

1

00:00:08.670 --> 00:00:31.959

Sharon Burney: Hi everyone, and welcome to the second webinar for those who've been invited to submit a final application to CLIR's *Digitizing Hidden Special Collections and Archives: Amplifying Unheard Voices* program this year. I'm Sharon Burney, a program officer of the CLIR's Grants Team. And today's session will be focused on assessing collections for digitization.

2

00:00:35.600 --> 00:00:42.549

Sharon Burney: We're going to start off with some introductions, first to this platform and then to our presenter for today's webinar.

3

00:00:44.290 --> 00:00:49.400

Sharon Burney: To start off. We want to share some notes about using zoom Zoom during this session.

4

00:00:49.570 --> 00:00:59.860

Sharon Burney: A live transcript is being generated if you like to utilize it. To turn on this feature, click the "CC Live Transcript" button at the bottom of your Zoom window.

5

00:01:00.020 --> 00:01:24.280

Sharon Burney: Attendees will be muted throughout the presentation. However, you're welcome to use chat for general introductions and conversation. If a chat box isn't already displaying on your zoom screen, you can hover your mouse toward the bottom of the screen and open it manually in order to send a message to everyone, remember to change the default from "All panelists" to "All panelists and attendees".

6

00:01:24.360 --> 00:01:28.570

Sharon Burney: We ask that you use a separate Q&A box to ask questions.

7

00:01:28.710 --> 00:01:42.690

Sharon Burney: This can be open the same way as the chat box. Because we have limited time together. Please feel free to follow along with questions as they come in and "upvote" them to let us know that it would be helpful to you to have them answered live.

8

00:01:42.710 --> 00:01:50.620

Sharon Burney: We'll do our best to answer all of the questions live, but we will also share responses to all questions alongside the recording.

9

00:01:50.670 --> 00:02:01.439

Sharon Burney: The session is being recorded. The slides, Q&A, recording, and transcript will be made available on the [Apply for an Award](#) page of our website in the next 2 weeks.

10

00:02:04.440 --> 00:02:16.279

Jane Larson (she/her): Before we move into our program content, we'd like to center today's conversation by acknowledging that, as residents of the US, we are speaking to you from unceded land of many indigenous peoples.

11

00:02:16.400 --> 00:02:25.130

Jane Larson (she/her): We recognize the longstanding history that has brought us to reside on this land, and we seek to have greater understanding of our place within that colonial history.

12

00:02:25.640 --> 00:02:32.130

Jane Larson (she/her): we denounce any acts of terrorism, colonialism, and environmental disregard, both past and present.

13

00:02:32.460 --> 00:02:41.019

Jane Larson (she/her): In particular, we would like to respectfully acknowledge and recognize the original and current caretakers of the land, water, and air where we live,

14

00:02:41.280 --> 00:02:45.180

Jane Larson (she/her): all indigenous peoples and all of their ancestors and descendants.

15

00:02:45.640 --> 00:02:53.640

Jane Larson (she/her): I'm grateful to have the opportunity to live and work on their homelands, and ask you to join me in acknowledging all indigenous communities.

16

00:02:53.920 --> 00:03:03.719

Jane Larson (she/her): May this acknowledgement in our ongoing work demonstrate CLIR's commitment to beginning the process of working to dismantle the ongoing legacies of settler colonialism.

17

00:03:03.940 --> 00:03:11.129

Jane Larson (she/her): If you haven't done so already, we'd love for you to introduce yourself in the chat and share a land acknowledgement for the area where you live, if you would like.

18

00:03:15.880 --> 00:03:40.820

Jada Wright-Greene: During today's session, our host will dig into one of the topics we mentioned last week: assessing collections for digitization. We've allotted 90 min total for today, and the speaker will pause at the halfway point at 2:45 ET for a quick break, continue with the presentation, and then move towards Q&A at approximately 3:05 ET.

19

00:03:40.920 --> 00:03:54.830

Jada Wright-Greene: We're hoping for an engaging time of conversation around, identifying and clarifying the collections and materials that will be nominated for digitization, accurately estimating volume of materials and digitization outputs,

20

00:03:54.840 --> 00:04:12.909

Jada Wright-Greene: making a competitive case for digitization. The plan is to reserve about 25 min for Q & A near the end of this session. Both the session, co-host and CLIR staff will be available to answer questions at that time. So, while you may post questions throughout using the Q&A box,

21

00:04:13.270 --> 00:04:23.499

Jada Wright-Greene: keep in mind that we may not get to every question as it's asked. Staff will be keeping track, and we'll circle back to unanswered questions during a dedicated Q&A.

22

00:04:28.520 --> 00:04:43.660

Jada Wright-Greene: We want to start with a quick poll to gauge how you're feeling about today's topic of assessing collections for digitization in today's session, you'll learn more about the skills and strategies you'll need to complete the the following application tasks:

23

00:04:44.170 --> 00:04:55.330

Jada Wright-Greene: Project Details - material task, which includes the list of collections upload and the representative samples upload and the tell the story task. And our poll has been launched.

24

00:05:09.350 --> 00:05:17.779

Alyson Pope: So it looks like we have responses from just about everyone. I'll give you guys a few more seconds before we go back to the share.

25

00:05:23.940 --> 00:05:24.890

Alyson Pope: Okay.

26

00:05:37.840 --> 00:05:59.649

Alyson Pope: Now, we're glad to turn things over to Dr. Mitch Frass, Senior Curator for Special Collections at the Kislak Center for Special Collections, Rare books, and Manuscripts at the University of Pennsylvania libraries will take just a moment to let Mitch share their slides and get oriented, and as they do that, please remember that we'll be breaking for 5 min at the at the halfway point. Thank you.

27

00:06:01.420 --> 00:06:05.120

Mitch Fraas: Hi, everyone just gonna share my slides here.

28

00:06:14.770 --> 00:06:15.790

Mitch Fraas: There we go.

29

00:06:18.940 --> 00:06:40.490

Mitch Fraas: Thanks so much to the CLIR staff and everyone for inviting me, and for bringing me into contact with all the applicants here today. It's really a privilege, a pleasure to talk to you about some of the issues around selecting and coming up with a cohesive argument over collections to digitize as part of the Hidden Collections project.

30

00:06:40.490 --> 00:06:49.179

Mitch Fraas: I'm probably a bit biased, but I think the the CLIR Hidden Collections Grant program, which has existed in various manifestations for over a decade now, is

31

00:06:49.330 --> 00:07:06.530

Mitch Fraas: pretty much one of the best things to happen in the field of cultural heritage and digitization. Anyone looking to highlight and engage with their holdings. This is just a godsend, and it's especially exciting to see how the program has really honed its focus in recent years on amplifying unheard voices.

32

00:07:06.580 --> 00:07:24.690

Mitch Fraas: You heard a little bit of my background there. I'm Senior Curator and Director of Special Collections and Research Services here at Kislak Center at Penn. I've been the PI, Penn lead on two successful Hidden Collections CLIR grants in the past.

33

00:07:24.970 --> 00:07:30.510

Mitch Fraas: And I'll put links in the chat right here, just so you can a

34

00:07:30.990 --> 00:07:34.299

Mitch Fraas: see that. And I've also

35

00:07:34.310 --> 00:07:37.020

Mitch Fraas: been really privileged to be a participant on

36

00:07:37.100 --> 00:07:48.429

Mitch Fraas: several other CLIR grants, including some done in collaboration with other institutions here in the area. And it's it's really been a

37

00:07:48.750 --> 00:07:50.809

Mitch Fraas: amazing privilege to work on these.

