
NASA
Ashok Srivastava, Ph.D.
Ames Research Center

1
00:00:16,189 --> 00:00:13,940
hi I'm Beth Dicky of the NASA Office of

2
00:00:17,990 --> 00:00:16,199
Communications welcome to the leading

3
00:00:20,269 --> 00:00:18,000
edge in aeronautics show where we take

4
00:00:22,070 --> 00:00:20,279
an in-depth look at aeronautics problems

5
00:00:23,540 --> 00:00:22,080
that NASA is working to solve through

6
00:00:24,800 --> 00:00:23,550
its own research or in collaboration

7
00:00:27,290 --> 00:00:24,810
with others

8
00:00:29,349 --> 00:00:27,300
today's topic is data mining and how it

9
00:00:32,150 --> 00:00:29,359
can be used to improve aviation safety

10
00:00:35,030 --> 00:00:32,160
modern airplanes produce volumes and

11
00:00:37,549 --> 00:00:35,040
volumes of data and NASA is working on

12
00:00:39,500 --> 00:00:37,559
ways to use that information to find

13
00:00:42,470 --> 00:00:39,510

problems before they lead to accidents

14
00:00:44,690 --> 00:00:42,480
to begin to understand what data mining

15
00:00:47,299 --> 00:00:44,700
is and how it can benefit the

16
00:00:49,510 --> 00:00:47,309
transportation system we spoke with two

17
00:00:52,790 --> 00:00:49,520
of America's experts on the subject

18
00:00:56,450 --> 00:00:52,800
Jhon has co-authored or co-edited

19
00:00:58,310 --> 00:00:56,460
several books 350 papers and organized

20
00:01:00,729 --> 00:00:58,320
more than a hundred conferences on the

21
00:01:02,869 --> 00:01:00,739
subject of data mining Stephen Boyd

22
00:01:05,630 --> 00:01:02,879
specializes in mathematical optimization

23
00:01:07,810 --> 00:01:05,640
of huge datasets he's been a visiting

24
00:01:10,250 --> 00:01:07,820
professor at nine prestigious

25
00:01:11,990 --> 00:01:10,260
universities around the world and holds

26
00:01:14,750 --> 00:01:12,000
an honorary Doctorate from the Royal

27
00:01:16,730 --> 00:01:14,760
Institute of Technology in Sweden both

28
00:01:18,710 --> 00:01:16,740
men have more than twenty thousand

29
00:01:25,130 --> 00:01:18,720
citations to their credit here's what

30
00:01:27,140 --> 00:01:25,140
they had to say data mine is the art of

31
00:01:29,390 --> 00:01:27,150
digging through mountains of information

32
00:01:31,760 --> 00:01:29,400
to find something useful when you're

33
00:01:34,460 --> 00:01:31,770
really not sure what you're looking for

34
00:01:36,679 --> 00:01:34,470
it is an exploration is a discovery

35
00:01:38,749 --> 00:01:36,689
process that's why I data mining some

36
00:01:41,960 --> 00:01:38,759
people call it knowledge discovery from

37
00:01:45,170 --> 00:01:41,970
beta J Han literally wrote the book on

38
00:01:48,020 --> 00:01:45,180

data mining concepts and techniques his

39
00:01:49,819 --> 00:01:48,030
textbook by that name is used by college

40
00:01:53,060 --> 00:01:49,829
students around the world there's so

41
00:01:55,219 --> 00:01:53,070
much data people want to dig in want to

42
00:01:58,520 --> 00:01:55,229
find patterns want to find it outliers

43
00:02:01,010 --> 00:01:58,530
one funded rules regularities this is

44
00:02:04,910 --> 00:02:01,020
very valuable because we are living in a

45
00:02:08,600 --> 00:02:04,920
in a data explosion age you know some

46
00:02:10,820 --> 00:02:08,610
people say we are no submerge the soaked

47
00:02:14,210 --> 00:02:10,830
with data but we still lack knowledge

48
00:02:16,220 --> 00:02:14,220
okay so if we want really getting

49
00:02:19,490 --> 00:02:16,230
knowledge to discover things useful we

50
00:02:21,830 --> 00:02:19,500
have to making use of it we all use data

51
00:02:23,990 --> 00:02:21,840
mining every day one example would be

52
00:02:27,530 --> 00:02:24,000
fraud detection for credit cards they're

53
00:02:29,750 --> 00:02:27,540
constantly monitoring transactions to

54
00:02:31,880 --> 00:02:29,760
see if their typical atypical or might

55
00:02:34,910 --> 00:02:31,890
indicate fraud another of course is

56
00:02:37,880 --> 00:02:34,920
recommendation engines go to Netflix you

57
00:02:40,460 --> 00:02:37,890
go to Amazon and you'll get sometimes

58
00:02:41,930 --> 00:02:40,470
laughable but often pretty good

59
00:02:45,229 --> 00:02:41,940
recommendations for things you might be

60
00:02:46,550 --> 00:02:45,239
interested in and of course I mean the

61
00:02:49,220 --> 00:02:46,560
biggest one by far would be Google

62
00:02:51,680 --> 00:02:49,230
Stanford engineering professor Stephen

63
00:02:53,990 --> 00:02:51,690

Boyd is a world-renowned authority on

64
00:02:57,320 --> 00:02:54,000
using mathematics to uncover information

65
00:02:59,660 --> 00:02:57,330
that may be hidden in huge data sets

66
00:03:02,360 --> 00:02:59,670
this is nothing anybody could ever do by

67
00:03:04,509 --> 00:03:02,370
hand of course and what you know and

68
00:03:06,710 --> 00:03:04,519
then buy a computer of course you will

69
00:03:09,140 --> 00:03:06,720
analyze this data and you'd want to

70
00:03:10,610 --> 00:03:09,150
tease out little patterns and things

71
00:03:12,110 --> 00:03:10,620
like that of course you have to be

72
00:03:15,500 --> 00:03:12,120
careful that you're not just picking up

73
00:03:16,790 --> 00:03:15,510
stuff that was just a fluke in the data

74
00:03:19,790 --> 00:03:16,800
so you want to be sure that what you're

75
00:03:22,009 --> 00:03:19,800
picking up is is real NASA they're

76
00:03:25,390 --> 00:03:22,019
studying how data mining techniques can

77
00:03:27,649 --> 00:03:25,400
be used to improve aviation safety

78
00:03:29,210 --> 00:03:27,659
collaborating with commercial carriers

79
00:03:31,490 --> 00:03:29,220
like Southwest Airlines

80
00:03:33,740 --> 00:03:31,500
NASA Aeronautics researchers are

81
00:03:36,589 --> 00:03:33,750
developing complex algorithms or

82
00:03:39,050 --> 00:03:36,599
computer programs that can help safety

83
00:03:41,930 --> 00:03:39,060
analysts find precursors or early

84
00:03:44,330 --> 00:03:41,940
indicators of safety concerns in tests

85
00:03:46,940 --> 00:03:44,340
using data from actual completed flights

86
00:03:49,070 --> 00:03:46,950
the algorithms have detected potential

87
00:03:52,160 --> 00:03:49,080
problems that might otherwise have been

88
00:03:54,199 --> 00:03:52,170

overlooked commercial airlines

89
00:03:56,780 --> 00:03:54,209
typically log data from hundreds of

90
00:03:59,869 --> 00:03:56,790
instruments at least once every second

91
00:04:02,119 --> 00:03:59,879
the flight data recorder stores numeric

92
00:04:04,580 --> 00:04:02,129
or symbolic readings on everything from

93
00:04:06,800 --> 00:04:04,590
altitude and airspeed to the positions

94
00:04:08,110 --> 00:04:06,810
of flight controllers such as wing flaps

95
00:04:10,149 --> 00:04:08,120
and landing gear

96
00:04:12,160 --> 00:04:10,159
to the settings of switches for the

97
00:04:15,100 --> 00:04:12,170
electrical and hydraulic systems when

98
00:04:15,550 --> 00:04:15,110
the airplane is in the air back on the

99
00:04:18,250 --> 00:04:15,560
ground

100
00:04:21,340 --> 00:04:18,260
pilot supplement the airborne data with

101
00:04:23,830 --> 00:04:21,350
written reports about the flight taken

102
00:04:26,409 --> 00:04:23,840
together all this data tells the story

103
00:04:28,750 --> 00:04:26,419
of what the airplane was doing but

104
00:04:31,360 --> 00:04:28,760
there's so much data it's been nearly

105
00:04:33,250 --> 00:04:31,370
impossible for airlines to do anything

106
00:04:36,219 --> 00:04:33,260
other than look back for the cause of

107
00:04:38,620 --> 00:04:36,229
something that's already happened the

108
00:04:40,659 --> 00:04:38,630
aviation safety challenge is to figure

109
00:04:44,230 --> 00:04:40,669
out how to use the data to look ahead

110
00:04:46,570 --> 00:04:44,240
and find precursors to problems we need

111
00:04:48,879 --> 00:04:46,580
to be able to sort through the data to

112
00:04:50,230 --> 00:04:48,889
see if there's anything unusual and then

113
00:04:52,840 --> 00:04:50,240

zoom in on those bits of information

114
00:04:55,450 --> 00:04:52,850
that are different to better understand

115
00:04:58,270 --> 00:04:55,460
them and we need to do it quickly and

116
00:05:00,189 --> 00:04:58,280
efficiently that you may want to get

117
00:05:02,189 --> 00:05:00,199
something I know you want to predict

118
00:05:06,850 --> 00:05:02,199
something we call predictive modeling

119
00:05:10,740 --> 00:05:06,860
the knowledge discovery the magic or the

120
00:05:14,110 --> 00:05:10,750
attractive things actually you work out

121
00:05:16,480 --> 00:05:14,120
your hypothesis or discover something

122
00:05:19,270 --> 00:05:16,490
and try to find something unknown

123
00:05:21,400 --> 00:05:19,280
predict something what this will do is

124
00:05:23,200 --> 00:05:21,410
it will look at lots of data where

125
00:05:25,360 --> 00:05:23,210
things are just off a little bit a human

126
00:05:27,159 --> 00:05:25,370
being wouldn't even notice these

127
00:05:29,170 --> 00:05:27,169
differences but to start integrating

128
00:05:31,870 --> 00:05:29,180
this across like two or three flights

129
00:05:33,760 --> 00:05:31,880
and something's out of whack is these

130
00:05:36,490 --> 00:05:33,770
are the kind of errors that things like

131
00:05:38,879 --> 00:05:36,500
this can catch and these could actually

132
00:05:41,469 --> 00:05:38,889
end up having a very valuable safety

133
00:05:42,879 --> 00:05:41,479
implications right this could this you

134
00:05:45,460 --> 00:05:42,889
could catch things like impending

135
00:05:48,310 --> 00:05:45,470
failure data mine is a lot like

136
00:05:50,680 --> 00:05:48,320
detective one you dig for clues what

137
00:05:57,160 --> 00:05:50,690
they uncover will help make your next

138
00:06:01,940 --> 00:05:59,630

and we want to say a special thank you

139
00:06:03,380 --> 00:06:01,950
to dr. Hahn and dr. Boyd for teaching us

140
00:06:05,390 --> 00:06:03,390
what data mining is and how it touches

141
00:06:07,580 --> 00:06:05,400
our lives every day in ways we may take

142
00:06:09,740 --> 00:06:07,590
for granted we also learned that we can

143
00:06:11,210 --> 00:06:09,750
use data mining to be safer in the air

144
00:06:14,150 --> 00:06:11,220
and that's what we're here to talk about

145
00:06:16,010 --> 00:06:14,160
today our first guest is Jeff Hamlet

146
00:06:17,810 --> 00:06:16,020
with Southwest Airlines you're the

147
00:06:19,310 --> 00:06:17,820
flight safety director there thanks for

148
00:06:21,260 --> 00:06:19,320
being with us thank you for inviting me

149
00:06:23,690 --> 00:06:21,270
back now we saw in that video that

150
00:06:26,420 --> 00:06:23,700
Southwest is collaborating with NASA to

151
00:06:28,610 --> 00:06:26,430
use data mining to improve its flight

152
00:06:30,890 --> 00:06:28,620
safety and its efficiency let's talk a

153
00:06:32,960 --> 00:06:30,900
little bit about how airlines do use

154
00:06:35,720 --> 00:06:32,970
data mining and why they do it well it's

155
00:06:37,820 --> 00:06:35,730
a very complex system Beth and probably

156
00:06:39,080 --> 00:06:37,830
the most complex system that that humans

157
00:06:40,360 --> 00:06:39,090
have ever built if you look at the

158
00:06:42,650 --> 00:06:40,370
Bureau of Transportation Statistics

159
00:06:45,770 --> 00:06:42,660
website you can see that the top 20

160
00:06:47,480 --> 00:06:45,780
carriers fly about 3/4 of a million

161
00:06:49,250 --> 00:06:47,490
flights per month and Southwest alone

162
00:06:52,910 --> 00:06:49,260
flies about a hundred thousand flights a

163
00:06:55,090 --> 00:06:52,920

month so we have a lot of data and we're

164
00:06:57,650 --> 00:06:55,100
using that to improve our safety

165
00:07:00,260 --> 00:06:57,660
customer service and efficiency I think

166
00:07:01,970 --> 00:07:00,270
we have a depiction of typical day in

167
00:07:05,510 --> 00:07:01,980
the life of the nation's air

168
00:07:06,830 --> 00:07:05,520
transportation system coming up on the

169
00:07:10,060 --> 00:07:06,840
screen in just a minute you're going to

170
00:07:13,580 --> 00:07:10,070
see a map of the United States with the

171
00:07:16,370 --> 00:07:13,590
large aircraft centers called out and

172
00:07:19,700 --> 00:07:16,380
there it is and you see all those ants

173
00:07:22,420 --> 00:07:19,710
there on the screen those are those are

174
00:07:24,680 --> 00:07:22,430
planes and you're going to see them

175
00:07:26,270 --> 00:07:24,690
picking up on the East Coast and then

176
00:07:27,650 --> 00:07:26,280
that whole yellow is going to move

177
00:07:30,020 --> 00:07:27,660
across to the west coast during the day

178
00:07:33,020 --> 00:07:30,030
that's an awful lot of airplanes it's it

179
00:07:34,610 --> 00:07:33,030
just must be impossible to keep track of

180
00:07:37,640 --> 00:07:34,620
all the problems that could occur well

181
00:07:40,330 --> 00:07:37,650
it is very complex and at the same time

182
00:07:43,010 --> 00:07:40,340
it's it's safer than it's ever been and

183
00:07:44,660 --> 00:07:43,020
that's why our partnership with NASA is

184
00:07:47,150 --> 00:07:44,670
very important because we need to have

185
00:07:50,120 --> 00:07:47,160
systematic ways to look into our data

186
00:07:52,220 --> 00:07:50,130
and find things that we're not currently

187
00:07:54,080 --> 00:07:52,230
looking for and put mitigations in place

188
00:07:56,600 --> 00:07:54,090

so that we do improve safety well what

189
00:07:58,870 --> 00:07:56,610
kinds of problems do you find right now

190
00:08:02,060 --> 00:07:58,880
we see a whole range of problems from

191
00:08:05,030 --> 00:08:02,070
pilot reports of problems specific

192
00:08:06,030 --> 00:08:05,040
problems at airports to issues in the

193
00:08:08,270 --> 00:08:06,040
data where there

194
00:08:12,620 --> 00:08:08,280
mechanical problems with aircraft or

195
00:08:15,960 --> 00:08:12,630
even specific arrival and approach

196
00:08:18,480 --> 00:08:15,970
abnormalities we don't see airplane

197
00:08:20,520 --> 00:08:18,490
crashes every day so obviously

198
00:08:22,350 --> 00:08:20,530
something's going right what what are

199
00:08:23,910 --> 00:08:22,360
the consequences of these problems well

200
00:08:26,670 --> 00:08:23,920
often there's there are no consequences

201
00:08:28,020 --> 00:08:26,680
and that's one of the the things that we

202
00:08:29,700 --> 00:08:28,030
want to continue or you know we're doing

203
00:08:32,909 --> 00:08:29,710
everything we can do to make the system

204
00:08:36,630 --> 00:08:32,919
as safe as possible back in 1931 dr.

