:10.290 --> 01:14:11.303
wastewater don't care

1618
01:14:11.303 --> 01:14:16.303
all of it, check box, all of it, please.

1619
01:14:20.360 --> 01:14:21.730
And then who's gonna collect that data.

1620
01:14:21.730 --> 01:14:22.710
It's going into, at least to CVC

1621
01:14:22.710 --> 01:14:24.361
let's make sure it's opened up, but how do we do more?

1622
01:14:24.361 --> 01:14:26.090
We need to start connecting that with other aspects.

1623
01:14:26.090 --> 01:14:31.090
We still barely understand the aspects

1624
01:14:33.650 --> 01:14:37.130
of mental health issues with regards to what COVID has done.

1625
01:14:37.130 --> 01:14:40.700

What is this done with aspects of other forms of addiction?

1626
01:14:40.700 --> 01:14:43.740
We're just beginning to scratch the surface.

1627
01:14:43.740 --> 01:14:45.640
Why aren't we asking those very specific

1628
01:14:45.640 --> 01:14:46.600
questions right now?

1629
01:14:46.600 --> 01:14:48.400
I know all of this community is

1630
01:14:48.400 --> 01:14:50.550
but where are other people asking that

1631
01:14:51.429 --> 01:14:53.030
and how are we actually making it tangible?

1632
01:14:53.030 --> 01:14:55.471
Those things need to come together.

1633
01:14:55.471 --> 01:14:58.360
Not just because co COVID we're only

1634
01:14:58.360 --> 01:14:59.592
at the halfway time,

1635
01:14:59.592 --> 01:15:02.270
we got to go a long way to go

1636
01:15:02.270 --> 01:15:03.510
before we're through this.

1637
01:15:03.510 --> 01:15:05.560
And then we got the next coronavirus coming

1638
01:15:05.560 --> 01:15:07.186
and this is still not pandemic flu.

1639
01:15:07.186 --> 01:15:11.121
So I look at this as what we've seen right now,

1640
01:15:11.121 --> 01:15:15.173
as the very tip of the iceberg of what we should have done

1641
01:15:15.173 --> 01:15:18.990
20 years ago to start putting in place.

1642
01:15:18.990 --> 01:15:21.250
And I know I'm preaching the choir here

1643
01:15:21.250 --> 01:15:22.965
but I'm emphasizing that to say

1644
01:15:22.965 --> 01:15:25.740
that what we should be really looking at

1645
01:15:25.740 --> 01:15:28.900
what I love for NIDA to be a champion

1646
01:15:28.900 --> 01:15:31.462
of and all this should be a champion

1647
01:15:31.462 --> 01:15:35.890
of his advocacy of what the country needs to do

1648
01:15:35.890 --> 01:15:38.774
for us to put the right infrastructure in place

1649
01:15:38.774 --> 01:15:42.190
so that we can actually build off of that

1650
01:15:42.190 --> 01:15:45.030

and kind of look back in five, 10 years and be like,

1651
01:15:45.030 --> 01:15:47.655
whew thank goodness we put that in place

1652
01:15:47.655 --> 01:15:51.440
because if it and boy was that hard lifting,

1653
01:15:51.440 --> 01:15:55.254
but wow, did we learn so much more?

1654
01:15:55.254 --> 01:15:58.670
Did we help change the trajectory because

1655
01:15:58.670 --> 01:15:59.910
of one small thing?

1656
01:15:59.910 --> 01:16:02.226
And I just want to emphasize this,

1657
01:16:02.226 --> 01:16:07.226
change at the federal national level that is scalable,

1658
01:16:08.840 --> 01:16:13.113
doesn't always have to happen because of a big bang project.

1659
01:16:14.792 --> 01:16:16.900
The fact that we have organ donations,

1660
01:16:16.900 --> 01:16:20.220
the way we do is literally

1661
01:16:20.220 --> 01:16:23.260
because a couple people got this tiny little bit

1662
01:16:23.260 --> 01:16:24.970
of this thing in there that said it should be done

1663
01:16:24.970 --> 01:16:26.083
at the DMV.

1664
01:16:27.460 --> 01:16:31.240
And so little things can have profound,

1665
01:16:31.240 --> 01:16:35.053
scalable impact given from the perch that we all occupy.

