WEBVTT

1
00:00:11.625 --> 00:00:15.295
Hello, I'm Llewellyn King, host of White House Chronicle.

2
00:00:15.465 --> 00:00:16.655
Thank you for coming along.

3
00:00:16.835 --> 00:00:21.575
I'm joined as always by the talented Adam Clayton Powell III,

4
00:00:21.915 --> 00:00:23.175
the co-host of the program.

5
00:00:23.755 --> 00:00:27.695
And today we're going to take a look at the impact of

6
00:00:28.655 --> 00:00:30.975
artificial intelligence on the law

7
00:00:31.275 --> 00:00:34.695
and other aspects of how we live today. To guide us

8
00:00:34.695 --> 00:00:35.775
through this jungle,

9
00:00:36.155 --> 00:00:38.935
we have, very fortunately, somebody

10
00:00:39.075 --> 00:00:40.855
who is a true expert on it.

11
00:00:41.245 --> 00:00:44.815
He's Roland Trope of the law firm of Trope

12
00:00:44.815 --> 00:00:47.655
and Schramm LLP in New York City.

13
00:00:48.035 --> 00:00:51.975
And he has been concentrating his career on cybersecurity

14
00:00:52.075 --> 00:00:54.455
and more recently on the impact

15
00:00:54.715 --> 00:00:57.895
of artificial intelligence in that arena.

16
00:00:58.125 --> 00:00:59.615
Welcome to the broadcast.

17
00:01:00.075 --> 00:01:05.015
And let me ask you, is cybersecurity impacted by

18
00:01:05.995 --> 00:01:08.815
AI and is it impacted for good or for bad?

19
00:01:10.065 --> 00:01:13.115
Well, they each have an effect on the other.

20
00:01:13.985 --> 00:01:17.235
Cybersecurity, if it's not maintained by the developers

21
00:01:17.255 --> 00:01:21.555
of AI, can adversely affect the training of a model. Think

22
00:01:21.555 --> 00:01:25.755
of feeding a model, uh, data that would cause it

23
00:01:25.755 --> 00:01:28.395
to be less accurate in its predictions.

24
00:01:29.135 --> 00:01:30.515
So in that regard, it's important

25
00:01:30.515 --> 00:01:32.995

for the developers throughout the training and

26
00:01:33.205 --> 00:01:35.835
after deployment of the model as they retrain it

27
00:01:35.835 --> 00:01:39.155
to keep it up to date, to keep the supply of data

28
00:01:39.215 --> 00:01:41.155
and the model itself cyber secure.

29
00:01:41.815 --> 00:01:45.235
The flip side of that is that AI makes it possible

30
00:01:45.575 --> 00:01:49.955
for adversaries to execute ever increasingly

31
00:01:50.065 --> 00:01:52.195
effective, inexpensive,

32
00:01:52.335 --> 00:01:55.435
and devastating cyber attacks.

33
00:01:56.005 --> 00:01:58.675
Think of the range of things that we now do with AI,

34
00:01:58.675 --> 00:02:00.035
ranging from deep fakes

35
00:02:00.575 --> 00:02:05.275
to execute fraudulent financial transactions to

36
00:02:05.565 --> 00:02:09.755
conducting attacks on various critical infrastructure.

37
00:02:09.935 --> 00:02:13.035
And apparently AI enables the adversary

38
00:02:13.575 --> 00:02:15.995
to not only operate faster than before,

39
00:02:16.055 --> 00:02:18.435
but to get into the target's decision loop

40
00:02:18.655 --> 00:02:22.035
and deprive them of the situation awareness they need in

41
00:02:22.035 --> 00:02:23.555
order to recover from the attack.

42
00:02:24.995 --> 00:02:26.635
Roland, you are on the Law

43
00:02:26.695 --> 00:02:28.595
and Artificial Intelligence Task Force

44
00:02:28.615 --> 00:02:30.355
of the American Bar Association.

45
00:02:30.935 --> 00:02:32.515
You and your colleagues

46
00:02:32.515 --> 00:02:36.985
have been discussing some recent announcements that AI,

47
00:02:37.325 --> 00:02:40.425
AI and AI vendors have been making about

48
00:02:41.455 --> 00:02:45.075
how their work affects you, other lawyers,

49
00:02:45.535 --> 00:02:47.155
law firms, and your clients.

50
00:02:48.975 --> 00:02:51.475

We have been trying to find out,

51
00:02:51.785 --> 00:02:56.435
what impact AI will have on law, the practice of law,

52
00:02:56.775 --> 00:02:58.595
the conduct of the judiciary.

53
00:02:59.225 --> 00:03:02.525
And what we're finding is that there's a number of pushing

54
00:03:02.665 --> 00:03:03.805
and pulling going on.

55
00:03:04.145 --> 00:03:06.285
The vendors are trying to push this bright,

56
00:03:06.295 --> 00:03:09.605
shiny power tool, as I call it, which is the way I refer

57
00:03:09.605 --> 00:03:12.085
to generative AI as opposed to conventional AI,

58
00:03:12.085 --> 00:03:14.245
which I consider a precision instrument.

59
00:03:15.075 --> 00:03:17.685
Clients are contacting their counsel,

60
00:03:17.785 --> 00:03:20.845
asking them in writing, tell us how you're going

61
00:03:20.905 --> 00:03:24.005
to get work done faster, more efficiently,

62
00:03:24.345 --> 00:03:26.045
and reduce our legal costs.

63
00:03:26.855 --> 00:03:30.685
Those cases have yet to be proven possible.

64
00:03:33.705 --> 00:03:35.965
I'm wondering, you taught I read

65
00:03:36.225 --> 00:03:39.165
and actually taught at West Point until last year.

66
00:03:39.555 --> 00:03:43.325
Were you teaching cybersecurity through AI or teaching AI?

67
00:03:43.785 --> 00:03:47.405
And what impact is it going to have on the military?

68
00:03:48.595 --> 00:03:51.125
I've taught at West Point since 1992.

