

1

00:00:07,220 --> 00:00:16,609

Thank you very much for inviting me here today. Um, I have to tell you there... there

2

00:00:16,609 --> 00:00:21,949

really was a funny story on the way here um, that - I had to sit out in the car and

3

00:00:21,949 --> 00:00:28,470

complete this - that there's a um, you know, how there... there's certain senators that go after

4

00:00:28,470 --> 00:00:35,680

research. And so there was a... starting this morning there was a um... one of the

5

00:00:35,680 --> 00:00:40,410

senators - I think it's Tom Coburn from Oklahoma - was going after a piece of our

6

00:00:40,410 --> 00:00:48,030

research, and he was saying, "here we are looking at rabbit massage. How silly." And

7

00:00:48,030 --> 00:00:54,580

um, so I... I had to... so we had to generate um, a response to that, you know, that um, attack

8

00:00:54,580 --> 00:01:00,710

um, this uh, this afternoon. And so, this is a piece of research that actually is about the

9

00:01:00,710 --> 00:01:05,509

effect of massage and its relaxation effects and its curative effects after

10

00:01:05,509 --> 00:01:11,110

intensive exercise, and they happen to use a rabbit model, you know. And so it

11

00:01:11,110 --> 00:01:15,980

really was a valuable piece of research.  
It was done by a, you know, by a sports

12

00:01:15,980 --> 00:01:21,640  
medicine person. And so I just thought it  
was interesting that, you know, here... how

13

00:01:21,640 --> 00:01:27,570  
crazy does it get when you actually have  
somebody going after um, rabbit massage.

14

00:01:27,570 --> 00:01:33,580  
That's the title of the piece that uh, we  
had to... so Channel 10 is calling my office

15

00:01:33,580 --> 00:01:42,220  
about rabbit massage. So uh, never... the  
craziness never stops. So... so I'm... I'm really glad to

16

00:01:42,220 --> 00:01:48,100  
be here today and I... I wanted to actually  
talk to you about the SHARE Project. And

17

00:01:48,100 --> 00:01:51,369  
I hope you're aware of this. If you're  
not aware of it, you will be in a minute.

18

00:01:51,369 --> 00:01:59,159  
Um, what is SHARE? It's really a... a... the higher education answer,  
really, to some of the

19

00:01:59,159 --> 00:02:05,060  
Open Access issues. And in February of  
2013, the Office of Science and

20

00:02:05,060 --> 00:02:09,509  
Technology Policy uh, John Holdren, issued a memo that um,

21

00:02:10,560 --> 00:02:16,770  
that I think sort of made us all sort  
of gasp - that every federal agency with

22

00:02:16,770 --> 00:02:21,920  
over a hundred million dollars in  
federal funds had to develop a draft

23

00:02:21,920 --> 00:02:27,610  
plan for public access of peer-reviewed  
publications. But this was the... the real...

24

00:02:27,610 --> 00:02:31,910  
the thing that actually made us gasp was...  
was digital data. I don't think it

25

00:02:31,910 --> 00:02:36,819  
actually said digital in the memo. It said data within six months. So  
all of those

26

00:02:36,819 --> 00:02:43,540  
agencies, by August, had to develop um, you know, plans. And I think  
they did, but the...

27

00:02:43,540 --> 00:02:48,480  
the... the... the incredible thing to me was  
that those plans weren't made public.

28

00:02:48,480 --> 00:02:54,080  
This is an Open Access initiative.  
They were not made public. And I think

29

00:02:54,080 --> 00:03:00,260  
the... the deal we planned was just made  
public here um, recently, and some of the

30

00:03:00,260 --> 00:03:07,269  
other agencies it's actually starting um,  
starting to come out now. But... so, there... there

31

00:03:07,269 --> 00:03:13,349  
were 23 agencies that actually had um, over a hundred million dollars  
in federal

32

00:03:13,349 --> 00:03:18,920

funding and they um... so those draft plans were developed um, and I think most of them

33

00:03:18,920 --> 00:03:23,340  
were submitted. Some of them were plans for plans um, but they were actually submitted

34

00:03:23,340 --> 00:03:28,170  
and they're starting to come out. And so what was spelled out in the memo was... was

35

00:03:28,170 --> 00:03:35,400  
four possibilities that an... an expansion of the existing NIH

36

00:03:35,400 --> 00:03:40,000  
PubMed Central. So NIH has been out there for a number of years with PubMed

37

00:03:40,000 --> 00:03:45,430  
Central. Um, an individual institutional repository approach, maybe at the PI

38

00:03:45,430 --> 00:03:52,170  
level. Um, the publishers rushed in and basically put forward an idea called

39

00:03:52,170 --> 00:03:56,799  
CHORUS, and I think some of the funding agencies are doing that. It's easy, and I

40

00:03:56,799 --> 00:04:01,450  
think that the publishers are saying, you know, it's not going to be... it's not going to be

41

00:04:01,450 --> 00:04:07,980  
very expensive, you know, and I don't believe that. Um, but... but essentially this is

42

00:04:07,980 --> 00:04:13,639  
what we focused on was this consortium of higher education institutions. This

43

00:04:13,639 --> 00:04:18,239

actually was... I don't think this was... this was not laid out in the memo, but we very

44

00:04:18,239 --> 00:04:22,590

quickly came together and s... and saying the data is ours.

45

00:04:22,590 --> 00:04:29,370

Um, the publications are from our data. The last thing I think we want to do is pay

46

00:04:29,370 --> 00:04:35,199

to have the... pay to hire the faculty, pay the faculty um, have them generate this

47

00:04:35,199 --> 00:04:40,270

great data, and then pay for it again to... to have it come back. So it was really felt as

48

00:04:40,270 --> 00:04:47,320

if the data is generated here in... in um, in... in higher ed and we need to figure out a way

49

00:04:47,320 --> 00:04:56,860

to actually comply with this new... new federal mandate. And so we came together

50

00:04:56,860 --> 00:05:03,470

around this... this... um, this thing called SHARE, and it is a partnership of higher

51

00:05:03,470 --> 00:05:07,820

education institutions. So... so you're familiar with ARL - that's the Association

52

00:05:07,820 --> 00:05:14,810

of Research um, Libraries. AAU is the Association of American Universities um, and

53

00:05:14,810 --> 00:05:20,180  
APLU - Association of Public and  
Land-grant Universities. So these are the

54

00:05:20,180 --> 00:05:26,370  
two major, higher ed institutions that  
are... that really speak on behalf of all the

55

00:05:26,370 --> 00:05:31,570  
universities, and this speaks on behalf of  
the libraries. So this was... we... we sort of came up

56

00:05:31,570 --> 00:05:37,330  
with a... a definition of what SHARE is,  
and it's a community facing initiative,

57

00:05:37,330 --> 00:05:41,590  
and where some, you know, many different  
disciplines come together around

58

00:05:41,590 --> 00:05:44,419  
um, information technology,

59

00:05:44,419 --> 00:05:49,210  
the library um, research officers like  
myself, digital repository managers,

60

00:05:49,210 --> 00:05:57,210  
scholarly societies, publishers, private  
funders uh, and... and more. So in the fall of

61

00:05:57,210 --> 00:06:03,630  
2013 there was a steering group put  
together um, to plan this SHARE project and...

