1
00:00:00,000 --> 00:00:07,633
[Title: Part 3: Data Management for Medical Researchers. Part 3 of 3.]

2
00:00:07,633 --> 00:00:12,500
All right so another aspect of storage

3
00:00:12,500 --> 00:00:15,000
that many people use are as we were

4
00:00:15,000 --> 00:00:16,033
talking a little bit as some of these

5
00:00:16,033 --> 00:00:18,333
cloud storage options. And while many

6
00:00:18,333 --> 00:00:20,266
people use cloud storage you should

7
00:00:20,266 --> 00:00:22,466
always be cautious when storing research

8
00:00:22,466 --> 00:00:24,900
data in the cloud. The first thing

9
00:00:24,900 --> 00:00:27,200
you want to do when using alternative

10
00:00:27,200 --> 00:00:30,466
cloud storage options are to check out

11
00:00:30,466 --> 00:00:32,766
the ownership and claim of ownership

12
00:00:32,766 --> 00:00:35,166
with different cloud storage data

13
00:00:35,166 --> 00:00:38,633
options. And second is it's often

14
00:00:38,633 --> 00:00:41,833
recommended to pick more than one

15
00:00:41,833 --> 00:00:44,033
ownership provider with cloud storage

16
00:00:44,033 --> 00:00:47,000
solutions in case as many people have

17
00:00:47,000 --> 00:00:49,333
often come across is there are cases

18
00:00:49,333 --> 00:00:52,500
where cloud storage options the

19
00:00:52,500 --> 00:00:55,366
ownership and of your data the owner of

20
00:00:55,366 --> 00:00:56,600
the cloud storage there have been

21
00:00:56,600 --> 00:00:59,666
bankruptcy issues that people

22
00:00:59,666 --> 00:01:01,700
have run across occasionally with cloud

23
00:01:01,700 --> 00:01:03,933
storage. Especially if it's a smaller

24
00:01:03,933 --> 00:01:07,266
provider. We have had people not

25
00:01:07,266 --> 00:01:08,833
necessarily here but different places

26
00:01:08,833 --> 00:01:13,366
come across. So in case your

27
00:01:13,366 --> 00:01:16,033
cloud storage provider goes down you

28
00:01:16,033 --> 00:01:17,600
often have you have more than one

29
00:01:17,600 --> 00:01:19,400
provider especially if you're looking

30
00:01:19,400 --> 00:01:22,466
outside of the university, it often as a

31
00:01:22,466 --> 00:01:26,600
safety solution is recommended. Another

32
00:01:26,600 --> 00:01:28,400
thing to discuss is where your data is

33
00:01:28,400 --> 00:01:30,233
stored at different parts of your

34
00:01:30,233 --> 00:01:32,900
workflow. And to document where you're

35
00:01:32,900 --> 00:01:34,566
storing your data during those different

36
00:01:34,566 --> 00:01:36,600
parts of the workflow. Where your data

37
00:01:36,600 --> 00:01:38,400
will be stored during data collection,

38
00:01:38,400 --> 00:01:40,933
processing, and analysis is crucial to

39
00:01:40,933 --> 00:01:42,266
have figured out at the beginning of

40
00:01:42,266 --> 00:01:44,133
your project. It's not something you want

41
00:01:44,133 --> 00:01:46,966
to decide during the process but rather

42
00:01:46,966 --> 00:01:48,600
to figure it out at the beginning and

43
00:01:48,600 --> 00:01:51,500
document that. It's especially important

44
00:01:51,500 --> 00:01:54,533
when working as part of a large team. All

45
00:01:54,533 --> 00:01:56,066
your folders where data will be stored

46
00:01:56,066 --> 00:01:58,300
should be indicated and where that data

47
00:01:58,300 --> 00:01:59,966
is stored at the end of the process

48
00:01:59,966 --> 00:02:02,333
should be figured out and documented as

49
00:02:02,333 --> 00:02:06,500
well. Here is an example of research data

50
00:02:06,500 --> 00:02:08,566
that was stored on one device

51
00:02:08,566 --> 00:02:11,566
and that device was stolen and all the

52
00:02:11,566 --> 00:02:14,100
raw data was lost. This is something that

53
00:02:14,100 --> 00:02:15,066
is not

54
00:02:15,066 --> 00:02:17,833
as common now but it has still happened.

55
00:02:17,833 --> 00:02:19,400
So it's something you just want to keep

56
00:02:19,400 --> 00:02:22,133
in mind that you don't want to be

57
00:02:22,133 --> 00:02:23,666
storing your data in one place only.