1666
01:16:36.660 --> 01:16:38.933
<v ->I believe Dr. Volkow has a comment.</v>

1667
01:16:40.909 --> 01:16:42.700
<v ->Is the day, thanks for a really terrific talk</v>

1668
01:16:42.700 --> 01:16:43.533
and dialogue.

1669
01:16:43.533 --> 01:16:45.163
And I had asked you the question on the COVID

1670
01:16:45.163 --> 01:16:48.130
because as we look forward, I mean,

1671
01:16:48.130 --> 01:16:49.880
you go on to learn from history.

1672
01:16:49.880 --> 01:16:53.230
And I guess that we didn't have the models

1673
01:16:53.230 --> 01:16:55.130
or all of the data that we wanted

1674
01:16:55.130 --> 01:16:58.430
but still there was data and get we neglected it.

1675
01:16:58.430 --> 01:16:59.650

And I think that in terms

1676
01:16:59.650 --> 01:17:01.958
of your question is how do we poised ourselves?

1677
01:17:01.958 --> 01:17:03.857
Because we have a responsibility

1678
01:17:03.857 --> 01:17:06.550
in terms of bringing forward what we think

1679
01:17:06.550 --> 01:17:10.780
are issues that we cannot afford ignore or neglect.

1680
01:17:10.780 --> 01:17:12.184
And in our case, it's of course the epidemic

1681
01:17:12.184 --> 01:17:15.524
of overdose deaths that we're observing

1682
01:17:15.524 --> 01:17:18.760
and for which we actually don't have accurate data.

1683
01:17:18.760 --> 01:17:20.919
So this has got to be very frustrating

1684
01:17:20.919 --> 01:17:23.690
but in more general terms, as I think

1685
01:17:23.690 --> 01:17:26.110
and I would love to hear your perspective.

1686
01:17:26.110 --> 01:17:28.153
It just strikes me and you say it doesn't need

1687
01:17:28.153 --> 01:17:31.800
to be these well-organized programs or projects.

1688
01:17:31.800 --> 01:17:33.840
And one of the most vulnerable things that I've

1689
01:17:33.840 --> 01:17:37.568
seen emerge is the science of citizen science.

1690
01:17:37.568 --> 01:17:41.030
And yet one of the infrastructures

1691
01:17:41.030 --> 01:17:44.318
that will allow us as a country advance much faster

1692
01:17:44.318 --> 01:17:47.480
and we don't have is an internet access to

1693
01:17:47.480 --> 01:17:49.440
everyone in the United States.

1694
01:17:49.440 --> 01:17:51.252
So my question is, because again,

1695
01:17:51.252 --> 01:17:53.900
what solutions are we bothering for?

1696
01:17:53.900 --> 01:17:55.590
How would you go about these?

1697
01:17:55.590 --> 01:17:59.223
What is your thinking on, on how to make this possible?

1698
01:18:00.510 --> 01:18:02.950
<v ->Yeah, the infrastructure of the internet</v>

1699
01:18:02.950 --> 01:18:06.508
is COVID is exposed how bad the digital divide is.

1700
01:18:06.508 --> 01:18:10.400

You know, in one of the most painful,

1701
01:18:10.400 --> 01:18:15.400
I spent the first hundred plus days of like a year ago

1702
01:18:15.730 --> 01:18:17.963
working on the California COVID response.

1703
01:18:17.963 --> 01:18:21.053
And some of the things that kind of came up were a lot

1704
01:18:21.053 --> 01:18:24.650
of people brought up technical solutions that really

1705
01:18:24.650 --> 01:18:27.163
only benefited one portion of the community.

1706
01:18:28.340 --> 01:18:30.363
The exposure notification apps.

1707
01:18:32.109 --> 01:18:33.950
These were fundamentally not going to be effective

1708
01:18:33.950 --> 01:18:35.958
for large parts of the population,

1709
01:18:35.958 --> 01:18:38.725
because they don't have devices

1710
01:18:38.725 --> 01:18:40.230
that are the most up to date.