205
00:08:38,969 --> 00:08:36,640
Heinrich developed this model for to

206
00:08:40,950 --> 00:08:38,979
describe or to sort of look at safety

207
00:08:43,140 --> 00:08:40,960
and complex systems and I don't think he

208
00:08:45,150 --> 00:08:43,150
anticipated the complexity of the modern

209
00:08:47,850 --> 00:08:45,160
aviation system but what it does tell us

210
00:08:49,770 --> 00:08:47,860
is that there are a lot of issues down

211
00:08:52,050 --> 00:08:49,780
at the base of the pyramid that we can

212
00:08:54,210 --> 00:08:52,060
take advantage of we can we can analyze

213
00:08:57,690 --> 00:08:54,220

that data to try to prevent things

214
00:08:58,980 --> 00:08:57,700
occurring farther up the pyramid all

215
00:09:01,740 --> 00:08:58,990
right so let's talk about the data you

216
00:09:03,600 --> 00:09:01,750
collect okay well we have three main

217
00:09:05,550 --> 00:09:03,610
safety programs we have a Aviation

218
00:09:07,890 --> 00:09:05,560
Safety Action Partnership Program

219
00:09:10,530 --> 00:09:07,900
we have a flight data analysis program

220
00:09:13,590 --> 00:09:10,540
and a traditional flight safety

221
00:09:15,540 --> 00:09:13,600
investigations program and do these all

222
00:09:18,210 --> 00:09:15,550
work in conjunction with each other what

223
00:09:20,580 --> 00:09:18,220
they do we have agreements between our

224
00:09:22,530 --> 00:09:20,590
pilot Association and the FAA and the

225
00:09:25,770 --> 00:09:22,540
company on how we govern this programs

226
00:09:28,380 --> 00:09:25,780
and we do gain some benefit from seeing

227
00:09:30,270 --> 00:09:28,390
something in one program and analyzing

228
00:09:32,040 --> 00:09:30,280
the data and other programs to find out

229
00:09:32,970 --> 00:09:32,050
more about a specific events all right

230
00:09:38,250 --> 00:09:32,980
well let's look at them individually

231
00:09:39,870 --> 00:09:38,260
talk about the ASAP ASAP is up as an

232
00:09:41,730 --> 00:09:39,880
agreement between the company and the

233
00:09:45,860 --> 00:09:41,740
pollow Association at Southwest and the

234
00:09:48,300 --> 00:09:45,870
FAA where pilots can report voluntarily

235
00:09:50,720 --> 00:09:48,310
safety issues that they that they

236
00:09:53,580 --> 00:09:50,730
experience while they're out flying the

237
00:09:55,410 --> 00:09:53,590
the FAA has made it possible and

238
00:09:57,030 --> 00:09:55,420

provided incentive for the pilots to

239
00:09:59,610 --> 00:09:57,040
report because if the pilot submits a

240
00:10:02,340 --> 00:09:59,620
report meets requirements for entering

241
00:10:03,960 --> 00:10:02,350
the program then they receive protection

242
00:10:05,940 --> 00:10:03,970
so the FAA has agreed to limit whatever

243
00:10:07,740 --> 00:10:05,950
certificate action may occur if there

244
00:10:09,660 --> 00:10:07,750
was a violation of a regulation and the

245
00:10:12,720 --> 00:10:09,670
benefit is we have a safe non-punitive

246
00:10:14,400 --> 00:10:12,730
environment for pilots to report and the

247
00:10:16,110 --> 00:10:14,410
company receives a benefit because we

248
00:10:18,090 --> 00:10:16,120
have a lot of information about events

249
00:10:19,830 --> 00:10:18,100
that we would otherwise not know about

250
00:10:23,130 --> 00:10:19,840
and we're able to take that

251
00:10:26,370 --> 00:10:23,140
and make systemic Corrections to improve

252
00:10:28,440 --> 00:10:26,380
our safety all right and what about you

253
00:10:30,690 --> 00:10:28,450
call it fadap the Federal Aviation

254
00:10:32,640 --> 00:10:30,700
Administration calls it aqua right

255
00:10:35,190 --> 00:10:32,650
flight operational quality assurance or

256
00:10:37,769 --> 00:10:35,200
flight data analysis program we have to

257
00:10:39,390 --> 00:10:37,779
be just a little bit different because

258
00:10:41,010 --> 00:10:39,400
that's exactly what we do we download

259
00:10:43,740 --> 00:10:41,020
data off of the aircraft every five to

260
00:10:45,630 --> 00:10:43,750
seven days from a memory card we put

261
00:10:48,360 --> 00:10:45,640
that into our database and our analysts

262
00:10:50,190 --> 00:10:48,370
go through the data and they determine

263
00:10:52,650 --> 00:10:50,200

what what exactly is going on with the

264
00:10:55,740 --> 00:10:52,660
airplane and that's the benefit of fadap

265
00:10:57,090 --> 00:10:55,750
is we that we analyze about 50,000

266
00:10:59,730 --> 00:10:57,100
flights a month so about half of our

267
00:11:01,860 --> 00:10:59,740
schedule we look at and we're able to

268
00:11:04,019 --> 00:11:01,870
tell pretty much what's going on with

269
00:11:05,610 --> 00:11:04,029
her with our aircraft and what was the

270
00:11:08,250 --> 00:11:05,620
third program that was flight safety

271
00:11:12,600 --> 00:11:08,260
investigations program and it's a very

272
00:11:14,310 --> 00:11:12,610
strong program the it's a forensic type

273
00:11:16,079 --> 00:11:14,320
program a traditional flight safety

274
00:11:18,450 --> 00:11:16,089
investigation program and if you think

275
00:11:21,390 --> 00:11:18,460
about that model that the pyramid the

276
00:11:23,519 --> 00:11:21,400
fadap and a set programs are looking at

277
00:11:25,380 --> 00:11:23,529
data that that we get from the bottom of

278
00:11:28,050 --> 00:11:25,390
that pyramid trying to prevent things

279
00:11:30,240 --> 00:11:28,060
farther up the pyramid but obviously we

280
00:11:32,250 --> 00:11:30,250
have a strong investigations program

281
00:11:34,530 --> 00:11:32,260
when we do need to take a look at that

282
00:11:36,500 --> 00:11:34,540
what's the ultimate goal that you have

283
00:11:38,700 --> 00:11:36,510
in collecting this information and

284
00:11:41,940 --> 00:11:38,710
assessing it well ultimately we're

285
00:11:43,710 --> 00:11:41,950
looking at improving safety through risk

286
00:11:46,020 --> 00:11:43,720
management so we want to identify

287
00:11:48,329 --> 00:11:46,030
hazards in the operation we want to

288
00:11:51,870 --> 00:11:48,339

analyze the risk that they present and

289
00:11:54,150 --> 00:11:51,880
then put mitigations in place to control

290
00:11:56,490 --> 00:11:54,160
that risk and improve our safety and

291
00:11:57,990 --> 00:11:56,500
efficiency as well all right now NASA

292
00:12:01,800 --> 00:11:58,000
administers something called the

293
00:12:06,449 --> 00:12:01,810
aviation safety reporting system for the

294
00:12:07,980 --> 00:12:06,459
FAA how does all of the the Pro how do

295
00:12:10,980 --> 00:12:07,990
all of the programs that you've told us

296
00:12:12,930 --> 00:12:10,990
about fit into that well except the

297
00:12:15,990 --> 00:12:12,940
pilot voluntary disclosure program

298
00:12:18,540 --> 00:12:16,000
really grew out of a SRS and a SRS is a

299
00:12:20,370 --> 00:12:18,550
program where any aviation professional

300
00:12:25,199 --> 00:12:20,380
in the country can submit a safety

301
00:12:27,390 --> 00:12:25,209
report and they are analyzed at NASA and

302
00:12:29,150 --> 00:12:27,400
the database is made available for other

303
00:12:31,230 --> 00:12:29,160
people to take a look at as well and

304
00:12:32,930 --> 00:12:31,240
virtually all of the airlines that have

305
00:12:35,000 --> 00:12:32,940
a SAP programs

306
00:12:36,500 --> 00:12:35,010
participate they by submitting a

307
00:12:39,050 --> 00:12:36,510
separate ports and into the ASRs

308
00:12:42,230 --> 00:12:39,060
database and Southwest does that yes and

309
00:12:44,570 --> 00:12:42,240
what do you get back well anybody we do

310
00:12:47,030 --> 00:12:44,580
in addition to the safety alerts that

311
00:12:49,840 --> 00:12:47,040
are produced by the analysts at a SRS

312
00:12:54,770 --> 00:12:49,850
when we initiate service into a new city

313
00:12:57,530 --> 00:12:54,780

we work with NASA to come up with a set

314
00:13:00,170 --> 00:12:57,540
of disclosure or the set of reports that

315
00:13:01,820 --> 00:13:00,180
pilots have submitted that may indicate

316
00:13:03,590 --> 00:13:01,830
issues that our pilots need to be aware

317
00:13:05,270 --> 00:13:03,600
of before they have an opportunity to

318
00:13:07,280 --> 00:13:05,280
fly there so we put together what we

319
00:13:09,410 --> 00:13:07,290
call a smart pack that's a synopsis of

320
00:13:11,870 --> 00:13:09,420
the events that were reported by other

321
00:13:14,060 --> 00:13:11,880
pilots and aviation professionals it's

322
00:13:18,260 --> 00:13:14,070
got air filled diagrams and charts and

323
00:13:20,540 --> 00:13:18,270
some overhead imagery so that we can

324
00:13:23,990 --> 00:13:20,550
provide our pilot with a first look that

325
00:13:26,270 --> 00:13:24,000
provides an opportunity for success all

326
00:13:28,520 --> 00:13:26,280
right so part of the point of the

327
00:13:30,350 --> 00:13:28,530
smartpak is to tell them about potential

328
00:13:32,090 --> 00:13:30,360
hazards let's talk some more about those

329
00:13:34,460 --> 00:13:32,100
what are some of the indicators and the

330
00:13:36,560 --> 00:13:34,470
markers that you have for hazards well

331
00:13:38,120 --> 00:13:36,570
our programs particularly fadap the

332
00:13:41,990 --> 00:13:38,130
flight data analysis program is set up

333
00:13:43,820 --> 00:13:42,000
to take a look at things exceedences

334
00:13:46,310 --> 00:13:43,830
that occur so we may set a threshold and

335
00:13:48,680 --> 00:13:46,320
our algorithms to capture when a certain

336
00:13:50,930 --> 00:13:48,690
air speed limit has been exceeded or a

337
00:13:52,430 --> 00:13:50,940
certain arrival profile we want to take

338
00:13:54,950 --> 00:13:52,440

a look at we'll design an algorithm to

339
00:13:57,650 --> 00:13:54,960
dig deeper into that so really it's

340
00:14:00,920 --> 00:13:57,660
based on looking at exceedences and it

341
00:14:02,690 --> 00:14:00,930
is a forensic type approach our a set

342
00:14:04,670 --> 00:14:02,700
program is the same type of thing where

343
00:14:07,520 --> 00:14:04,680
pilots or reporting things that they

344
00:14:09,590 --> 00:14:07,530
observe out in the operation how would a

345
00:14:12,290 --> 00:14:09,600
predictive insight that you could gain

346
00:14:15,170 --> 00:14:12,300
from data mining help you in that in

347
00:14:19,040 --> 00:14:15,180
those kinds of situations well we're

348
00:14:20,720 --> 00:14:19,050
looking at ways to find events and

349
00:14:22,310 --> 00:14:20,730
issues that were not currently aware of

350
00:14:23,600 --> 00:14:22,320
or not we're not currently looking for

351
00:14:25,870 --> 00:14:23,610
and that's why the partnership with NASA

352
00:14:28,550 --> 00:14:25,880
is so important we want to be able to

353
00:14:30,170 --> 00:14:28,560
find out what the hazards are measured

354
00:14:33,260 --> 00:14:30,180
the risk and then put those mitigations

355
00:14:36,670 --> 00:14:33,270
in place have you ever been able to

356
00:14:39,980 --> 00:14:36,680
detect a problem an anomaly before a

357
00:14:42,680 --> 00:14:39,990
pilot has well what happened nASA has

358
00:14:45,080 --> 00:14:42,690
done some very interesting work on fuel