58
00:02:23,666 --> 00:02:27,866
Another thing to consider is to save if

59
00:02:27,866 --> 00:02:30,033
you're still using a physical device is

60
00:02:30,033 --> 00:02:34,266
to save multiple copies of your data and

61
00:02:34,266 --> 00:02:38,100
disperse them geographically. Especially

62
00:02:38,100 --> 00:02:40,766
if there's something like environmental

63
00:02:40,766 --> 00:02:42,566
disasters. And again if we're not

64
00:02:42,566 --> 00:02:46,400
necessarily talking so much about cloud

65
00:02:46,400 --> 00:02:50,200
storage but rather physical devices if

66
00:02:50,200 --> 00:02:52,233
you're storing. When storing your data

67
00:02:52,233 --> 00:02:54,500
there should also be several reasons why

68
00:02:54,500 --> 00:02:55,933
you should be there are several reasons

69
00:02:55,933 --> 00:02:56,933
why you should be concerned about

70
00:02:56,933 --> 00:03:00,466
security. So let's go over those. If

71
00:03:00,466 --> 00:03:02,200
you're collecting personal health

72
00:03:02,200 --> 00:03:04,133
information you want to make sure you're

73
00:03:04,133 --> 00:03:05,733
keeping your data is secure in a way

74
00:03:05,733 --> 00:03:08,166
that's HIPAA compliant. So you got to

75
00:03:08,166 --> 00:03:12,833
keep in mind compliance with HIPAA. If

76
00:03:12,833 --> 00:03:16,133
you're working on patients or patents

77
00:03:16,133 --> 00:03:18,033
or commercial data you want to ensure

78
00:03:18,033 --> 00:03:20,433
that that too is protected and secure

79
00:03:20,433 --> 00:03:23,266
from theft and/or damage. And if you're

80
00:03:23,266 --> 00:03:25,000
concerned just about the

81
00:03:25,000 --> 00:03:27,933
intellectual property you've worked hard

82
00:03:27,933 --> 00:03:30,000
to collect your information and data and

83
00:03:30,000 --> 00:03:31,366
you don't want to lose all your hard

84
00:03:31,366 --> 00:03:34,000
work. So security is just something very

85
00:03:34,000 --> 00:03:38,900
very important to keep in mind. All right

86
00:03:38,900 --> 00:03:41,266
so when thinking about security make

87
00:03:41,266 --> 00:03:43,533
sure you're working with IT to

88
00:03:43,533 --> 00:03:45,700
understand what they can do for you

89
00:03:45,700 --> 00:03:49,200
about your about security. And what extra

90
00:03:49,200 --> 00:03:50,933
steps they can help you with to make

91
00:03:50,933 --> 00:03:53,433
sure you're keeping your data safe. On

92
00:03:53,433 --> 00:03:54,866
your own there are several things that

93
00:03:54,866 --> 00:03:58,233
you can do. Add passwords to files or

94
00:03:58,233 --> 00:04:00,533
folders to keep them safe. Locking your

95
00:04:00,533 --> 00:04:02,500
machine when you walk away to make sure

96
00:04:02,500 --> 00:04:05,100
that nobody can just walk up and access

97
00:04:05,100 --> 00:04:06,533
your information. Just things that you

98
00:04:06,533 --> 00:04:08,400
can just do on a daily basis to make

99
00:04:08,400 --> 00:04:09,866
sure that nobody can access your

100
00:04:09,866 --> 00:04:13,666
information. Have other researchers who

101
00:04:13,666 --> 00:04:15,800
work with your data sign use agreements.

102
00:04:15,800 --> 00:04:17,533
These agreements will make sure your

103
00:04:17,533 --> 00:04:19,300
collaborators are held accountable for

104
00:04:19,300 --> 00:04:21,300
what they can do with your data. Use

105
00:04:21,300 --> 00:04:23,400
agreements can also restrict what

106
00:04:23,400 --> 00:04:24,766
someone can or cannot

107
00:04:24,766 --> 00:04:27,800
do with your data. And when storage is

108
00:04:27,800 --> 00:04:30,500
needed for personally identifiable data

109
00:04:30,500 --> 00:04:32,200
it's important to consider a number of

110
00:04:32,200 --> 00:04:35,366
factors due to HIPPA. From a data

111
00:04:35,366 --> 00:04:37,833
management perspective HIPPA indicates

112
00:04:37,833 --> 00:04:39,866
that you should have documentation for

113
00:04:39,866 --> 00:04:42,633
who is responsible for managing stored

114
00:04:42,633 --> 00:04:44,900
data at the original site during

115
00:04:44,900 --> 00:04:48,933
transfer, and at the new

116
00:04:48,933 --> 00:04:51,700
storage site. And indicated who the

117
00:04:51,700 --> 00:04:53,733
authorized users are at the beginning of

118
00:04:53,733 --> 00:04:56,400
the project is best to go over during

119
00:04:56,400 --> 00:04:58,500
your standard procedures. So just make

120
00:04:58,500 --> 00:05:00,800
sure you have all of this documented. So

121
00:05:00,800 --> 00:05:02,933
as you're getting here documentation is

122
00:05:02,933 --> 00:05:05,566
key with everything we're doing here.

123
00:05:05,566 --> 00:05:08,600
Just from storage, security, and with

124
00:05:08,600 --> 00:05:12,766
HIPAA requirements. And HIPPA states that

125
00:05:12,766 --> 00:05:15,133
PHI data requires end-to-end encryption

126
00:05:15,133 --> 00:05:17,566
meaning that data must be encrypted at

127
00:05:17,566 --> 00:05:20,100
its original location, the connection

128
00:05:20,100 --> 00:05:21,900
that data must travel through to get to

129
00:05:21,900 --> 00:05:23,600
its destination must also be encrypted,

130
00:05:23,600 --> 00:05:27,000
and the data must remain encrypted at

131
00:05:27,000 --> 00:05:29,233
in its other

132
00:05:29,233 --> 00:05:31,933
storage location. The encryption key must

133
00:05:31,933 --> 00:05:37,233
be stored separately. And keeping HIPAA

134
00:05:37,233 --> 00:05:39,066
at on a portable device is not

135
00:05:39,066 --> 00:05:42,566
recommended. So keep that all in

136
00:05:42,566 --> 00:05:45,900
mind when we're dealing with HIPAA data.

137
00:05:45,900 --> 00:05:50,300
And if you're using cloud storage you

138
00:05:50,300 --> 00:05:51,733
need to make sure you're following all

139
00:05:51,733 --> 00:05:53,933
HIPAA regulations and cloud storage is

140
00:05:53,933 --> 00:05:55,966
allowed with HIPAA. Cloud storage is

141
00:05:55,966 --> 00:05:57,566
allowed under HIPAA, provided that your

142
00:05:57,566 --> 00:05:59,733
covered entity or business associate

143
00:05:59,733 --> 00:06:02,300
enters into a HIPAA compliant business

144
00:06:02,300 --> 00:06:05,166
agreement with the CSP that will be

145
00:06:05,166 --> 00:06:06,933
creating receiving maintaining or

146
00:06:06,933 --> 00:06:09,066
transmitting electronic protected health

147
00:06:09,066 --> 00:06:11,733
information on its behalf and otherwise

148
00:06:11,733 --> 00:06:15,566
complies with HIPAA rules. So while you

149
00:06:15,566 --> 00:06:18,533
couldn't use a personal Google Drive for

150
00:06:18,533 --> 00:06:22,700
storing that HIPAA information

151
00:06:22,700 --> 00:06:28,233
you could however use a work

152
00:06:28,233 --> 00:06:31,066
Google Drive if you entered into a BAA

153
00:06:31,066 --> 00:06:34,400
with Google Drive for work. So if that

154
00:06:34,400 --> 00:06:36,500
was a professional

155
00:06:36,500 --> 00:06:40,033
business agreement with that provider. So

156
00:06:40,033 --> 00:06:41,766
something like that it just have to

157
00:06:41,766 --> 00:06:45,666
depends on what your provider allowed. So

158
00:06:45,666 --> 00:06:47,600
learn more about cloud storage and HIPAA

159
00:06:47,600 --> 00:06:49,966
we're gonna send out a URL that will go

160
00:06:49,966 --> 00:06:52,633
over all of these details. Alright

161
00:06:52,633 --> 00:06:55,000
let's get away from storage and go into

162
00:06:55,000 --> 00:06:56,666
a little bit more about preservation

163
00:06:56,666 --> 00:06:58,200
which will be a little less technical.