1711
01:18:40.230 --> 01:18:42.160
There are surveillance questions, there's all sorts

1712
01:18:42.160 --> 01:18:45.564
of things that people hadn't thought about.

1713
01:18:45.564 --> 01:18:49.950
And one of the takeaways that I realized

1714
01:18:49.950 --> 01:18:51.960
that I personally took away is

1715
01:18:51.960 --> 01:18:56.960
that many people who work on technology haven't really lived

1716
01:18:58.530 --> 01:19:00.470
in the communities that are most impacted

1717
01:19:00.470 --> 01:19:02.223
by these aspects.

1718
01:19:04.554 --> 01:19:05.387
And I've talked to a lot of technologist

1719
01:19:05.387 --> 01:19:07.210
friends who are from the classic Apple

1720
01:19:07.210 --> 01:19:10.280
and Googles and Amazons, all the big companies.

1721
01:19:10.280 --> 01:19:12.310
And I said, great, like, let's suppose we have

1722
01:19:12.310 --> 01:19:14.370
exposure notification, but they live

1723
01:19:14.370 --> 01:19:16.229
in a multi-generational household

1724
01:19:16.229 --> 01:19:19.060
or they're migrant workers.

1725
01:19:19.060 --> 01:19:22.080

Who's going to pay for them to isolate, like what

1726
01:19:22.080 --> 01:19:23.730
like you pull on it a little bit,

1727
01:19:24.782 --> 01:19:26.120
like this collapses, right?

1728
01:19:26.120 --> 01:19:29.013
Like what's the next level of answer.

1729
01:19:30.843 --> 01:19:33.230
And so that needs to be a holistic policy process

1730
01:19:33.230 --> 01:19:37.102
and problem, the other one, which is there,

1731
01:19:37.102 --> 01:19:38.763
and it sounds trivial.

1732
01:19:42.505 --> 01:19:46.050
Like I spent a good chunk of my time building data

1733
01:19:46.050 --> 01:19:47.710
dictionaries because we got all this data

1734
01:19:47.710 --> 01:19:49.071
coming in and there was no data dictionary

1735
01:19:49.071 --> 01:19:49.950
so nobody would use the data.

1736
01:19:49.950 --> 01:19:53.280
And so there's a lot of this very simple stuff

1737
01:19:56.090 --> 01:19:57.281
that isn't there, the digital divide one that you're

1738
01:19:57.281 --> 01:19:58.114
pointing out, but I think it was the most problematic is

1739
01:19:58.114 --> 01:19:59.120
on education.

1740
01:19:59.120 --> 01:20:02.350
And the aspects of this as like people just

1741
01:20:02.350 --> 01:20:04.831
don't have access to the internet.

1742
01:20:04.831 --> 01:20:06.000
Then we spent a lot of our policy time.

1743
01:20:06.000 --> 01:20:08.250
Actually this goes back to Clinton's

1744
01:20:09.370 --> 01:20:10.905
initial initiatives, I'm trying to wire schools

1745
01:20:10.905 --> 01:20:13.840
but we've found is like wiring schools,

1746
01:20:13.840 --> 01:20:16.810
you've got to wire people wherever they're going to be.

1747
01:20:16.810 --> 01:20:20.120
And I think that that cellular connectivity

1748
01:20:20.120 --> 01:20:23.020
other aspects are showing that they're promises there.

1749
01:20:23.020 --> 01:20:25.596
But I think we need to really ask ourselves was

1750
01:20:25.596 --> 01:20:28.540

going to look like over the next decade because

1751
01:20:28.540 --> 01:20:31.121
the technologies that are about to shift quite a bit

1752
01:20:31.121 --> 01:20:36.121
and what we need is actually a much bigger plan.

1753
01:20:36.427 --> 01:20:39.860
And the model that I actually think is the best is

1754
01:20:39.860 --> 01:20:44.300
what does it mean to wire and connect tribal lands.