38

00:07:51.870 --> 00:07:56.100

Mitch Fraas: The two that I've been primarily working on are Manuscripts of the Muslim World

39

00:07:56.170 --> 00:08:02.169

Mitch Fraas: and the Health of the New Nation, both big collaborative grants largely out of Philadelphia.

40

00:08:02.200 --> 00:08:12.650

Mitch Fraas: and which produced, I think, very different outputs. One focused on a lot of early, modern and pre modern manuscripts, and the other on a lot of archival and

41

00:08:12.800 --> 00:08:14.410

Mitch Fraas: nineteenth century material.

42

00:08:14.530 --> 00:08:28.749

Mitch Fraas: I've also worked on, not as a lead, but in sort of consulting roles on a large digitization project here at Penn around Mary Anderson's papers and helped out on a couple of other grants, including one to digitize records from

43

00:08:28.920 --> 00:08:32.709

Mitch Fraas: some of Philadelphia's houses of worship. So it's a

44

00:08:33.030 --> 00:08:49.669

Mitch Fraas: the CLIR Hidden Collections project has played a big role in my history as a librarian here at Penn. Also, I should give a shout out to CLIR for starting me on my librarian journey. I was a CLIR postdoc back in 2011, which was my first full time library job. So

45

00:08:49.670 --> 00:09:12.650

Mitch Fraas: I can't say enough good things about CLIR. My own academic background is in the history of South Asia, but in the past decade, I've really enjoyed working on collections here, ranging from medieval manuscripts to the Afrofuturists on Ra. And I imagine all of you here today have a similarly broad range of interests across different time periods and subjects and institutions.

46

00:09:12.680 --> 00:09:16.420

Mitch Fraas: I should say, I'm acutely aware that I'm coming from

47

00:09:16.460 --> 00:09:35.829

Mitch Fraas: a relatively well-resourced, university library setting, which is different than I think a lot of applicants might be coming from. And I think that can result in some significant blind spots or biases in how we talk about selecting material or even in the digitization process itself, and I'm sort of hopeful that

48

00:09:35.830 --> 00:09:50.009

Mitch Fraas: my remarks won't be too predicated on any particular resource dependency. But please do put in the QA, if there are questions about how to do things in different sort of institutional or resource environments, I think that would be

49

00:09:50.040 --> 00:10:10.440

Mitch Fraas: very helpful. I have, though, in the past, on, on CLIR grants, worked with a variety of institutions ranging from large, but not well resourced historical societies to small colleges,

private organizations, to, you know, individual houses of worship who just happen to have records in their basement. So if you know, I've sort of seen the gamut of

50

00:10:10.440 --> 00:10:20.480

Mitch Fraas: different kinds of digitization structures. So you know, this is something that I think we can maybe talk about more in the Q&A if you have, you know, particular

51

00:10:20.500 --> 00:10:32.510

Mitch Fraas: concerns about scale and scope. I'm gonna be talking today about assessing collections for digitization and in a couple of key areas.

52

00:10:33.410 --> 00:10:46.370

Mitch Fraas: First telling the story of the materials you want to digitize, listing the materials that you're proposing the formats and amounts to be digitized, and also estimating outputs and amounts

53

00:10:46.430 --> 00:11:15.500

Mitch Fraas: and then the sort of representative samples you'd be submitting. I'm taking this outline from the general CLIR application, which you'll have seen. And these are sort of some of the main areas from that CLIR application that you know, you'll be expected to complete, and it makes a decent framework to talk about some of these materials. There's a lot to cover and I know I'll probably leave some things out, so please do bring things up in the chat or in the Q&A as they do.

54

00:11:15.670 --> 00:11:26.479

Mitch Fraas: So I imagine everyone who's at this point in the CLIR process, you know, you've done this sort of pre-application. You're you're at this point of putting together a big package, and you're probably

55

00:11:27.060 --> 00:11:36.119

Mitch Fraas: thinking about and wading into the many pages of technical details, budgets, lists, and other documents which can, I think, at least to me, when I was doing it, seemed sort of overwhelming.

56

00:11:36.360 --> 00:11:50.970

Mitch Fraas: But often, I would sort of lose sight of the the core part of the grant, which is really, or really any grant, which is the story that you're telling of, like, why, why are we even doing this before you get to the lists of files, and you know letters of support and all that.

57

00:11:51.030 --> 00:12:16.870

Mitch Fraas: And I think that story, and the way material support it, is at the heart of this, and it makes a good application and one that is going to be a successful project. And it's not just a story you're telling to the reviewers of the grant in order to get the grant. But a story you're telling to yourself as the applicant. And then the communities that you know you are serving. So this sort of narrative of why it is we're we're doing this, and and why these particular

58

00:12:17.150 --> 00:12:18.260

Mitch Fraas: materials.

59

00:12:18.480 --> 00:12:25.510

Mitch Fraas: And I was going to share, I think, one thing from one of our successful

60

00:12:25.940 --> 00:12:33.710

Mitch Fraas: applications in just a second, which I think helps show a little about how you can, how you can narrativize this.

61

00:12:34.060 --> 00:12:36.280

Mitch Fraas: One,

62

00:12:36.450 --> 00:12:48.279

Mitch Fraas: I think important point here that is just cribbed from CLIR itself, but I think important for the amplifying hidden voices particular focus of these grant programs is

63

00:12:48.490 --> 00:12:52.200

Mitch Fraas: always remembering the public knowledge, community centered access.

64

00:12:52.310 --> 00:13:00.729

Mitch Fraas: you know, broad representation, authentic partnerships, and sustainable infrastructures that CLIR is looking for coming out of these grant projects. And

65

00:13:01.150 --> 00:13:03.369

Mitch Fraas: trying to to think about as you're

66

00:13:03.600 --> 00:13:08.529

Mitch Fraas: as your writing your statement of why this material and why now?

67

00:13:08.560 --> 00:13:11.790

Mitch Fraas: Trying to answer each of those questions in

68

00:13:11.890 --> 00:13:28.279

Mitch Fraas: both the broad sense and then, in the particular sense of not just is this a good idea generally, but why this certain number of boxes or this set of materials? You know? What exactly is this gonna contribute? And sort of reminding yourself, how does what I'm choosing here to digitize, help meet each of these goals?

69

00:13:29.990 --> 00:13:36.969

Mitch Fraas: I have cribbed this from one of our projects, we, a colleague of mine and and several others

70

00:13:37.080 --> 00:13:47.569

Mitch Fraas: did a major hidden collections grant with our Marion Anderson collection here. We have probably over 500 boxes of her papers

71

00:13:47.700 --> 00:14:05.200

Mitch Fraas: including recordings. You know, everything you can possibly imagine. And we wanted to apply for a grant to to make this more available to the public. And I particularly like this statement. And the simplicity of the justification that I've put on the screen here.

72

00:14:05.200 --> 00:14:23.719

Mitch Fraas: I should say I did not write this. I think my colleagues here have distilled a highly complex set of collection materials. You know, it's over 500 boxes, huge range of stuff, different formats, and they've distilled that around you know, a particular figure. Say, you know, Marian Anderson. Into really 3 concise

73

00:14:24.200 --> 00:14:48.350

Mitch Fraas: points. You know, we're going to digitize broadly these things: we're gonna digitize her private sound recordings, her audio interviews and transcripts or recital programs, more than 3 things, diaries and notebooks and scrapbooks. Of all this massive stuff, 500 bliss boxes like these are the things we're gonna do from this figure and here's why. I've just put that in yellow. I think I like that formulation. You know, now that they're digitized.