164
00:06:58,200 --> 00:06:59,933
And I will jumble a little less.

165
00:06:59,933 --> 00:07:03,466
Difference is a little bit are basically

166
00:07:03,466 --> 00:07:04,800
it's important to understand the

167
00:07:04,800 --> 00:07:08,266
differences first off, storage is mostly

168
00:07:08,266 --> 00:07:11,733
about your storage preservation is going

169
00:07:11,733 --> 00:07:13,966
to be going over how to make sure your

170
00:07:13,966 --> 00:07:17,600
information is not just stored but also

171
00:07:17,600 --> 00:07:20,800
accessible long-term. So just because

172
00:07:20,800 --> 00:07:21,800
we're storing our data doesn't

173
00:07:21,800 --> 00:07:23,733
necessarily mean that you're going to be

174
00:07:23,733 --> 00:07:26,200
able to access it long-term. Preservation

175
00:07:26,200 --> 00:07:28,666
focuses on making sure your data will be

176
00:07:28,666 --> 00:07:31,966
available and available in a way that

177
00:07:31,966 --> 00:07:33,733
you can use it in the same way that you

178
00:07:33,733 --> 00:07:36,633
were able to collect it. So looking back

179
00:07:36,633 --> 00:07:39,166
on your previous research you want it to

180
00:07:39,166 --> 00:07:41,733
others to be able to still

181
00:07:41,733 --> 00:07:46,033
access it. And as we move forward

182
00:07:46,033 --> 00:07:47,600
preservation will be increasingly

183
00:07:47,600 --> 00:07:50,633
important with more sharing agreements

184
00:07:50,633 --> 00:07:53,200
and sharing requirements emerging from

185
00:07:53,200 --> 00:07:55,400
funders and publishers. So the first

186
00:07:55,400 --> 00:07:57,066
thing we're going to talk about is our

187
00:07:57,066 --> 00:08:00,200
hardware obsolescence. Preservation helps

188
00:08:00,200 --> 00:08:02,933
you helps protect you from hardware

189
00:08:02,933 --> 00:08:05,266
obsolescence. And you will want to avoid

190
00:08:05,266 --> 00:08:08,000
avoid scenarios where all your data is

191
00:08:08,000 --> 00:08:10,566
saved on in an old format like a Jaz

192
00:08:10,566 --> 00:08:13,533
drive, only not

193
00:08:13,533 --> 00:08:17,300
to be able to open it later because you

194
00:08:17,300 --> 00:08:19,533
know Jaz drives are no longer supported.

195
00:08:19,533 --> 00:08:21,566
So you always want to migrate to a new

196
00:08:21,566 --> 00:08:23,766
hardware format so that your data will

197
00:08:23,766 --> 00:08:26,600
be able to be available long term. If

198
00:08:26,600 --> 00:08:29,233
you're using a proprietary or specific

199
00:08:29,233 --> 00:08:32,900
hardware format have a plan for how you

200
00:08:32,900 --> 00:08:36,100
can migrate it to a more universal

201
00:08:36,100 --> 00:08:38,566
format. And like hardware obsolescence

202
00:08:38,566 --> 00:08:40,866
there's also software obsolescence.

203
00:08:40,866 --> 00:08:42,700
So you want to be able to think about

204
00:08:42,700 --> 00:08:44,600
how you can save your data to open

205
00:08:44,600 --> 00:08:46,000
software formats.

206
00:08:46,000 --> 00:08:48,966
If you're using a proprietary software

207
00:08:48,966 --> 00:08:52,566
in the lab or homegrown software you

208
00:08:52,566 --> 00:08:54,833
created to collect data there may not be

209
00:08:54,833 --> 00:08:59,866
a way for others to access it. So

210
00:08:59,866 --> 00:09:02,200
moving data to open formats will also

211
00:09:02,200 --> 00:09:03,866
protect you from going back to old data

212
00:09:03,866 --> 00:09:05,833
opening it only to find that it's

213
00:09:05,833 --> 00:09:08,266
unreadable. Because of this it's

214
00:09:08,266 --> 00:09:10,300
important to just to make a

215
00:09:10,300 --> 00:09:12,833
distinction between how you collect data

216
00:09:12,833 --> 00:09:15,166
versus how you will disseminate it. You

217
00:09:15,166 --> 00:09:17,700
may still need to collect data in one

218
00:09:17,700 --> 00:09:19,366
way using specific hardware or software

219
00:09:19,366 --> 00:09:21,966
but you should have a plan to how you

220
00:09:21,966 --> 00:09:24,000
will transform it into a format that

221
00:09:24,000 --> 00:09:26,566
will be suitable for collection,

222
00:09:26,566 --> 00:09:28,900
collaboration, and dissemination. So the

223
00:09:28,900 --> 00:09:30,600
difference between collecting it one way

224
00:09:30,600 --> 00:09:32,533
and being able to disseminate it another

225
00:09:32,533 --> 00:09:35,700
way. So our best practices for

226
00:09:35,700 --> 00:09:38,300
preservation is to save your data on

227
00:09:38,300 --> 00:09:41,400
preservation formats. So there's a few

228
00:09:41,400 --> 00:09:44,666
gold there are four gold standard

229
00:09:44,666 --> 00:09:46,500
preservation formats that we're gonna

230
00:09:46,500 --> 00:09:49,233
talk about here. So the first and they

231
00:09:49,233 --> 00:09:50,733
can the good thing about these

232
00:09:50,733 --> 00:09:53,800
gold standard formats are that

233
00:09:53,800 --> 00:09:57,300
they can be used and viewed on basically

234
00:09:57,300 --> 00:10:01,766
any operating system with any kind of

235
00:10:01,766 --> 00:10:05,166
software. So we've got the XML extensible

236
00:10:05,166 --> 00:10:07,900
markup language and this is used to

237
00:10:07,900 --> 00:10:11,366
ensure simplicity, generability, and

238
00:10:11,366 --> 00:10:14,200
usability across the internet and can be

239
00:10:14,200 --> 00:10:16,633
used to save documents or web service

240
00:10:16,633 --> 00:10:20,400
content. The CSV (Comma separated values)

241
00:10:20,400 --> 00:10:22,833
is an ideal way to save spreadsheets in

242
00:10:22,833 --> 00:10:25,066
preservation formats.

243
00:10:25,066 --> 00:10:28,833
We've got the PDF and the nice thing

244
00:10:28,833 --> 00:10:31,366
about PDFs is they can be used to save

245
00:10:31,366 --> 00:10:34,500
documents in perpetuity.