62

00:06:03,630 --> 00:06:08,220  
was just almost exactly a year ago. I got  
the call to be sort of the lone research

63

00:06:08,220 --> 00:06:14,760  
officer on that group. And um, I really  
didn't know much about um, Open Access,

64

00:06:14,760 --> 00:06:21,030

didn't know much about um... I found out I didn't know much about the library um, when

65

00:06:21,030 --> 00:06:25,830

I actually joined this. And the acronyms were somewhat overwhelming, but I have my

66

00:06:25,830 --> 00:06:29,900

own acronyms in my field so, you know, it's kind of the same thing. So we all

67

00:06:29,900 --> 00:06:34,880

have our acronyms. So this... this steering committee was... it's composed of mostly

68

00:06:34,880 --> 00:06:35,750

librarians.

69

00:06:35,750 --> 00:06:39,820

There's some IT experts on it, and then there's a... a representative -

70

00:06:39,820 --> 00:06:44,580

one of the vice presidents from AAU, one of the vice presidents from A... APLU, and

71

00:06:44,580 --> 00:06:51,740

I'm... I'm the... the research officer. So there were working groups formed - this was um, um,

72

00:06:51,740 --> 00:06:56,250

a little bit later in the process - but these working groups have been very, very

73

00:06:56,250 --> 00:07:00,860

important in this last year. There's a technical working group that... again, I can't

74

00:07:00,860 --> 00:07:05,560

understand what they do and I... and we'll...  
we'll get into all those. It's very IT

75

00:07:05,560 --> 00:07:15,970  
focused. The repository community,  
workflow - um, how we oops... how we essentially um, work

76

00:07:15,970 --> 00:07:19,990  
through the flow of information - and  
you'll see what I'm talking about in a

77

00:07:19,990 --> 00:07:20,530  
minute -

78

00:07:20,530 --> 00:07:29,820  
rights, and... and communications. So... so the problem statement in  
really how to... how

79

00:07:29,820 --> 00:07:34,660  
to sort of characterize SHARE - and I... I  
hope you'll... you'll see that it... it's really

80

00:07:34,660 --> 00:07:39,720  
difficult to keep up with the release of  
publications. And this is from... from the

81

00:07:39,720 --> 00:07:43,440  
researcher from the field standpoint.  
It's difficult to keep up with

82

00:07:43,440 --> 00:07:48,140  
publications, datasets, and other research  
outputs. You know, there's no central

83

00:07:48,140 --> 00:07:54,560  
repository for an abstract, for a poster,  
for a presentation that's made at a

84

00:07:54,560 --> 00:07:58,890  
meeting. So if you're trying to follow a  
field and really keep up to the minute,

85

00:07:58,890 --> 00:08:04,010  
you really can't do that very easily  
because there's all of these pieces of...

86

00:08:04,010 --> 00:08:09,260  
of information. It may take a year to get...  
get a paper out there, and by then some

87

00:08:09,260 --> 00:08:12,540  
of these... some of these pieces of  
information are actually quite old. So

88

00:08:12,540 --> 00:08:18,340  
there's no single, structured way to  
report a research output release in a

89

00:08:18,340 --> 00:08:24,080  
timely um, manner that's sort of out there for everybody to see. And  
so the concern,

90

00:08:24,080 --> 00:08:29,630  
actually, that I brought to the table really  
is that there's um... my concern is if we have

91

00:08:29,630 --> 00:08:35,870  
23 funding agencies and each of them  
comes up with their own plan, I'm really

92

00:08:35,870 --> 00:08:39,589  
concerned that there are researchers  
that are going to have to comply with

93

00:08:39,589 --> 00:08:41,979  
widely differing sets of

94

00:08:41,979 --> 00:08:45,870  
rules. So think about the person that has  
an NIH grant, a... a

95

00:08:46,529 --> 00:08:52,290  
Department of Energy grant, and an NSF

grant, and all of those rules are widely

96

00:08:52,290 --> 00:08:58,050  
different. They publish a paper and... or  
make a... a meeting presentation or they're

97

00:08:58,050 --> 00:09:02,370  
an astronomer and they... and they  
generate a... a s... a set of research data, and all

98

00:09:02,370 --> 00:09:05,290  
of a sudden they've got to do three  
different things with that data. I mean that

99

00:09:05,290 --> 00:09:10,019  
is, you know, there's administrative  
burden now on the researcher. I can only

100

00:09:10,019 --> 00:09:18,380  
imagine the outcry when they have to do  
something like this. So the... the problem... the

101

00:09:18,380 --> 00:09:24,510  
first project that SHARE is addressing -  
and there are four steps in the SHARE process

102

00:09:24,510 --> 00:09:28,940  
and I... I... I hope you'll see that we're  
pretty excited about this project - but

103

00:09:28,940 --> 00:09:36,160  
the first... the first project is... is  
notification. So how do you know that

104

00:09:36,160 --> 00:09:42,510  
something has happened out there in the...  
in the research um, arena? You know, let's say

105

00:09:42,510 --> 00:09:47,600  
you're following cancer. How do you know  
that something has actually happened -

106

00:09:47,600 --> 00:09:52,269  
that there's been American Association  
for Cancer Research meeting and there's

107

00:09:52,269 --> 00:09:56,829  
something, and you're interested in  
leukemia and that there's a... there's...

108

00:09:56,829 --> 00:10:02,959  
there's been a series of very important  
papers um, or presentations on leukemia? So

109

00:10:02,959 --> 00:10:08,730  
the... the... the... the notification system is  
really directed at that. So it aims to

110

00:10:08,730 --> 00:10:14,829  
notify the stakeholders, and we, the  
public, are stakeholders. When a res...

111

00:10:14,829 --> 00:10:19,370  
real... research release event occurs -  
that can be a publication, that can be a

112

00:10:19,370 --> 00:10:26,600  
presentation, that can be release of data.  
Um, it really lets us, as stakeholders, know

113

00:10:26,600 --> 00:10:32,750  
in a timely, structured, and comprehensive  
sort of manner. Now you can imagine that

114

00:10:32,750 --> 00:10:38,760  
really takes a lot of um, a lot of effort.  
And you'll see actually what's... what's

115

00:10:38,760 --> 00:10:44,680  
being done to... to address that um, just that issue. So here's the  
challenge, I think,

116

00:10:44,680 --> 00:10:49,430

put very succinctly - it's the  
preservation of research output, it's the

117

00:10:49,430 --> 00:10:53,059  
access... it's the public access to that, and it's...

118

00:10:53,059 --> 00:10:59,959  
and... and importantly it's the reuse because I think as we enter into  
an area of data

119

00:10:59,959 --> 00:11:06,439  
mining um, that becomes very important. How do you link - if we keep  
with the cancer

120

00:11:06,439 --> 00:11:12,699  
example - how do you link this publication  
on leukemia with this presentation on

121

00:11:12,699 --> 00:11:17,409  
leukemia by two very different authors,  
maybe even two different countries, how

122

00:11:17,409 --> 00:11:21,269  
do you link them both together? So  
there's some really exciting things

123

00:11:21,269 --> 00:11:27,909  
coming out around data mining um, that  
allows you to do that. So when we look at

124

00:11:27,909 --> 00:11:34,379  
the challenges here um, we have really  
several challenges. We have challenges of

125

00:11:34,379 --> 00:11:41,379  
infrastructure, of having, let's say, data  
storage capabilities um, around this issue.