74

00:14:48.510 --> 00:15:02.839

Mitch Fraas: These materials will allow. X group of people, a community, whoever in this case, we were saying, you know, academic researchers, you know, digitizing these things will allow X group of people to do Y, you know.

75

00:15:03.210 --> 00:15:10.549

Mitch Fraas: What will, having these online and either available to the public or available to a smaller set of people. You know, what what are you

76

00:15:10.680 --> 00:15:17.259

Mitch Fraas: really trying to achieve? You know, this is not just about, you know, I think it would not be,

77

00:15:17.350 --> 00:15:38.590

Mitch Fraas: you know, I was thinking back to some of the big micro filming projects of the seventies, where the justification was, well, you know, we'll get every issue of, you know, the Philadelphia Inquirer on microfilm like that's the justification. Got every single one of them. That's enough. That's all you need to know. I wanna know who's gonna use it, but we gotta do it. In this case, I think there's a lot sharper focus in that,

78

00:15:38.590 --> 00:16:01.520

Mitch Fraas: you know, making a clear justification for what having, you know, a particular set of materials online is going to do for whoever you identify as a community. And I think in this case, you know, we highlight academic researchers. For the CLIR grant, I think there are so many different communities that could be the possible audience here. Thinking through and then highlighting specifically for the reviewers

79

00:16:01.620 --> 00:16:04.869

Mitch Fraas: who is it you're focused on as your audience?

80

00:16:05.410 --> 00:16:06.770

Mitch Fraas: I have,

81

00:16:07.030 --> 00:16:10.840

Mitch Fraas: you know, thinking through how you select.

82

00:16:11.730 --> 00:16:18.299

Mitch Fraas: This is something I'll be covering at length. But just to sort of dive in, I wanted to highlight the importance of

83

00:16:18.390 --> 00:16:19.870

Mitch Fraas: thinking through,

84

00:16:20.390 --> 00:16:31.450

Mitch Fraas: you know, once you've sort of conceptualized: okay, we've got Mary Anderson, a really important historical figure for a lot of reasons. We've got 500 boxes of stuff. We have sort of a vague idea of categories, you know.

85

00:16:31.560 --> 00:16:34.799

Mitch Fraas: How are we going to proceed from here? And

86

00:16:34.970 --> 00:17:01.780

Mitch Fraas: I think this is an exercise that is is very helpful, and one that I'll talk about sort of in terms of community engagement, but on the the bottom left there, you can see one set of boxes from her collection in the upper right, a different set of boxes, you know it's huge amount of material, and the bottom left is boxes of awards she received. She received hundreds of awards over the course of her lifetime, you know these are little like statuettes, or little plaques, or little wooden things.

87

00:17:01.810 --> 00:17:09.040

Mitch Fraas: Tons of them. And the upper right is a listing from our sort of inventory, or finding aid of the collection

88

00:17:09.079 --> 00:17:19.350

Mitch Fraas: of her personal notebooks, journals, diaries, and other writings. And I'm sort of setting up a probably a false dichotomy here, but

89

00:17:19.460 --> 00:17:24.280

Mitch Fraas: in this case it was easy for us to look at 500 boxes and say,

90

00:17:24.450 --> 00:17:45.660

Mitch Fraas: okay, we've only got a certain capacity in the CLIR grant. You know, it's not millions of dollars. We wanna focus on things that are gonna have the highest research impact. That was research and classroom impact were some of our big deliverables. And then, as a third, you know, the arts community and the Philadelphia community writ large. And we thought, you know, if you go through this

91

00:17:45.930 --> 00:17:50.059

Mitch Fraas: big list of what we have, it's so clear to us that

92

00:17:50.120 --> 00:17:58.549

Mitch Fraas: her diaries, notebooks, and journals created by her in her own handwriting, detailing her life, are must have. These are

93

00:17:58.620 --> 00:18:04.540

Mitch Fraas: sources that were only available in the reading room before that, which we could make available to a wide audience.

94

00:18:04.610 --> 00:18:29.380

Mitch Fraas: And equally clear, I think, was the argument that taking pictures with a camera of a bunch of awards was not a great use of our resources for the particular audience we were dealing with, that, you know, proposing to CLIR. You know, we're gonna do this big Mary Anderson digitization project. And you know, we're gonna start with the awards. That's what we're gonna do, I think would be not a very compelling

95

00:18:29.860 --> 00:18:46.089

Mitch Fraas: kind of application. It would be hard to explain how these awards would, you know, pictures of the awards from even 2 angles or something, would contribute to, you know, greater public knowledge and engagement with the life of this great musical artist.

96

00:18:46.100 --> 00:18:59.280

Mitch Fraas: This is not to say that we couldn't construct an argument for why for a different community or a different purpose, and say, you know, we think that 3D modeling of these awards serve some purpose in mapping out,

97

00:18:59.400 --> 00:19:04.719

Mitch Fraas: you know, sort of respectability politics, you know, who's getting what award. I think there are some arguments to be made there, but

98

00:19:04.820 --> 00:19:25.219

Mitch Fraas: you know, we made a choice. I think all of the applicants here are having to make tough choices about what you want to do and for what audience and we chose to pick out, you know, the primary sort of manuscript writing material. She did her personal recordings and the sort of programs and ephemeral records of her different performances. So, we

99

00:19:25.980 --> 00:19:37.449

Mitch Fraas: took 500 boxes of stuff and distilled it into a much smaller amount of things, and created a sort of argument around why this would contribute to our particular community of researchers

100

00:19:37.680 --> 00:19:44.589

Mitch Fraas: increase in engagement with the material and certainly, by default, access through making this available online.

101

00:19:47.160 --> 00:19:50.090

Mitch Fraas: The next thing

102

00:19:50.710 --> 00:20:06.069

Mitch Fraas: I have here is the importance of access and engagement, and making an argument for your selections. And really how difficult selection could be sometimes, and it's not always as self-evident as someone like Mary Anderson, where you're presented with, in our case, 500 boxes of pretty well described stuff.

103

00:20:06.210 --> 00:20:08.460

Mitch Fraas: And you're just choosing like what is the best

104

00:20:08.530 --> 00:20:18.590

Mitch Fraas: you know, most important for our audience. Sometimes, and this is, I'm using a case from the Manuscripts of the Muslim world grant we did, you're really thinking about

105

00:20:18.630 --> 00:20:23.509

Mitch Fraas: just the raw goal of increasing access to a given set of materials

106

00:20:23.530 --> 00:20:42.710

Mitch Fraas: and making sure, not just that these materials are available, but that these are of cultural importance to a particular community, and that they are being described and made available in an accurate and respectful way. And that in itself, I think, can go a long way to making an argument for access. I have an example here.

107

00:20:42.750 --> 00:20:53.969

Mitch Fraas: On the left and sort of bottom, right are what the descriptions of some of our Islamic manuscripts here at Penn looked like before the CLIR grant.

108

00:20:54.100 --> 00:20:55.100

Mitch Fraas: I think

109

00:20:55.450 --> 00:21:12.340

Mitch Fraas: I don't know if needs saying, but I should say that describing a manuscript Quran as being held in a quote "beige shoe box," or that a copy of Al Jazuli's compilation of prayers would bear the outdated and often offensive description of Muhammadan book of prayers on its front there in guilt.