246
00:10:34,500 --> 00:10:36,733
The one thing about PDFs to keep in mind

247
00:10:36,733 --> 00:10:39,366
is generally you can't really edit them.

248
00:10:39,366 --> 00:10:43,366
So you're not really going to be able to

249
00:10:43,366 --> 00:10:45,300
go back and edit them but it is way to

250
00:10:45,300 --> 00:10:48,666
get a way to capture them in a lasting

251
00:10:48,666 --> 00:10:50,833
format. So you won't be able to change

252
00:10:50,833 --> 00:10:52,633
them but you can freeze a document in a

253
00:10:52,633 --> 00:10:55,133
PDF format. And then you have TIFF - tagged

254
00:10:55,133 --> 00:10:56,966
image file format - and it's the gold

255
00:10:56,966 --> 00:10:58,933
standard for saving images.

256
00:10:58,933 --> 00:11:01,600
And they are already preservation ready

257
00:11:01,600 --> 00:11:03,533
format and they can ensure the quality

258
00:11:03,533 --> 00:11:06,566
of an image over time. So those are

259
00:11:06,566 --> 00:11:09,900
typically four of our gold

260
00:11:09,900 --> 00:11:11,300
the gold standard for preservation.

261
00:11:11,300 --> 00:11:13,800
format. So we didn't really talk about

262
00:11:13,800 --> 00:11:16,600
this question but looking at this list

263
00:11:16,600 --> 00:11:18,866
what would you guys say would be the

264
00:11:18,866 --> 00:11:21,433
file format for an open data for

265
00:11:21,433 --> 00:11:25,433
documents? You would probably for

266
00:11:25,433 --> 00:11:29,633
when looking at a basic document

267
00:11:29,633 --> 00:11:31,966
you would probably want to use a .txt

268
00:11:31,966 --> 00:11:35,100
file. Because it's just an open it's an

269
00:11:35,100 --> 00:11:40,100
open software. That said it's not always

270
00:11:40,100 --> 00:11:43,033
possible to use a preservation format so

271
00:11:43,033 --> 00:11:45,700
when at all possible use one

272
00:11:45,700 --> 00:11:48,100
but if you can't keep in mind that some

273
00:11:48,100 --> 00:11:52,100
files are more ready available than others.

274
00:11:52,100 --> 00:11:55,800
Microsoft Excel or Microsoft Word is

275
00:11:55,800 --> 00:11:58,233
probably going to be around a lot longer

276
00:11:58,233 --> 00:12:01,733
than some more obscure lab-created

277
00:12:01,733 --> 00:12:05,233
software. So use your best judgment when

278
00:12:05,233 --> 00:12:08,266
deciding what to use for preservation

279
00:12:08,266 --> 00:12:11,100
format. So if you've created your own

280
00:12:11,100 --> 00:12:13,733
homegrown software and you don't know if

281
00:12:13,733 --> 00:12:15,500
you're going to be around or you know

282
00:12:15,500 --> 00:12:17,500
that you were going off to the middle of

283
00:12:17,500 --> 00:12:19,466
nowhere and nobody's going to be able to

284
00:12:19,466 --> 00:12:22,166
reach you, but you know your data is gonna

285
00:12:22,166 --> 00:12:24,400
have to exist

286
00:12:24,400 --> 00:12:28,066
forever, see if you can you know transfer

287
00:12:28,066 --> 00:12:29,933
it to a format that other people will be

288
00:12:29,933 --> 00:12:32,600
able to access. Like Microsoft

289
00:12:32,600 --> 00:12:34,233
Excel if that's the only thing that you

290
00:12:34,233 --> 00:12:37,533
could use. That said it's not always

291
00:12:37,533 --> 00:12:40,133
possible but do your best to find

292
00:12:40,133 --> 00:12:41,500
something more stable.

293
00:12:41,500 --> 00:12:46,200
Chances are Excel, Word, Microsoft Office

294
00:12:46,200 --> 00:12:50,966
suite is going to be around. We're all

295
00:12:50,966 --> 00:12:54,700
gonna have problems if they disappear in

296
00:12:54,700 --> 00:12:57,833
the near future.

297
00:12:57,833 --> 00:13:00,066
Remember if your if your data is

298
00:13:00,066 --> 00:13:03,733
irreplaceable and you can't

299
00:13:03,733 --> 00:13:05,566
collect it again or it take took years

300
00:13:05,566 --> 00:13:09,733
to collect, it's best to make sure you're

301
00:13:09,733 --> 00:13:11,733
saving it in preservation format

302
00:13:11,733 --> 00:13:13,733
so it's available and accessible for a

303
00:13:13,733 --> 00:13:16,233
long-term. So just kind of think about

304
00:13:16,233 --> 00:13:20,300
what you're collecting, figure out you

305
00:13:20,300 --> 00:13:22,700
know figure out what it means to you

306
00:13:22,700 --> 00:13:25,033
when you are trying to preserve it and

307
00:13:25,033 --> 00:13:27,966
make those decisions for yourself. Other

308
00:13:27,966 --> 00:13:30,000
ways to protect your data from

309
00:13:30,000 --> 00:13:31,966
degradation and ensure its preservation

310
00:13:31,966 --> 00:13:35,400
is to try to avoid

311
00:13:35,400 --> 00:13:38,600
encrypting or compressing your data when

312
00:13:38,600 --> 00:13:41,633
possible. Now this is not always possible.

313
00:13:41,633 --> 00:13:43,866
Encryption is useful for keeping your

314
00:13:43,866 --> 00:13:45,866
data safe but encrypting data can also

315
00:13:45,866 --> 00:13:48,500
make it difficult to access your data

316
00:13:48,500 --> 00:13:51,500
later on if you haven't kept good

317
00:13:51,500 --> 00:13:54,266
documentation. Key word: documentation

318
00:13:54,266 --> 00:13:56,066
another word that's coming up a lot here

319
00:13:56,066 --> 00:13:58,933
on how to access your encrypted data.

320
00:13:58,933 --> 00:14:01,533
When compressing your data you risk

321
00:14:01,533 --> 00:14:04,433
damaging your data each time as it can

322
00:14:04,433 --> 00:14:07,133
remove the quality and integrity of your

323
00:14:07,133 --> 00:14:09,666
data over time. Always keep an original

324
00:14:09,666 --> 00:14:12,333
data set, make a copy, and then compress

325
00:14:12,333 --> 00:14:15,800
or encrypt that data to send to others.