126

00:11:41,379 --> 00:11:47,939  
We have workflow um, issues and we also  
have policy issues, right? We have

127

00:11:47,939 --> 00:11:54,369  
publishers that hold copyrights to  
articles. Um, and so there's... there's all... all

128

00:11:54,369 --> 00:12:00,109  
three of these really we feel forms a, you know, forms sort of a... a  
triangle here around

129

00:12:00,109 --> 00:12:08,279  
issues. So in the infrastructure category  
there's varied repository platforms.

130

00:12:08,279 --> 00:12:14,039  
They have var... varied capabilities, and  
they have various standards and um, and... and

131

00:12:14,039 --> 00:12:19,129  
protocols. In the... in the policy area,  
which it looks like you... maybe you talked

132

00:12:19,129 --> 00:12:20,799  
about earlier today,

133

00:12:20,799 --> 00:12:27,099  
public access, Open Access, copyright um,  
copyright policies, and data management

134

00:12:27,099 --> 00:12:34,119  
plans um, and sharing plans. And finally in... in um, the workflow -  
this is something that

135

00:12:34,119 --> 00:12:40,289  
I'm particularly concerned about - we have  
multiple siloed systems. So you have

136

00:12:40,289 --> 00:12:49,229  
um, a... a... you... you have tremendous administrative burden right  
now on faculty members on um, any level of

137

00:12:49,229 --> 00:12:54,289

um, of researcher. So workflow - how do we get all of these systems to actually talk to

138

00:12:54,289 --> 00:12:57,289  
each other?

139

00:12:58,150 --> 00:13:03,610  
So knowing who is... who is producing what and under... and under whose sort of auspices are

140

00:13:03,610 --> 00:13:09,860  
control, I think it's very critical to um,  
all sorts of stakeholders. But... by the

141

00:13:09,860 --> 00:13:15,550  
funding agencies, by... by our campuses, by  
sponsored research offices um, the

142

00:13:15,550 --> 00:13:21,070  
government agencies, even tenure and  
promotion committees um, who want access to

143

00:13:21,070 --> 00:13:25,240  
this, you know, the latest, you know,  
information. Repository managers here in

144

00:13:25,240 --> 00:13:32,150  
the library and the entire um, the... the entire research community.  
So these are the

145

00:13:32,150 --> 00:13:38,520  
share components um, that were actually put together and were  
conceptualized in

146

00:13:38,520 --> 00:13:43,840  
December of 2013. I just laid out the  
case for the notification system. That's

147

00:13:43,840 --> 00:13:50,010  
what... that's sort of step one. But these  
are the four steps in the SHARE program.

148

00:13:50,010 --> 00:13:59,100

The second is um, is... is a registry, which is really curated metadata. So it's a... it's...

149

00:13:59,100 --> 00:14:06,070

it's an entire 8 registration around what is included in the notification system.

150

00:14:06,070 --> 00:14:13,850

Um, there's a discovery layer of being able to go in from the outside and say tell

151

00:14:13,850 --> 00:14:19,910

me everything about leukemia. And then finally there's a mining and... and reuse

152

00:14:19,910 --> 00:14:27,030

phase around um, mining that, looking... digging down into that, you know, into that

153

00:14:27,030 --> 00:14:33,190

category of leukemia, for example, and then uh, mining... asking a specific question -

154

00:14:33,190 --> 00:14:38,360

therapeutic um, you know, therapeutic approaches to leukemia - what's the very

155

00:14:38,360 --> 00:14:43,390

latest in terms of therapy? And finally reuse of that data. So where are we right

156

00:14:43,390 --> 00:14:50,010

now? So we're at the notification stage - um, sort of halfway through uh, development of

157

00:14:50,010 --> 00:14:54,680

the notification service, and then on our way really to... to registry. So this is

158

00:14:54,680 --> 00:14:59,930

probably, you know, two, three year process  
that actually is... is going to happen. Which

159

00:14:59,930 --> 00:15:05,500  
actually is fair... fairly light speed if you...  
if you think about it compared to

160

00:15:05,500 --> 00:15:08,250  
um, essentially, to... to where we are right now.

161

00:15:08,250 --> 00:15:21,160  
So in... here we are in early 2014. We  
got a um, a joint a... a... a grant jointly from IMLS,

162

00:15:21,160 --> 00:15:26,340  
the Institute for... for Museum and Library  
Science, and the Sloan Foundation for a

163

00:15:26,340 --> 00:15:31,690  
million dollars to develop the  
notification service and really kind of

164

00:15:31,690 --> 00:15:35,480  
a long-term vision for SHARE. You know, and... and there's been a lot  
of discussion

165

00:15:35,480 --> 00:15:45,680  
around what should SHARE be. Um, is  
it a, you know, is it a stand-alone institute, is it d... is it

166

00:15:45,680 --> 00:15:51,010  
virtual, does it sort of exist virtually in all  
of our institutions, or what actually

167

00:15:51,010 --> 00:15:55,150  
should it be? You know, and I think there  
will be at some point... I think there will

168

00:15:55,150 --> 00:15:59,470  
be... you know, when the grant funding runs  
out, you know, there will be a charge

169

00:15:59,470 --> 00:16:05,660  
probably back to the institutions. We'll  
have to bear that charge um, of... if this

170

00:16:05,660 --> 00:16:12,220  
notification system, if... if this registry  
is really useful to all of us. So I think

171

00:16:12,220 --> 00:16:19,900  
this... this next year becomes very important in demonstration of the  
value of SHARE to

172

00:16:19,900 --> 00:16:24,910  
the entire research community. So we got  
this million dollars. This time frame is

173

00:16:24,910 --> 00:16:30,600  
from March 2014 through September 2015.  
Um, and... and I think if there's... if there's

174

00:16:30,600 --> 00:16:34,630  
good progress, there will probably more  
grants out there. But grants aren't meant to

175

00:16:34,630 --> 00:16:40,420  
be forever, you know. And so there will be  
time when the grants, however long that

176

00:16:40,420 --> 00:16:41,390  
is, when they run out.