69
00:03:51.945 --> 00:03:54.965
For part of that time I was teaching

70
00:03:54.965 --> 00:03:56.365
defense contracting issues,

71
00:03:56.425 --> 00:03:57.805
but for the last 20 years,

72
00:03:58.315 --> 00:03:59.605
I've been teaching in the Department

73
00:03:59.605 --> 00:04:01.565
of Electrical Engineering and Computer Science,

74
00:04:02.105 --> 00:04:05.285
and they wanted me to talk about intellectual property,

75
00:04:05.395 --> 00:04:08.565

copyright, trademark, patents, trade secrets.

76
00:04:09.105 --> 00:04:10.365
And they particularly wanted that

77
00:04:10.365 --> 00:04:13.285
because at that time there was a considerable problem with

78
00:04:13.885 --> 00:04:16.965
students, including students at the military academies,

79
00:04:17.605 --> 00:04:20.045
downloading without authorization copyrighted works.

80
00:04:20.865 --> 00:04:23.085
10 years ago, the department asked me,

81
00:04:23.085 --> 00:04:25.805
would you shift from giving examples of music

82
00:04:26.145 --> 00:04:29.765
and images to giving examples related to AI

83
00:04:30.075 --> 00:04:33.085
because the cadets are going to increasingly need

84
00:04:33.085 --> 00:04:34.285
to know about AI.

85
00:04:34.665 --> 00:04:37.445
And so since then, my lectures tend to divide up

86
00:04:37.445 --> 00:04:42.405
between talking about ai, its capabilities, its limitations,

87
00:04:42.665 --> 00:04:46.165
its risks, and the intersection it has with copyright.

88
00:04:47.265 --> 00:04:48.625
Roland, you just made an interesting distinction

89
00:04:48.625 --> 00:04:51.825
between two different kinds of AI, one precise

90
00:04:51.965 --> 00:04:55.265
and one, uh, maybe we could say is more creative.

91
00:04:55.475 --> 00:04:56.585
Could you go into that a bit?

92
00:04:57.865 --> 00:05:00.425
A study published in Lancet,

93
00:05:00.785 --> 00:05:03.625
a distinguished European medical journal, back in August

94
00:05:03.645 --> 00:05:07.905
of 2023, found in its study that the use

95
00:05:07.905 --> 00:05:10.105
of AI to enhance the reading

96
00:05:10.645 --> 00:05:15.185
of mammograms enabled the detection of 20% more cancers

97
00:05:15.835 --> 00:05:16.865
years in advance

98
00:05:16.885 --> 00:05:20.905
of when two trained radiographers could have done it

99
00:05:21.205 --> 00:05:24.065
and without increasing the number of false positives.

100
00:05:24.405 --> 00:05:27.145

And when I go into radiology clinics in Manhattan,

101
00:05:27.235 --> 00:05:28.745
there are usually these ceiling

102
00:05:28.885 --> 00:05:33.225
to floor banners urging women to opt in to use AI

103
00:05:33.765 --> 00:05:35.665
for the reading of those results.

104
00:05:38.025 --> 00:05:40.725
AI in that regard is a very precise instrument.

105
00:05:40.785 --> 00:05:43.125
You are relying on it for extraordinary outcomes.

106
00:05:43.475 --> 00:05:45.445
It's used by astronomers, scientists,

107
00:05:45.725 --> 00:05:47.885
engineers in those very precise ways.

108
00:05:49.235 --> 00:05:51.765
It's a prediction machine when it does that,

109
00:05:52.185 --> 00:05:55.125
and it predicts in radiography two kinds of prediction.

110
00:05:55.625 --> 00:05:57.685
One is it looks at the pixel level,

111
00:05:57.685 --> 00:06:00.285
which is why it can see things that we can't

112
00:06:00.445 --> 00:06:03.045
'cause we cannot look at, at the pixel level,

113
00:06:03.705 --> 00:06:07.455
it can see emergent patterns of pixels

114
00:06:07.845 --> 00:06:10.095
that it can predict, already

115
00:06:10.775 --> 00:06:12.855
identify a tumor or an emerging one.

116
00:06:13.235 --> 00:06:17.255
But it can also, more recently it's been able to say

117
00:06:17.965 --> 00:06:19.295
with the thousands

118
00:06:19.295 --> 00:06:23.895
and millions of, uh, exams that it has looked at,

119
00:06:24.475 --> 00:06:27.615
it is now capable of predicting certain areas

120
00:06:27.625 --> 00:06:29.455
where the cancers will will emerge

121
00:06:29.455 --> 00:06:31.135
and where they need to be monitored.

122
00:06:31.605 --> 00:06:35.815
Compare that now to a subset of AI, which is generative AI.

123
00:06:36.065 --> 00:06:37.735
Generative AI is

124
00:06:38.375 --> 00:06:41.215
a tool in which the user enters a prompt

125
00:06:41.435 --> 00:06:44.855

and the prompt can be voice generated, it can be typing in

126
00:06:44.855 --> 00:06:48.575
of text, dropping in of an image that gets fed into a model,

127
00:06:48.635 --> 00:06:50.775
and the model then generates an output

128
00:06:51.115 --> 00:06:52.575
and the output is new.

129
00:06:53.445 --> 00:06:57.595
It's predicting the answer that the user requested.

130
00:06:57.895 --> 00:07:01.355
And what's important is to realize that generative ai,

131
00:07:01.705 --> 00:07:05.955
even though its output when it's text mimics conversation,

132
00:07:07.005 --> 00:07:10.345
it doesn't understand language, it doesn't think,

133
00:07:11.255 --> 00:07:14.115
and it is only predicting the next word

134
00:07:14.135 --> 00:07:16.715
or the next string of words based on the quality

135
00:07:16.775 --> 00:07:17.995
of the input you give it.

136
00:07:18.015 --> 00:07:20.075
And you can give it repeated inputs

137
00:07:20.075 --> 00:07:22.675
or you can give it longer and longer strings of inputs

138
00:07:22.855 --> 00:07:25.315
to refine what it generates out.