359
00:14:46,550 --> 00:14:45,090
burn and by taking a look at nominal

360
00:14:49,040 --> 00:14:46,560
fuel burns between

361
00:14:50,210 --> 00:14:49,050
city pairs and then analyzing what

362
00:14:53,870 --> 00:14:50,220
happens in different environmental

363
00:14:55,610 --> 00:14:53,880

environmental conditions we can gain

364
00:14:58,010 --> 00:14:55,620
some efficiency of that but going

365
00:14:59,930 --> 00:14:58,020
further into that information we can

366
00:15:02,630 --> 00:14:59,940
even compare the performance of each

367
00:15:06,050 --> 00:15:02,640
engine on the same aircraft and detect

368
00:15:07,640 --> 00:15:06,060
small divergences in parameters and that

369
00:15:10,040 --> 00:15:07,650
may lead to another problem so that's

370
00:15:11,540 --> 00:15:10,050
kind of what we're hoping for - in our

371
00:15:13,490 --> 00:15:11,550
partnership with NASA is to be able to

372
00:15:15,380 --> 00:15:13,500
take a forward-looking approach and find

373
00:15:18,320 --> 00:15:15,390
small things that pilots may not

374
00:15:20,840 --> 00:15:18,330
necessarily detect and fix them before

375
00:15:22,550 --> 00:15:20,850
they become a service interruption

376
00:15:24,530 --> 00:15:22,560
alright and and what else do you think

377
00:15:31,700 --> 00:15:24,540
you can glean from the collaboration

378
00:15:36,890 --> 00:15:31,710
with the agency we are we've got that

379
00:15:40,220 --> 00:15:36,900
partnership that we I'm guessing that

380
00:15:42,710 --> 00:15:40,230
it's that that that it's a systemic kind

381
00:15:44,630 --> 00:15:42,720
of oh well yes we want to develop a

382
00:15:46,820 --> 00:15:44,640
systematic approach to our data because

383
00:15:49,040 --> 00:15:46,830
we have so much data we want to be able

384
00:15:50,870 --> 00:15:49,050
to take a look and and find out those

385
00:15:53,090 --> 00:15:50,880
issues that that we currently weren't

386
00:15:55,730 --> 00:15:53,100
looking for because we really don't want

387
00:15:57,290 --> 00:15:55,740
to miss anything in the data and this

388
00:15:59,120 --> 00:15:57,300

kind of thing can help you prevent the

389
00:16:01,670 --> 00:15:59,130
delays and cancelations candidate that's

390
00:16:04,070 --> 00:16:01,680
true there's as we get safer we have to

391
00:16:06,620 --> 00:16:04,080
go farther and farther up the stream of

392
00:16:08,390 --> 00:16:06,630
data and what we find is that there are

393
00:16:09,290 --> 00:16:08,400
efficiency benefits and customer service

394
00:16:12,470 --> 00:16:09,300
benefits as well

395
00:16:16,520 --> 00:16:12,480
all right great now the open source

396
00:16:18,320 --> 00:16:16,530
website nASA has called - link this is a

397
00:16:21,470 --> 00:16:18,330
place where people can go and download

398
00:16:23,540 --> 00:16:21,480
software for data mining applications

399
00:16:25,610 --> 00:16:23,550
and what not have his Southwest taking

400
00:16:29,240 --> 00:16:25,620
advantage of that yes we worked with the

401
00:16:31,010 --> 00:16:29,250
analyst from NASA and taken advantage a

402
00:16:32,030 --> 00:16:31,020
couple of those algorithms and the

403
00:16:35,810 --> 00:16:32,040
things that we have learned from

404
00:16:38,900 --> 00:16:35,820
analyzing sets of data involving over

405
00:16:41,990 --> 00:16:38,910
7,000 flights have been incorporated in

406
00:16:42,680 --> 00:16:42,000
our daily review of flight data all

407
00:16:45,470 --> 00:16:42,690
right

408
00:16:51,019 --> 00:16:45,480
I'm curious if you can tell me a story

409
00:16:53,720 --> 00:16:51,029
about data mining saving the day well my

410
00:16:56,380 --> 00:16:53,730
first experience where we really I

411
00:16:58,880 --> 00:16:56,390
really got excited about data mining was

412
00:16:59,910 --> 00:16:58,890
when we first started looking at

413
00:17:01,829 --> 00:16:59,920

high-energy approach

414
00:17:04,740 --> 00:17:01,839
in our flight data in the fadap program

415
00:17:06,449 --> 00:17:04,750
and we saw a number of instances where

416
00:17:09,120 --> 00:17:06,459
we had high energy approaches at certain

417
00:17:10,620 --> 00:17:09,130
airports but in our ASA pilot reporting

418
00:17:13,079 --> 00:17:10,630
data we really didn't see a lot of

419
00:17:16,079 --> 00:17:13,089
pilots talking about that so we took a

420
00:17:18,929 --> 00:17:16,089
we got an application from NASA that was

421
00:17:20,579 --> 00:17:18,939
developed by dr. Michael McDreamy called

422
00:17:22,769 --> 00:17:20,589
paralog and we were able to take a

423
00:17:24,720 --> 00:17:22,779
really good report from a pilot about a

424
00:17:26,819 --> 00:17:24,730
high-energy approach and search for

425
00:17:28,919 --> 00:17:26,829
other examples of that same situation

426
00:17:31,230 --> 00:17:28,929
and what we found is very interesting

427
00:17:33,389 --> 00:17:31,240
because we discovered that there were

428
00:17:35,730 --> 00:17:33,399
airspeed and altitude requirements on

429
00:17:37,950 --> 00:17:35,740
arrival that were creating a very busy

430
00:17:40,470 --> 00:17:37,960
environment in the cockpit and pilots

431
00:17:42,330 --> 00:17:40,480
often just forgot to switch the radio

432
00:17:45,779 --> 00:17:42,340
and talk to tower so they were reporting

433
00:17:47,310 --> 00:17:45,789
it as a communication issue but we were

434
00:17:48,870 --> 00:17:47,320
able to take that information and go

435
00:17:50,519 --> 00:17:48,880
back to air traffic control and work

436
00:17:53,909 --> 00:17:50,529
with them to modify the way they bring

437
00:17:55,710 --> 00:17:53,919
us into airports to provide a better

438
00:17:58,110 --> 00:17:55,720

arrival so that's a that's a good

439
00:17:59,940 --> 00:17:58,120
example of how you're sort of shifting

440
00:18:02,299 --> 00:17:59,950
from that forensic approach to the

441
00:18:04,590 --> 00:18:02,309
prognostic kind of approach to

442
00:18:06,960 --> 00:18:04,600
predictive analysis limit corrections

443
00:18:11,700 --> 00:18:06,970
and and helping things in the future

444
00:18:15,570 --> 00:18:11,710
become more safe all right how would you

445
00:18:18,870 --> 00:18:15,580
use data mining to let's say monitor the

446
00:18:20,340 --> 00:18:18,880
health of an airplane well we can take

447
00:18:23,730 --> 00:18:20,350
advantage of the redundancy of the

448
00:18:25,889 --> 00:18:23,740
system so for instance we have systems

449
00:18:27,480 --> 00:18:25,899
that are supposed to perform in a very

450
00:18:29,490 --> 00:18:27,490
similar manner we can look for small

451
00:18:31,379 --> 00:18:29,500
differences in the way those systems are

452
00:18:35,159 --> 00:18:31,389
performing in individual parameters and

453
00:18:36,539 --> 00:18:35,169
then we can we can possibly determine

454
00:18:38,250 --> 00:18:36,549
there's an issue before the pilot

455
00:18:40,289 --> 00:18:38,260
becomes aware of the problem and alert

456
00:18:43,080 --> 00:18:40,299
maintenance so that they can correct it

457
00:18:45,930 --> 00:18:43,090
before it becomes interruption I think

458
00:18:48,269 --> 00:18:45,940
you had told me earlier about an example

459
00:18:51,060 --> 00:18:48,279
of using using this kind of information

460
00:18:54,600 --> 00:18:51,070
to work on engine cooling operations

461
00:18:56,639 --> 00:18:54,610
that's true well and and fuel is is very

462
00:18:59,159 --> 00:18:56,649
expensive so as we were looking at ways

463
00:19:01,620 --> 00:18:59,169

that we may be able to save a little

464
00:19:02,629 --> 00:19:01,630
fuel on the taxi in after landing we

465
00:19:05,399 --> 00:19:02,639
also needed to take into consideration

466
00:19:07,230 --> 00:19:05,409
engine cooling and there's a requirement

467
00:19:09,930 --> 00:19:07,240
to cool the engines for one to three

468
00:19:11,850 --> 00:19:09,940
minutes after you land and once that

469
00:19:13,410 --> 00:19:11,860
condition is satisfied then we can turn

470
00:19:16,860 --> 00:19:13,420
off one engine to save a little

471
00:19:18,870 --> 00:19:16,870
but fuel as we taxi in and it's because

472
00:19:20,730 --> 00:19:18,880
if you it is important to get that

473
00:19:22,320 --> 00:19:20,740
adequate cooling because the long-term

474
00:19:23,790 --> 00:19:22,330
impact of taking good care of the

475
00:19:26,370 --> 00:19:23,800
engines is increased engines life

476
00:19:28,440 --> 00:19:26,380
because over time if the engine is not

477
00:19:30,060 --> 00:19:28,450
allowed to cool properly the oil jets

478
00:19:32,340 --> 00:19:30,070
that feed the lubrication to the

479
00:19:34,050 --> 00:19:32,350
bearings can become clogged so we want

480
00:19:35,280 --> 00:19:34,060
to be very careful about that and

481
00:19:36,960 --> 00:19:35,290
improve the long-term health of the

482
00:19:38,790 --> 00:19:36,970
engines but at the same time if we can

483
00:19:42,570 --> 00:19:38,800
it take advantage of some fuel

484
00:19:45,660 --> 00:19:42,580
efficiency then that's great as well do

485
00:19:46,020 --> 00:19:45,670
you fly yes how is data mining help to

486
00:19:48,000 --> 00:19:46,030
you

487
00:19:50,040 --> 00:19:48,010
well that specific example of engine

488
00:19:51,750 --> 00:19:50,050

cooling has really allowed me to go in

489
00:19:53,520 --> 00:19:51,760
and develop a you know different

490
00:19:56,310 --> 00:19:53,530
strategies on how I monitor the engines

491
00:19:57,450 --> 00:19:56,320
after I land so that that I can take

492
00:19:58,950 --> 00:19:57,460
better care of the engines and get a

493
00:20:00,420 --> 00:19:58,960
little fuel efficiency but beyond that

494
00:20:03,360 --> 00:20:00,430
our analysts have done some really good

495
00:20:06,690 --> 00:20:03,370
work on arrival paths at specific

496
00:20:08,340 --> 00:20:06,700
airports specifically El Paso is an

497
00:20:13,800 --> 00:20:08,350
example because of the high terrain to

498
00:20:16,170 --> 00:20:13,810
the west and the requires a higher

499
00:20:17,670 --> 00:20:16,180
altitude on arrival and our analysts did

500
00:20:19,290 --> 00:20:17,680
some interesting analysis where they

501
00:20:21,420 --> 00:20:19,300
laid down tracks to show the difference

502
00:20:23,040 --> 00:20:21,430
between an optimum profile and maybe one

503
00:20:25,530 --> 00:20:23,050
that's a little too tight and so that

504
00:20:28,680 --> 00:20:25,540
has really influenced the way I approach

505
00:20:29,970 --> 00:20:28,690
that Airport and provides a better more

506
00:20:32,040 --> 00:20:29,980
comfortable ride for the customers as

507
00:20:34,380 --> 00:20:32,050
well as an optimum path all right well

508
00:20:36,060 --> 00:20:34,390
great Jeff thanks for coming along and

509
00:20:39,750 --> 00:20:36,070
giving us a little bit of insight into

510
00:20:42,540 --> 00:20:39,760
Southwest's air safety operations we're

511
00:20:44,820 --> 00:20:42,550
so glad thank you for the join us it's

512
00:20:46,770 --> 00:20:44,830
interesting to see how nasa's data

513
00:20:48,810 --> 00:20:46,780

mining expertise is being applied in

514
00:20:51,360 --> 00:20:48,820
industry in a moment we'll take a closer