177

00:16:41,390 --> 00:16:46,890  
We have to have a sustainability  
plan. So in developing this notification

178

00:16:46,890 --> 00:16:52,170  
system, we came across the Center for  
Open Science, which is a Charlottesville,

179

00:16:52,170 --> 00:16:59,120  
Virginia-based, nonprofit start-up, and

they are um... it's an amazing little company

180

00:16:59,120 --> 00:17:04,970  
that is growing very fast, and they  
really have very much of an Open Access

181

00:17:04,970 --> 00:17:10,380  
sort of mindset, if you will. And they... so  
we really felt there was a lot of

182

00:17:10,380 --> 00:17:17,980  
mission alignment with um, with the SHARE um, concept. And there...  
this is their tagline

183

00:17:17,980 --> 00:17:23,020  
or their... their motto: "we foster openness  
integrity and reproducibility of

184

00:17:23,020 --> 00:17:27,579  
scientific research." And I'm sure you're ve... very much aware of  
this reproducibility

185

00:17:27,579 --> 00:17:32,350  
issue. This has become, particularly in  
biomedical data, this has become a big

186

00:17:32,350 --> 00:17:38,880  
issue in terms of um, a reproducibility. So I... I was just at a... at  
a SHARE meeting last week

187

00:17:38,880 --> 00:17:42,940  
in Washington D.C., and the Center for Open Science is a very  
impressive

188

00:17:42,940 --> 00:17:49,500  
organization. They have ten million  
dollars for five years to um, develop their

189

00:17:49,500 --> 00:17:54,710  
business model. Um, they're very well-funded and they have something  
like fifty

190

00:17:54,710 --> 00:18:00,280  
employees. They've been in... in business I,  
think, a year. Um, and it's... it's impressive

191

00:18:00,280 --> 00:18:05,710  
what they um, actually are able to do. So  
Center for Open Science - keep your eye on um,

192

00:18:05,710 --> 00:18:12,720  
actually the Center for Open Science. So the types of research release  
events - so back

193

00:18:12,720 --> 00:18:17,970  
to... kind of back to the notification - I'm going to dig down a  
little bit into the

194

00:18:17,970 --> 00:18:24,880  
notification service because I think this is  
something that um, really will be of... a great

195

00:18:24,880 --> 00:18:28,200  
deal of use to the research community,  
particularly to the people

196

00:18:28,760 --> 00:18:34,840  
generating the data and to the people  
utilizing that data. So, you know, k... kind of

197

00:18:34,840 --> 00:18:39,670  
at the ver... at the most obvious level um,  
articles are going to be... published

198

00:18:39,670 --> 00:18:41,320  
articles are going to be in

199

00:18:41,320 --> 00:18:46,509  
that notification service. Preprints - so I  
think that is something that you really

200

00:18:46,509 --> 00:18:49,509  
can't so much get your hands on right

now.

201

00:18:50,240 --> 00:18:54,539

A few journals... a few Open Access journals will put those out there, but this will

202

00:18:54,539 --> 00:19:01,250

be essentially most everything, and datasets. Now I... I sort of had heartburn

203

00:19:01,250 --> 00:19:07,820

when I first read this memo because, you know, datasets to me, you know, I... thi... this is not

204

00:19:07,820 --> 00:19:12,230

meant to be, you know, the latest experiment by your graduate student.

205

00:19:12,230 --> 00:19:15,899

That's, I think, how some people interpreted this, but this might be the

206

00:19:15,899 --> 00:19:21,259

data that accompanies, let's say, figure one in a manuscript - something that has been

207

00:19:21,259 --> 00:19:27,409

peer reviewed. If you look at it fields like astronomy, for example, there's

208

00:19:27,409 --> 00:19:33,179

huge data sets out there. That's a field that has come together a single platform

209

00:19:33,179 --> 00:19:39,110

um, and that's really advanced that... the field of astronomy by all coming

210

00:19:39,110 --> 00:19:43,490

together and deciding what software we're going to use, what platforms we're

211  
00:19:43,490 --> 00:19:49,970  
going to use. And so they really have an  
agreed-upon data set um, in the field of

212  
00:19:49,970 --> 00:19:53,940  
astronomy. So that's the kind of data  
sets we're talking about that have been...

213  
00:19:53,940 --> 00:19:59,350  
really that have been reviewed and that  
are sort of widely accepted. And in the

214  
00:19:59,350 --> 00:20:04,779  
data that accompanies sort of re... peer  
reviewed um, journals and things like that.

215  
00:20:04,779 --> 00:20:11,210  
But what we're... what we're proposing in  
the... in the notification service is a... is an

216  
00:20:11,210 --> 00:20:15,690  
additional layer of data like  
presentations, like posters that you

217  
00:20:15,690 --> 00:20:21,519  
would do at a, you know, at a national meeting. Data management plans  
have also been

218  
00:20:21,519 --> 00:20:28,730  
sort of thrown out there. Um, programming  
projects, researcher profiles, and more. So

219  
00:20:28,730 --> 00:20:35,340  
it really is proposed to be a very rich  
data set. And the interesting thing, there

220  
00:20:35,340 --> 00:20:41,070  
was a um, there was a pilot that the Center  
for Open Science did with five

221  
00:20:41,070 --> 00:20:46,259

institutions and they were hoping to get,  
I think, five thousand pieces of data. They got

222

00:20:46,259 --> 00:20:52,019  
forty thousand pieces of data in that first,  
you know, sort of five institution um,

223

00:20:52,930 --> 00:20:58,660  
you know, pilot. So who's interested in this,  
you know, in these research release events?

224

00:20:58,660 --> 00:21:07,200  
Who are the consumers here? Um, certainly  
like... like Wright State - the... the campus repository

225

00:21:07,200 --> 00:21:12,200  
would be interested, particularly not just  
in the research that's done here, but the

226

00:21:12,200 --> 00:21:16,730  
research that's, you know, done elsewhere.  
Sponsored research offices - I'm very

227

00:21:16,730 --> 00:21:23,180  
interested in, you know, that sort of... of data, particularly for my  
own institution, and

228

00:21:23,180 --> 00:21:31,360  
funders are extremely interested. So  
NSF and NIH wants to know, you

229

00:21:31,360 --> 00:21:37,570  
know, what... what is... what... what are we paying for? You know,  
what's the, you know, what's the

230

00:21:37,570 --> 00:21:43,040  
output of what we've provided all these...  
all these dollars for? You know, and so

231

00:21:43,040 --> 00:21:50,080  
um, those are at least three of the um, three of the people that are  
interested, but

232

00:21:50,080 --> 00:21:55,960  
publishers would probably be interested  
in these sorts of things. Um, current

233

00:21:55,960 --> 00:22:00,860  
research information systems um, and  
particularly other researchers and the

234

00:22:00,860 --> 00:22:06,530  
public. So I mean I... I... I have a hard time  
sort of envisioning, you know, sort of the...

235

00:22:06,530 --> 00:22:13,900  
the Twitter public going in here and...  
this is highly technical stuff, you know, and

236

00:22:13,900 --> 00:22:20,210  
sort of mucking around in... in a lot of  
these um, a lot of these sort of data sets.

237

00:22:20,210 --> 00:22:25,690  
But, you know, maybe so um, maybe some... some retired scientist would  
love to do this,

238

00:22:25,690 --> 00:22:30,520  
but, you know, I think the... the average um, the average person I  
think might have a little hard

239

00:22:30,520 --> 00:22:37,490  
time doing that. So this is sort of the  
way it looks right now. Um, there's just... I

240

00:22:37,490 --> 00:22:43,160  
mean data is going sporadically, you know,  
to um, all over the place. You know, funders

241

00:22:43,160 --> 00:22:49,700  
get um, preprints sent to them as part of  
progress reports. Um, you know, sponsored

242

00:22:49,700 --> 00:22:55,830  
research offices get... get articles as  
part of progress reports. Um, repositories

243  
00:22:55,830 --> 00:23:02,000  
might get preprints and articles. So  
it's just... so... so this is what uh, what the

244  
00:23:02,000 --> 00:23:03,460  
situation looks like today.