139
00:07:25.695 --> 00:07:30.515
But because of the significant, uh, drawbacks

140
00:07:31.265 --> 00:07:34.875
such as the ability to generate inaccurate information,

141
00:07:35.525 --> 00:07:38.315
false information, misleading information,

142
00:07:38.725 --> 00:07:42.115
incomplete information, information that's not up to date,

143
00:07:43.485 --> 00:07:46.185
it should always be remembered as something

144
00:07:46.185 --> 00:07:47.825
that's operating off of a model

145
00:07:48.605 --> 00:07:51.705
and is as often said, all models are wrong,

146
00:07:51.895 --> 00:07:53.145
some models are useful.

147
00:07:53.405 --> 00:07:56.945
And if you keep that in mind, you won't fall prey

148
00:07:57.365 --> 00:07:59.785
to thinking that it thinks. If you think

149
00:07:59.785 --> 00:08:00.905
it thinks, think again.

150
00:08:01.605 --> 00:08:03.905

And the reason I'm saying that is

151
00:08:03.905 --> 00:08:07.425
that the developers have increasingly used language

152
00:08:07.565 --> 00:08:11.465
to encourage we, the tool users to think that the tool

153
00:08:12.045 --> 00:08:16.425
is not only conversational mimicking, but is an agent.

154
00:08:17.085 --> 00:08:19.985
And so they'll use terms like hallucination

155
00:08:20.085 --> 00:08:21.825
to describe an inaccurate output.

156
00:08:22.175 --> 00:08:24.425
It's not hallucinating, that's a metaphor,

157
00:08:24.645 --> 00:08:27.345
but it's a metaphor that's anthropomorphic.

158
00:08:27.925 --> 00:08:31.745
And the more anthropomorphic terms that the developer uses

159
00:08:32.085 --> 00:08:36.385
and can encourage the public, consumer users, to adopt,

160
00:08:36.925 --> 00:08:38.905
the more we lose sight of the boundary

161
00:08:38.905 --> 00:08:40.985
between the tool and the tool user.

162
00:08:41.405 --> 00:08:43.785
And I think that's a dangerous thing to allow to happen

163
00:08:45.335 --> 00:08:46.675
In this book.

164
00:08:47.075 --> 00:08:49.875
Roland, Security in the Cyber Age

165
00:08:49.875 --> 00:08:52.515
and the two authors are Derek S. Reveron

166
00:08:52.775 --> 00:08:55.685
and John E. Savage, they worried about back doors.

167
00:08:56.465 --> 00:08:58.245
Uh, is that your experience?

168
00:08:58.505 --> 00:09:00.365
Is that in accordance with, uh,

169
00:09:00.635 --> 00:09:02.445
your general set of concerns?

170
00:09:03.285 --> 00:09:05.165
I haven't concentrated on back doors.

171
00:09:05.355 --> 00:09:09.205
What I focus on are the ways in which information can

172
00:09:09.235 --> 00:09:10.405
leak out of the model.

173
00:09:11.225 --> 00:09:15.955
And when you put in a prompt, there was a tendency when

174
00:09:16.585 --> 00:09:19.355
Chat GPT was first released, people put in all kinds

175
00:09:19.355 --> 00:09:20.755

of personal information

176
00:09:21.015 --> 00:09:22.235
and delighted in the

177
00:09:22.755 --> 00:09:25.035
seemingly personal responses they were getting.

178
00:09:26.055 --> 00:09:29.515
But what they often didn't realize is that some

179
00:09:29.515 --> 00:09:32.435
of the developers reserve a right in their terms of use,

180
00:09:32.435 --> 00:09:36.275
which nobody reads, to review the prompts

181
00:09:36.575 --> 00:09:38.035
to review the conversations.

182
00:09:38.545 --> 00:09:41.355
They don't know that the default in many of these is

183
00:09:41.375 --> 00:09:45.075
to preserve the conversation, so-called, the,

184
00:09:45.995 --> 00:09:48.055
the exchanges between the user's prompt

185
00:09:48.075 --> 00:09:49.415
and the model's output,

186
00:09:49.995 --> 00:09:52.855
and by saving those to use those to train the model.

187
00:09:53.195 --> 00:09:56.495
But if you use that personal information to train the model,

188
00:09:56.955 --> 00:09:59.255
we have found that the models sometimes leak

189
00:09:59.255 --> 00:10:01.615
that personal information into the

190
00:10:01.675 --> 00:10:03.095
output for a different user.

191
00:10:04.075 --> 00:10:06.175
So imagine if you were entering information

192
00:10:06.175 --> 00:10:07.975
that was confidential to a client,

193
00:10:07.975 --> 00:10:09.175
which you should never do,

194
00:10:10.155 --> 00:10:12.655
and the model then, uh, incorporated

195
00:10:12.655 --> 00:10:15.015
that in its training set.

196
00:10:15.015 --> 00:10:18.255
There have been instances where it has output certain

197
00:10:18.255 --> 00:10:21.455
information that should only have gone back to the user

198
00:10:21.835 --> 00:10:24.515
who put it in. That kind of leaking,

199
00:10:24.535 --> 00:10:26.675
it is the kind that concerns me

200
00:10:26.675 --> 00:10:29.515

because lawyers have, above all

201
00:10:29.515 --> 00:10:33.435
of their various ethical obligations, to keep client

202
00:10:34.225 --> 00:10:36.835
disclosures, information, confidential.

203
00:10:37.575 --> 00:10:40.635
And if we use any tool that puts that at risk,

204
00:10:41.215 --> 00:10:43.515
that's an unreasonable risk for us to be taking

205
00:10:43.515 --> 00:10:47.755
because our clients ultimately don't hire us for our ability

206
00:10:47.855 --> 00:10:49.875
to use a tool like gen AI.

207
00:10:50.305 --> 00:10:51.755
They hire us for our judgment.

208
00:10:52.845 --> 00:10:56.825
For the benefit of our listeners on Sirius XM Radio's POTUS

209
00:10:56.895 --> 00:11:00.545
Channel 124, you are listening to White House Chronicle.

210
00:11:00.765 --> 00:11:02.985
I'm Llewellyn King, executive producer

211
00:11:03.085 --> 00:11:07.465
and host, co-host Adam Clayton Powell III joins me.