326
00:14:15,800 --> 00:14:18,733
Finally with preservation it's important

327
00:14:18,733 --> 00:14:20,233
to know who owns the data you've

328
00:14:20,233 --> 00:14:23,033
collected and you can't assume that you

329
00:14:23,033 --> 00:14:25,100
own your data. So before you share it,

330
00:14:25,100 --> 00:14:28,300
collaborate, or delete data it's

331
00:14:28,300 --> 00:14:29,433
important to check funder or

332
00:14:29,433 --> 00:14:31,600
institutional policies to understand

333
00:14:31,600 --> 00:14:34,800
your user rights. Now we're gonna pass

334
00:14:34,800 --> 00:14:37,333
back on to Tom to talk a little bit

335
00:14:37,333 --> 00:14:40,766
about providing access. And it's the

336
00:14:40,766 --> 00:14:43,166
final stretch. I promise to make this

337
00:14:43,166 --> 00:14:45,566
quick. I'm just curious does anybody know

338
00:14:45,566 --> 00:14:47,333
what the actual university policy is for

339
00:14:47,333 --> 00:14:50,666
data ownership? They own it. If you go

340
00:14:50,666 --> 00:14:52,533
to the research homepage they do have a

341
00:14:52,533 --> 00:14:55,400
policy there. And so that's important to

342
00:14:55,400 --> 00:14:56,966
know if you're going to be leaving to go

343
00:14:56,966 --> 00:14:58,500
somewhere else and you cannot

344
00:14:58,500 --> 00:15:01,000
automatically assume you will be able to

345
00:15:01,000 --> 00:15:02,633
take your data with you. You probably

346
00:15:02,633 --> 00:15:04,600
will. I don't think it's hard and fast.

347
00:15:04,600 --> 00:15:06,566
But the basic policy is that the

348
00:15:06,566 --> 00:15:09,266
university does own data produced using

349
00:15:09,266 --> 00:15:11,466
their facilities. So it's a good thing to

350
00:15:11,466 --> 00:15:15,600
know. Ok access and standards. Again

351
00:15:15,600 --> 00:15:18,333
access is the issue here. Can be it's one

352
00:15:18,333 --> 00:15:19,700
thing to store it, it's another thing

353
00:15:19,700 --> 00:15:22,200
that people can actually get to it. So

354
00:15:22,200 --> 00:15:24,133
where are you going to put it? Most

355
00:15:24,133 --> 00:15:25,566
people choose repositories.

356
00:15:25,566 --> 00:15:27,800
There's two basic types of repositories.

357
00:15:27,800 --> 00:15:29,433
There's a cross-disciplinary adoptable

358
00:15:29,433 --> 00:15:31,600
repository you can put anything in there

359
00:15:31,600 --> 00:15:34,333
it is not subject specific. We'll talk

360
00:15:34,333 --> 00:15:35,900
about a couple of these that are very

361
00:15:35,900 --> 00:15:38,800
popular. Probably more in interest to

362
00:15:38,800 --> 00:15:40,833
this group would be discipline specific

363
00:15:40,833 --> 00:15:43,933
repositories. Dealing with specific areas

364
00:15:43,933 --> 00:15:46,233
of research. The nice thing about that is

365
00:15:46,233 --> 00:15:47,700
other people who are doing the research

366
00:15:47,700 --> 00:15:49,733
probably know to go there to get the

367
00:15:49,733 --> 00:15:51,566
data. So it makes it a little bit easier

368
00:15:51,566 --> 00:15:53,366
to find for other people doing the same

369
00:15:53,366 --> 00:15:55,766
kind of research. One of the nice things

370
00:15:55,766 --> 00:15:58,866
is the NIH has a nice list on their

371
00:15:58,866 --> 00:16:01,900
webpage of data repositories dealing

372
00:16:01,900 --> 00:16:05,000
with medical research issues. If you just

373
00:16:05,000 --> 00:16:06,500
go to their webpage do a keyword search

374
00:16:06,500 --> 00:16:09,200
repositories you'll get this list. Also

375
00:16:09,200 --> 00:16:10,666
the slideshows will be available so you

376
00:16:10,666 --> 00:16:12,666
can see the link. So you can see they do

377
00:16:12,666 --> 00:16:14,900
a really good job there's over 60 listed

378
00:16:14,900 --> 00:16:16,966
here. And they get very specific you've

379
00:16:16,966 --> 00:16:19,366
got Cancer Nanotechnology

380
00:16:19,366 --> 00:16:23,033
Library, the Cancer Imaging Archive. All

381
00:16:23,033 --> 00:16:24,200
sorts of things.

382
00:16:24,200 --> 00:16:26,900
i-Genes, peptides...And it will give you

383
00:16:26,900 --> 00:16:28,966
the repository name, a description of

384
00:16:28,966 --> 00:16:31,000
what's in there. They will give you the

385
00:16:31,000 --> 00:16:33,133
submission policies for how to put it in

386
00:16:33,133 --> 00:16:35,633
what they require when you do deposit

387
00:16:35,633 --> 00:16:38,733
data. And then how to access the data. So

388
00:16:38,733 --> 00:16:40,233
this would probably be the most

389
00:16:40,233 --> 00:16:41,833
interesting place for you guys to go. We

390
00:16:41,833 --> 00:16:43,600
will talk about some other options but

391
00:16:43,600 --> 00:16:45,400
just keep in mind the NIH has kind of

392
00:16:45,400 --> 00:16:48,066
done this for you. See if there's a place

393
00:16:48,066 --> 00:16:50,166
that's appropriate for your research

394
00:16:50,166 --> 00:16:55,300
data on their list. R3data.org is

395
00:16:55,300 --> 00:16:58,066
a very popular research data repository.

396
00:16:58,066 --> 00:16:59,266
I was actually gonna bring up their page

397
00:16:59,266 --> 00:17:00,700
but since we're having trouble showing

398
00:17:00,700 --> 00:17:03,566
other web pages now I'm not going to do

399
00:17:03,566 --> 00:17:07,366
that. There is a link to our re3data on

400
00:17:07,366 --> 00:17:10,400
my research guide. Hopefully when we're

401
00:17:10,400 --> 00:17:12,066
done here I can get that page up but I

402
00:17:12,066 --> 00:17:13,066
don't want to interrupt the slideshow

403
00:17:13,066 --> 00:17:16,600
right now. Re3data is not

404
00:17:16,600 --> 00:17:18,866
subject specific. They have their own

405
00:17:18,866 --> 00:17:22,100
metadata standards. Very easy to use they

406
00:17:22,100 --> 00:17:25,200
are fee based. The feed the varies

407
00:17:25,200 --> 00:17:28,033
depending upon how much data you're

408
00:17:28,033 --> 00:17:30,500
inputting. They've got policies on their

409
00:17:30,500 --> 00:17:32,900
home page you can look at. Also OpenDoar

410
00:17:32,900 --> 00:17:34,333
the directory of open access

411
00:17:34,333 --> 00:17:38,166
repositories. They've got a

412
00:17:38,166 --> 00:17:42,900
open-access, non discipline-specific. So

413
00:17:42,900 --> 00:17:45,300
these are two just popular places you

414
00:17:45,300 --> 00:17:49,066
can google them up, or get there the

415
00:17:49,066 --> 00:17:52,200
links off the slide here. Figshare.