245  
00:23:03,460 --> 00:23:08,100  
But here's what we're actually  
proposing: that the SHARE notification

246  
00:23:08,100 --> 00:23:14,210  
service really be the... sort of the  
central clearing house for all of this,

247  
00:23:14,210 --> 00:23:21,809  
you know, data, preprints, articles um, and then back out to... to um,  
funders, sponsored

248  
00:23:21,809 --> 00:23:26,600  
research offices, and... and repositories, as  
well as the public, as well as um, as other

249  
00:23:26,600 --> 00:23:33,980  
researchers. So it's... it's trying to  
organize this um, this flow of information,

250  
00:23:33,980 --> 00:23:43,990  
if you will. So I'd... I'd encourage you to um, visit the SHARE wiki  
at... it's called at... at... @ GITHUB.

251  
00:23:43,990 --> 00:23:51,970  
And um, I think this is um, a... it's an open source development  
process right now. And you

252  
00:23:51,970 --> 00:23:56,919  
are welcome to visit the GITHUB  
repository and kind of follow this SHARE

253

00:23:56,919 --> 00:24:02,500

project as it, you know, as it develops, and as, you know, as sources of data are

254

00:24:02,500 --> 00:24:09,760

identified and data flows into this. I don't think the um, I don't think the forty

255

00:24:09,760 --> 00:24:12,750

thousand pieces of data are in there yet for...

256

00:24:12,750 --> 00:24:19,190

for public consumption but... but hopefully they will be. So this um, so I think the

257

00:24:19,190 --> 00:24:23,299

benefits of this... some additional benefits of this notification system -

258

00:24:23,299 --> 00:24:29,020

remember this is step one of four - is to track and connect a PI with an award

259

00:24:29,020 --> 00:24:35,529

and the research output. So right now I have a hard time in my office um, tracking

260

00:24:35,529 --> 00:24:43,320

John Smith - his... his award that he has from NIH with his publications and all

261

00:24:43,320 --> 00:24:46,710

of the research output surrounding that award. It would be a wonderful

262

00:24:46,710 --> 00:24:52,549

thing if I could do that because a lot of... a lot of my job is... is about promoting

263

00:24:52,549 --> 00:24:57,500

research. It's about going out and  
talking - just like this - talking about, you

264

00:24:57,500 --> 00:25:01,700

know, research um, off campus, on campus, and it would

265

00:25:01,700 --> 00:25:09,710

be very helpful to know the entire um,  
entire library of that... of that data. And

266

00:25:09,710 --> 00:25:13,480

it's also, from the library standpoint, to  
enable repository managers to identify

267

00:25:13,480 --> 00:25:16,440

manuscripts for deposit and university

268

00:25:16,440 --> 00:25:19,800

and funding agency grant  
administrators to determine compliance.

269

00:25:19,800 --> 00:25:25,420

You know, compliance is... is not something I think that a lot of  
faculty, you know,

270

00:25:25,420 --> 00:25:30,140

really it's not sort of their central uh,  
concern. It does have to be some of us...

271

00:25:30,140 --> 00:25:36,770

some of us administrators - it does have  
to be a uh, concern we have, and uh, maintaining

272

00:25:36,770 --> 00:25:45,240

compliance with public access policies.  
And I think that... that... that the benefits to um, to

273

00:25:45,240 --> 00:25:51,320

research officers - it does optimize the  
interaction among um, the... the researcher,

274

00:25:51,320 --> 00:25:57,300  
the sponsored um, the... the... the senior research officer, and  
the... and the granting agency.

275  
00:25:57,300 --> 00:26:03,710  
So I think it... it indirectly, I think, can  
help generate more research on campus. It

276  
00:26:03,710 --> 00:26:09,250  
can help funding, I think. It does create  
an organic link between the grant

277  
00:26:09,250 --> 00:26:14,200  
compliance and the analytics um, for tenure and promotion. So I think  
it does help

278  
00:26:14,200 --> 00:26:20,570  
faculty make uh, climb that... that promotion ladder. And it does help  
institutions with Open

279  
00:26:20,570 --> 00:26:28,630  
Access policies to better track  
participation on... on Open Access. So here

280  
00:26:28,630 --> 00:26:35,790  
are um, where data is being harvested from  
right now. And maybe um, as... I think some of you

281  
00:26:35,790 --> 00:26:40,250  
are library folks and this... some of these  
might mean more to you than they do to

282  
00:26:40,250 --> 00:26:48,810  
me um, but cl... clinical trials I know about. So the... the... the  
data is coming from um, clinical trials,

283  
00:26:48,810 --> 00:26:55,530  
Department of Energy SciTech and PAGES um, PLoS - the Public Library  
of Science. Um,

284  
00:26:56,550 --> 00:27:02,000  
these are some of the institutions -

University of California is participating in

285

00:27:02,000 --> 00:27:09,210

the um, in the pilot. So U.C. eScholarship...  
Wayne State is participating in the

286

00:27:09,210 --> 00:27:14,860

um, in the pilots of the Wayne State Digital  
Commons. So is Virginia Tech - that's what Vtech

287

00:27:14,860 --> 00:27:24,559

Works. So NLM - National Library of  
Medicine... oops, I keep doing the wrong thing here. National Library  
of Medicine um, um, is...

288

00:27:24,559 --> 00:27:32,090

is contributing PubMed Central. CrossRef,  
arXiv, and um, DataONE. So there... there's a lot of

289

00:27:32,090 --> 00:27:39,309

data being harvested from uh, a number of um, sources, if you will.

290

00:27:41,409 --> 00:27:50,549

So as we um, even in the... in the fall of 2014, which was sort of  
halfway through at

291

00:27:50,549 --> 00:27:55,690

this point, between now and the end of  
the year there will be further expansion

292

00:27:55,690 --> 00:28:02,350

of the prototypes. So there will be ten more  
campuses um, I think, from... campuses using

293

00:28:02,350 --> 00:28:05,010

DuraSpace, camp... campuses using bepress, with the

294

00:28:05,010 --> 00:28:11,529

inclusion of more data sets in addition  
of a... of DMPTool data. Um, and there's going to be

295

00:28:11,529 --> 00:28:17,919  
a "push API" utilized - a push strategy, and  
some further subscription methods

296  
00:28:17,919 --> 00:28:22,700  
actually to the... to the prototype. So this  
prototype is going to be expanded to, now, 15

297  
00:28:22,700 --> 00:28:30,529  
campuses an... to see what um, see what actually comes out of that.  
And so as.. as we

298  
00:28:30,529 --> 00:28:38,299  
look at... at governance for SHARE - that's  
a big issue. Um, and... the... the... the governance was... was...