515
00:20:53,730 --> 00:20:51,370
look at the work nasa's doing to do

516
00:20:55,680 --> 00:20:53,740
research in data mining for aviation

517
00:20:57,390 --> 00:20:55,690
safety but first we want to give you an

518
00:20:59,340 --> 00:20:57,400
opportunity to get acquainted with one

519
00:21:01,200 --> 00:20:59,350
of our four aeronautics research centers

520
00:23:02,940 --> 00:21:01,210
the Dryden Flight Research Center in

521
00:23:24,720 --> 00:23:20,330
you

522
00:23:26,759 --> 00:23:24,730
aeronautics research discussion program

523
00:23:29,789 --> 00:23:26,769
brought to you by NASA our topic today

524
00:23:33,320 --> 00:23:29,799
is data mining and how NASA research is

525
00:23:36,149 --> 00:23:33,330
helping to make air safety better

526
00:23:38,039 --> 00:23:36,159
leading this section of our discussion

527
00:23:40,080 --> 00:23:38,049
is NASA's associate administrator for

528
00:23:42,029 --> 00:23:40,090
aeronautics research jae-hwan shin and

529
00:23:45,509 --> 00:23:42,039
he'll be speaking with Ashok Srivastav

530
00:23:48,509 --> 00:23:45,519
ah who is the project manager for this

531
00:23:50,700 --> 00:23:48,519
is a long title so correct me if I get

532
00:23:52,950 --> 00:23:50,710
it wrong system-wide safety and

533
00:23:55,799 --> 00:23:52,960
assurance technologies in NASA's

534
00:23:59,009 --> 00:23:55,809
aviation safety program dr. Srivastava

535
00:24:01,619 --> 00:23:59,019
is also the leader of the intelligent

536
00:24:04,320 --> 00:24:01,629
data understanding group at NASA's Ames

537
00:24:07,139 --> 00:24:04,330
Research Center his group performs data

538
00:24:09,419 --> 00:24:07,149

mining research in a number of areas in

539
00:24:12,239 --> 00:24:09,429
aviation safety and application domains

540
00:24:15,299 --> 00:24:12,249
such as earth science to study global

541
00:24:16,739 --> 00:24:15,309
climate processes and astrophysics to

542
00:24:18,479 --> 00:24:16,749
help characterize the large-scale

543
00:24:21,090 --> 00:24:18,489
structure of the universe he's the

544
00:24:23,129 --> 00:24:21,100
author of many research papers and on

545
00:24:25,289 --> 00:24:23,139
data mining machine learning and text

546
00:24:27,539 --> 00:24:25,299
mining and currently editing two books

547
00:24:29,700 --> 00:24:27,549
man when do you find time to work on

548
00:24:32,639 --> 00:24:29,710
data mining applications in astronomy

549
00:24:33,149 --> 00:24:32,649
and health management dr. Shin thank you

550
00:24:35,940 --> 00:24:33,159
Beth

551
00:24:38,190 --> 00:24:35,950
Chuck it's good to see you again and

552
00:24:41,430 --> 00:24:38,200
thank you for being here with us we just

553
00:24:44,249 --> 00:24:41,440
heard from captain Jeff Hamlet about the

554
00:24:46,859 --> 00:24:44,259
progress Southwest Airlines making using

555
00:24:49,049 --> 00:24:46,869
NASA's technologies to address their

556
00:24:52,710 --> 00:24:49,059
maintenance in not only maintenance but

557
00:24:56,310 --> 00:24:52,720
efficiency interest you are here to talk

558
00:24:58,440 --> 00:24:56,320
about NASA's interest in data mining so

559
00:25:01,560 --> 00:24:58,450
why don't we start talking about data

560
00:25:03,539 --> 00:25:01,570
mining in general terms well thank you

561
00:25:05,039 --> 00:25:03,549
Jay one and thank you for inviting me to

562
00:25:06,210 --> 00:25:05,049
speak here it's really a pleasure to be

563
00:25:08,879 --> 00:25:06,220

here

564
00:25:11,099 --> 00:25:08,889
NASA as you know has some of the largest

565
00:25:14,190 --> 00:25:11,109
and most complex datasets in the world

566
00:25:17,249 --> 00:25:14,200
to give you an idea for instance in the

567
00:25:19,470 --> 00:25:17,259
earth sciences we have satellites

568
00:25:20,849 --> 00:25:19,480
orbiting the Earth constantly taking

569
00:25:25,139 --> 00:25:20,859
measurements at very high resolution

570
00:25:27,509 --> 00:25:25,149
sometimes 250 meters square resolution

571
00:25:28,910 --> 00:25:27,519
across the whole globe and all this data

572
00:25:31,280 --> 00:25:28,920
is streaming in to

573
00:25:33,890 --> 00:25:31,290
systems and it needs to be analyzed and

574
00:25:36,380 --> 00:25:33,900
understood if you look at space systems

575
00:25:38,510 --> 00:25:36,390
for instance we have telescopes looking

576
00:25:40,580 --> 00:25:38,520
out into space taking data at an

577
00:25:43,100 --> 00:25:40,590
unprecedented rate sometimes people say

578
00:25:45,320 --> 00:25:43,110
in tens of terabytes or even petabytes

579
00:25:47,960 --> 00:25:45,330
of data is accumulating from some of

580
00:25:49,550 --> 00:25:47,970
these systems if you look within the

581
00:25:51,830 --> 00:25:49,560
Earth's atmosphere at aeronautical

582
00:25:54,290 --> 00:25:51,840
applications we just heard from Jeff

583
00:25:56,150 --> 00:25:54,300
saying that their major carriers

584
00:25:58,700 --> 00:25:56,160
accumulating a hundred thousand flights

585
00:26:01,190 --> 00:25:58,710
per month within the United States

586
00:26:03,770 --> 00:26:01,200
there's about 10 million flights per

587
00:26:05,720 --> 00:26:03,780
year all of this data that's

588
00:26:08,300 --> 00:26:05,730

accumulating needs to be analyzed and

589
00:26:10,520 --> 00:26:08,310
understood to improve quality of life

590
00:26:12,470 --> 00:26:10,530
for the public and what we're doing is

591
00:26:14,570 --> 00:26:12,480
developing tools and techniques so that

592
00:26:17,420 --> 00:26:14,580
people for example in the safety

593
00:26:19,790 --> 00:26:17,430
industry can take these algorithms take

594
00:26:21,800 --> 00:26:19,800
these computer methods apply it to data

595
00:26:24,050 --> 00:26:21,810
and understand something useful that

596
00:26:26,630 --> 00:26:24,060
could have some good benefit from the

597
00:26:28,640 --> 00:26:26,640
safety perspective great it is

598
00:26:31,520 --> 00:26:28,650
mind-boggling to think about that amount

599
00:26:34,130 --> 00:26:31,530
of data so what particular technology

600
00:26:37,850 --> 00:26:34,140
are we developing and what is its

601
00:26:41,450 --> 00:26:37,860
relevance to flight polling well one of

602
00:26:43,250 --> 00:26:41,460
the key issues that comes up in the area

603
00:26:45,950 --> 00:26:43,260
of aviation safety and trying to

604
00:26:50,750 --> 00:26:45,960
understand some of these issues is that

605
00:26:52,910 --> 00:26:50,760
we have small events that are occurring

606
00:26:55,550 --> 00:26:52,920
for example in the national air space

607
00:26:58,640 --> 00:26:55,560
and these small events when taken

608
00:27:00,920 --> 00:26:58,650
together can form a bigger picture of

609
00:27:02,690 --> 00:27:00,930
some key issues that that are occurring

610
00:27:05,330 --> 00:27:02,700
and we need to develop algorithms that

611
00:27:07,520 --> 00:27:05,340
can actually sort through and understand

612
00:27:10,550 --> 00:27:07,530
that so that an analyst can really make

613
00:27:12,170 --> 00:27:10,560

use of it and make good decisions one of

614
00:27:16,700 --> 00:27:12,180
the key issues that we face is that

615
00:27:18,860 --> 00:27:16,710
these anomalies can really start to

616
00:27:22,550 --> 00:27:18,870
creep up and as they creep up they can

617
00:27:26,240 --> 00:27:22,560
lead to bigger problems in the future so

618
00:27:29,450 --> 00:27:26,250
what sort of information are we talking

619
00:27:32,600 --> 00:27:29,460
about here and how much of it is out

620
00:27:35,210 --> 00:27:32,610
there well again if you go back to this

621
00:27:37,700 --> 00:27:35,220
example we have about 10 million flights

622
00:27:40,230 --> 00:27:37,710
occurring within the United States each

623
00:27:42,690 --> 00:27:40,240
airplane might be outfitted to

624
00:27:45,780 --> 00:27:42,700
obtain maybe a few hundred parameters of

625
00:27:48,120 --> 00:27:45,790
data each parameter is sampled once per

626
00:27:50,610 --> 00:27:48,130
second so this is an astronomical amount

627
00:27:53,580 --> 00:27:50,620
of information on top of that we have

628
00:27:55,169 --> 00:27:53,590
text reports and these text reports are

629
00:27:57,419 --> 00:27:55,179
written by pilots they're written by

630
00:27:59,730 --> 00:27:57,429
co-pilots or other members of the flight

631
00:28:01,669 --> 00:27:59,740
crew we're trying to take all this

632
00:28:04,680 --> 00:28:01,679
information together in order to

633
00:28:06,450 --> 00:28:04,690
discover an anomaly sometimes this is

634
00:28:08,640 --> 00:28:06,460
like looking for a needle in a haystack

635
00:28:10,799 --> 00:28:08,650
which means that we're really trying to

636
00:28:13,799 --> 00:28:10,809
figure out if there's something small

637
00:28:16,919 --> 00:28:13,809
that's occurring that could lead to a

638
00:28:20,820 --> 00:28:16,929

bigger problem I think we have a slide

639
00:28:22,980 --> 00:28:20,830
on the idea that in the past we were

640
00:28:24,720 --> 00:28:22,990
looking at forensic information and I

641
00:28:27,270 --> 00:28:24,730
think that's very important we should

642
00:28:30,570 --> 00:28:27,280
continue to do this where you basically

643
00:28:32,460 --> 00:28:30,580
look for the last accident what happened

644
00:28:34,799 --> 00:28:32,470
why it happened figure all of that out

645
00:28:37,110 --> 00:28:34,809
and then and then try and make sure that

646
00:28:39,360 --> 00:28:37,120
doesn't occur again what we're moving to

647
00:28:41,549 --> 00:28:39,370
though though is a more prognostic

648
00:28:44,370 --> 00:28:41,559
approach and in the prognostic approach

649
00:28:46,650 --> 00:28:44,380
it's really about looking at the data

650
00:28:48,570 --> 00:28:46,660
and what it's saying at this moment and

651
00:28:52,230 --> 00:28:48,580
then making a prediction into the future

652
00:28:54,600 --> 00:28:52,240
and saying 10 minutes from now 10 hours

653
00:28:58,169 --> 00:28:54,610
from now 10 days from now something bad

654
00:29:00,180 --> 00:28:58,179
could happen and so we need to to make

655
00:29:02,280 --> 00:29:00,190
the right decision so we're moving from

656
00:29:04,290 --> 00:29:02,290
forensics to prognostics which is I

657
00:29:08,630 --> 00:29:04,300
think an important development in the

658
00:29:12,060 --> 00:29:08,640
industry so your needle in haystack

659
00:29:15,810 --> 00:29:12,070
analogy is is very intriguing can you

660
00:29:18,000 --> 00:29:15,820
tell us more about that sure when we're

661
00:29:21,210 --> 00:29:18,010
searching for something that's rare or

662
00:29:23,370 --> 00:29:21,220
something that is a precursor to a

663
00:29:25,350 --> 00:29:23,380

failure one of the key issues is that we

664
00:29:28,230 --> 00:29:25,360
need to validate it and make sure that

665
00:29:30,690 --> 00:29:28,240
that thing that we're discovering isn't

666
00:29:33,120 --> 00:29:30,700
just another needle in the haystack sit

667
00:29:34,740 --> 00:29:33,130
sit another piece of hay I should say in

668
00:29:38,130 --> 00:29:34,750
the haystack that it's really something

669
00:29:40,169 --> 00:29:38,140
different and we're fortunate that dr.