1755
01:20:44.300 --> 01:20:47.420
If we can figure out how to wire tribal lands

1756
01:20:47.420 --> 01:20:49.020
we're gonna figure this out now,

1757
01:20:51.601 --> 01:20:52.434
why am I not giving a concrete proposal is honestly,

1758
01:20:52.434 --> 01:20:53.860
because I think we need to take an approach

1759
01:20:53.860 --> 01:20:56.854
that we call scout and scale is

1760
01:20:56.854 --> 01:20:59.900
that we need to deploy funds and for a number

1761
01:20:59.900 --> 01:21:02.032
of experiments to take place around the country

1762
01:21:02.032 --> 01:21:03.350
figure out which ones are working, take the best

1763
01:21:03.350 --> 01:21:06.580
of them within a span of three years, and then figure

1764
01:21:06.580 --> 01:21:09.376
out based on that, what is the scalable approach

1765
01:21:09.376 --> 01:21:11.020
that it needs to be deployed across the country?

1766
01:21:11.020 --> 01:21:13.060
Because if we try to say just here's how the

1767
01:21:13.060 --> 01:21:15.910
from the top down approach and figure it out.

1768
01:21:15.910 --> 01:21:17.530
I know I appreciate the choir

1769
01:21:17.530 --> 01:21:18.590
but this is the classic.

1770
01:21:18.590 --> 01:21:19.460
Like we just kind of come down

1771
01:21:19.460 --> 01:21:21.090
from the top and then the local community's like

1772
01:21:21.090 --> 01:21:23.240
that's not gonna work for ours.

1773
01:21:23.240 --> 01:21:25.190
And so we need to actually make this a little bit

1774
01:21:25.190 --> 01:21:26.323
more organic,

1775
01:21:27.549 --> 01:21:30.150

but kind of make the soil more fertile

1776
01:21:30.150 --> 01:21:31.770
for the experiments to take place

1777
01:21:31.770 --> 01:21:32.603
and then figure out

1778
01:21:32.603 --> 01:21:35.703
what's going to kind of take action and to be effective.

1779
01:21:37.370 --> 01:21:39.460
<v ->And so how can we help as an agency?</v>

1780
01:21:39.460 --> 01:21:41.910
Because when there, and I also think that, I mean

1781
01:21:43.531 --> 01:21:44.550
I completely agree that you want to do this sort of

1782
01:21:44.550 --> 01:21:47.560
in a way, large pilot projects that will allow us to

1783
01:21:47.560 --> 01:21:48.880
understand what works the best

1784
01:21:48.880 --> 01:21:50.680
because there's not going to be one solution

1785
01:21:50.680 --> 01:21:52.730
but how do we start?

1786
01:21:52.730 --> 01:21:55.110
How do we contribute for that to happen?

1787
01:21:55.110 --> 01:21:56.910
Because this is at the essence

1788
01:21:56.910 --> 01:21:58.650
on the one hand, as you say, location

1789
01:21:58.650 --> 01:22:01.987
but also for being able to obtain data.

1790
01:22:01.987 --> 01:22:05.000
And so it's like the soil

1791
01:22:05.000 --> 01:22:09.923
from which data science can really blossom.

1792
01:22:11.060 --> 01:22:11.893
<v ->Yeah.</v>

1793
01:22:11.893 --> 01:22:13.190
So thanks for bringing this up.

1794
01:22:13.190 --> 01:22:17.000
So that very genuine answer is we have to

1795
01:22:17.000 --> 01:22:18.523
become better communicators.

1796
01:22:19.990 --> 01:22:21.530
I think we're really good

1797
01:22:21.530 --> 01:22:24.820
at communicating oftentimes to our peers and our colleagues.

1798
01:22:24.820 --> 01:22:27.410
We are not effective at the level

1799
01:22:27.410 --> 01:22:31.870
we need to be with society and broader policy makers.

1800
01:22:31.870 --> 01:22:34.960

So, what does that translate into

1801
01:22:35.821 --> 01:22:39.718
it actually is communicating in more novel methods.

1802
01:22:39.718 --> 01:22:44.718
And what I mean by that is not journal articles.

1803
01:22:46.520 --> 01:22:48.050
And I'm not saying at all

1804
01:22:48.050 --> 01:22:51.860
like get rid of journal, this is an ant conversation.

1805
01:22:51.860 --> 01:22:52.693
It is.

1806
01:22:52.693 --> 01:22:56.400
How do we, whether it's writing books, doing podcasts

1807
01:22:56.400 --> 01:23:01.400
blog posts, tech, docs, Twitter, I don't know.