416
00:17:52,200 --> 00:17:54,300
Figshare's another open repository. The good

417
00:17:54,300 --> 00:17:55,800
thing about Figshare is it's a great

418
00:17:55,800 --> 00:17:58,900
place for open source code. They do a lot

419
00:17:58,900 --> 00:18:01,566
of other deposit other datasets there

420
00:18:01,566 --> 00:18:03,866
but it tends to be a very popular site

421
00:18:03,866 --> 00:18:06,100
for open source code. So as Sarah

422
00:18:06,100 --> 00:18:08,333
mentioned earlier, if you can use open

423
00:18:08,333 --> 00:18:10,500
source software for your data

424
00:18:10,500 --> 00:18:14,000
manipulation, data analysis. a final

425
00:18:14,000 --> 00:18:16,966
product whatever, that always helps other

426
00:18:16,966 --> 00:18:18,766
people who want to use the data later on

427
00:18:18,766 --> 00:18:20,133
because they will probably be able to

428
00:18:20,133 --> 00:18:22,233
get to the software easier

429
00:18:22,233 --> 00:18:24,233
Figshare is a great place for that.

430
00:18:24,233 --> 00:18:26,400
They're very good at keeping things

431
00:18:26,400 --> 00:18:29,333
updated when a new

432
00:18:29,333 --> 00:18:30,633
(whoops sorry)

433
00:18:30,633 --> 00:18:33,266
when a new version of the software gets

434
00:18:33,266 --> 00:18:34,933
uploaded, it will be immediately there

435
00:18:34,933 --> 00:18:37,766
will be a notice about it. So a good

436
00:18:37,766 --> 00:18:41,733
place to keep in mind. And again the

437
00:18:41,733 --> 00:18:44,133
issue here is access versus

438
00:18:44,133 --> 00:18:45,800
meaningful access. You put it up

439
00:18:45,800 --> 00:18:48,233
somewhere that's great. It's somewhere

440
00:18:48,233 --> 00:18:49,733
how do people get to it?

441
00:18:49,733 --> 00:18:51,766
Meaningful access means that you've got

442
00:18:51,766 --> 00:18:54,400
some metadata that will be

443
00:18:54,400 --> 00:18:57,566
encoded with the file when you upload it.

444
00:18:57,566 --> 00:19:00,433
So that it is documented and people can

445
00:19:00,433 --> 00:19:01,200
get to it.

446
00:19:01,200 --> 00:19:03,933
Most repository sites do have metadata

447
00:19:03,933 --> 00:19:07,233
standards on the site that will usually

448
00:19:07,233 --> 00:19:08,966
be a very obvious link. You can just go

449
00:19:08,966 --> 00:19:10,600
there and see what metadata standards

450
00:19:10,600 --> 00:19:12,133
they offer. Usually it will be

451
00:19:12,133 --> 00:19:13,266
discipline-specific.

452
00:19:13,266 --> 00:19:15,600
There are also different disciplines

453
00:19:15,600 --> 00:19:18,700
specific metadata standard sites. There

454
00:19:18,700 --> 00:19:20,333
is a link to that on my research guide

455
00:19:20,333 --> 00:19:22,000
which hopefully I can show you later.

456
00:19:22,000 --> 00:19:23,566
And it's the kind of thing where you can

457
00:19:23,566 --> 00:19:25,800
just google it up very easily. Metadata

458
00:19:25,800 --> 00:19:27,900
standards biochemistry whatever. And

459
00:19:27,900 --> 00:19:30,900
you'll find those sites. So the point is

460
00:19:30,900 --> 00:19:33,266
try to get some well-documented metadata

461
00:19:33,266 --> 00:19:36,000
associated with your file so that other

462
00:19:36,000 --> 00:19:40,066
people can find it. A good example of

463
00:19:40,066 --> 00:19:41,733
people who are really pretty stringent

464
00:19:41,733 --> 00:19:45,700
about their metadata standards is the

465
00:19:45,700 --> 00:19:47,266
NIH National Heart Lung and Blood

466
00:19:47,266 --> 00:19:50,033
Institute. The data repository

467
00:19:50,033 --> 00:19:51,966
actually started back in 1975 so they've

468
00:19:51,966 --> 00:19:52,300
been doing this

469
00:19:52,300 --> 00:19:54,866
for a long time. Their guidelines state

470
00:19:54,866 --> 00:19:57,466
and this is a quote: "The documentation

471
00:19:57,466 --> 00:19:59,833
must be sufficiently complete such that

472
00:19:59,833 --> 00:20:02,233
a person responsible for determining use

473
00:20:02,233 --> 00:20:05,166
and selection of biospecimens could tell

474
00:20:05,166 --> 00:20:07,466
how the study was conducted, how the

475
00:20:07,466 --> 00:20:10,033
specimens were prepared and stored, and

476
00:20:10,033 --> 00:20:15,000
how associated data was collected." So

477
00:20:15,000 --> 00:20:17,666
they'll have a list here of the study

478
00:20:17,666 --> 00:20:20,166
names. You can just bring those up.

479
00:20:20,166 --> 00:20:22,400
They're pretty strict about these

480
00:20:22,400 --> 00:20:24,600
standards and they're so strict that

481
00:20:24,600 --> 00:20:26,100
they've actually got a biorepository

482
00:20:26,100 --> 00:20:28,100
guide to building biospecimen collection

483
00:20:28,100 --> 00:20:30,133
so they will tell you exactly what they

484
00:20:30,133 --> 00:20:32,533
want to have. Basically their standard is

485
00:20:32,533 --> 00:20:34,200
that somebody who is not associated with

486
00:20:34,200 --> 00:20:36,300
the experiment can go find the data as

487
00:20:36,300 --> 00:20:38,833
easy as someone who's associated with the

488
00:20:38,833 --> 00:20:41,466
experiment. This is the only one I know

489
00:20:41,466 --> 00:20:43,033
where they actually have a guide for

490
00:20:43,033 --> 00:20:44,400
this. I'm assuming again that this will

491
00:20:44,400 --> 00:20:47,033
become more common and it will be

492
00:20:47,033 --> 00:20:50,700
research area dependable as well. Tools.