670
00:29:42,480 --> 00:29:40,179
nikunj Oso who's a member of my group at

671
00:29:43,770 --> 00:29:42,490
NASA Ames Research Center is an expert

672
00:29:45,540 --> 00:29:43,780
in the field of data mining and machine

673
00:29:47,820 --> 00:29:45,550
learning and we're going to have a video

674
00:29:49,610 --> 00:29:47,830
now in which he discusses some of the

675
00:29:52,279 --> 00:29:49,620
applications of data mining in

676
00:29:55,370 --> 00:29:52,289
anomaly detection so if we can roll that

677
00:29:57,970 --> 00:29:55,380
video I lead a team of seven of us that

678
00:30:02,169 --> 00:29:57,980
look at data from commercial aviation

679
00:30:05,060 --> 00:30:02,179
and develop tools and technologies for

680
00:30:07,539 --> 00:30:05,070
data mining on that data for those so

681
00:30:10,490 --> 00:30:07,549
that analysts can use that to find

682
00:30:12,950 --> 00:30:10,500
operationally significant anomalies if

683
00:30:15,380 --> 00:30:12,960
our algorithms identify something as

684
00:30:18,200 --> 00:30:15,390
being anomalous that's not necessarily

685
00:30:21,710 --> 00:30:18,210
bad or necessarily dangerous it's just

686
00:30:23,750 --> 00:30:21,720
unusual but those are the candidates

687
00:30:26,450 --> 00:30:23,760
then that we would present to a domain

688
00:30:28,639 --> 00:30:26,460

expert and say that maybe ask if if

689
00:30:30,830 --> 00:30:28,649
maybe some of those are operationally

690
00:30:33,230 --> 00:30:30,840
significant we at least assume that that

691
00:30:36,110 --> 00:30:33,240
overwhelming majority of flights that

692
00:30:38,230 --> 00:30:36,120
are similar to one another are normal

693
00:30:40,639 --> 00:30:38,240
and so there probably are not

694
00:30:43,190 --> 00:30:40,649
operationally significant or dangerous

695
00:30:45,950 --> 00:30:43,200
events occurring within those flights I

696
00:30:48,830 --> 00:30:45,960
often like to give the example that is

697
00:30:50,630 --> 00:30:48,840
probably fairly popular is looking at

698
00:30:52,820 --> 00:30:50,640
the in case of supermarkets you look at

699
00:30:54,590 --> 00:30:52,830
so-called market basket analysis so you

700
00:30:57,440 --> 00:30:54,600
look at customers and what their buying

701
00:30:59,779 --> 00:30:57,450
patterns are and you might then try to

702
00:31:02,000 --> 00:30:59,789
group them in different ways and then

703
00:31:03,409 --> 00:31:02,010
you know basically act based on those

704
00:31:06,139 --> 00:31:03,419
groups rather than based on the

705
00:31:07,850 --> 00:31:06,149
individuals and in some way we we do

706
00:31:09,500 --> 00:31:07,860
that here as well so within aviation

707
00:31:11,510 --> 00:31:09,510
safety like I was mentioning we

708
00:31:13,070 --> 00:31:11,520
identified that large block of normal

709
00:31:16,730 --> 00:31:13,080
flights but then we say okay here are

710
00:31:18,169 --> 00:31:16,740
some groups of abnormal flights and then

711
00:31:20,810 --> 00:31:18,179
you try to learn

712
00:31:22,159 --> 00:31:20,820
think about those particular those

713
00:31:24,560 --> 00:31:22,169

particular anomalies trying to determine

714
00:31:27,200 --> 00:31:24,570
are they operationally significant or

715
00:31:30,039 --> 00:31:27,210
are they just statistically unusual so

716
00:31:33,710 --> 00:31:30,049
essentially we ask the question what

717
00:31:36,169 --> 00:31:33,720
data is kind of most unusual relative to

718
00:31:38,389 --> 00:31:36,179
the rest we definitely try to make our

719
00:31:40,549 --> 00:31:38,399
algorithms useful to analyst so in fact

720
00:31:44,509 --> 00:31:40,559
we have deployed some of our algorithms

721
00:31:46,730 --> 00:31:44,519
at FAA and at some airlines and they

722
00:31:49,279 --> 00:31:46,740
have tested these algorithms and given

723
00:31:52,279 --> 00:31:49,289
us feedback on how well our out of

724
00:31:54,230 --> 00:31:52,289
them's are doing and also how we can

725
00:31:57,200 --> 00:31:54,240
make them to better make them more

726
00:31:59,419 --> 00:31:57,210
efficient help them you know find the

727
00:32:05,060 --> 00:31:59,429
events that they're interested in more

728
00:32:09,480 --> 00:32:06,660
so shall

729
00:32:12,360 --> 00:32:09,490
nikunj ii gave us good understanding of

730
00:32:16,610 --> 00:32:12,370
what needles were looking for in the

731
00:32:19,470 --> 00:32:16,620
haystack so to speak but he mentioned

732
00:32:21,720 --> 00:32:19,480
showing the resorts to domain experts

733
00:32:24,450 --> 00:32:21,730
who are they and what do they think

734
00:32:26,820 --> 00:32:24,460
about needles that we're finding well

735
00:32:28,530 --> 00:32:26,830
one of the most interesting aspects in

736
00:32:31,170 --> 00:32:28,540
my opinion of data mining is that we

737
00:32:33,420 --> 00:32:31,180
don't do it in a vacuum we actually do

738
00:32:35,880 --> 00:32:33,430

it hand in hand with domain experts like

739
00:32:38,880 --> 00:32:35,890
Jeff or other people who are really

740
00:32:40,560 --> 00:32:38,890
interested and knowledgeable about the

741
00:32:43,740 --> 00:32:40,570
operations that we're trying to analyze

742
00:32:46,380 --> 00:32:43,750
so in the next video you're going to see

743
00:32:48,330 --> 00:32:46,390
two people who we work with closely one

744
00:32:50,610 --> 00:32:48,340
is a gentleman named Brian Matthews

745
00:32:55,230 --> 00:32:50,620
Brian is a member of my group and his

746
00:32:57,540 --> 00:32:55,240
job is to develop and then validate the

747
00:33:00,570 --> 00:32:57,550
algorithms that we have unreal data and

748
00:33:02,490 --> 00:33:00,580
so he goes through and he sees the

749
00:33:06,240 --> 00:33:02,500
nature of the results that are coming up

750
00:33:09,630 --> 00:33:06,250
but we also have a captain a retired

751
00:33:11,610 --> 00:33:09,640
captain of a triple7 who works with us

752
00:33:14,250 --> 00:33:11,620
for the last many years and his name is

753
00:33:16,980 --> 00:33:14,260
Bob Lawrence so Brian and Bob worked

754
00:33:19,440 --> 00:33:16,990
together in order to look at the results

755
00:33:20,970 --> 00:33:19,450
that are coming off of the computer

756
00:33:23,790 --> 00:33:20,980
programs the algorithms that we're

757
00:33:25,440 --> 00:33:23,800
developing and then Bob can give us some

758
00:33:27,600 --> 00:33:25,450
insight into the operational

759
00:33:29,790 --> 00:33:27,610
significance because what we want to do

760
00:33:32,040 --> 00:33:29,800
is really discover information that's

761
00:33:34,800 --> 00:33:32,050
useful for the flying public useful for

762
00:33:36,450 --> 00:33:34,810
the carrier's useful for the FAA and so

763
00:33:39,000 --> 00:33:36,460

this is the way we go through and and

764
00:33:41,880 --> 00:33:39,010
use our domain experts so the next video

765
00:33:43,710 --> 00:33:41,890
will show you Brian and Bob interacting

766
00:33:45,960 --> 00:33:43,720
and some of the results that they

767
00:33:48,870 --> 00:33:45,970
they've gotten I'm retired

768
00:33:52,820 --> 00:33:48,880
airline captain Bob Lawrence I flew for

769
00:33:55,640 --> 00:33:52,830
33 years for a major US airline

770
00:33:57,680 --> 00:33:55,650
and subsequently for eight years managed

771
00:34:00,380 --> 00:33:57,690
a NASA project called aviation

772
00:34:02,660 --> 00:34:00,390
performance measuring systems and I'm

773
00:34:06,800 --> 00:34:02,670
currently working with NASA's data

774
00:34:10,460 --> 00:34:06,810
mining group as a domain expert to help

775
00:34:13,280 --> 00:34:10,470
relate their scientific results to the

776
00:34:15,620 --> 00:34:13,290
world of commercial aviation transport

777
00:34:18,350 --> 00:34:15,630
my name is Brian Matthews I've been at

778
00:34:21,740 --> 00:34:18,360
NASA aim for about nine years I've

779
00:34:24,410 --> 00:34:21,750
recently started working in the aviation

780
00:34:27,770 --> 00:34:24,420
safety program where a fantastic to find

781
00:34:31,970 --> 00:34:27,780
anomalous events in enormous amount of

782
00:34:33,260 --> 00:34:31,980
aviation data I work with Bob and once

783
00:34:35,510 --> 00:34:33,270
we find the anomalies that we've

784
00:34:38,090 --> 00:34:35,520
detected with our algorithms we use

785
00:34:40,820 --> 00:34:38,100
Bob's expertise to determine these

786
00:34:43,910 --> 00:34:40,830
anomalies are significant operational

787
00:34:45,910 --> 00:34:43,920
significant for the airlines well there

788
00:34:47,780 --> 00:34:45,920

are basically two kinds of data

789
00:34:52,160 --> 00:34:47,790
currently that we're working with

790
00:34:54,620 --> 00:34:52,170
numeric data and text data text data of

791
00:34:57,640 --> 00:34:54,630
course is pilot reports incident reports

792
00:35:00,890 --> 00:34:57,650
flight safety reports the numeric data

793
00:35:04,040 --> 00:35:00,900
comes from a flight data recorder which

794
00:35:07,180 --> 00:35:04,050
is measuring things such as switch

795
00:35:11,090 --> 00:35:07,190
positions on the control panel

796
00:35:13,240 --> 00:35:11,100
warning modes the position of the flight

797
00:35:16,130 --> 00:35:13,250
instruments themselves altimeter x'