1808
01:23:01.570 --> 01:23:05.240
We have to try a much bigger

1809
01:23:05.240 --> 01:23:09.317
for actually engaging the public, helping them.

1810
01:23:09.317 --> 01:23:11.080
We need to give them stories.

1811
01:23:11.080 --> 01:23:13.590
We need to help them see why this is important

1812
01:23:13.590 --> 01:23:15.390
help them advocate write.

1813
01:23:15.390 --> 01:23:17.298
Op-eds do all these things.

1814
01:23:17.298 --> 01:23:18.140
And some of this feels

1815
01:23:18.140 --> 01:23:22.280
it feels weird because we're all as you know, for those

1816
01:23:22.280 --> 01:23:26.670
especially that are civil servants in the career force.

1817
01:23:26.670 --> 01:23:28.020
There's challenges with that.

1818
01:23:28.020 --> 01:23:29.153
There are real challenges with that.

1819
01:23:29.153 --> 01:23:31.927
I don't want to dismiss that, but we have to get out there.

1820
01:23:31.927 --> 01:23:34.640
And I think Francis actually is a role model on this

1821
01:23:34.640 --> 01:23:36.460
of being out there and communicating

1822
01:23:36.460 --> 01:23:38.780
but we need Tony is to Tony Fowchee, but more

1823
01:23:38.780 --> 01:23:42.370
we need more of that communication.

1824
01:23:42.370 --> 01:23:43.490
We need to engage.

1825
01:23:43.490 --> 01:23:45.528

And NIH has really good.

1826
01:23:45.528 --> 01:23:47.730
This we need to, I need you all to engage more

1827
01:23:47.730 --> 01:23:52.410
with the Hill and also States state leadership.

1828
01:23:52.410 --> 01:23:54.540
We forget, we need to find a way to engage

1829
01:23:54.540 --> 01:23:58.360
with the governors and the mayors and that communication.

1830
01:23:58.360 --> 01:24:01.922
Once we start to do that, the dialogue will follow.

1831
01:24:01.922 --> 01:24:04.980
I have found almost always when we do a little bit

1832
01:24:04.980 --> 01:24:07.390
of that storytelling, you open a door and you

1833
01:24:07.390 --> 01:24:10.980
start to get advocacy advocates that start carrying us

1834
01:24:10.980 --> 01:24:13.650
in a different methodology

1835
01:24:13.650 --> 01:24:16.080
that allows us to engage with the public

1836
01:24:16.080 --> 01:24:20.160
and the place where we see that the most.

1837
01:24:20.160 --> 01:24:22.410
And if you kind of look at these stories over time

1838
01:24:22.410 --> 01:24:24.850
whether it's cancer or anything saying things

1839
01:24:24.850 --> 01:24:27.970
other things I have come a belief that is a community

1840
01:24:29.021 --> 01:24:31.030
organizing efforts that allow that make

1841
01:24:32.498 --> 01:24:36.153
the messaging to be more effective for dispersal of ideas.

1842
01:24:42.280 --> 01:24:43.113
<v ->And I think that,</v>

1843
01:24:43.113 --> 01:24:45.830
and I was looking at you just got three minutes

1844
01:24:45.830 --> 01:24:47.815
got another aspect that we have been

1845
01:24:47.815 --> 01:24:51.180
all of our sensitize has been the whole notion

1846
01:24:51.180 --> 01:24:52.680
of the health disparities

1847
01:24:52.680 --> 01:24:55.750
which drives on into socio-economical differences

1848
01:24:55.750 --> 01:24:57.717
poorer and poorer location is a fundamental

1849
01:24:57.717 --> 01:25:01.313
in all of these distinctions.

1850
01:25:02.411 --> 01:25:04.072

And as, I mean, I made so very clear with COVID

1851
01:25:04.072 --> 01:25:06.586
but did you think about it?

1852
01:25:06.586 --> 01:25:09.367
I mean, with the access to the internet and

1853
01:25:09.367 --> 01:25:13.340
with all of the advances in technology that have developed

1854
01:25:13.340 --> 01:25:15.635
with COVID who communicate with one another

1855
01:25:15.635 --> 01:25:17.410
we now have the opportunity

1856
01:25:18.785 --> 01:25:20.700
of providing high quality education to everyone.