212
00:11:07.765 --> 00:11:09.945
Our guest today is Roland Trope,

213
00:11:10.345 --> 00:11:13.465
a partner in the New York City offices of Trope

214
00:11:13.805 --> 00:11:15.665
and Schramm LLP

215
00:11:16.005 --> 00:11:19.585
and a former adjunct professor at the United States Military

216
00:11:19.655 --> 00:11:23.585
Academy at West Point. White House Chronicle

217
00:11:23.685 --> 00:11:27.265
airs nationwide on PBS and public educational

218
00:11:27.325 --> 00:11:29.265
and government access channels.

219
00:11:30.185 --> 00:11:31.985
Subscribe for full episodes

220
00:11:31.985 --> 00:11:34.665
of White House Chronicle each week as a podcast

221
00:11:35.285 --> 00:11:37.345
on your favorite audio platform.

222
00:11:37.595 --> 00:11:41.425
Watch this and previous episodes on the White House Chronicle

223
00:11:41.815 --> 00:11:44.665
website, whchronicle.com.

224
00:11:45.325 --> 00:11:49.465
For more information on co-host Adam Clayton Powell III's,

225
00:11:49.915 --> 00:11:52.425

University of Southern California project,

226
00:11:52.735 --> 00:11:54.685
contact him at

227
00:11:54.685 --> 00:11:57.965
truevote@usc.edu.

228
00:11:58.545 --> 00:12:01.485
You can now find me on Substack.

229
00:12:03.465 --> 00:12:08.205
As the various companies are training, uh,

230
00:12:08.415 --> 00:12:11.005
their generative AI uh, tools,

231
00:12:12.555 --> 00:12:16.655
can they now reach into data sets, into data

232
00:12:16.725 --> 00:12:20.735
that we may have in the cloud that we may have, um, uh,

233
00:12:20.735 --> 00:12:22.295
we'll use an example perhaps, uh,

234
00:12:22.295 --> 00:12:25.095
having inadvertently made public something which is private.

235
00:12:25.595 --> 00:12:27.575
Is, is this an area that we need to worry about?

236
00:12:28.705 --> 00:12:32.615
Apple last week announced that it was now going

237
00:12:32.635 --> 00:12:37.175
to be asking users to opt in to allow it to sample

238
00:12:38.135 --> 00:12:40.295
portions of their emails to train

239
00:12:41.845 --> 00:12:44.665
its Apple Intelligence, its AI machine.

240
00:12:45.135 --> 00:12:48.705
What it seems to suggest is that they create synthetic data,

241
00:12:48.735 --> 00:12:52.185
data that a model would generate, for example,

242
00:12:52.405 --> 00:12:55.065
of expressions in an email like I'm going

243
00:12:55.065 --> 00:12:56.145
out to play tennis.

244
00:12:57.045 --> 00:13:01.785
And then it embeds that in a kind of coding, it will send

245
00:13:01.785 --> 00:13:03.905
that to the users who opt in.

246
00:13:05.235 --> 00:13:09.215
The user's device is now coded to take

247
00:13:09.215 --> 00:13:10.575
that embedding

248
00:13:10.575 --> 00:13:13.895
and compare it to email expressions that have

249
00:13:14.125 --> 00:13:15.455
that user has created

250
00:13:16.355 --> 00:13:20.135

and to then send back whether there is a match.

251
00:13:20.275 --> 00:13:23.375
Has the user used that expression recently?

252
00:13:23.795 --> 00:13:27.615
In which case knowing that somehow is information

253
00:13:27.615 --> 00:13:29.175
Apple wants to train its model.

254
00:13:29.195 --> 00:13:30.215
And Apple says, see,

255
00:13:30.515 --> 00:13:33.615
that's differential privacy when we've introduced some

256
00:13:33.615 --> 00:13:34.695
noise into the signal,

257
00:13:35.035 --> 00:13:36.495
And that protects your privacy.

258
00:13:36.995 --> 00:13:39.055
But what they're not explaining clearly enough is

259
00:13:39.055 --> 00:13:40.255
there's a privacy budget.

260
00:13:41.275 --> 00:13:44.295
The more noise you introduce, the poorer the results,

261
00:13:45.035 --> 00:13:46.255
but the greater the privacy.

262
00:13:47.235 --> 00:13:50.935
So Apple's not letting us know exactly how often is it going

263
00:13:50.935 --> 00:13:52.255
to send this back to the cloud

264
00:13:52.795 --> 00:13:55.735
and how easily will somebody else be able to

265
00:13:56.325 --> 00:13:58.655
reverse engineer and get access to it.

266
00:13:58.675 --> 00:14:00.295
Now, Apple promises that won't happen,

267
00:14:00.635 --> 00:14:03.135
but this morning my phone,

268
00:14:03.465 --> 00:14:06.895
which I bought recently, an Apple iPhone, just did an update

269
00:14:07.595 --> 00:14:11.135
and it told me that about these new additions

270
00:14:11.395 --> 00:14:12.655
to Apple Intelligence.

271
00:14:13.075 --> 00:14:15.695
And I went in and I started to read the explanations.

272
00:14:16.145 --> 00:14:18.335
After a half hour, I was still struggling not only

273
00:14:18.335 --> 00:14:20.775
to understand them—and Apple is quite capable

274
00:14:20.835 --> 00:14:24.495
of writing clearly, this to me was frustrating.

275
00:14:24.955 --> 00:14:28.335

But I suddenly realized there were numerous settings on my

276
00:14:28.335 --> 00:14:32.575
phone that now in default mode seemed to opt me in.

277
00:14:33.275 --> 00:14:36.535
For example, that if I had only short expressions

278
00:14:36.645 --> 00:14:40.815
that I was exchanging with Siri, its AI agent,

279
00:14:41.105 --> 00:14:43.375
those would be processed and stored on my phone,

280
00:14:43.475 --> 00:14:47.495
but longer ones would be processed up in Apple's Cloud.