798
00:35:18,860 --> 00:35:16,140
airspeed the positions of the flight

799
00:35:21,650 --> 00:35:18,870
controls whether it's the aircraft yoke

800
00:35:24,290 --> 00:35:21,660
or the throttles or the landing gear or

801
00:35:25,760 --> 00:35:24,300
the flaps or the spoilers we can deal

802
00:35:28,610 --> 00:35:25,770
with up to millions of flights a year

803
00:35:30,440 --> 00:35:28,620
from some airlines it's such an

804
00:35:31,850 --> 00:35:30,450
interesting exercise when Brian and I

805
00:35:34,370 --> 00:35:31,860
sit down in front of a computer screen

806
00:35:37,160 --> 00:35:34,380
and look at the anomalies that the

807
00:35:38,030 --> 00:35:37,170
program has come up with and then we try

808
00:35:40,849 --> 00:35:38,040
to

809
00:35:42,650 --> 00:35:40,859
explaining what was going on what's the

810
00:35:46,599 --> 00:35:42,660
story here what's what story is being

811
00:35:51,260 --> 00:35:46,609
told and so we'll select different

812
00:35:54,080 --> 00:35:51,270
parameters to look at and we'll go

813
00:35:56,750 --> 00:35:54,090

through and try to explain what it is

814
00:35:58,310 --> 00:35:56,760
that was going on and sometimes it turns

815
00:36:01,250 --> 00:35:58,320
out that what was happening was a

816
00:36:03,290 --> 00:36:01,260
completely normal thing but then once in

817
00:36:05,540 --> 00:36:03,300
a while we get one of those gems which

818
00:36:08,540 --> 00:36:05,550
is very interesting to an airline

819
00:36:11,630 --> 00:36:08,550
analyst and we hope that they will find

820
00:36:13,910 --> 00:36:11,640
our software useful in the future it's

821
00:36:17,930 --> 00:36:13,920
like playing detective or you're not

822
00:36:19,820 --> 00:36:17,940
able to interview the airline pilot but

823
00:36:22,099 --> 00:36:19,830
you have all this data that you can sift

824
00:36:24,560 --> 00:36:22,109
through and display different modes of

825
00:36:26,300 --> 00:36:24,570
the flight and then expertise who's

826
00:36:28,040 --> 00:36:26,310
who's been there you can kind of

827
00:36:30,109 --> 00:36:28,050
determine what was actually happening

828
00:36:34,130 --> 00:36:30,119
here in that flight from the point of

829
00:36:36,020 --> 00:36:34,140
view of a pilot I am painfully aware of

830
00:36:38,870 --> 00:36:36,030
the fact that I have only personally

831
00:36:41,240 --> 00:36:38,880
been exposed to a small part of the

832
00:36:43,640 --> 00:36:41,250
operation myself I find it very

833
00:36:47,780 --> 00:36:43,650
comforting to know that there are people

834
00:36:51,410 --> 00:36:47,790
out there who are constructing the big

835
00:36:55,400 --> 00:36:51,420
picture and then analyzing it to pick

836
00:36:57,740 --> 00:36:55,410
out what could possibly happen most of

837
00:37:00,470 --> 00:36:57,750
what could possibly happen has never

838
00:37:03,200 --> 00:37:00,480

happened to me personally but it's

839
00:37:07,640 --> 00:37:03,210
happened but when you add together all

840
00:37:09,530 --> 00:37:07,650
of the anomalous events and you

841
00:37:17,160 --> 00:37:09,540
construct a big picture out of that then

842
00:37:23,200 --> 00:37:20,769
great joke I'd say that's really good

843
00:37:25,779 --> 00:37:23,210
teamwork right there

844
00:37:27,999 --> 00:37:25,789
I think we're getting good education of

845
00:37:31,289 --> 00:37:28,009
what we are looking for and how it

846
00:37:34,390 --> 00:37:31,299
relates to real life situations but

847
00:37:39,069 --> 00:37:34,400
could you say more about the complexity

848
00:37:41,979 --> 00:37:39,079
of the task sure what we've noticed in

849
00:37:44,019 --> 00:37:41,989
the pursuit of data mining in the area

850
00:37:46,569 --> 00:37:44,029
of aviation safety is that we're trying

851
00:37:48,670 --> 00:37:46,579
to look for what we call precursors or

852
00:37:50,680 --> 00:37:48,680
these are small events that could be

853
00:37:53,170 --> 00:37:50,690
occurring at different times so

854
00:37:54,940 --> 00:37:53,180
something happens now and then something

855
00:37:57,190 --> 00:37:54,950
happens a little later at different

856
00:37:59,289 --> 00:37:57,200
places different locations and we're

857
00:38:01,630 --> 00:37:59,299
trying to pull all of this together in

858
00:38:03,819 --> 00:38:01,640
order to obtain a big picture of what's

859
00:38:07,029 --> 00:38:03,829
going on I think Bob touched on it very

860
00:38:09,999 --> 00:38:07,039
nicely that any one person gets only a

861
00:38:11,950 --> 00:38:10,009
sliver of information a sliver of

862
00:38:13,690 --> 00:38:11,960
experience about what's going on in the

863
00:38:15,999 --> 00:38:13,700

National Airspace but what we're trying

864
00:38:18,249 --> 00:38:16,009
to do is understand it in all of its

865
00:38:20,729 --> 00:38:18,259
complexity and that's a difficult task

866
00:38:24,970 --> 00:38:20,739
that requires algorithms that can scale

867
00:38:26,289 --> 00:38:24,980
which means analyze and and and process

868
00:38:29,170 --> 00:38:26,299
massive amounts of data

869
00:38:31,029 --> 00:38:29,180
it requires algorithms that can process

870
00:38:33,819 --> 00:38:31,039
different kinds of information that's

871
00:38:36,400 --> 00:38:33,829
coming at different times it also

872
00:38:38,799 --> 00:38:36,410
requires the ability to know when you're

873
00:38:41,229 --> 00:38:38,809
right and when you're not wrong and also

874
00:38:43,509 --> 00:38:41,239
some measure of confidence and so all of

875
00:38:45,910 --> 00:38:43,519
these things taken together if you think

876
00:38:47,559 --> 00:38:45,920
about it becomes very complex but in my

877
00:38:49,150 --> 00:38:47,569
opinion that's where the reward lies

878
00:38:51,460 --> 00:38:49,160
that's where what makes this such an

879
00:38:54,069 --> 00:38:51,470
exciting field because we're really

880
00:38:57,069 --> 00:38:54,079
doing something that has a lot of impact

881
00:38:59,440 --> 00:38:57,079
potential positive impact for the flying

882
00:39:02,559 --> 00:38:59,450
public so it's it's very exciting

883
00:39:05,079 --> 00:39:02,569
great I think it's not going to be very

884
00:39:07,690 --> 00:39:05,089
difficult to imagine that this kind of

885
00:39:09,940 --> 00:39:07,700
technology can benefit not only air

886
00:39:12,190 --> 00:39:09,950
transportation system but also space

887
00:39:15,999 --> 00:39:12,200
systems and I understand you actually

888
00:39:17,920 --> 00:39:16,009

had such a case to help space shuttle

889
00:39:20,620 --> 00:39:17,930
program could you say more about there

890
00:39:22,539 --> 00:39:20,630
sure you know a couple of years ago I

891
00:39:24,900 --> 00:39:22,549
remember it was on Valentine's Day it

892
00:39:26,910 --> 00:39:24,910
was a Saturday my cell phone rang

893
00:39:29,490 --> 00:39:26,920
it was a colleague of mine from Kennedy

894
00:39:31,860 --> 00:39:29,500
Space Center who had called up to find

895
00:39:34,680 --> 00:39:31,870
out whether or not we could apply some

896
00:39:37,500 --> 00:39:34,690
data mining techniques to discover a

897
00:39:40,410 --> 00:39:37,510
potential anomaly on shuttle mission and

898
00:39:42,600 --> 00:39:40,420
so this was just before STS 119 and I

899
00:39:44,580 --> 00:39:42,610
think on the monitor you can see the

900
00:39:48,300 --> 00:39:44,590
launch of of that particular mission

901
00:39:50,460 --> 00:39:48,310
what we did is develop some algorithms

902
00:39:52,200 --> 00:39:50,470
that could be applied to the main

903
00:39:55,320 --> 00:39:52,210
propulsion system of the Space Shuttle

904
00:39:57,330 --> 00:39:55,330
and to detect anomalies in one of the

905
00:39:59,310 --> 00:39:57,340
hydrogen lines that was there and it

906
00:40:01,740 --> 00:39:59,320
turns out that we were able to detect it

907
00:40:04,590 --> 00:40:01,750
and that information went up into the

908
00:40:06,570 --> 00:40:04,600
Flight Readiness review through the NASA

909
00:40:08,610 --> 00:40:06,580
engineering and Safety Center to the

910
00:40:10,620 --> 00:40:08,620
Flight Readiness review and I think they

911
00:40:12,720 --> 00:40:10,630
used some of that information in making

912
00:40:14,730 --> 00:40:12,730
a determination about the next launch so

913
00:40:17,790 --> 00:40:14,740

we were really pleased to see how data

914
00:40:20,790 --> 00:40:17,800
mining could in short order affect an

915
00:40:25,920 --> 00:40:20,800
important mission like STS 119 that's

916
00:40:29,700 --> 00:40:25,930
wonderful so in the case of STS 119 you

917
00:40:33,150 --> 00:40:29,710
receive a request to look into something

918
00:40:36,720 --> 00:40:33,160
but if we were going through airline

919
00:40:41,030 --> 00:40:36,730
commercial airline data and say you ran

920
00:40:44,220 --> 00:40:41,040
into something noteworthy how can we

921
00:40:46,740 --> 00:40:44,230
inform about the data and what can we do

922
00:40:48,870 --> 00:40:46,750
about it well if we see something in

923
00:40:52,320 --> 00:40:48,880
data using one of our algorithms we

924
00:40:55,170 --> 00:40:52,330
notify the airline and and they then go

925
00:40:57,300 --> 00:40:55,180
and do further analysis determine

926
00:40:59,220 --> 00:40:57,310
whether or not what we've discovered is

927
00:41:01,440 --> 00:40:59,230
useful and then they might take some

928
00:41:04,140 --> 00:41:01,450
implementation make some implementation

929
00:41:05,760 --> 00:41:04,150
based on that but I think it's important

930
00:41:08,130 --> 00:41:05,770
to remember that NASA is developing

931
00:41:11,010 --> 00:41:08,140
tools and technologies so that others

932
00:41:13,830 --> 00:41:11,020
like the FAA like the carriers other

933
00:41:16,310 --> 00:41:13,840
safety analysts can really go through

934
00:41:18,960 --> 00:41:16,320
and analyze and manage some of that data

935
00:41:21,570 --> 00:41:18,970
another aspect of the issue is that we

936
00:41:24,570 --> 00:41:21,580
really don't have identified information

937
00:41:26,340 --> 00:41:24,580
in other words our data doesn't say the

938
00:41:29,510 --> 00:41:26,350

tail number of the aircraft it doesn't

939
00:41:32,760 --> 00:41:29,520
identify pilots or give really other

940
00:41:34,710 --> 00:41:32,770
identifying information and so what

941
00:41:37,200 --> 00:41:34,720
happens is that we use the data to

942
00:41:38,290 --> 00:41:37,210
validate the algorithms to check to make

943
00:41:40,690 --> 00:41:38,300
sure that they're do

944
00:41:43,090 --> 00:41:40,700
something useful but then we transition

945
00:41:44,650 --> 00:41:43,100
it over to carriers such as Southwest

946
00:41:46,810 --> 00:41:44,660
were delighted to be working with them

947
00:41:48,760 --> 00:41:46,820
because then they can take it do further

948
00:41:50,200 --> 00:41:48,770
validation and then operationalize that

949
00:41:52,540 --> 00:41:50,210
they've already done that in the past

950
00:41:54,120 --> 00:41:52,550
and I think that with our partnership we

951
00:41:58,450 --> 00:41:54,130
hope to do more of that in the future

952
00:42:00,400 --> 00:41:58,460
great I think folks will agree that this

953
00:42:03,670 --> 00:42:00,410
is a very exciting area to improve

954
00:42:06,310 --> 00:42:03,680
aviation safety but we actually do quite

955
00:42:09,160 --> 00:42:06,320
more than just data mining right in

956
00:42:11,770 --> 00:42:09,170
aviation safety could you say more in

957
00:42:13,870 --> 00:42:11,780
broader terms or what we are doing you

958
00:42:15,970 --> 00:42:13,880
know data mining is really the tip of

959
00:42:18,250 --> 00:42:15,980
the iceberg when it comes to aviation

960
00:42:20,500 --> 00:42:18,260
safety there's so much going on within

961
00:42:23,650 --> 00:42:20,510
nasa within the aviation safety program

962
00:42:26,010 --> 00:42:23,660
for example there's significant research

963
00:42:28,270 --> 00:42:26,020

going on in the area of icing

964
00:42:30,850 --> 00:42:28,280
understanding how icing affects both

965
00:42:33,250 --> 00:42:30,860
general aviation aircraft but also

966
00:42:35,080 --> 00:42:33,260
commercial aircraft understanding

967
00:42:39,250 --> 00:42:35,090
lightning and the effects that lightning

968
00:42:41,350 --> 00:42:39,260
can have on both types of aircraft I

969
00:42:43,600 --> 00:42:41,360
think is another important area of

970
00:42:45,790 --> 00:42:43,610
research we have research going on in an

971
00:42:48,280 --> 00:42:45,800
area called vehicle health management

972
00:42:50,560 --> 00:42:48,290
which is really looking at the detection

973
00:42:53,830 --> 00:42:50,570
and prediction of adverse events on a

974
00:42:55,690 --> 00:42:53,840
single aircraft that's very important we

975
00:42:58,030 --> 00:42:55,700
also have work going on in the human

976
00:43:00,130 --> 00:42:58,040
factors area where we try to understand

977
00:43:04,360 --> 00:43:00,140
how humans and these complex machines

978
00:43:06,640 --> 00:43:04,370
and then setzen and and so many of these

979
00:43:08,440 --> 00:43:06,650
machines interact in the seamless

980
00:43:10,810 --> 00:43:08,450
fashion so there's a lot of work going

981
00:43:12,910 --> 00:43:10,820
on in a number of areas in aviation

982
00:43:16,120 --> 00:43:12,920
safety and data mining is one of those

983
00:43:18,180 --> 00:43:16,130
areas which is hopefully going to have

984
00:43:22,630 --> 00:43:18,190
an important impact in the future

985
00:43:24,430 --> 00:43:22,640
good I I feel very fortunate and I think

986
00:43:27,400 --> 00:43:24,440
I'm speaking for everyone in NASA

987
00:43:31,240 --> 00:43:27,410
Aeronautics community to have you as one

988
00:43:33,520 --> 00:43:31,250