1857
01:25:20.700 --> 01:25:23.280
It will be a challenge to create programs that are

1858
01:25:23.280 --> 01:25:24.620
actually going to be salient

1859
01:25:24.620 --> 01:25:26.950
for children and keep their attention.

1860
01:25:26.950 --> 01:25:29.070
But if we can do it for entertainment

1861
01:25:29.070 --> 01:25:31.320
we can do it for inpatient.

1862
01:25:31.320 --> 01:25:33.500
And so we have the means, right?

1863
01:25:33.500 --> 01:25:36.219
To modify these terrible health disparities.

1864
01:25:36.219 --> 01:25:38.533
And again, I think

1865
01:25:38.533 --> 01:25:40.920
that the way what you are saying is we need to

1866
01:25:40.920 --> 01:25:44.740
communicate better and to basically help guide the points

1867
01:25:44.740 --> 01:25:46.890
about why this is so fundamental

1868
01:25:46.890 --> 01:25:50.520
to be able to provide everyone with these access

1869
01:25:50.520 --> 01:25:53.837
to communicating, obtaining data, sending information

1870
01:25:53.837 --> 01:25:56.823
and being part of those, that giving information.

1871
01:25:58.449 --> 01:25:59.430
So I got the message

1872
01:25:59.430 --> 01:26:01.683
but I think at that, thanks you very much.

1873
01:26:03.287 --> 01:26:04.570
It has been terrific having on the

1874
01:26:04.570 --> 01:26:06.670
we would love to steal your brain

1875
01:26:06.670 --> 01:26:08.583

because you're always going to be

1876
01:26:08.583 --> 01:26:11.040
around how to get these things going.

1877
01:26:11.040 --> 01:26:13.990
And as you saw at the essence of data, science

1878
01:26:13.990 --> 01:26:17.502
is the painting data from very different perspectives.

1879
01:26:17.502 --> 01:26:20.713
And you have given us a very novel and unique one.

1880
01:26:23.052 --> 01:26:25.265
<v ->I guess, maybe close with this as it should is like</v>

1881
01:26:25.265 --> 01:26:30.265
first part is the project that

1882
01:26:31.430 --> 01:26:35.940
I didn't have time to run through at the white house

1883
01:26:35.940 --> 01:26:38.500
which is the one that I wish we had kind of started with

1884
01:26:38.500 --> 01:26:40.450
which is actually the issue of poverty.

1885
01:26:41.340 --> 01:26:43.580
You know, whether it's data on evictions

1886
01:26:43.580 --> 01:26:44.800
which doesn't exist

1887
01:26:44.800 --> 01:26:47.690
there's no real good data on evictions, criminal justice.

1888
01:26:47.690 --> 01:26:50.250
They all kind of are around the whole idea of poverty.

1889
01:26:50.250 --> 01:26:54.052
And what is the impact of that health disparities?

1890
01:26:54.052 --> 01:26:55.783
You've kind of seen

1891
01:26:55.783 --> 01:26:56.910
that the hurt classic Raj Chetty work

1892
01:26:56.910 --> 01:26:58.020
others there's so much federal data

1893
01:26:58.020 --> 01:27:00.670
so much state city data that could be opened

1894
01:27:00.670 --> 01:27:03.650
up pulled together that would give us phenomenal insights

1895
01:27:03.650 --> 01:27:05.390
on these different disparities.

1896
01:27:05.390 --> 01:27:06.617
And what would that start to look like?

1897
01:27:06.617 --> 01:27:08.599
And especially if it was funded appropriately

1898
01:27:08.599 --> 01:27:11.060
it's been started funded

1899
01:27:11.966 --> 01:27:12.979
by philanthropic efforts, Arnold foundation, others Gates

1900
01:27:12.979 --> 01:27:14.947

but that's not sufficient and it's not sustainable.

1901
01:27:14.947 --> 01:27:18.645
And we need that level of funding to have to be put

1902
01:27:18.645 --> 01:27:23.433
in place gun violence is in there as well.