281
00:14:47.725 --> 00:14:49.295
Well now if it's stored up in the cloud,

282
00:14:49.365 --> 00:14:51.245
what assurances do I have

283
00:14:51.435 --> 00:14:53.125
that they're not getting access to

284
00:14:53.125 --> 00:14:54.885
something like that? And when I tried

285
00:14:54.885 --> 00:14:56.285
to turn them off individually,

286
00:14:56.645 --> 00:14:58.285
I found in the end the only way I could

287
00:14:58.885 --> 00:15:00.445
seemingly protect myself from some

288
00:15:00.445 --> 00:15:03.685
of these leakages would be to turn off Siri entirely

289
00:15:03.865 --> 00:15:06.725
and thereby deprive myself of that tool.

290
00:15:07.305 --> 00:15:10.325
So if I wanna learn about the tool, I have

291
00:15:10.325 --> 00:15:11.725
to buy into the convenience

292
00:15:12.105 --> 00:15:14.365
and a reduction in my privacy budget.

293
00:15:15.105 --> 00:15:17.165
If you move down the legal scale

294
00:15:17.265 --> 00:15:21.245
to your local neighborhood lawyer who does DUI,

295
00:15:21.845 --> 00:15:24.485
conveyancing, divorces, that kind of thing,

296
00:15:24.795 --> 00:15:29.525
they rely very heavily on AI to produce documents,

297
00:15:30.065 --> 00:15:33.525
uh, which they would've composed years ago themselves,

298
00:15:34.305 --> 00:15:35.925
but which now they find easier,

299
00:15:36.305 --> 00:15:38.045
I'm told, having talked to some of them,

300
00:15:38.535 --> 00:15:41.845

where they can get a, a lot of the boilerplate, if you will,

301
00:15:42.225 --> 00:15:43.485
uh, straight off the web

302
00:15:43.505 --> 00:15:48.365
and straight out of AI with a few prompts, um, to um,

303
00:15:48.865 --> 00:15:50.685
uh, personalize it for the case.

304
00:15:51.065 --> 00:15:54.765
But there's no defense against that material being used

305
00:15:55.465 --> 00:15:58.925
and those court cases in which it is used.

306
00:15:59.665 --> 00:16:02.405
How does the, how do you know its veracity?

307
00:16:02.545 --> 00:16:03.765
How do you know its origin?

308
00:16:04.585 --> 00:16:07.805
And we did see a case in New York some time ago

309
00:16:08.295 --> 00:16:12.805
where a whole brief was concocted by AI with

310
00:16:13.445 --> 00:16:15.685
citations that were phony.

311
00:16:16.225 --> 00:16:18.805
Uh, how desperate is this situation?

312
00:16:19.015 --> 00:16:20.725
Early use of generative AI,

313
00:16:20.745 --> 00:16:23.645
and the case you referred to Mata versus Avianca,

314
00:16:23.645 --> 00:16:26.725
which was decided in the New York Southern District, uh,

315
00:16:26.865 --> 00:16:30.125
and as you rightly said, involved citation of cases

316
00:16:30.155 --> 00:16:32.325
that turned out as the judge called them bogus,

317
00:16:32.765 --> 00:16:34.125
opposing counsel couldn't find them,

318
00:16:34.125 --> 00:16:35.205
the judge couldn't find them.

319
00:16:35.585 --> 00:16:38.285
And when the lawyers were initially questioned, they even

320
00:16:39.125 --> 00:16:41.125
insisted that the cases were real

321
00:16:41.385 --> 00:16:42.765
and that made the case worse.

322
00:16:43.385 --> 00:16:45.445
But we've had that kind of problem

323
00:16:46.905 --> 00:16:49.395
ever since lawyers started practicing law.

324
00:16:49.735 --> 00:16:51.995
If you cite a case that you haven't read,

325
00:16:52.695 --> 00:16:54.275

you're not performing your duties

326
00:16:54.815 --> 00:16:57.555
and you're failing to fulfill your obligations

327
00:16:57.695 --> 00:16:58.795
as an officer of the court.

328
00:16:59.815 --> 00:17:01.675
And whether someone, you know,

329
00:17:01.675 --> 00:17:04.115
pulled a case from a law library that somebody then handed

330
00:17:04.135 --> 00:17:05.875
to them and said, this, this is on point,

331
00:17:05.905 --> 00:17:07.635
just cite it, you can't do that.

332
00:17:08.375 --> 00:17:10.035
And the mistakes that have been made,

333
00:17:10.625 --> 00:17:14.125
almost without exception, in the use of generative AI

334
00:17:14.315 --> 00:17:18.285
that created bogus citations is that the lawyers who

335
00:17:18.905 --> 00:17:23.085
put the citations into their materials didn't read the case.

336
00:17:23.785 --> 00:17:26.965
And sometimes I find that opposing counsel in litigation

337
00:17:27.475 --> 00:17:29.445
will have put in a case that works

338
00:17:29.465 --> 00:17:30.845
for my side more than theirs,

339
00:17:30.845 --> 00:17:32.365
and I realize somebody doing the

340
00:17:32.565 --> 00:17:34.125
research didn't look at the case.

341
00:17:34.825 --> 00:17:38.925
So it's not that AI causes us to make the mistakes.

342
00:17:39.185 --> 00:17:44.165
The duties haven't changed by the adding of a new tool.

343
00:17:44.385 --> 00:17:46.285
The addition of fax machines

344
00:17:46.345 --> 00:17:48.885
or computers did not change our duties.

345
00:17:49.595 --> 00:17:53.285
What they change is how we make sure we fulfill them.

346
00:17:53.865 --> 00:17:56.805
And the fact that something can conveniently print out

347
00:17:56.805 --> 00:18:00.125
something that resembles a case should not lead you

348
00:18:00.125 --> 00:18:03.485
by its convenience to fail to do the hard work

349
00:18:03.545 --> 00:18:06.525
of reading the case, seeing if it's relevant,

350
00:18:06.705 --> 00:18:08.765

if it's relevant, does it help your client

351
00:18:08.945 --> 00:18:10.045
in fact decide it?

352
00:18:10.585 --> 00:18:13.445
And is that going to be useful to that judge?

353
00:18:13.625 --> 00:18:16.605
Now, in doing that, by the way, I, I have a theory

354
00:18:16.625 --> 00:18:18.685
of mind just as you do, and Adam does.