of our members to make this very very

989
00:43:36,640 --> 00:43:33,530
important progress in technical

990
00:43:42,280 --> 00:43:36,650
development but I'm curious what led you

991
00:43:44,680 --> 00:43:42,290
to NASA well many years ago I was

992
00:43:47,950 --> 00:43:44,690
contacted by the chief of information

993
00:43:50,080 --> 00:43:47,960
technology at NASA Ames Research Center

994
00:43:53,140 --> 00:43:50,090
and he said that if I came to NASA

995
00:43:54,940 --> 00:43:53,150
I would be able to set the strategic

996
00:43:58,150 --> 00:43:54,950
direction for data mining within the

997
00:44:00,460 --> 00:43:58,160
agency and since I was a child I have to

998
00:44:02,530 --> 00:44:00,470
say that NASA and the problems that it

999
00:44:05,860 --> 00:44:02,540
has have really captured my imagination

1000
00:44:08,430 --> 00:44:05,870
they're amazing fascinating problems and

1001
00:44:10,510 --> 00:44:08,440
to be a part of that I think was really

1002
00:44:14,680 --> 00:44:10,520
something that I couldn't turn down

1003
00:44:16,540 --> 00:44:14,690
after coming here and then realizing

1004
00:44:20,340 --> 00:44:16,550
that we could work in aviation safety

1005
00:44:22,690 --> 00:44:20,350
and having a potential impact there for

1006
00:44:25,150 --> 00:44:22,700
millions if not billions of people

1007
00:44:27,340 --> 00:44:25,160
around the world I thought this was an

1008
00:44:29,740 --> 00:44:27,350
amazing opportunity you know when I was

1009
00:44:31,600 --> 00:44:29,750
a child I used to go with my father to

1010
00:44:33,130 --> 00:44:31,610
the airport because he used to travel a

1011
00:44:36,430 --> 00:44:33,140
lot and I would see these airplanes

1012
00:44:38,290 --> 00:44:36,440
taking off and it it had left such an

1013
00:44:40,510 --> 00:44:38,300

impression on me that I wanted to be

1014
00:44:42,520 --> 00:44:40,520
part of that and I'm really delighted to

1015
00:44:45,130 --> 00:44:42,530
be here at NASA and part of the aviation

1016
00:44:47,800 --> 00:44:45,140
safety program we're really very

1017
00:44:51,100 --> 00:44:47,810
delighted as well Ashoka thank you very

1018
00:44:53,710 --> 00:44:51,110
much for a very exciting and in light

1019
00:44:56,680 --> 00:44:53,720
enlightening education that you have

1020
00:45:00,670 --> 00:44:56,690
given us today and I think this is a

1021
00:45:03,690 --> 00:45:00,680
great example with your lead in this

1022
00:45:06,120 --> 00:45:03,700
work and also Captain Hamlet's

1023
00:45:09,520 --> 00:45:06,130
collaboration with NASA

1024
00:45:11,530 --> 00:45:09,530
what are such a great example that we

1025
00:45:14,530 --> 00:45:11,540
are working together to help our flying

1026
00:45:16,420 --> 00:45:14,540
public so thank for your leadership and

1027
00:45:19,810 --> 00:45:16,430
thank you so much for everything you're

1028
00:45:21,400 --> 00:45:19,820
doing thank you for the opportunity next

1029
00:45:23,380 --> 00:45:21,410
we'll take some questions from our

1030
00:45:24,880 --> 00:45:23,390
audience but first we want to get you

1031
00:45:27,130 --> 00:45:24,890
acquainted with another one of our

1032
00:45:42,690 --> 00:45:27,140
Aeronautics research centers the Langley

1033
00:45:51,480 --> 00:45:46,960
before flights there are ideas before

1034
00:45:55,120 --> 00:45:51,490
missions answers and before exploration

1035
00:45:58,450 --> 00:45:55,130
comes innovation it's the way NASA works

1036
00:46:01,270 --> 00:45:58,460
and for nearly 100 years a first stop

1037
00:46:04,080 --> 00:46:01,280
for answers and ideas the Langley

1038
00:46:07,000 --> 00:46:04,090

Research Center the birthplace of NASA

1039
00:46:10,240 --> 00:46:07,010
this is where the aviation industry we

1040
00:46:12,490 --> 00:46:10,250
rely on was created where astronauts

1041
00:46:15,430 --> 00:46:12,500
trained and the space program was

1042
00:46:18,849 --> 00:46:15,440
launched and we're right now climate

1043
00:46:21,490 --> 00:46:18,859
change is being tracked from space today

1044
00:46:24,310 --> 00:46:21,500
the Langley Research Center jump starts

1045
00:46:27,340 --> 00:46:24,320
what's next for NASA and what makes life

1046
00:46:30,070 --> 00:46:27,350
better for all of us here

1047
00:46:32,650 --> 00:46:30,080
the scientists engineers and creative

1048
00:46:35,950 --> 00:46:32,660
problem-solvers have innovation in their

1049
00:46:38,920 --> 00:46:35,960
DNA and tackle some of the most complex

1050
00:46:42,190 --> 00:46:38,930
challenges of our time like designing

1051
00:46:44,890 --> 00:46:42,200
greener quieter aircraft and a safe

1052
00:46:47,440 --> 00:46:44,900
interstate Skyway developing the

1053
00:46:49,830 --> 00:46:47,450
earliest robotic spacecraft and paving

1054
00:46:52,450 --> 00:46:49,840
the way for planetary exploration

1055
00:46:54,940 --> 00:46:52,460
uncovering how our atmosphere is

1056
00:46:57,400 --> 00:46:54,950
changing and how that impacts everything

1057
00:47:00,250 --> 00:46:57,410
from air travel to climate change and

1058
00:47:04,120 --> 00:47:00,260
making renewable energy sources viable

1059
00:47:05,400 --> 00:47:04,130
not just vital Langley's is a vision of

1060
00:47:08,220 --> 00:47:05,410
safer fast

1061
00:47:10,800 --> 00:47:08,230
more affordable airplanes of personal

1062
00:47:14,330 --> 00:47:10,810
journeys into space of a world where

1063
00:47:20,310 --> 00:47:14,340

climate change is no longer a threat

1064
00:47:24,840 --> 00:47:20,320
it's a vision based on answers and ideas

1065
00:47:33,510 --> 00:47:24,850
and with those Langley is creating the

1066
00:47:37,770 --> 00:47:35,400
you're watching the leading edge and

1067
00:47:39,359 --> 00:47:37,780
Aeronautics research discussion program

1068
00:47:42,030 --> 00:47:39,369
brought to you by NASA we've been

1069
00:47:44,400 --> 00:47:42,040
discussing data mining today and NASA's

1070
00:47:46,980 --> 00:47:44,410
work to improve to enable the airlines

1071
00:47:49,380 --> 00:47:46,990
to improve their aviation safety and the

1072
00:47:51,600 --> 00:47:49,390
safety of your flight our panelists

1073
00:47:54,270 --> 00:47:51,610
today are Jeff Hamlet of Southwest

1074
00:47:56,550 --> 00:47:54,280
Airlines and Ashok Srivastav of NASA's

1075
00:47:58,109 --> 00:47:56,560
aviation safety program now is the

1076
00:48:00,570 --> 00:47:58,119
opportunity for members of our audience

1077
00:48:02,820 --> 00:48:00,580
to ask questions we'll start with

1078
00:48:05,280 --> 00:48:02,830
questions from those here at the

1079
00:48:07,260 --> 00:48:05,290
headquarters auditorium and then from

1080
00:48:09,660 --> 00:48:07,270
employees participating remotely around

1081
00:48:15,720 --> 00:48:09,670
our centers anyone here in the audience

1082
00:48:18,270 --> 00:48:15,730
with a question right here wait wait for

1083
00:48:19,560 --> 00:48:18,280
please wait for the mic thank you have a

1084
00:48:22,140 --> 00:48:19,570
question you've talked a lot about

1085
00:48:24,420 --> 00:48:22,150
integrating information from text

1086
00:48:25,859 --> 00:48:24,430
sources with quantitative data can you

1087
00:48:27,570 --> 00:48:25,869
tell us more about how you do that it

1088
00:48:29,850 --> 00:48:27,580

seems like so many fields could use

1089
00:48:34,410 --> 00:48:29,860
methods to integrate narrative data with

1090
00:48:37,609 --> 00:48:34,420
quantitative data sure one way that we

1091
00:48:39,510 --> 00:48:37,619
do this is we have developed of a

1092
00:48:42,300 --> 00:48:39,520
methodology or an approach a

1093
00:48:45,210 --> 00:48:42,310
mathematical approach which treats all

1094
00:48:47,280 --> 00:48:45,220
data as one depending independent of

1095
00:48:50,400 --> 00:48:47,290
where it comes from so it could take

1096
00:48:52,470 --> 00:48:50,410
text it could take numbers you could

1097
00:48:55,980 --> 00:48:52,480
take others the kinds of data structures

1098
00:48:58,380 --> 00:48:55,990
like trees and graphs and networks and

1099
00:49:00,450 --> 00:48:58,390
it combines that all into a single

1100
00:49:02,820 --> 00:49:00,460
mathematical framework and then you turn

1101
00:49:05,010 --> 00:49:02,830
a crank and when you turn the crank you

1102
00:49:07,550 --> 00:49:05,020
get out a prediction algorithm you can

1103
00:49:10,500 --> 00:49:07,560
get out an anomaly detection algorithm

1104
00:49:12,030 --> 00:49:10,510
classification and so forth so that's

1105
00:49:14,970 --> 00:49:12,040
the approach that we're taking currently

1106
00:49:16,560 --> 00:49:14,980
there are some top people in the field

1107
00:49:18,630 --> 00:49:16,570
of machine learning and data mining from

1108
00:49:21,090 --> 00:49:18,640
around the country that are exploring

1109
00:49:22,500 --> 00:49:21,100
complementary approaches but this is one

1110
00:49:24,840 --> 00:49:22,510
of the approaches that we're taking

1111
00:49:27,060 --> 00:49:24,850
right now and we're hoping to see in

1112
00:49:30,599 --> 00:49:27,070
partnership with Southwest some benefit

1113
00:49:35,790 --> 00:49:30,609

coming out of this research here in the

1114
00:49:39,090 --> 00:49:35,800
auditorium there's one over there my

1115
00:49:41,810 --> 00:49:39,100
question is for dr. Srivastava you had

1116
00:49:44,300 --> 00:49:41,820
mentioned in your explanation of

1117
00:49:47,640 --> 00:49:44,310
analyzing the data that you are able to

1118
00:49:52,920 --> 00:49:47,650
make some predictions about when

1119
00:49:54,510 --> 00:49:52,930
airline or airplane parts may fail such

1120
00:49:57,240 --> 00:49:54,520
as may be and you've said into your

1121
00:49:59,910 --> 00:49:57,250
example ten minutes ten hours you know

1122
00:50:01,590 --> 00:49:59,920
etc etc what is the range that you can

1123
00:50:04,260 --> 00:50:01,600
predict in the future from immediate

1124
00:50:07,070 --> 00:50:04,270
future to you know and maybe even years

1125
00:50:10,260 --> 00:50:07,080
to come and in that means what is your

1126
00:50:12,900 --> 00:50:10,270
percentage error what is the accuracy or

1127
00:50:17,220 --> 00:50:12,910
precision of your predictions those are

1128
00:50:20,910 --> 00:50:17,230
great questions so indicate depends on

1129
00:50:22,620 --> 00:50:20,920
the system and the subsystems that we're

1130
00:50:24,840 --> 00:50:22,630
talking about on a particular aircraft

1131
00:50:26,870 --> 00:50:24,850
we're looking at it's important to

1132
00:50:29,550 --> 00:50:26,880
remember that we're looking at

1133
00:50:31,620 --> 00:50:29,560
information from components of an

1134
00:50:34,440 --> 00:50:31,630
aircraft all the way up to the national

1135
00:50:37,050 --> 00:50:34,450
airspace so because of that it changes

1136
00:50:39,930 --> 00:50:37,060
the range of our predictions and the way

1137
00:50:42,660 --> 00:50:39,940
we measure accuracy so for example if

1138
00:50:43,940 --> 00:50:42,670

you take a specific component of an

1139
00:50:46,710 --> 00:50:43,950
aircraft for example an

1140
00:50:49,080 --> 00:50:46,720
electromechanical actuator some of our

1141
00:50:51,570 --> 00:50:49,090
partners at NASA Ames Research Center

1142
00:50:55,050 --> 00:50:51,580
developing technologies that can pretty

1143
00:50:57,240 --> 00:50:55,060
accurately predict when that actuator is

1144
00:50:59,970 --> 00:50:57,250
going to fail now it turns out actuators

1145
00:51:02,730 --> 00:50:59,980
don't fail very often maybe I've heard

1146
00:51:04,740 --> 00:51:02,740
after millions of hours of operation

1147
00:51:07,470 --> 00:51:04,750
there might be a failure so these are

1148
00:51:09,750 --> 00:51:07,480
extremely rare events what we're looking

1149
00:51:11,880 --> 00:51:09,760
at though and data mining is really the

1150
00:51:14,490 --> 00:51:11,890
confluence of a number of different

1151
00:51:16,410 --> 00:51:14,500
factors that could lead to a problem so

1152
00:51:18,750 --> 00:51:16,420
it might be that the component is a

1153
00:51:21,270 --> 00:51:18,760
little off on an aircraft it might be

1154
00:51:22,770 --> 00:51:21,280
that the aircraft in front is having a

1155
00:51:26,340 --> 00:51:22,780
little bit of a problem it might be that

1156
00:51:28,500 --> 00:51:26,350
the pilot has some fatigue and taking

1157
00:51:33,420 --> 00:51:28,510
all these things into account what is

1158
00:51:36,480 --> 00:51:33,430
the likelihood of a of a problem

1159
00:51:38,430 --> 00:51:36,490
so measuring accuracy in that context

1160
00:51:40,860 --> 00:51:38,440
and itself is difficult so it's hard to

1161
00:51:43,410 --> 00:51:40,870
say what our accuracy or precision or

1162
00:51:44,370 --> 00:51:43,420
recall is in those terms but that's

1163
00:51:47,970 --> 00:51:44,380

something that we're actively

1164
00:51:50,370 --> 00:51:47,980
researching alright I've got a question

1165
00:51:53,070 --> 00:51:50,380
for Jeff I I was intrigued by your

1166
00:51:54,510 --> 00:51:53,080
discussion of the three data collection

1167
00:51:57,360 --> 00:51:54,520
programs can you tell us a little bit

1168
00:51:59,790 --> 00:51:57,370
more about how those programs interact

1169
00:52:01,260 --> 00:51:59,800
and or how they may or may not work

1170
00:52:04,080 --> 00:52:01,270
together sure

1171
00:52:05,130 --> 00:52:04,090
all those programs are for the ASAP and

1172
00:52:06,780 --> 00:52:05,140
the fadap programs are voluntary

1173
00:52:09,120 --> 00:52:06,790
programs that are really formed on the

1174
00:52:11,670 --> 00:52:09,130
basis of a very strong relationship with

1175
00:52:13,350 --> 00:52:11,680
our pilot Association and one of the

1176
00:52:15,420 --> 00:52:13,360
things that we really work hard to do is