110

00:21:12.460 --> 00:21:16.060

Mitch Fraas: Pretty easy to argue that this is not

111

00:21:16.300 --> 00:21:24.820

Mitch Fraas: a way that people are gonna engage with this material. This is not providing, you know, respectful and accurate representation of these materials.

112

00:21:24.900 --> 00:21:31.389

Mitch Fraas: They're certainly hidden in that they are, you know, described only in paper and not digital. But even if these were

113

00:21:31.760 --> 00:21:43.240

Mitch Fraas: sort of electronic records, they would be equally, I think, hidden. And we wanted to, you know, in this case, sort of transform these manuscripts from

114

00:21:43.350 --> 00:21:45.180

Mitch Fraas: sort of like

115

00:21:45.200 --> 00:22:00.140

Mitch Fraas: a **curious** in a collection of a large university to something that was engaging with a community of scholars, and I think it really helped us make a justification for how intervention, through digitization and cataloging, would be a really worthwhile goal for CLIR to fund.

116

00:22:00.190 --> 00:22:17.670

Mitch Fraas: You'll see the end result of the grant is the sort of tiny text there in the bluish box is what our catalog records look now with Arabic script, thorough description, and you know proper terms that are findable. Many of these now, I was just looking at our stats today, viewed

117

00:22:17.750 --> 00:22:26.000

Mitch Fraas: 12 or 16,000 times since they went up in 2021. I think. So, you know, an amount of engagement, you know,

118

00:22:26.330 --> 00:22:29.879

Mitch Fraas: logarithmically more than they were getting in the beige shoe box.

119

00:22:32.340 --> 00:22:38.429

Mitch Fraas: On that point, I wanted to talk about consultation and selection. And,

120

00:22:38.500 --> 00:22:41.840

Mitch Fraas: you know, really the the role that

121

00:22:42.200 --> 00:22:47.250

Mitch Fraas: plays. I'll put up here, from the manuscripts the Muslim World grant, just a selection of the people

122

00:22:47.370 --> 00:22:54.450

Mitch Fraas: we worked with in developing the project. I think everyone knows, you know, most successful grants won't involve, just like

123

00:22:54.560 --> 00:22:59.129

Mitch Fraas: the person who's applying, or the curator, whoever the person who had the idea,

124

00:22:59.200 --> 00:23:20.369

Mitch Fraas: sort of they're off the top of the head notion of like, yeah, let's digitize this like, I just feel like it. Almost all these projects require consultation and careful thought to determine, even if the idea is originally the the work of one person, the actual selecting of materials, the nitty gritty, you know, figuring out what to include, what not to include, is,

125

00:23:20.470 --> 00:23:25.019

Mitch Fraas: I think, best done by a large group of people in consultation.

126

00:23:25.050 --> 00:23:29.700

Mitch Fraas: I should say, you know in that Mary Anderson's Awards example,

127

00:23:29.970 --> 00:23:40.559

Mitch Fraas: you know, I can imagine, if we were not considering a research audience, having an audience of singers and people involved in Philadelphia arts who might have had a very, you know, different idea of

128

00:23:40.660 --> 00:23:45.980

Mitch Fraas: what was in and what was out. I think there'd be a lot of overlap. But

129

00:23:46.320 --> 00:23:53.660

Mitch Fraas: thinking through who you'd like to serve with this grant, helps determine, that also getting a

130

00:23:53.840 --> 00:24:03.520

Mitch Fraas: array of people with different let's say, disciplinary interests ranging in sort of career phases, as well. Some of our most active

131

00:24:03.850 --> 00:24:11.410

Mitch Fraas: participants sort of conceptualizing the grant and figuring out material, were graduate students who had a lot of enthusiasm and new ideas.

132

00:24:11.460 --> 00:24:22.500

Mitch Fraas: To give an example of what came out of this consultative project, we knew we wanted to digitize, this is the sort of idea was, we'll digitize all of the Islamic manuscripts in Philadelphia and

133

00:24:22.850 --> 00:24:27.970

Mitch Fraas: maybe New York. After talking with a lot of people,

134

00:24:28.010 --> 00:24:54.139

Mitch Fraas: we really refined the scope, and it expanded it to be not just manuscripts with Islamic content or religious manuscripts, but rather from Islamic cultures broadly. So there was a big request, from the people we were talking with, for Christian Coptic manuscripts from Egypt, Hindu materials written in Persian and South Asia. Likewise we were able to include materials in Arabic from the Caribbean, from enslaved people in the eighteenth century,

135

00:24:54.140 --> 00:25:03.160

Mitch Fraas: and talking with our sort of constituents of researchers and scholars, really helped us make the case. For why

136

00:25:03.290 --> 00:25:06.559

Mitch Fraas: our scope of these manuscripts, not those manuscripts,

137

00:25:06.580 --> 00:25:15.940

Mitch Fraas: would make sense. And it was nice to be able to say, you know, we can back this up with, you know, direct conversations with people in the field.

138

00:25:16.270 --> 00:25:18.399

Mitch Fraas: This is not to say, I think,

139

00:25:18.940 --> 00:25:27.039

Mitch Fraas: you know, to try to combat the bias of just academic study, which was certainly one of the main purposes of this particular grant I worked on

140

00:25:27.650 --> 00:25:28.700

Mitch Fraas:

141

00:25:28.770 --> 00:25:45.350

Mitch Fraas: Through the course of this, we realized and saw that different audiences in a different kind of grant might have had very different justifications and talked to very different people. Kudos to the Free Library of Philadelphia, who was one of our partners who decided a year or so into the grant to consult

142

00:25:45.350 --> 00:26:05.060

Mitch Fraas: and talk with local Philadelphia mosques and members of the Muslim community in Philadelphia about what they would like to see from these now digitized objects. And you know, I can imagine if that was our aim from the beginning, you know, maybe a focus on more devotional manuscripts, a focus on transcription. You know, I can imagine this

143

00:26:05.140 --> 00:26:07.029

Mitch Fraas: having an argument driven

144

00:26:07.320 --> 00:26:21.300

Mitch Fraas: grant proposal in a slightly different direction. So I think both of those are equally valid, and you know, especially in the amplifying hidden voices grant application, you know, and the criteria that CLIR sets forth.

145

00:26:21.410 --> 00:26:31.610

Mitch Fraas: Being very clear about who you've consulted and what kind of audience you're imagining, I think makes a huge difference in how this will be read, and how the community also will receive it. You know, is this

146

00:26:31.640 --> 00:26:44.809

Mitch Fraas: primarily something we've talked only with tenured professors about, and what they really wanna do? Or is it something with a different kind of audience, and that you've taken into consideration when selecting which things are in and which things are out?

147

00:26:44.860 --> 00:27:03.079

Mitch Fraas: I think in our case, classroom use and academic study was a perfectly justifiable sort of audience and goal. But I think that is one of very many possible outcomes of these grants. And I think the more creative and consultative you can get with

148

00:27:03.190 --> 00:27:05.629

Mitch Fraas: how you're feeding materials the better.

149

00:27:07.070 --> 00:27:10.610

Mitch Fraas: So formats and amounts to be digitized. Some of the

150

00:27:10.750 --> 00:27:12.299

Mitch Fraas: getting into more of the

151

00:27:12.320 --> 00:27:18.069

Mitch Fraas: particular questions, how do you do this work of selection? You know,

152

00:27:18.190 --> 00:27:25.480

Mitch Fraas: it's one thing to sort of talk with your constituents. In the case of the manuscripts of the Muslim world, we had sort of little

153

00:27:25.500 --> 00:27:30.560

Mitch Fraas: get together symposium beforehand before we wrote things up to talk about things.