355
00:18:19.025 --> 00:18:20.805
We all do. That's the way we communicate.

356
00:18:21.105 --> 00:18:23.685
We build a model of what we think the other person will

357
00:18:23.835 --> 00:18:28.125
respond to and we predict and then generate utterances.

358
00:18:28.555 --> 00:18:31.165
When you tell me that lots of smaller

359
00:18:31.815 --> 00:18:36.085
firms are using AI to draft, to summarize meetings,

360
00:18:36.625 --> 00:18:39.325
I'd like to offer a counterpoint to that if I may.

361
00:18:39.625 --> 00:18:41.845
And in offering it, I'm not criticizing them,

362
00:18:41.875 --> 00:18:43.405
it's not my place to do that,

363
00:18:43.865 --> 00:18:46.325
but I take a different position on these things.

364
00:18:47.025 --> 00:18:50.815
The use of every tool involves a trade off.

365
00:18:51.935 --> 00:18:54.875
The addition of faxes, for example, enabled us

366
00:18:54.935 --> 00:18:59.745
to quickly get correspondence to and from our clients,

367
00:18:59.745 --> 00:19:03.385
but to and from our clients' counterparties, sometimes in

368
00:19:03.585 --> 00:19:07.825
negotiations, sometimes in, you know, disagreements

369
00:19:07.825 --> 00:19:09.105
that led up to litigation.

370
00:19:09.615 --> 00:19:10.785
What was the trade off?

371
00:19:11.205 --> 00:19:12.785
It meant that people thought they had

372
00:19:12.785 --> 00:19:15.505
to respond faster when responding slower

373
00:19:15.825 --> 00:19:18.705
would've enabled them to think more carefully.

374
00:19:18.705 --> 00:19:20.785
And I've been in meetings where I told a client,

375
00:19:20.785 --> 00:19:22.785

it doesn't matter that they just faxed you.

376
00:19:23.155 --> 00:19:27.105
Let's wait a while. Think over the best response

377
00:19:27.105 --> 00:19:29.585
because we're not just writing to them.

378
00:19:30.425 --> 00:19:33.745
Whenever you're writing correspondence in a potential

379
00:19:33.745 --> 00:19:36.865
dispute, you are writing to a future audience, the court

380
00:19:36.865 --> 00:19:40.265
that may look at it someday and look at it in hindsight,

381
00:19:40.285 --> 00:19:41.905
and it better look good that way.

382
00:19:42.965 --> 00:19:45.985
That's not the kind of thinking that anybody

383
00:19:46.445 --> 00:19:48.345
who doesn't practice law will tend to do.

384
00:19:48.645 --> 00:19:51.945
But if you're a lawyer, you cannot delegate that kind

385
00:19:51.945 --> 00:19:53.465
of thinking to AI.

386
00:19:53.655 --> 00:19:55.265
It's only a prediction machine.

387
00:19:55.535 --> 00:20:00.465
It's not making the judgment of what would be effective

388
00:20:00.965 --> 00:20:02.985
to the other side to perhaps get them

389
00:20:03.325 --> 00:20:04.945
to back down from this dispute,

390
00:20:05.005 --> 00:20:08.345
and if it doesn't, that if read later by a judge

391
00:20:08.445 --> 00:20:10.185
or a jury, they will say,

392
00:20:10.215 --> 00:20:12.545
that didn't antagonize the situation.

393
00:20:12.805 --> 00:20:14.025
You were trying to work it out.

394
00:20:14.195 --> 00:20:16.145
Which positions my client better.

395
00:20:17.275 --> 00:20:20.055
To put it differently, to the extent that lawyers

396
00:20:21.735 --> 00:20:24.965
allow themselves to delegate to AI,

397
00:20:25.565 --> 00:20:28.365
to generative AI in particular, tasks

398
00:20:28.515 --> 00:20:32.285
that seem mundane, summaries of meetings, writing of emails,

399
00:20:32.685 --> 00:20:34.085
drafting of memorandum,

400
00:20:34.555 --> 00:20:36.685

they're ignoring several important facts

401
00:20:36.875 --> 00:20:40.445
that OpenAI itself put in what's called a system card.

402
00:20:40.995 --> 00:20:44.565
They warned about the over-reliance on these tools

403
00:20:45.065 --> 00:20:48.165
and the better they get, the more you over-rely on them.

404
00:20:48.165 --> 00:20:50.525
And when you do, when you trust them too much,

405
00:20:51.775 --> 00:20:54.915
you stop practicing skills that require practice

406
00:20:54.935 --> 00:20:56.515
to maintain the skill.

407
00:20:56.895 --> 00:21:00.075
If you don't write and you write a lot, your ability

408
00:21:00.175 --> 00:21:04.635
to write degrades. Young lawyers, if they are turning

409
00:21:04.815 --> 00:21:08.995
to generative AI to teach, to write memoranda,

410
00:21:09.145 --> 00:21:10.675
they're not getting the experience

411
00:21:10.675 --> 00:21:11.915
of writing the first draft.

412
00:21:12.375 --> 00:21:13.475
And in my experience,

413
00:21:13.475 --> 00:21:15.235
and I'm not saying this holds for everybody,

414
00:21:16.295 --> 00:21:20.035
the most creative critical thinking I do on a matter is the

415
00:21:20.035 --> 00:21:22.275
first draft of the memo, the first draft

416
00:21:22.295 --> 00:21:25.195
of the correspondence, the revising of it.

417
00:21:26.035 --> 00:21:29.695
And those skills have to be practiced by young lawyers.

418
00:21:30.085 --> 00:21:32.175
They need to be practiced by senior lawyers.

419
00:21:32.355 --> 00:21:33.455
We cannot sit back

420
00:21:33.455 --> 00:21:36.095
because writing is a lifelong apprenticeship.

421
00:21:36.725 --> 00:21:39.265
And if you don't practice it, if you turn it over to AI,

422
00:21:39.565 --> 00:21:41.025
you're saying, well, look at what I got.

423
00:21:41.545 --> 00:21:42.865
I have more time now,

424
00:21:43.885 --> 00:21:47.625
but it's going to be time that I'm not going

425
00:21:47.625 --> 00:21:50.625

to be using developing the skill.