493
00:20:50,700 --> 00:20:53,233
Tools you can use to help with your data

494
00:20:53,233 --> 00:20:56,133
management. The biggest most popular one

495
00:20:56,133 --> 00:20:59,133
is the DMP tool. Anybody actually use

496
00:20:59,133 --> 00:21:00,600
this yet, I'm curious?

497
00:21:00,600 --> 00:21:03,533
Okay, wonderful. So this has been up for

498
00:21:03,533 --> 00:21:05,366
about four or five years. It is basically

499
00:21:05,366 --> 00:21:07,200
just an online tool which will walk you

500
00:21:07,200 --> 00:21:09,133
through the process of creating a data

501
00:21:09,133 --> 00:21:12,000
management plan. One of the nice things

502
00:21:12,000 --> 00:21:14,533
about it is it will have links to the

503
00:21:14,533 --> 00:21:16,533
funding agencies so you can find out

504
00:21:16,533 --> 00:21:19,433
what they require. It will have templates

505
00:21:19,433 --> 00:21:21,533
for each funding agency, so if you're

506
00:21:21,533 --> 00:21:23,900
doing NIH or NSF or whatever and

507
00:21:23,900 --> 00:21:25,266
sometimes there's sub-disciplines in

508
00:21:25,266 --> 00:21:26,666
there NSF has a lot of sub disciplines

509
00:21:26,666 --> 00:21:28,633
for different areas of engineering.

510
00:21:28,633 --> 00:21:30,733
Because they might require different

511
00:21:30,733 --> 00:21:33,600
data elements and different information

512
00:21:33,600 --> 00:21:36,166
in their data management plan. So you'll

513
00:21:36,166 --> 00:21:38,700
be able to go find out exactly what the

514
00:21:38,700 --> 00:21:40,233
funding agency that you're dealing with

515
00:21:40,233 --> 00:21:42,866
requires. They will have a template there

516
00:21:42,866 --> 00:21:45,100
when you bring the template up they will

517
00:21:45,100 --> 00:21:46,833
walk you through the process of filling

518
00:21:46,833 --> 00:21:49,966
out each field. You can cut and paste you

519
00:21:49,966 --> 00:21:51,900
can share your account with other people.

520
00:21:51,900 --> 00:21:53,033
If you are creating a data

521
00:21:53,033 --> 00:21:54,666
management plan with someone else on the

522
00:21:54,666 --> 00:21:57,466
team you can add them to there, you can

523
00:21:57,466 --> 00:21:59,866
both review it at the same time. They

524
00:21:59,866 --> 00:22:01,900
will have previously published data

525
00:22:01,900 --> 00:22:04,366
management plans that people have agreed

526
00:22:04,366 --> 00:22:05,700
to make public. So if you want to see

527
00:22:05,700 --> 00:22:06,000
what someone

528
00:22:06,000 --> 00:22:07,600
else looks like for a certain field

529
00:22:07,600 --> 00:22:09,900
you can go see what they did. Again the

530
00:22:09,900 --> 00:22:11,333
important thing to remember about a DMP

531
00:22:11,333 --> 00:22:13,500
is it does not have to be long. In fact

532
00:22:13,500 --> 00:22:15,433
they usually want it short. Two pages or

533
00:22:15,433 --> 00:22:17,700
less. It's just important that you

534
00:22:17,700 --> 00:22:19,800
provide the elements of information that

535
00:22:19,800 --> 00:22:21,600
they are interested in having. You wanna

536
00:22:21,600 --> 00:22:22,866
make sure that when you don't want to

537
00:22:22,866 --> 00:22:24,566
have I actually had a call from somebody

538
00:22:24,566 --> 00:22:28,100
over in the Materials Research Center

539
00:22:28,100 --> 00:22:30,700
that had (or the Energy Conversion Center

540
00:22:30,700 --> 00:22:33,333
I'm sorry) who had submitted it had

541
00:22:33,333 --> 00:22:35,400
submitted a grant application did not

542
00:22:35,400 --> 00:22:37,466
supply the DMP. They rejected it. Sent it

543
00:22:37,466 --> 00:22:39,233
back - I got a panicked call "What do I do?"

544
00:22:39,233 --> 00:22:41,833
Took two days to fix the problem, no big

545
00:22:41,833 --> 00:22:43,466
deal but you do not want to have be in

546
00:22:43,466 --> 00:22:47,466
that situation. This is available off of

547
00:22:47,466 --> 00:22:50,066
my research guide, once again I hope I

548
00:22:50,066 --> 00:22:53,700
can show you. If you just go to that link

549
00:22:53,700 --> 00:22:55,833
as long as you're on a University of

550
00:22:55,833 --> 00:22:58,500
Delaware campus network - so you're on the

551
00:22:58,500 --> 00:23:01,200
UD network - when you click on that get

552
00:23:01,200 --> 00:23:03,766
started link it will then bring up a

553
00:23:03,766 --> 00:23:05,833
drop-down menu of all the participating

554
00:23:05,833 --> 00:23:07,966
institutions. University of Delaware

555
00:23:07,966 --> 00:23:09,766
will be one of those on there. Simply

556
00:23:09,766 --> 00:23:12,233
select it say "ok". It will then throw you

557
00:23:12,233 --> 00:23:14,700
into your CAS-authentification. That's

558
00:23:14,700 --> 00:23:16,700
how you get in. Once you're in you've got

559
00:23:16,700 --> 00:23:18,566
an account created, you can start

560
00:23:18,566 --> 00:23:20,800
creating plans, you can save them as you

561
00:23:20,800 --> 00:23:24,566
go along. They're very good at updating

562
00:23:24,566 --> 00:23:26,266
when new templates came out. You'll see

563
00:23:26,266 --> 00:23:29,000
there's a new template for what DOT. And

564
00:23:29,000 --> 00:23:32,733
NASA funded research. So they'll have

565
00:23:32,733 --> 00:23:34,966
News, they'll have old DMPs you can look

566
00:23:34,966 --> 00:23:36,400
at, they'll have the

567
00:23:36,400 --> 00:23:39,133
requirements by agency and some private

568
00:23:39,133 --> 00:23:40,833
institutions as well. Private grant

569
00:23:40,833 --> 00:23:44,066
funders. Again if you need to find out

570
00:23:44,066 --> 00:23:46,400
what the requirements are for a DNP for

571
00:23:46,400 --> 00:23:48,866
an agency or public funder that you are

572
00:23:48,866 --> 00:23:50,666
dealing with and you cannot find it

573
00:23:50,666 --> 00:23:52,766
contact us we'll see if we can help

574
00:23:52,766 --> 00:23:56,266
track it down. This is just an example of

575
00:23:56,266 --> 00:23:58,400
a new template notification that they

576
00:23:58,400 --> 00:24:00,366
put up for the genomic data sharing

577
00:24:00,366 --> 00:24:03,500
policy. Again they're very good on the

578
00:24:03,500 --> 00:24:05,600
DMP tool about keeping things very up to

579
00:24:05,600 --> 00:24:08,000
date. We've had no complaints about that.