154

00:27:30.700 --> 00:27:43.069

Mitch Fraas: But largely, you know, that's at sort of a high level. So you've decided we're gonna do everything from Egypt, regardless of religious background. Good, all very good. But who's how do you actually do the the selection of the actual things,

155

00:27:43.120 --> 00:27:54.759

Mitch Fraas: coming up with the amounts to digitize the outcomes, the tough, tough questions about, you know, we've got 50 manuscripts and money to do 40 of them.

156

00:27:54.800 --> 00:28:00.470

Mitch Fraas: So this is, I think, the sort of meaty problem, given that many of you already have identified a sort of

157

00:28:00.500 --> 00:28:07.900

Mitch Fraas: core set, or have some idea of a set of resources. You know, how you how you begin to to drill down on that.

158

00:28:07.930 --> 00:28:12.770

Mitch Fraas: I've taken a snippet here from the CLIR application, on the left.

159

00:28:12.870 --> 00:28:17.249

Mitch Fraas: which I think kudos to the CLIR staff. Really really well.

160

00:28:17.490 --> 00:28:19.299

Mitch Fraas: written and very

161

00:28:20.160 --> 00:28:26.610

Mitch Fraas: clear. Oh. should remove that. Very evident. And I think

162

00:28:27.300 --> 00:28:45.989

Mitch Fraas: these are really the things that you should be asking yourself throughout the process of doing this selection, and then the actual picking of material. You know, it's all well and good to have the sort of high, fluting ideas about a big concept. But you know, identifying the actual material formats of things,

163

00:28:46.300 --> 00:29:03.379

Mitch Fraas: their history of their place in your institution, where in the world they're from, what date they're from, you know what kind of conditions they're in, their storage, and just the sheer, how many things. You'd be surprised how many times especially this is true in libraries.

164

00:29:03.470 --> 00:29:07.659

Mitch Fraas: Discussions at the high level get derailed when it comes time to be like, okay, so

165

00:29:07.860 --> 00:29:15.940

Mitch Fraas: how many of these VHS cassettes do we have? And they're like, well, I don't know a lot. This is the time to actually make those decisions and come up with a number.

166

00:29:17.550 --> 00:29:18.370

Mitch Fraas: So

167

00:29:19.040 --> 00:29:32.090

Mitch Fraas: this is always my temptation when people ask like, so you're yo're interested in the X collection, or whatever. How big is it? What do you want to digitize? Yeah. Yeah, 12 boxes. It's 12 boxes. That's what we want to do.

168

00:29:32.120 --> 00:29:34.729

Mitch Fraas: We want to digitize those 12 boxes of the collection.

169

00:29:34.950 --> 00:29:37.569

Mitch Fraas: Sadly,

170

00:29:37.580 --> 00:29:49.640

Mitch Fraas: that doesn't doesn't really work most of the time. There's nothing quite as scary as, especially in my case, a donor saying, I've got 50 boxes from the Community Action League. Yeah, you don't even need, it's just a lot of different material. It's all great. It's good.

171

00:29:49.670 --> 00:29:58.640

Mitch Fraas: You know, like, never know what's in there, and for the purposes of making a really strong application with a coherent narrative and justification,

172

00:29:58.850 --> 00:30:08.460

Mitch Fraas: this is where getting into the like, what is it that we're actually digitizing? And the answer is not boxes. Each of these boxes, obviously, holds all manner of surprises.

173

00:30:08.510 --> 00:30:09.840

Mitch Fraas: and I think,

174

00:30:10.250 --> 00:30:15.139

Mitch Fraas: oftentimes, one of the the hard parts of doing these applications is actually the hands dirty,

175

00:30:15.220 --> 00:30:17.359

Mitch Fraas: sifting through things to see like

176

00:30:17.630 --> 00:30:25.459

Mitch Fraas: what's in there and realizing that the inventory from 1999 taped to the front of the box is completely wrong. And you've gotta rethink everything.

177

00:30:25.510 --> 00:30:45.039

Mitch Fraas: So you know, actually having an opportunity to go through, look at some general categories of material and begin sorting, aggregating them in your head of like, okay. You know, we've got these things, which we wanna work with, these things, which I think are a second priority. And then these other things. I don't think we make sense to apply for

178

00:30:45.150 --> 00:30:51.000

Mitch Fraas: You know what's in those boxes. I think if we had a show of hands oops

179

00:30:51.140 --> 00:30:52.390

Mitch Fraas: wrong slide.

180

00:30:52.540 --> 00:31:02.789

Mitch Fraas: If we had a show of hands, there we go. A lot of you would say you've got printed material of some sort in what you're applying to to digitize newspapers, magazines, printed books,

181

00:31:02.910 --> 00:31:06.990

Mitch Fraas: which, sort of the bread and butter of of

182

00:31:07.120 --> 00:31:11.399

Mitch Fraas: digitization projects. And since the beginning of digitization,

183

00:31:11.740 --> 00:31:20.360

Mitch Fraas: this is something that we have always struggled with, where I am of prioritizing and figuring out

184

00:31:20.970 --> 00:31:27.740

Mitch Fraas: how to balance sort of unique material against important material in context.

185

00:31:27.810 --> 00:31:31.980

Mitch Fraas: What do I mean by that? I would say it is,

186

00:31:32.480 --> 00:31:36.620

Mitch Fraas: using the example from the Marian Anderson papers here,

187

00:31:36.870 --> 00:31:45.709

Mitch Fraas: it would not be compelling, I'm speaking for CLIR. CLIR should in the chat just be like this guy is crazy. I don't think it would be a very compelling application, to say, you know,

188

00:31:45.980 --> 00:31:51.310

Mitch Fraas: Marian Anderson was mentioned in the New York Times 800 times during her career.

189

00:31:51.330 --> 00:32:09.460

Mitch Fraas: As a result, we want to digitize the New York Times from 1940 to 1960. I think it is not controversial to say that's not a very compelling vision for digitization. The New York Times. Well represented online, albeit for a fee. So you could make an argument about public access, some rights issues there. But

190

00:32:09.650 --> 00:32:21.159

Mitch Fraas: you know, it's not a unique resource, it's a widely known thing, it's got enormous digitization coverage, and to find 800 articles on Marian Anderson by digitizing

191

00:32:21.520 --> 00:32:24.730

Mitch Fraas: 50,000 papers is not a very, you know, you get the point,

192

00:32:24.800 --> 00:32:31.279

Mitch Fraas: a good use of resources. On the other hand, in our case, you know, we have

193

00:32:31.490 --> 00:32:39.420

Mitch Fraas: a couple scrapbooks of newspaper clippings covering her career from all manner of of magazines, newspapers, etc., and we sort of

194

00:32:39.820 --> 00:32:41.710

Mitch Fraas: waffled on this like, yes,

195

00:32:41.790 --> 00:32:51.180

Mitch Fraas: you could actually search the New York Times database online and find many of these, even the Afro American there is covered in some online databases. But

196

00:32:51.460 --> 00:32:59.249

Mitch Fraas: despite that, we think that her scrapbooking practices and the ability of researchers to see all those clippings in one place

197

00:32:59.310 --> 00:33:00.740

Mitch Fraas: by flipping through

198

00:33:00.890 --> 00:33:22.720

Mitch Fraas: the scrapbooks made a very strong case for why we should digitize it. And not say, you know this, this content isn't unique. This clipping from the Afro American here is not unique. Many research libraries and public libraries across the country have copies of the Afro American and we'll have that exact thing, but they won't have it readily clipped out and you know, available