426
00:21:51.165 --> 00:21:54.025
And if I could double back to that privacy example

427
00:21:54.055 --> 00:21:55.705
with Apple, remember I gave us the

428
00:21:55.705 --> 00:21:57.025
example, I'm going out to play tennis.

429
00:21:57.335 --> 00:21:59.065
What if the utterance was, I'm going out

430
00:21:59.065 --> 00:22:01.265
to join the Hands Off protest.

431
00:22:02.085 --> 00:22:06.145
Do I want that sample going up to Apple's computers?

432
00:22:06.565 --> 00:22:10.905
And more importantly, if I'm a lawyer, I cannot take

433
00:22:12.115 --> 00:22:16.415
the chance that any of my emails to my clients will go

434
00:22:16.415 --> 00:22:19.015
to Apple because I have a duty to protect

435
00:22:19.015 --> 00:22:20.055
that confidentiality.

436
00:22:20.695 --> 00:22:24.615
I also have that duty to make sure that any staff working

437
00:22:24.675 --> 00:22:27.895
for me who are under my supervision, I'm required

438
00:22:28.235 --> 00:22:30.695
to supervise them to the same level

439
00:22:30.955 --> 00:22:33.375
of honoring my obligations to my clients.

440
00:22:34.035 --> 00:22:36.495
So that's something I will not

441
00:22:37.275 --> 00:22:39.295
in the near foreseeable future

442
00:22:39.475 --> 00:22:41.855
and probably beyond that opt into.

443
00:22:42.075 --> 00:22:44.575
And what I'm worried about are the default settings

444
00:22:44.585 --> 00:22:45.735
where I don't even know

445
00:22:45.885 --> 00:22:47.775
that they're using it without making

446
00:22:47.775 --> 00:22:48.975
it clear to me that they are.

447
00:22:49.415 --> 00:22:50.615
I was talking to Evan Thomas,

448
00:22:50.795 --> 00:22:53.775
the very distinguished former Washington bureau chief

449
00:22:53.795 --> 00:22:57.415
of the Newsweek who went off to teach university,

450
00:22:58.275 --> 00:23:00.695

and I asked him whether he was still teaching

451
00:23:00.715 --> 00:23:02.375
and he said no, he had stopped

452
00:23:02.725 --> 00:23:07.215
because journalism had changed because of, uh, social media

453
00:23:07.795 --> 00:23:09.295
and the journalism he knew

454
00:23:09.295 --> 00:23:12.695
and the journalism he was teaching may long may no longer be

455
00:23:12.695 --> 00:23:14.095
as relevant as it was.

456
00:23:14.955 --> 00:23:17.935
And, uh, I wonder how true that is in law.

457
00:23:18.355 --> 00:23:22.895
Uh, do you find that today's lawyers, the old school

458
00:23:22.955 --> 00:23:27.735
of instructors in the law schools are maybe teaching a form

459
00:23:27.735 --> 00:23:29.855
of the law that is not being practiced

460
00:23:29.915 --> 00:23:32.135
or the practices in the law

461
00:23:32.485 --> 00:23:34.375
that are no longer being followed

462
00:23:34.685 --> 00:23:36.655
because of the new technology?

463
00:23:38.975 --> 00:23:41.615
I haven't been back to the Yale Law School in a long time,

464
00:23:42.635 --> 00:23:45.485
so I don't think I can comment on

465
00:23:45.795 --> 00:23:47.365
what the law professors are doing.

466
00:23:47.385 --> 00:23:49.765
But I can look at what young lawyers are doing

467
00:23:50.425 --> 00:23:52.245
and what I am finding increasingly,

468
00:23:52.265 --> 00:23:53.965
and I found this years ago

469
00:23:53.965 --> 00:23:55.525
before there was generative AI,

470
00:23:55.835 --> 00:24:00.205
there's less interest in doing hard work of reviewing,

471
00:24:00.515 --> 00:24:02.805
analyzing, and most importantly writing.

472
00:24:03.385 --> 00:24:06.155
And if you don't write, you don't think

473
00:24:06.175 --> 00:24:07.715
as clearly as you do otherwise.

474
00:24:07.905 --> 00:24:10.395
It's very important that the courts throughout this country

475
00:24:11.125 --> 00:24:13.435

write opinions, or in the,

476
00:24:13.815 --> 00:24:16.035
and when they write them, sometimes they find

477
00:24:16.385 --> 00:24:19.995
that the decision won't write, it just, it doesn't work out.

478
00:24:19.995 --> 00:24:21.955
They realize there's a flaw in the logic.

479
00:24:22.945 --> 00:24:26.675
Similarly, if they had turned it over to generative AI,

480
00:24:26.985 --> 00:24:29.995
that Fourth Circuit decision recently by, that was written

481
00:24:30.015 --> 00:24:32.435
by Judge Wilkerson I think on the Fourth Circuit,

482
00:24:33.055 --> 00:24:36.675
um, when you read it, it's a masterclass in communicating

483
00:24:36.855 --> 00:24:40.275
to multiple audiences in different

484
00:24:41.745 --> 00:24:43.585
dictions all in one opinion.

485
00:24:44.325 --> 00:24:46.945
You can't turn that over to generative AI.

486
00:24:47.005 --> 00:24:49.825
Now, law schools are increasingly teaching classes

487
00:24:50.405 --> 00:24:53.225
in cybersecurity and in generative AI,

488
00:24:53.245 --> 00:24:55.825
but what I wonder is, are they teaching the students

489
00:24:56.685 --> 00:24:58.905
to do some of the fundamental things you should do

490
00:24:58.905 --> 00:25:00.505
before you use a tool?

491
00:25:01.085 --> 00:25:02.985
Are they reading the terms of use?

492
00:25:03.525 --> 00:25:06.065
Now, you might say as well, why?

493
00:25:06.185 --> 00:25:08.225
I mean that, that contract everybody clicks.

494
00:25:08.245 --> 00:25:11.625
Yes, but the terms of use are a wonderful place to find out

495
00:25:11.875 --> 00:25:14.145
where the developer thinks the limitations

496
00:25:14.245 --> 00:25:16.465
and risks are in the use of the tool.