299  
00:28:38,299 --> 00:28:43,000  
was just changed. And so there... there...  
there was a steering committee to start

300  
00:28:43,000 --> 00:28:48,899  
with and that has morphed into um, an  
advisory board and exec... thi... this is the

301  
00:28:48,899 --> 00:28:53,870  
workhorse committee. And Tyler Walters  
from Virginia Tech has been appointed,

302  
00:28:53,870 --> 00:28:54,800  
actually, the...

303  
00:28:54,800 --> 00:28:59,820  
the Director. So these... these working  
groups have become community working

304  
00:28:59,820 --> 00:29:06,530  
groups um, and so this is - not sure whether you care so much about  
this - but um, this is sort of

305  
00:29:06,530 --> 00:29:11,480  
how it's... how it's governed. And so here  
is what the whole thing sort of looks

306

00:29:11,480 --> 00:29:15,420  
like going forward. So the SHARE  
notification service um,

307

00:29:15,420 --> 00:29:22,350  
leads to the SHARE registry, which is,  
really um, the organization of the

308

00:29:22,350 --> 00:29:27,610  
notification service - the... kind of a  
listing of all of the data that is in

309

00:29:27,610 --> 00:29:31,330  
there. It's sort of like a table of  
contents, if you will, and it is... it's

310

00:29:31,330 --> 00:29:41,150  
really meant to be the metadata around um, around the uh, uh,  
notification service. And... and then

311

00:29:41,150 --> 00:29:48,360  
the... the SHARE discovery layer um, is really searchable and I like  
the friendly term.

312

00:29:48,360 --> 00:29:56,390  
Uh, it's... it's supposed to be uh, a friendly...  
friendly and searchable. And then finally um,

313

00:29:56,390 --> 00:30:08,160  
well this... this just shows you who it's for. (cough) And finally the  
aggregation and... and reuse um, to be able to go in and actually

314

00:30:08,160 --> 00:30:15,680  
ask specific questions and look at what  
is um, what can actually be done with this

315

00:30:15,680 --> 00:30:22,740  
data. You know, what is the um, taking big  
data sets from one field um, looking at,

316

00:30:22,740 --> 00:30:28,180  
you know, another field and going like this. So that's... that's the  
real um, I think, the real

317  
00:30:28,180 --> 00:30:32,690  
power of this sort of data. Now you can,  
you know, it doesn't take too long to

318  
00:30:32,690 --> 00:30:37,510  
realize there could be some... there... there  
could be some mischief go on here. You

319  
00:30:37,510 --> 00:30:43,770  
know, we have people who really don't  
know um, the... the ins and outs of this data - how

320  
00:30:43,770 --> 00:30:47,500  
it was collected. You know, right now we  
have problems with... with

321  
00:30:47,500 --> 00:30:53,930  
irreproducibility, and so s... sometimes that  
relates to dif... widely differing

322  
00:30:53,930 --> 00:30:59,410  
conditions that are used and... and how data is collected. So I can  
see that... that that's

323  
00:30:59,410 --> 00:31:02,740  
going to be an immediate problem we have  
in that this paper doesn't agree with

324  
00:31:02,740 --> 00:31:05,710  
this paper and you have a third paper  
out there of...

325  
00:31:05,710 --> 00:31:11,700  
of um, extrapolating on why that is. So I  
think this aggregation and reuse, while it has

326  
00:31:11,700 --> 00:31:18,600  
tremendous possibilities, also has some

tremendous risks as well. So um, that in a

327

00:31:18,600 --> 00:31:26,409

nutshell is the SHARE system, and um, um, hopefully you've... you've um, learned something about the SHARE system,

328

00:31:26,409 --> 00:31:36,769

but I'll be... I'll be happy to take questions um, at this point. How about questions around SHARE?

329

00:31:39,780 --> 00:31:42,702

So... (inaudible)

330

00:31:42,702 --> 00:31:51,280

Bepress is looking for partners for one of the repository communities that's original partners with SHARE

331

00:31:51,280 --> 00:31:56,920

What do you see...uhm, as (inaudible)

332

00:31:56,920 --> 00:32:02,380

How does an institution become involved

333

00:32:02,380 --> 00:32:06,780

I'm not saying we can't become a partner at this prototype stage, but down the road

334

00:32:06,780 --> 00:32:11,180

as these... as this becomes more readily available

335

00:32:11,180 --> 00:32:12,900

and... and... (inaudible)

336

00:32:12,900 --> 00:32:20,340

How do you... how do you see institutions... (inaudible)

337

00:32:20,350 --> 00:32:26,539

So I... I see it largely through the library at this point. I mean, I see the library as really

338

00:32:26,539 --> 00:32:31,429  
critical. You know, I see that... that... that  
really the... there really needs to be a

339  
00:32:31,429 --> 00:32:40,980  
threefold partnership here, but... probably  
between the um, the CIO's office um, the research

340  
00:32:40,980 --> 00:32:46,850  
office, and the library really, I think,  
need to work very closely together um, on

341  
00:32:46,850 --> 00:32:52,980  
this. So I see that... that maybe the library  
becomes the central repository for

342  
00:32:52,980 --> 00:32:59,250  
institutional data, you know. Now is that a um, I don't know the  
Wright State Library

343  
00:32:59,250 --> 00:33:05,169  
system - is there an institutional  
repository here? Are the faculty um, do they

344  
00:33:05,169 --> 00:33:10,960  
put their publications in there now? So  
is that required that they do that? So

345  
00:33:10,960 --> 00:33:22,450  
they do it voluntarily? Is it just  
publications? It... it's a wide um, breadth of materials that we have  
in our repository from special collections and archives materials,

346  
00:33:22,450 --> 00:33:36,400  
to administrative documents from our faculty senate, to scholarship.  
Yeah. So, I mean, I see institutional repositories as being critical  
to this um, but I think it's... it's...

347  
00:33:36,400 --> 00:33:43,120  
thi... this is not meant to um, to actually be  
cumbersome. It's not meant that PIs have

348

00:33:43,120 --> 00:33:49,000  
to physically send things in. I mean,  
there is an automatic harvest from... from

349  
00:33:49,000 --> 00:33:55,120  
publishers, from um, you know, from... from  
repositories. So, you know, I would see

350  
00:33:55,120 --> 00:34:03,060  
institutional repositories being queried um,  
publishers um, meetings, you know. So these...

351  
00:34:03,060 --> 00:34:08,970  
but... but it's meant to be very automatic because one thing we... we  
realized is that if PIs

352  
00:34:08,970 --> 00:34:13,380  
have to go in and enter this stuff it's  
just not going to happen, you know. So

353  
00:34:13,380 --> 00:34:37,619  
there has to be a lot of um, input into the  
system. And is SHARE going to uh, um, create standards for metadata or  
(inaudible)

354  
00:34:37,619 --> 00:34:43,419  
Yeah. (inaudible) Dr. Whitacre um, Cleveland State is just asking  
you... for you to repeat the question

355  
00:34:43,429 --> 00:34:52,609  
since the audience is not mic'd. Oh, sorry... okay... sorry. So... so  
the um, the... the question was... was around um, will... will

356  
00:34:52,609 --> 00:34:59,120  
there be sort of a standardization of... of... of metadata. And my  
understanding is that SHARE is actually

357  
00:34:59,140 --> 00:35:06,809  
going to, you know, suggest some standards for... for that... that  
metadata because it is

358  
00:35:06,809 --> 00:35:15,489

different than um, kind of the... the standard today. So that's my... that's my understanding, but again I'm

359

00:35:15,530 --> 00:35:29,770

not a library person; I'm a mouse doctor, so. (laugh) We have a question from Cleveland State. They're asking um, this is actually uh, Lisa Macklin who was a... a presenter earlier today.