199

00:33:22.750 --> 00:33:27.710

Mitch Fraas: in the sort of context that we do. So, I would say the same,

200

00:33:27.840 --> 00:33:35.830

Mitch Fraas: in my opinion, applies to a lot of different printed material. Making tough decisions on, you know, is this printed material

201

00:33:36.080 --> 00:33:38.810

Mitch Fraas: unique? And part of the argument,

202

00:33:39.040 --> 00:33:46.910

Mitch Fraas: you know, amplifying that I'm trying to do here, and if not, if it is from sort of mass media that's well represented, is it,

203

00:33:47.150 --> 00:33:50.250

Mitch Fraas: though Part of some sort of compilation or

204

00:33:50.480 --> 00:33:56.719

Mitch Fraas: practice that makes it particularly interesting or useful to your

205

00:33:57.270 --> 00:34:05.440

Mitch Fraas: target community to have all in one place or accessible and think through that way. Especially,

206

00:34:05.520 --> 00:34:19.900

Mitch Fraas: in a world like this one, where, between printed books, newspapers, newspapers, and magazines, the extant coverage online is remarkably rich. There's huge, mass digitization of these 3 categories

207

00:34:20.040 --> 00:34:22.800

Mitch Fraas: and some careful thinking about

208

00:34:23.460 --> 00:34:32.229

Mitch Fraas: uniqueness and non duplication is probably helpful. You'd be surprised how often, you know, even happens here, people will say like, Oh, you know, we should really digitize this Philadelphia

209

00:34:32.270 --> 00:34:39.549

Mitch Fraas: Journal or magazine. It's like published by a Penn department of chemistry. Who's going to have that? And you're like, Oh, wait! University of Michigan already, did it?

210

00:34:39.900 --> 00:34:48.030

Mitch Fraas: Great? Wow! So you know, figuring out some of these priorities can help in the printed realm.

211

00:34:49.060 --> 00:34:51.350

Mitch Fraas: I think this next section here,

212

00:34:51.650 --> 00:35:01.929

Mitch Fraas: you know, unpublished manuscript type, script material, also a bread and butter of past CLIR grants is what many of you probably will be proposing to digitize.

213

00:35:02.060 --> 00:35:07.899

Mitch Fraas: And you know this is often unique material already. So, you don't have that sort of

214

00:35:08.030 --> 00:35:18.350

Mitch Fraas: trade off. You know, it's just sort of self-evidently, the only access to this is going to be through your images often. So, it has that sort of case checked off.

215

00:35:18.680 --> 00:35:19.809

Mitch Fraas: But I think

216

00:35:19.870 --> 00:35:26.369

Mitch Fraas: in this selection here, I've seen internally, at least when we do our own digitization, people get tripped up around

217

00:35:26.630 --> 00:35:32.339

Mitch Fraas: trying to figure out viable, I'd say granularity or,

218

00:35:32.660 --> 00:35:37.069

Mitch Fraas: you know, what is the user really going to see here. To give an example,

219

00:35:37.190 --> 00:35:39.480

Mitch Fraas: you know, if you're proposing, digitizing

220

00:35:39.870 --> 00:35:53.879

Mitch Fraas: this kind of material from your city's parks department. You know, displaying to users a thousand digitized pages of financial minutes. Could be useful, but think of that against something that's like,

221

00:35:53.890 --> 00:36:04.679

Mitch Fraas: you know, described as a memo on the costs of changing signage at the former Robert E. Lee Park, 2 pages 1983. Like that level of granularity, obviously does a different amount of work than

222

00:36:04.820 --> 00:36:15.980

Mitch Fraas: dumping a thousand pages of financial minutes into one big PDF or something into one big bucket. Let's put it that way. Other times, you know, you could make a convincing case that

223

00:36:16.310 --> 00:36:18.969

Mitch Fraas: you know, the the volume is what

224

00:36:19.110 --> 00:36:30.710

Mitch Fraas: really is going to drive engagement and research, that, having all of the financial minutes of your particular city's finance department is an incredible you know, window on a

225

00:36:30.760 --> 00:36:33.289

Mitch Fraas: a particular research area or

226

00:36:33.310 --> 00:36:39.289

Mitch Fraas: kind of civic engagement, and even though it'll be available at sort of a broad level like,

227

00:36:39.820 --> 00:36:46.910

Mitch Fraas: you know, financial minutes 1978, you know, not hugely searchable. But despite that, you know that this would be providing,

228

00:36:47.070 --> 00:36:49.300

Mitch Fraas: you know, an incredible level of

229

00:36:49.540 --> 00:36:58.750

Mitch Fraas: availability for citizen journalists or others. Or you know, I could see arguments being made there, and oftentimes, you know, if you don't have the level of description at that granularity,

230

00:36:59.100 --> 00:37:13.269

Mitch Fraas: there's a lot of cases to be made to put things on online at the best you can do, really? But I think in selecting, if you have options, that's not always the case, but if you have options in these kinds of materials,

231

00:37:14.820 --> 00:37:19.599

Mitch Fraas: choosing based on your audience, what kinds of levels of granularity to to digitize as

232

00:37:19.850 --> 00:37:22.560

Mitch Fraas: because you're not going to have,

233

00:37:22.770 --> 00:37:23.810

Mitch Fraas: you know, it's

234

00:37:24.180 --> 00:37:36.690

Mitch Fraas: sort of as self intuitive as you know, people will sort of, going back to this, you know, sort of silly New York Times example. There's sort of an intuitive browsing of that. People understand the logic of newspapers to have a given year and a month and a page.

235

00:37:36.700 --> 00:37:56.039

Mitch Fraas: You know, it's browsable. A 1,000 pages of financial minutes often do not have any of that inherent browsability. Similarly, I think many of you probably will be proposing some amount of digitization, of photographs, slides, and negatives.

236

00:37:56.150 --> 00:38:12.440

Mitch Fraas: I should say I could have lumped photographs, newspapers, manuscripts altogether in some ways, because I think the format of the resulting digital file that at the end of the day will look pretty similar. I think you know, in terms of technical standards,

237

00:38:12.450 --> 00:38:25.499

Mitch Fraas: most people digitizing manuscripts or a photograph. are gonna output a a TIF file that is largeish and then maybe derivative files for people to view on their website

238

00:38:25.640 --> 00:38:32.780

Mitch Fraas: Could go either way. And so, in a way, the the original format, then, is just sort of transferred to one big image file

239

00:38:33.080 --> 00:38:39.330

Mitch Fraas: and I think these are very understandable. You know, you've got one photograph. You output a TIF

240

00:38:39.420 --> 00:38:49.799

Mitch Fraas: file. This is a little less complicated than some of the other media formats. In the case, of photographs, negative slides all that stuff, though I think there's a separate

241

00:38:50.120 --> 00:38:51.660

Mitch Fraas: kind of danger.

242

00:38:51.690 --> 00:38:59.689

Mitch Fraas: Which is sort of the shock and awe problem of large photographic collections. Thinking of some of ours here, and

243

00:39:00.130 --> 00:39:04.810

Mitch Fraas: it can seem super appealing. I'm thinking of some of our more obscure ones, you know.

244

00:39:04.830 --> 00:39:07.740

Mitch Fraas: The incredible,

245

00:39:07.800 --> 00:39:21.350

Mitch Fraas: you know, 17,000 photos of historic cat figurines does a certain kind of work. You just got so many photos accessible to the public for the first time. You want a cat figurine photo. You've got it. Truly comprehensive. All of that is going in the application.