580
00:24:08,000 --> 00:24:12,300
Github. Github is another good place to

581
00:24:12,300 --> 00:24:14,733
go for actually, Github's another

582
00:24:14,733 --> 00:24:18,000
place for software. As long as well as

583
00:24:18,000 --> 00:24:19,300
Figshare.

584
00:24:19,300 --> 00:24:21,966
In fact, Github is probably better. So the

585
00:24:21,966 --> 00:24:23,400
nice thing about Github is they will

586
00:24:23,400 --> 00:24:25,100
give you notifications once things gets

587
00:24:25,100 --> 00:24:27,800
updated. It will tell you how to upload

588
00:24:27,800 --> 00:24:31,100
things, how to share things. This would be

589
00:24:31,100 --> 00:24:33,166
a notice of something that one changed

590
00:24:33,166 --> 00:24:35,166
file, two additions, two deletions. They're

591
00:24:35,166 --> 00:24:37,900
very good at letting you know about that.

592
00:24:37,900 --> 00:24:40,566
There's also links to Figshare and Github

593
00:24:40,566 --> 00:24:43,133
off of my research guide. A good

594
00:24:43,133 --> 00:24:45,900
place to go for open source software.

595
00:24:45,900 --> 00:24:48,400
Lab notebooks. If you do not want to use

596
00:24:48,400 --> 00:24:51,300
traditional lab notebooks. This is

597
00:24:51,300 --> 00:24:52,700
another interesting thing I found out from

598
00:24:52,700 --> 00:24:55,333
the from the civil engineering people, a

599
00:24:55,333 --> 00:24:56,566
lot of them are still using that but a

600
00:24:56,566 --> 00:24:57,833
lot of people don't like having paper in

601
00:24:57,833 --> 00:24:59,366
the lab anymore. You might spill on it

602
00:24:59,366 --> 00:25:02,033
might do things. So there are some online

603
00:25:02,033 --> 00:25:04,400
lab notebooks "lab archives", which is

604
00:25:04,400 --> 00:25:06,200
pretty much a general-purpose site. "Labguru"

605
00:25:06,200 --> 00:25:08,700
which does really inventory

606
00:25:08,700 --> 00:25:10,700
management for equipment and processes.

607
00:25:10,700 --> 00:25:13,700
And "Benchling" which is more specifically

608
00:25:13,700 --> 00:25:16,300
for DNA tool integration. There's other

609
00:25:16,300 --> 00:25:18,000
ones available these just tend to be the

610
00:25:18,000 --> 00:25:20,066
most popular. Again if you don't feel

611
00:25:20,066 --> 00:25:22,566
like keeping traditional lab notebooks...

612
00:25:22,566 --> 00:25:24,566
Another interesting thing I noticed with the

613
00:25:24,566 --> 00:25:25,933
civil engineering people is, and I'm not

614
00:25:25,933 --> 00:25:27,666
picking on them they're just people I

615
00:25:27,666 --> 00:25:28,866
just did this project with so I know

616
00:25:28,866 --> 00:25:30,666
about it, a lot of them just have a

617
00:25:30,666 --> 00:25:32,366
little file cabinets stacked up with

618
00:25:32,366 --> 00:25:35,000
this stuff, you know. That's saving it

619
00:25:35,000 --> 00:25:37,566
in a way, but if you want to go online

620
00:25:37,566 --> 00:25:39,500
there are options so just be aware of

621
00:25:39,500 --> 00:25:43,266
that. So what have we concluded? Plan your

622
00:25:43,266 --> 00:25:45,233
data management plan before starting the

623
00:25:45,233 --> 00:25:46,633
research. When we do the general

624
00:25:46,633 --> 00:25:47,900
workshops this is something we talked

625
00:25:47,900 --> 00:25:49,633
about at the very beginning and we actually

626
00:25:49,633 --> 00:25:51,200
have a slide of the data management

627
00:25:51,200 --> 00:25:53,533
cycle. It's very easy to think that you

628
00:25:53,533 --> 00:25:54,766
don't have to think about these issues

629
00:25:54,766 --> 00:25:55,666
till the end.

630
00:25:55,666 --> 00:25:57,433
Oh no I've done the research and I've

631
00:25:57,433 --> 00:25:59,433
got this data, and where am I going to

632
00:25:59,433 --> 00:26:02,066
put it? It's a really good idea to start

633
00:26:02,066 --> 00:26:04,033
planning this before your

634
00:26:04,033 --> 00:26:06,666
grant. Start thinking how what am I going

635
00:26:06,666 --> 00:26:08,200
to do? What am I going to need? What are

636
00:26:08,200 --> 00:26:09,366
the funding requirements going to be?

637
00:26:09,366 --> 00:26:11,100
Where am I thinking of publishing? What

638
00:26:11,100 --> 00:26:13,400
what might their requirements be?

639
00:26:13,400 --> 00:26:16,033
If you have a team, again assign somebody

640
00:26:16,033 --> 00:26:17,600
to be in charge of all this so you're

641
00:26:17,600 --> 00:26:18,900
not wondering who's taking care of

642
00:26:18,900 --> 00:26:21,766
things. So take care of it throughout the

643
00:26:21,766 --> 00:26:23,666
research process. It's not just something

644
00:26:23,666 --> 00:26:24,866
you need to be thinking about at the end.

645
00:26:24,866 --> 00:26:28,566
And do not ignore the march toward more

646
00:26:28,566 --> 00:26:30,800
data management and sharing both

647
00:26:30,800 --> 00:26:35,400
from grant providers and publishers. That

648
00:26:35,400 --> 00:26:37,466
is it. I really want to thank you for

649
00:26:37,466 --> 00:26:40,966
letting us come down here.