497
00:25:17.125 --> 00:25:19.425
And they make a habit in those instances

498
00:25:19.765 --> 00:25:22.225
of transferring those risks to the user.

499
00:25:22.365 --> 00:25:25.665
So that tells you, oh, that's something

500
00:25:26.255 --> 00:25:28.385

that the developer has identified.

501
00:25:28.385 --> 00:25:32.025
And if I can give you one example, Open AI's latest terms

502
00:25:32.205 --> 00:25:34.665
of use for its latest version

503
00:25:34.665 --> 00:25:39.465
of Chat GPT effective December 2024 states

504
00:25:39.465 --> 00:25:42.385
that when you use our services, you understand

505
00:25:42.405 --> 00:25:43.705
and agree among other things

506
00:25:44.095 --> 00:25:48.145
that you must not use any output relating to a person

507
00:25:48.325 --> 00:25:50.585
for any purpose that could have a legal

508
00:25:50.965 --> 00:25:54.785
or material impact on that person, such as making credit,

509
00:25:54.895 --> 00:25:57.265
educational insurance, legal, medical,

510
00:25:57.365 --> 00:25:59.305
or other important decisions about them.

511
00:25:59.605 --> 00:26:03.625
So if a lawyer were to use that publicly available model,

512
00:26:04.135 --> 00:26:06.865
they would, for work on their client,

513
00:26:07.095 --> 00:26:10.185
they would be violating the terms of use. That's something

514
00:26:10.185 --> 00:26:11.265
that's useful to know about

515
00:26:11.325 --> 00:26:12.545
before you start to do that.

516
00:26:13.425 --> 00:26:16.545
I would like to quote, uh, Nora Ephron, the great journalist

517
00:26:16.965 --> 00:26:19.345
who said The hard thing about writing is writing.

518
00:26:20.445 --> 00:26:24.105
And, uh, I tend to think it never gets easier or harder.

519
00:26:24.735 --> 00:26:29.345
It's a constant thing. When I find, uh, what happens with

520
00:26:30.205 --> 00:26:33.705
AI, at least when I've tried using it, is it

521
00:26:34.295 --> 00:26:35.305
homogenizes things.

522
00:26:35.965 --> 00:26:40.425
It produces a kind of, uh, you know, standard fare, uh,

523
00:26:40.805 --> 00:26:43.845
uh, that is not satisfying as literature

524
00:26:44.425 --> 00:26:46.445
or as good red herring, if you will.

525
00:26:47.215 --> 00:26:49.465

Professor David Mindell, who teaches

526
00:26:49.965 --> 00:26:53.265
aeronautical engineering at MIT told us once,

527
00:26:53.925 --> 00:26:56.145
AI is really good at bad writing

528
00:26:56.245 --> 00:26:57.665
and really poor at good writing.

529
00:26:58.245 --> 00:27:02.025
And one of the things to keep in mind, take writers

530
00:27:02.205 --> 00:27:05.225
who have a particular gift that they have had to develop

531
00:27:05.335 --> 00:27:09.385
with tremendous practice, poetry, poets can do things

532
00:27:09.385 --> 00:27:10.425
that AI can't.

533
00:27:10.425 --> 00:27:15.265
And one of those is to intelligently use borrowed utterance,

534
00:27:15.325 --> 00:27:18.385
not training set utterance, but borrowed utterance.

535
00:27:18.385 --> 00:27:19.785
Now let me give you one example.

536
00:27:21.525 --> 00:27:25.305
The English poet John Milton wrote Samson Agonistes,

537
00:27:25.665 --> 00:27:29.065
a drama about Samson's experience

538
00:27:29.315 --> 00:27:31.065
after his eyes had been put out.

539
00:27:31.765 --> 00:27:35.105
And early in, in the drama, the line appears, oh, dark,

540
00:27:35.335 --> 00:27:38.145
dark, dark, amid the blaze of noon.

541
00:27:38.205 --> 00:27:41.465
And this is a soliloquy by Samson describing

542
00:27:41.575 --> 00:27:44.425
what he's now experiencing, uh, not being able

543
00:27:44.425 --> 00:27:45.545
to see the light of day.

544
00:27:46.165 --> 00:27:49.665
The poet TS Eliot when writing East Coker, one

545
00:27:49.665 --> 00:27:51.545
of the four quartets that he composed

546
00:27:51.725 --> 00:27:54.345
during World War II while residing in London

547
00:27:54.405 --> 00:27:56.825
during the Blitz, decided he wanted

548
00:27:56.825 --> 00:28:01.525
to include in there a description of the Blitz.

549
00:28:01.985 --> 00:28:05.365
And one of those descriptions talked about people going into

550
00:28:05.365 --> 00:28:06.845

the underground to seek shelter.

551
00:28:07.305 --> 00:28:09.885
So he borrowed that line from Milton,

552
00:28:10.225 --> 00:28:13.485
put it into the opening of a stanza in East Coker,

553
00:28:13.585 --> 00:28:16.645
and it says, oh, dark, dark.

554
00:28:17.075 --> 00:28:20.925
They all go into the dark, the vacant into the vacant.

555
00:28:21.555 --> 00:28:23.965
What a brilliant use of language

556
00:28:24.025 --> 00:28:27.645
to describe going into the blacked out underground.

557
00:28:28.515 --> 00:28:30.335
But if you knew the original lines,

558
00:28:30.675 --> 00:28:33.575
you are also hearing the resonance of what Milton did

559
00:28:33.575 --> 00:28:36.975
with it and how Eliot is talking about the kind

560
00:28:36.975 --> 00:28:40.735
of blindness that you feel in that fearful time.

561
00:28:41.755 --> 00:28:43.695
That's something generative AI can't do,

562
00:28:43.755 --> 00:28:45.215
but really good poets can.

563
00:28:45.215 --> 00:28:46.775
And Eliot was a master at it.

564
00:28:46.995 --> 00:28:48.295
That's our show for today.

565
00:28:48.355 --> 00:28:51.575
We will be back, in the light, next week. Cheers.