246

00:39:21.420 --> 00:39:29.070

Mitch Fraas: But thinking of that like, how how does that help your argument and your narrative for why you're doing this project

247

00:39:29.360 --> 00:39:40.690

Mitch Fraas: in a way versus, say, digitizing a hundred photographs of individually described members of like the plaintiffs and lawyers in a landmark chemical pollution suit from 1987, where you've got

248

00:39:40.700 --> 00:39:44.269

Mitch Fraas: lots of good detail on the people in the photos. It's not a huge number. But

249

00:39:44.640 --> 00:39:50.300

Mitch Fraas: you know you're actually able to say something about them. Again, you know. I

250

00:39:50.350 --> 00:39:55.420

Mitch Fraas: see arguments for both the 17,000 cat photos and the 100 highly specific photos.

251

00:39:55.530 --> 00:40:06.410

Mitch Fraas: But with very different use cases, and different sort of perhaps community feedback on why you would or would not want that kind of comprehensiveness or not.

252

00:40:06.840 --> 00:40:17.899

Mitch Fraas: Also, something to think about with this kind of media, was just having a conversation their day, when our digitization studio outputted entirely negative TIF images.

253

00:40:18.260 --> 00:40:28.349

Mitch Fraas: Which is, you know, correct. You know they're taking pictures of negatives through a special machine. It is very hard for users to engage with negative images. It is very easy to flip them to positive

254

00:40:28.370 --> 00:40:30.210

Mitch Fraas: image.

255

00:40:30.220 --> 00:40:42.230

Mitch Fraas: And folks and photographers will say, it's subjective, you know, when you're making positive images, there's some lots of choices involved. Think that's fair. But in terms of researcher access, I feel like

256

00:40:42.620 --> 00:40:52.320

Mitch Fraas: browsing positive images is so much easier than negative images. Likewise, thinking with your community about if you've got contact sheets

257

00:40:52.610 --> 00:40:56.329

Mitch Fraas: versus individual photos, you know. Are there arguments to be made for

258

00:40:56.440 --> 00:41:04.430

Mitch Fraas: getting to do more digitization of contact sheets? Sort of raw by the end versus the sort of high resolution on that particular.

259

00:41:04.630 --> 00:41:08.700

Mitch Fraas: Again, depends on what you're trying to depict and what kind of

260

00:41:08.850 --> 00:41:14.309

Mitch Fraas: impact you're looking to make for whatever community you're digitizing for.

261

00:41:14.380 --> 00:41:30.909

Mitch Fraas: In our case, recently, we decided to do contact sheets over individual photographs because we imagined the main use case was browsing and seeing a broad range of somewhat similar photographers' photos rather than scrolling through individual 10,000 photos. I think I

262

00:41:31.330 --> 00:41:33.639

Mitch Fraas: can make those arguments.

263

00:41:33.940 --> 00:41:47.240

Mitch Fraas: Tapes, cassettes, discs, media obviously output different file formats, and in many cases these can be quite large. I'll talk about that in a little bit. One thing, I should say here, I know there's gonna be

264

00:41:47.390 --> 00:42:06.380

Mitch Fraas: other sessions dealing with rights, but with all of these categories, especially audio, though, and video, thinking through rights implications. Obviously, there's a session on this, because you could talk for hours. It is important. I'd also add that captioning and accessibility here for audio and video

265

00:42:06.980 --> 00:42:08.300

Mitch Fraas: is something that

266

00:42:08.320 --> 00:42:14.779

Mitch Fraas: you really would want to think about why you're taking a particular stance. If you're saying, you know, we're gonna send these out to a vendor

267

00:42:15.230 --> 00:42:33.340

Mitch Fraas: and receive audio files back, and where you are going to put them online raw, and here's why. For resource reasons for main audience reasons, these are musical and abstract, and don't have words. You know what? What are the reasons? You're making choices around whether to provide captioning or not.

268

00:42:33.520 --> 00:42:40.559

Mitch Fraas: Also thinking through risks on recoverability. This is a photo of a disc that we sent off

269

00:42:40.740 --> 00:42:50.470

Mitch Fraas: and paid a lot of money to find out that it is irretrievable. It's delaminated and actually looks really good in this photo. It does not look good in real life.

270

00:42:50.480 --> 00:42:58.549

Mitch Fraas: And I think you know, before proposing to digitize, you know, a hundred oral history discs that have never been listened to

271

00:42:58.660 --> 00:43:12.210

Mitch Fraas: and finding out later that they're unusable, you know, doing a quick sample with a vendor, someone knowledgeable, to make sure that at least some of these are viable is really important

272

00:43:12.270 --> 00:43:25.039

Mitch Fraas: before making a major grant commitment to something. And also to understand some of the complications. There, sampling can be a fantastic way to feel better about embarking on one of these large media projects.

273

00:43:25.240 --> 00:43:32.920

Mitch Fraas: I think before a break I'll make a note on granularity. I've sort of touched on this.

274

00:43:33.170 --> 00:43:37.169

Mitch Fraas: I see this making a big difference in

275

00:43:37.390 --> 00:43:44.200

Mitch Fraas: the way different grants make different arguments for why you're selecting particular material

276

00:43:44.290 --> 00:44:00.900

Mitch Fraas: and you want to think about how your argument or your story for why digitizing this versus that, changes between, say, you know, digitizing all of boxes 1-20 of manuscript material from the papers of the Society for Owl Studies, I don't know,

277

00:44:00.910 --> 00:44:20.949

Mitch Fraas: versus digitizing a selection of single elaborate amuletic manuscripts like the one on the left there owned by people in owl studies. I don't know. They say different things, they serve different audiences, and they will have very different engagement. That individual amulet there, it's actually about this big.

278

00:44:21.090 --> 00:44:29.629

Mitch Fraas: you know, it's gonna see a lot of individual use. Probably used outside the context. This is from the manuscripts of the Muslim World Project.

279

00:44:29.700 --> 00:44:49.689

Mitch Fraas: Based on our download, like probably most used outside of that context. So someone is not looking at all of our manuscripts. They're going and looking at this one manuscript in particular. Whereas with those boxes of owl studies, which is made up, I don't know why I just chose that, you can imagine that most researchers or

280

00:44:49.780 --> 00:45:00.460

Mitch Fraas: interested parties working on that are gonna look at more than one thing in there, you know they're in it for the owls. They're gonna start browsing through that collection and really using all different pieces of it, you know.

281

00:45:00.770 --> 00:45:03.790

Mitch Fraas: And as you make decisions about, you know,

282

00:45:04.270 --> 00:45:25.219

Mitch Fraas: how much of a coherent collection you're digitizing versus how much of individual things, I think, taking into account that that use case is super important, and thinking through what you gain and lose in terms of access, engagement, community responsibility, as you choose, discrete objects versus comprehensive holes. So

283

00:45:25.370 --> 00:45:35.419

Mitch Fraas: with that we will break briefly to catch your breath, and then I'll finish off with some thoughts about estimation and representative samples.

284

00:45:39.290 --> 00:45:43.710

Alyson Pope: We'll take a quick 5 min break, and we'll see you back at 2:50.

285

00:46:11.890 --> 00:46:13.529

Sharon Burney: Whenever you're ready, Mitch.

286

00:46:14.280 --> 00:46:15.300

Mitch Fraas: Oh, thanks so much.

287

00:46:15.620 --> 00:46:19.830

Mitch Fraas: Right. I hope everyone is

288

00:46:20.050 --> 00:46:21.820

Mitch Fraas: rested.