1177
00:52:17,400 --> 00:52:15,430
preserve the D identification of events

1178
00:52:20,130 --> 00:52:17,410
so one of the things that we're not able

1179
00:52:22,950 --> 00:52:20,140
to do is take a pilot disclosure from

1180
00:52:25,290 --> 00:52:22,960
ASAP and then go into fadap and look for

1181
00:52:28,560 --> 00:52:25,300
a specific flight that man may have had

1182
00:52:30,420 --> 00:52:28,570
that same profile but what we can do is

1183
00:52:34,050 --> 00:52:30,430
take the concept of the event that

1184
00:52:36,300 --> 00:52:34,060
occurred and develop an algorithm to go

1185
00:52:38,070 --> 00:52:36,310
look for and in an aggregate sense look

1186
00:52:41,370 --> 00:52:38,080
for those types of events that occur in

1187
00:52:43,020 --> 00:52:41,380
the data and both of those it works both

1188
00:52:44,730 --> 00:52:43,030

ways where we can see something in fadap

1189
00:52:47,190 --> 00:52:44,740
I think I talked about that earlier and

1190
00:52:48,660 --> 00:52:47,200
we can go search our database an ASAP to

1191
00:52:50,910 --> 00:52:48,670
get the explanation of why that's

1192
00:52:52,500 --> 00:52:50,920
occurring and then we can know exactly

1193
00:52:54,720 --> 00:52:52,510
what's happening with the aircraft

1194
00:52:56,550 --> 00:52:54,730
through our fadap program all right and

1195
00:52:58,350 --> 00:52:56,560
we do have some questions that have been

1196
00:53:01,860 --> 00:52:58,360
submitted to us offline and this is a

1197
00:53:05,670 --> 00:53:01,870
this is an interesting one I guess for a

1198
00:53:08,370 --> 00:53:05,680
stroke could you use Google to mine the

1199
00:53:10,920 --> 00:53:08,380
data you get from airplanes at least the

1200
00:53:13,410 --> 00:53:10,930
written reports you know it's a question

1201
00:53:15,630 --> 00:53:13,420
that I get a lot regarding Google

1202
00:53:17,880 --> 00:53:15,640
because Google is a ubiquitous everybody

1203
00:53:20,040 --> 00:53:17,890
uses that and people can reasonably say

1204
00:53:22,020 --> 00:53:20,050
can you use this to analyze text reports

1205
00:53:23,910 --> 00:53:22,030
and I think the answer the short answer

1206
00:53:26,160 --> 00:53:23,920
is yes you could use it to look at the

1207
00:53:28,320 --> 00:53:26,170
text reports and get some idea of what's

1208
00:53:30,330 --> 00:53:28,330
in there it's basically a search engine

1209
00:53:32,450 --> 00:53:30,340
technology applied to text which makes

1210
00:53:34,890 --> 00:53:32,460
perfect sense what we're doing is

1211
00:53:37,050 --> 00:53:34,900
looking at not only the text but also

1212
00:53:39,300 --> 00:53:37,060
the numeric data coming off the aircraft

1213
00:53:41,430 --> 00:53:39,310

and so far as I know that's not

1214
00:53:43,230 --> 00:53:41,440
something that Google does at the moment

1215
00:53:45,810 --> 00:53:43,240
but you never know they could get

1216
00:53:47,970 --> 00:53:45,820
involved we have had a partnership with

1217
00:53:49,710 --> 00:53:47,980
Google for many years in some of these

1218
00:53:52,440 --> 00:53:49,720
areas and some of the technologies that

1219
00:53:53,910 --> 00:53:52,450
we've developed in analyzing numeric

1220
00:53:56,040 --> 00:53:53,920
data has been done in that partnership

1221
00:54:02,140 --> 00:53:56,050
all right we have any more questions

1222
00:54:06,789 --> 00:54:04,210
I think this is a question for both

1223
00:54:09,190 --> 00:54:06,799
guests I get this vision when you're

1224
00:54:10,870 --> 00:54:09,200
describing the process of Jodie Foster

1225
00:54:12,430 --> 00:54:10,880
in the movie contact you know sitting on

1226
00:54:13,960 --> 00:54:12,440
the hood of the car with headphones and

1227
00:54:15,700 --> 00:54:13,970
just listening and listening and waiting

1228
00:54:17,349 --> 00:54:15,710
and waiting forever and ever so I was

1229
00:54:18,970 --> 00:54:17,359
just curious like when you're looking

1230
00:54:21,339 --> 00:54:18,980
for that needle in a haystack what does

1231
00:54:23,380 --> 00:54:21,349
it look like when you find it and then

1232
00:54:25,359 --> 00:54:23,390
what is the handoff like when you give

1233
00:54:26,829 --> 00:54:25,369
that information I know you said just

1234
00:54:29,109 --> 00:54:26,839
joke you make it available to all the

1235
00:54:31,630 --> 00:54:29,119
airlines but what is that transition or

1236
00:54:35,579 --> 00:54:31,640
handoff or hey we found something look

1237
00:54:40,710 --> 00:54:35,589
like from NASA to someone like Southwest

1238
00:54:43,539 --> 00:54:40,720

so when we find an issue in the data

1239
00:54:45,760 --> 00:54:43,549
oftentimes it turns out to be perfectly

1240
00:54:47,920 --> 00:54:45,770
known everybody in this in the industry

1241
00:54:49,450 --> 00:54:47,930
knows about it and we didn't and that's

1242
00:54:51,069 --> 00:54:49,460
actually a good thing because that's

1243
00:54:52,420 --> 00:54:51,079
kind of validating the algorithm it's

1244
00:54:55,059 --> 00:54:52,430
saying that it found something useful

1245
00:54:57,089 --> 00:54:55,069
and so we might talk to Bob we might

1246
00:55:00,250 --> 00:54:57,099
talk to Jeff and get that kind of a

1247
00:55:02,620 --> 00:55:00,260
response sometimes what we find frankly

1248
00:55:05,109 --> 00:55:02,630
is useless it's something that's just

1249
00:55:06,970 --> 00:55:05,119
noise or a spurious results and so Bob

1250
00:55:09,160 --> 00:55:06,980
or Jeff tells us usually in a nice way

1251
00:55:10,599 --> 00:55:09,170
that hey this isn't really useful but

1252
00:55:13,150 --> 00:55:10,609
sometimes we find something that's

1253
00:55:15,819 --> 00:55:13,160
actually useful and there was a nice

1254
00:55:18,279 --> 00:55:15,829
example of this and working with

1255
00:55:20,500 --> 00:55:18,289
Southwest a couple of years ago where we

1256
00:55:22,089 --> 00:55:20,510
found something actually they found it

1257
00:55:24,220 --> 00:55:22,099
using our algorithms and then I think

1258
00:55:26,349 --> 00:55:24,230
the transition was very smooth they were

1259
00:55:28,329 --> 00:55:26,359
able to take that and operationalize

1260
00:55:30,099 --> 00:55:28,339
that maybe Jeff can talk a little bit

1261
00:55:32,470 --> 00:55:30,109
more about that process well and that's

1262
00:55:34,660 --> 00:55:32,480
exactly right because if we do if we do

1263
00:55:36,670 --> 00:55:34,670

find something often it may be just a

1264
00:55:38,500 --> 00:55:36,680
parameter validation issue it could be

1265
00:55:39,819 --> 00:55:38,510
just a problem with a sensor on the

1266
00:55:42,670 --> 00:55:39,829
aircraft that's delivering bad

1267
00:55:45,130 --> 00:55:42,680
information but in this case we were

1268
00:55:47,289 --> 00:55:45,140
able to work with the algorithm that was

1269
00:55:49,870 --> 00:55:47,299
written by NASA and then take it and

1270
00:55:52,000 --> 00:55:49,880
sort of customize it so that we could

1271
00:55:54,460 --> 00:55:52,010
use it within our normal library of

1272
00:55:56,500 --> 00:55:54,470
algorithms that we use for looking at

1273
00:55:58,690 --> 00:55:56,510
data on a daily basis and give some

1274
00:56:02,620 --> 00:55:58,700
benefit so it does take some transition

1275
00:56:05,349 --> 00:56:02,630
and - and tailoring for our purposes one

1276
00:56:06,730 --> 00:56:05,359
of the things I'm particularly excited

1277
00:56:09,099 --> 00:56:06,740
about is that the work that we're doing

1278
00:56:10,990 --> 00:56:09,109
is for the public's benefit and so we

1279
00:56:12,700 --> 00:56:11,000
have a website called - link you can

1280
00:56:15,280 --> 00:56:12,710
type it into your favorite search engine

1281
00:56:17,440 --> 00:56:15,290
and download the algorithms that were

1282
00:56:19,990 --> 00:56:17,450
talking about right now download all the

1283
00:56:21,940 --> 00:56:20,000
papers about it ask questions you can

1284
00:56:24,130 --> 00:56:21,950
post a blog there if you want to you can

1285
00:56:25,600 --> 00:56:24,140
create a project and study it for

1286
00:56:27,940 --> 00:56:25,610
yourself this is one of the ways that

1287
00:56:30,610 --> 00:56:27,950
we're trying to validate our results so

1288
00:56:33,370 --> 00:56:30,620

that people in the public can benefit

1289
00:56:35,440 --> 00:56:33,380
from the work that we're doing but we

1290
00:56:40,150 --> 00:56:35,450
don't have the URL with it and I think

1291
00:56:40,780 --> 00:56:40,160
it's - link dot AR c dot nasa dot govt

1292
00:56:44,020 --> 00:56:40,790
yeah

1293
00:56:47,020 --> 00:56:44,030
- link NASA gov although I type it into

1294
00:56:50,530 --> 00:56:47,030
a search engine to get the to get the

1295
00:56:53,380 --> 00:56:50,540
link myself and what you can find there

1296
00:56:55,540 --> 00:56:53,390
is applications of these things to other

1297
00:56:57,550 --> 00:56:55,550
problems so we've had people in the

1298
00:56:59,470 --> 00:56:57,560
medical industry community for instance

1299
00:57:01,060 --> 00:56:59,480
say hey you know that problem that

1300
00:57:02,500 --> 00:57:01,070
you're doing with the Airlines it sounds

1301
00:57:05,320 --> 00:57:02,510
a lot like what we want to do with

1302
00:57:07,570 --> 00:57:05,330
pacemakers so can we use your algorithms

1303
00:57:09,640 --> 00:57:07,580
and I say sure go to - link download

1304
00:57:11,980 --> 00:57:09,650
them let let us know if you use them and

1305
00:57:14,140 --> 00:57:11,990
how it works out so the scope here is

1306
00:57:15,940 --> 00:57:14,150
enormous and the problems are

1307
00:57:19,870 --> 00:57:15,950
interesting and the technology is

1308
00:57:23,730 --> 00:57:19,880
fascinating alright anyone else here

1309
00:57:28,690 --> 00:57:27,250
yeah questions for both if you look out

1310
00:57:31,240 --> 00:57:28,700
two or three years word where do you

1311
00:57:33,070 --> 00:57:31,250
think will be both on a technical basis

1312
00:57:36,340 --> 00:57:33,080
and an operational basis and using these

1313
00:57:38,920 --> 00:57:36,350

kinds of technologies I think it's hard

1314
00:57:40,510 --> 00:57:38,930
to put a timeframe on that one of the

1315
00:57:44,170 --> 00:57:40,520
limitations that we have in terms of

1316
00:57:47,020 --> 00:57:44,180
doing more real-time interventions is

1317
00:57:48,640 --> 00:57:47,030
the delivery of data so for instance we

1318
00:57:51,070 --> 00:57:48,650
download our data every five to seven

1319
00:57:55,480 --> 00:57:51,080
days and we plug it into the database

1320
00:57:57,910 --> 00:57:55,490
via a memory card and so by the time we

1321
00:58:00,340 --> 00:57:57,920
have access to information off of the

1322
00:58:03,030 --> 00:58:00,350
airplane it's typically an event has

1323
00:58:05,710 --> 00:58:03,040
come and gone so we have to work on

1324
00:58:09,220 --> 00:58:05,720
delivering that data either on a daily

1325
00:58:13,720 --> 00:58:09,230
basis or even after every flight and so

1326
00:58:17,230 --> 00:58:13,730
that and I think what our job is to sort

1327
00:58:18,640 --> 00:58:17,240
of justify the safety benefit because

1328
00:58:20,410 --> 00:58:18,650
ultimately we know that when we make

1329
00:58:22,170 --> 00:58:20,420
safety improvements that's better for

1330
00:58:24,520 --> 00:58:22,180
efficiency and customer service as well

1331
00:58:27,220 --> 00:58:24,530
all right and we have time for one more

1332
00:58:31,120 --> 00:58:27,230
question has to be really quick go ahead

1333
00:58:38,510 --> 00:58:36,160
umeela that's a great question I I

1334
00:58:40,730 --> 00:58:38,520
entered the Reserve Officer Training

1335
00:58:42,620 --> 00:58:40,740
Corps in the Air Force

1336
00:58:45,860 --> 00:58:42,630
back when I was in college at Oklahoma

1337
00:58:47,750 --> 00:58:45,870
State University and had an opportunity

1338
00:58:49,780 --> 00:58:47,760

to go to pilot training after I

1339
00:58:52,130 --> 00:58:49,790
graduated from college so I did my

1340
00:58:54,530 --> 00:58:52,140
initial flying and training in the Air

1341
00:58:57,110 --> 00:58:54,540
Force and then after I finished my

1342
00:58:59,240 --> 00:58:57,120
career in the military on active duty

1343
00:59:02,180 --> 00:58:59,250
I joined Southwest Airlines and became

1344
00:59:04,880 --> 00:59:02,190
an airline pilot all right well you're

1345
00:59:05,690 --> 00:59:04,890
an inspiration to us all all right and

1346
00:59:07,430 --> 00:59:05,700
that's gonna have to be our last

1347
00:59:09,830 --> 00:59:07,440
question we've had a great discussion

1348
00:59:12,110 --> 00:59:09,840
today but we're out of time on behalf of

1349
00:59:15,080 --> 00:59:12,120
associate administrator jaywen shin at

1350
00:59:17,420 --> 00:59:15,090
NASA thanks to our guests Jeff hamlets

1351
00:59:19,640 --> 00:59:17,430
of Southwest Airlines Ashok's for

1352
00:59:21,890 --> 00:59:19,650
mustapha of NASA's Ames Research Center

1353
00:59:24,770 --> 00:59:21,900
and to our guests who joined us Don

1354
00:59:26,990 --> 00:59:24,780
video nikunj OSA Bob Lawrence Brian

1355
00:59:29,840 --> 00:59:27,000
Matthews and data mining experts

1356
00:59:31,910 --> 00:59:29,850
Jay way Hahn and Stephen Boyd you can

1357
00:59:36,800 --> 00:59:31,920
learn more about nasa's data mining work

1358
00:59:38,510 --> 00:59:36,810
by visiting wwsz a Ovie and i want to

1359
00:59:40,490 --> 00:59:38,520
remind you that Ashok and Jeff will be

1360
00:59:42,950 --> 00:59:40,500
our guests for a live web chat this

1361
00:59:45,380 --> 00:59:42,960
afternoon at 2:00 Eastern Time bring

1362
00:59:47,570 --> 00:59:45,390
your questions and go to nasa.gov click

1363
00:59:49,250 --> 00:59:47,580

on the link for a chat in the NASA