360

00:35:29,770 --> 00:35:42,480

What should libraries do to prepare for the implementation of the notification system? That's a great question. Um, you know what, I... I... I think that if... if

361

00:35:42,510 --> 00:35:48,630

libraries today don't have that linkage with the... with the research office and

362

00:35:48,630 --> 00:35:54,549

the um, and the CIO I would think that would be... that would be a good system to sort

363

00:35:54,549 --> 00:36:02,790

of setup at this point. So, I mean, at... at... at Ohio State we have um, uh, a... a working group and

364

00:36:02,790 --> 00:36:09,020

we... we meet periodically between the head of the library, the CIO, and myself. And we

365

00:36:09,020 --> 00:36:13,150

have some folks that actually are designated in our organizations to

366

00:36:13,150 --> 00:36:18,700

actually carry this forward, you know. So we... we meet about SHARE, we meet about data

367

00:36:18,700 --> 00:36:23,839

management plans, we meet about, you know, anything related to... and it's... it's amazing

368

00:36:23,839 --> 00:36:31,900

to me how much we overlap in, you know, in areas um, between the three organizations.

369

00:36:31,900 --> 00:36:39,880

Um, and it's... it's um, in... in our place the CIO is... his office is right next to the head of the library.

370

00:36:39,880 --> 00:36:47,000

His... his office actually is in the library. And so at um, that's a... that's a terrific hook-up. So

371

00:36:47,000 --> 00:36:51,530

there, they talk all the time, and bringing me in from the research office

372

00:36:51,530 --> 00:37:07,040

was um, sort of the third... the third part of that triangle. There's another question from Cleveland State. You talk about the tremendous risks to reuse data, but you don't seem to have anything

373

00:37:07,040 --> 00:37:17,700

to substantiate that. Can you give some examples that show how sharing data has been a risk and how people will muck around with it? (laugh)

374

00:37:17,700 --> 00:37:28,460

Well, I... I don't have any examples - that's a good point. Um, it's really just the... the um, what if, you know, sort of... of... I... I've... I've... I've

375

00:37:28,470 --> 00:37:34,380

seen a lot of um, a lot of criticism about data that can't be reproduced. There was

376

00:37:34,380 --> 00:37:41,490

a very public paper that actually came out of... more of a... a lay press article about

377

00:37:41,490 --> 00:37:48,730

um, I think... I think it was in The Economist. It was... it was the, you know, how, particularly

378

00:37:48,730 --> 00:37:52,050  
biomedical data can't... can't be  
reproduced. And this was about clinical

379  
00:37:52,050 --> 00:37:58,620  
trial data. Um, and so I think it's, you know,  
when I'm... I'm projecting. This is purely

380  
00:37:58,620 --> 00:38:05,510  
my um, my opinion, but when you have people that aren't necessarily  
trained in these

381  
00:38:05,510 --> 00:38:14,980  
areas starting to interpret data um, and  
interpret um, you know, some... some data that, I mean,

382  
00:38:14,980 --> 00:38:20,360  
we've all seen data that is, you know,  
there... there could be some leaps made um, in

383  
00:38:20,360 --> 00:38:25,060  
conclusions by scientists and we've all  
seen that in... in various papers. And I'm...

384  
00:38:25,060 --> 00:38:31,830  
I'm... I'm thinking that... that the public, you know... some...  
someone that's not trained in analysis

385  
00:38:31,830 --> 00:38:39,170  
of clinical trials can essentially take  
a conclusion at face value, and, say, take

386  
00:38:39,170 --> 00:38:44,350  
this conclusion that doesn't match th...  
this conclusion and um, and kind of run with

387  
00:38:44,350 --> 00:38:51,310  
that ball. And so it's purely conjecture on  
my part. I don't have a... a... a good example.

388  
00:38:57,960 --> 00:39:45,160

Other questions? So you want to run out there and put all your papers into SHARE? (laugh) Well I guess I was... I will... I will ask another one. Um, so (inaudible)

389

00:39:45,230 --> 00:39:52,680

Well, I mean, I think one of the things actually that... that... that I've sort of taken upon myself as the

390

00:39:52,680 --> 00:39:59,660

research officer um, is really to... to start talking this up amongst the faculty, you

391

00:39:59,660 --> 00:40:03,570

know, because I think it's... it's important that they realize the possibilities

392

00:40:03,570 --> 00:40:07,750

here, and maybe to have some... what... what we realized at the SHARE meeting

393

00:40:07,750 --> 00:40:13,750

last week was that we need use cases. This was a... I... I had to ask uh, the definition of

394

00:40:13,750 --> 00:40:20,840

that, so. I mean it's really more of an example of what... what is a um, uh, an example to a

395

00:40:20,840 --> 00:40:22,280

researcher of

396

00:40:22,280 --> 00:40:30,119

what this can mean to their career, of... of having um, of be... being able to... to... to establish

397

00:40:30,119 --> 00:40:34,550

collaborations, for example - how important this is to have your... have your work out

398

00:40:34,550 --> 00:40:38,920

there. You know, I think there's... there's some faculty I know at my own institution who

399

00:40:38,920 --> 00:40:44,119

will probably view this is a threat to them. You know, they... they want to... they

400

00:40:44,119 --> 00:40:49,070

want their own group of collaborators and they don't essentially want their

401

00:40:49,070 --> 00:40:57,760

um, articles really spr... spread around. Um, and, I mean, there's... there's some risk of public

402

00:40:57,760 --> 00:41:04,000

criticism of... of those articles, but I think it's um, and... and this is totally voluntary.

403

00:41:04,000 --> 00:41:07,850

This is not something that, you know, you don't have to put your poster in there,

404

00:41:07,850 --> 00:41:13,890

you don't have to put your abstracts in there, but I... I think it will be clear, I

405

00:41:13,890 --> 00:41:18,480

think, to those people that do embrace this that there are real advantages. So...