1903
01:27:25.150 --> 01:27:29.560
The other part I think of this is I just want to

1904
01:27:29.560 --> 01:27:32.723
say thank you to all of you,

1905
01:27:35.136 --> 01:27:37.569
that one of the things I think that is important

1906
01:27:37.569 --> 01:27:38.551
to call it is, with the data science skills

1907
01:27:38.551 --> 01:27:41.460
you can pretty much go do anything you want

1908
01:27:41.460 --> 01:27:44.000
but you're choosing to do this.

1909
01:27:44.000 --> 01:27:47.963
And in my opinion, there is no more noble profession

1910
01:27:47.963 --> 01:27:50.543
than the ones that you all are on.

1911
01:27:51.840 --> 01:27:54.110
Being able to work with NIH

1912
01:27:54.110 --> 01:27:57.018
being able to work on this mission, you fundamentally are

1913
01:27:57.018 --> 01:28:01.260
at the helm of where the nation is going

1914
01:28:01.260 --> 01:28:03.908
on the most important issues of our day

1915
01:28:03.908 --> 01:28:08.908
and what you have the ability to do what your charter is

1916
01:28:09.640 --> 01:28:14.388
what your mission is, is the thing that I think when

1917
01:28:14.388 --> 01:28:17.060
the country doesn't always appreciate it

1918
01:28:17.060 --> 01:28:19.227
doesn't always get to see until much later.

1919
01:28:19.227 --> 01:28:24.220
And they often, when they don't have it

1920
01:28:24.220 --> 01:28:27.490
secretary Carter used to say, still says

1921
01:28:27.490 --> 01:28:28.910
is security is like air.

1922
01:28:28.910 --> 01:28:31.462
You only know you need it when you don't have it.

1923
01:28:31.462 --> 01:28:34.312
That's been exposed during COVID.

1924
01:28:34.312 --> 01:28:37.960
And so I think you all are the air, you're the oxygen

1925
01:28:37.960 --> 01:28:40.627

for our society, for what our kids are going to need.

1926
01:28:40.627 --> 01:28:43.160
And so I just want to express my gratitude

1927
01:28:43.160 --> 01:28:45.323
and thanks for what y'all do.

1928
01:28:47.060 --> 01:28:48.210
<v ->Thank you so much, DJ.</v>

1929
01:28:48.210 --> 01:28:50.570
This has been amazing talking to you and hearing all

1930
01:28:50.570 --> 01:28:52.420
about your career and hearing their insights.

1931
01:28:52.420 --> 01:28:54.740
And I think it's great for us to kind of learn how

1932
01:28:54.740 --> 01:28:57.380
we can use these insights to bring more data science

1933
01:28:57.380 --> 01:28:58.640
to addiction research.

1934
01:28:58.640 --> 01:29:00.310
So I really appreciate your time

1935
01:29:00.310 --> 01:29:01.840
and I want to thank the audience as well.

1936
01:29:01.840 --> 01:29:02.850
I'm sorry we didn't get a chance

1937
01:29:02.850 --> 01:29:03.700
to get to all the questions.

1938
01:29:03.700 --> 01:29:05.260
I think some of them were kind of touched on some

1939
01:29:05.260 --> 01:29:07.950
of the previous questions and I also want to thank again to

1940
01:29:07.950 --> 01:29:10.900
at T team that helped organize this and the technical team.

1941
01:29:11.825 --> 01:29:12.900
And we do have another seminar next Monday

1942
01:29:12.900 --> 01:29:14.988
at 9:00 AM on the side is going up.

1943
01:29:14.988 --> 01:29:17.450
We have Kirk borne from Booz Allen

1944
01:29:17.450 --> 01:29:20.240
and Martin Paulus who Laureate brain Institute.

1945
01:29:20.240 --> 01:29:21.980
So I hope you'll tune in and information

1946
01:29:21.980 --> 01:29:24.677
on the next two seminars after that it will come in.

1947
01:29:25.948 --> 01:29:26.927
So thanks again for your attention.

1948
01:29:26.927 --> 01:29:29.277
And this recording will be made available soon.

1949
01:29:30.260 --> 01:29:31.093
Bye everyone.