406

00:41:18,480 --> 00:41:23,180

so sort of step one that... that I'm doing, and Ellen I'd encourage you to do that too,

407

00:41:23,180 --> 00:41:28,000

is to really... is to really talk to researchers about this because I think the

408

00:41:28,000 --> 00:41:32,740

more we talk about it and the more we, you know, make it sort of less

409

00:41:32,740 --> 00:41:40,740

threatening um, the better because I think it really is um, I mean, it's something that... that

410

00:41:40,740 --> 00:41:45,580

OSTP has mandated, I mean, it's coming, you know, whether we essentially embrace it

411

00:41:45,580 --> 00:41:51,590

or not. You know, this sort of OSTP sort of mindset and... and Congress - let's not

412

00:41:51,590 --> 00:41:55,279

forget where this came from - was really Congress

413

00:41:55,279 --> 00:42:01,869

saying, you know, we spend all this money, you know. You scientists really have, you know,

414

00:42:01,869 --> 00:42:07,009

accountability to us, you know, about what you spent the money on. You know, and you

415

00:42:07,009 --> 00:42:12,299

can't... you can't really argue with that, you know, sort of logic. So it is coming

416

00:42:12,299 --> 00:42:31,099

um, and I think we need to, you know, kind of get our faculty ready for this. Yes? Is there any traction emerging around the idea that these... trying to harmonize the departments of 23 different funding (inaudible)

417

00:42:31,109 --> 00:42:44,869

Thanks for... yeah, thanks for bringing that up. So what... what it's looking like - and not all of them are out at this point, which just boggles my mind - it's Open Access - um, is there will be a

418

00:42:44,869 --> 00:42:50,549

clustering of, you know, of funding agencies. So, for example, the... all the

419

00:42:50,549 --> 00:42:56,789  
public health service agencies will  
probably use PubMed, you know. And so, you

420  
00:42:56,789 --> 00:43:00,950  
know, I don't how PubMed is going to deal  
with data, though, is the issue. It's set

421  
00:43:00,950 --> 00:43:04,980  
for publications. I, you know, I have to put  
my own publications in PubMed and - it's

422  
00:43:04,980 --> 00:43:11,359  
PubMed Central - and it's not... it's not the  
most um, user-friendly system - multiple steps

423  
00:43:11,359 --> 00:43:18,460  
and, you know, I'm hoping that this will  
enable um... I forgot to repeat the question,

424  
00:43:18,460 --> 00:43:26,579  
sorry. Um, I'm just not... not sure, I mean, I'm... I'm hoping PubMed  
will actually, you know,

425  
00:43:26,579 --> 00:43:34,670  
take the hint and make this a simpler  
system, but I think there are some um, s... some of the

426  
00:43:34,670 --> 00:43:40,930  
agencies are really getting together. So  
there may be um, a mixture of four or five

427  
00:43:40,930 --> 00:43:47,849  
rather than 23. So, I mean, it's... it's not  
like one, but it's better. And I think it...

428  
00:43:47,849 --> 00:43:55,069  
they um, some of these clusters may be um, closer to what some faculty  
do. I mean, so... so... so it may link

429  
00:43:55,069 --> 00:44:02,989

closer to what, you know, the um, the spectrum of what a single faculty member may... may

430

00:44:02,989 --> 00:44:07,609  
have funding from. Yes? I remember reading somewhere

431

00:44:07,610 --> 00:44:21,090  
that each campus would have to designate a repository. Now is that from... that might have been from the White House itself or perhaps from SHARE, but I guess it's sort of... I'm not clear on will SHARE

432

00:44:21,100 --> 00:44:30,960  
just serve as the repository for this material or... or... So... so the question is will... will SHARE serve as the repository or will there be an institutional um, you know, an

433

00:44:30,980 --> 00:44:37,470  
individual institutional repository. That... that was, I think, the first volley was, you know, every campus is

434

00:44:37,470 --> 00:44:41,770  
going to have to have their own repository and I think it's important for the non-

435

00:44:41,770 --> 00:44:45,400  
librarians in the room to realize that, you know, there's a huge spectrum out

436

00:44:45,400 --> 00:44:50,680  
there of institutions that have really robust repositories and then

437

00:44:50,680 --> 00:44:55,370  
institutions that don't, you know, that may have nothing in terms of repository.

438

00:44:55,370 --> 00:45:01,200  
So I think that was the first um, wave out there was that... that was one of the

439

00:45:01,200 --> 00:45:06,350  
solutions that was put out in the

Holdren memo was that every institution is

440

00:45:06,350 --> 00:45:08,100  
going to have to deal with this on their  
own.

441

00:45:08,100 --> 00:45:12,360  
They're going to have to put all their  
data in, all their publications, you know,

442

00:45:12,360 --> 00:45:17,160  
I mean, it... the whole concept of SHARE  
wasn't on the table when the February

443

00:45:17,160 --> 00:45:21,370  
2013 memo came out. But that was one of  
the ways to deal with this, that you put

444

00:45:21,370 --> 00:45:28,050  
your publications in, you put your data  
in, you know, only locally. And I think

445

00:45:28,050 --> 00:45:33,350  
that... that would be a huge mistake. That  
wouldn't give you the power of the

446

00:45:33,350 --> 00:45:36,350  
um, aggregation.

447

00:45:39,320 --> 00:46:22,840  
Great questions. Um, (inaudible) Are those kinds of systems going to  
play a role?

448

00:46:22,930 --> 00:46:30,300  
So, I mean, I see... so the... so the question is um, are these um,  
products out there like from

449

00:46:30,300 --> 00:46:32,980  
publishers like SciVal from Elsevier. Are

450

00:46:32,980 --> 00:46:40,400

they going to play a role in this um, in the SHARE system. I don't see a... a central

451

00:46:40,400 --> 00:46:45,309  
role of something like that because they are... tend to align with publishers,

452

00:46:45,309 --> 00:46:51,109  
you know, like SciVal is um, is Elsevier's, but you know, I... because I think they... they have a

453

00:46:51,109 --> 00:46:56,500  
different mission, you know, so they actually do look at things like

454

00:46:56,500 --> 00:47:02,140  
establishing collaborations, you know. They map collaborations um, through

455

00:47:02,140 --> 00:47:09,550  
published articles. So I think you... you can um, you can do some of the same things. I

456

00:47:09,550 --> 00:47:14,079  
mean, there's not a complete overlap at all, but... but it breaks down when you get

457

00:47:14,079 --> 00:47:20,730  
to data. So I think what SHARE brings that these other commercial products

458

00:47:20,730 --> 00:47:26,000  
don't is access to the data that actually is produced here, you know, and

459

00:47:26,000 --> 00:47:30,170  
is... it is produced on all of our campuses. So, you know, I think, I mean, things like

460

00:47:30,170 --> 00:47:37,109  
academic analytics and um, and SciVal and Scopus and, you know, things like that I think will

461

00:47:37,109 --> 00:47:50,500

be... will be important input resources, but  
the data still is the issue. I think we...

462

00:47:50,500 --> 00:47:56,520

we actually made a huge um, step forward when we sort of got to the  
point that it must be s...

463

00:47:56,520 --> 00:48:02,990

that data must be somehow reviewed. It's  
not just random, you know, it's not just sort of

464

00:48:02,990 --> 00:48:07,030

random. If... if I think about putting your  
lab notebook in there that a graduate

465

00:48:07,030 --> 00:48:11,030

student did it's sort of horrifying. (laugh)

466

00:48:11,069 --> 00:48:14,069

You don't want to do that.

467

00:48:17,420 --> 00:48:20,600

Other questions?

468

00:48:22,770 --> 00:48:27,250

Any more questions from up the road? I think we're clear (inaudible)

469

00:48:27,250 --> 00:48:31,020

Okay. Thank you very much. I really  
appreciated your questions and your

470

00:48:31,020 --> 00:48:33,940

interest. Thanks. (applause)