289

00:46:21.890 --> 00:46:31.520

Mitch Fraas: And thanks for joining on an afternoon. I know a lot of us are in Zoom a lot, so it's always humbling when people choose to to hear you on Zoom.

290

00:46:32.120 --> 00:46:32.870

Mitch Fraas:

291

00:46:34.250 --> 00:46:37.710

Mitch Fraas: I'm transitioning here to this sort of really nitty, gritty stuff

292

00:46:38.260 --> 00:46:45.400

Mitch Fraas: through the lens of, I think we're gonna talk about some of these like CLIR templates in a second, but

293

00:46:45.590 --> 00:46:51.380

Mitch Fraas: one of the things that I found helpful to sort materials in my mind at least of how

294

00:46:51.460 --> 00:46:54.150

Mitch Fraas: how we wanted to make an argument around them.

295

00:46:54.240 --> 00:47:02.129

Mitch Fraas: And a big thing in my mind is sort of authorship and origin of materials and making a case around them.

296

00:47:02.300 --> 00:47:03.210

Mitch Fraas: A.

297

00:47:03.510 --> 00:47:09.240

Mitch Fraas: Here on the left and right are items from the Mary Anderson papers here.

298

00:47:09.590 --> 00:47:10.800

Mitch Fraas: And

299

00:47:10.870 --> 00:47:13.190

Mitch Fraas: they're from the same collection.

300

00:47:13.380 --> 00:47:32.640

Mitch Fraas: They're not in the same box, but they're in the same milieu. They're two very fascinating documents. One is a photo of Anderson with LBJ and Lady Bird Johnson there at the White House taken by the White House photographer and the one on the left is a page from Anderson's Journal when she was in South Korea.

301

00:47:33.190 --> 00:47:37.630

Mitch Fraas: These are both important historic documents.

302

00:47:37.950 --> 00:47:53.649

Mitch Fraas: They have very different origins. And I think that helps me make some decisions. On the left is something that she created herself. She wrote that in her own hand, and she kept it with her journals and diaries over her whole life.

303

00:47:53.780 --> 00:48:06.589

Mitch Fraas: The White House photograph we know less about. It's not a photograph she took. Obviously she's in it. Taken by a professional photographer. It is in her collection with other materials and photographs

304

00:48:06.840 --> 00:48:14.440

Mitch Fraas: of her. She almost certainly had this. In fact, we sort of know it was in her possession. But we know less about,

305

00:48:14.510 --> 00:48:16.340

Mitch Fraas: you know,

306

00:48:16.790 --> 00:48:36.220

Mitch Fraas: where it fits into her thinking. So I, for a lot of reasons, you know, you'll have the whole right session. By the way, White House photographers are really jealous about copyright, you think they'd be products of the Federal government. It's complicated. But in any event, when you have these different kinds of authorships, we decided in our

307

00:48:36.300 --> 00:48:58.479

Mitch Fraas: CLIR grant, around Anderson, to really focus on things she herself created. We've separately digitized photographs of her. It's a really different project, but also very worthwhile. But in our particular one, we wanted to make a case around doing her own writings, her own recordings of herself that she did at home, and it helped sort of focus in our mind like, you know, what

308

00:48:58.770 --> 00:49:00.760

Mitch Fraas: what we wanted to do

309

00:49:00.790 --> 00:49:03.960

Mitch Fraas: and why, it would matter that this isn't just

310

00:49:03.980 --> 00:49:27.070

Mitch Fraas: anything in the world about her or of her, it also had more or less a single authorship, and since Penn has rights to Anderson's intellectual property it made it a lot easier, also. So, two nice things there. So I think, if it helps ever to sort material, at least intellectually, by some of these criteria. I found it

311

00:49:27.350 --> 00:49:33.390

Mitch Fraas: helpful in selecting. And now listing.

312

00:49:33.870 --> 00:49:35.110

Mitch Fraas: One of the

313

00:49:35.300 --> 00:49:56.909

Mitch Fraas: big, I think, disconnects often between cultural heritage institutions and researchers or community members is, I think, we rightly, I'm gonna assert that, we have a kind of obsession with format, and you'll see that in the CLIR application as well. You're asked to list, you know, what format are these things? You know, what kind of files are you producing? I was even thinking about our Mary Anderson site. This is the old one.

314

00:49:56.910 --> 00:50:18.429

Mitch Fraas: Broke everything up by format. You've got your scrapbooks, your notebooks and diaries, your programs, your photographs, recordings, your interview transcripts. They're separate, right. They're different things. They had to go to different digitization studios, some of these different file types, for the recordings. There's different kinds of descriptive metadata for a photograph versus a scrapbook. This makes a lot of library sense.

315

00:50:18.760 --> 00:50:26.219

Mitch Fraas: I think I've been convinced that it does not always make sense for researchers, members of the community, people at large

316

00:50:26.430 --> 00:50:33.670

Mitch Fraas: who might be interested in, you know, Marion Anderson and religious education through song.

317

00:50:33.770 --> 00:50:47.009

Mitch Fraas: That has nothing to do with format. You know, it could be in her scrapbooks, could be in a recording, could be in a program, you know, it's everywhere. And I think this is where you know hard decisions about how you organize your work,

318

00:50:47.430 --> 00:50:50.440

Mitch Fraas: you know, are sometimes trade offs that,

319

00:50:50.510 --> 00:50:57.930

Mitch Fraas: you know, having to be granular about media type and format is important for a lot of technical reasons, but how can you mitigate that

320

00:50:58.010 --> 00:50:59.330

Mitch Fraas: risk of sort of

321

00:50:59.900 --> 00:51:05.520

Mitch Fraas: separating out everything into its own bucket. I don't have any great answers except to

322

00:51:05.790 --> 00:51:11.009

Mitch Fraas: provide enough context and explanation and user engagement that people know

323

00:51:11.130 --> 00:51:17.360

Mitch Fraas: that they can peruse or engage with a collection beyond just particular format

324

00:51:17.630 --> 00:51:24.450

Mitch Fraas: niches. Sort of a soap box there. But thinking ahead to the nitty gritty,

325

00:51:24.530 --> 00:51:37.390

Mitch Fraas: this is a scary slide. I'm sorry. This is the CLIR template. I remember filling this out. It's actually not that many fields, but I found it really useful to look at before

326

00:51:38.210 --> 00:51:44.159

Mitch Fraas: attempting my rough drafts of listings. Because even knowing I would have to answer these questions,

327

00:51:44.480 --> 00:52:04.130

Mitch Fraas: affected what I was sort of selecting. Knowing that you're gonna have to say something about the rights holder, knowing that you're gonna have to say material types, current storage condition, definitely influenced my thinking of like, oh, wait, you know, actually, let's not do the Mary Anderson cassettes that are in that box marked "moldy." Not a true story.

328

00:52:04.790 --> 00:52:07.870

Mitch Fraas: No moldy cassettes, but it does happen with other collections,

329

00:52:08.070 --> 00:52:23.379

Mitch Fraas: you know. So I, even though it's sort of dry, the template, I think, is super interesting to begin to consider, and thinking of integrity now rather than later matters a lot. And I'll give you an example of

330

00:52:23.420 --> 00:52:29.980

Mitch Fraas: why. So, this is too small, but this is this is the the actual template I filled out for

331

00:52:29.990 --> 00:52:32.300

Mitch Fraas: our contribution to the Health of a New Nation.

332

00:52:32.500 --> 00:52:44.660

Mitch Fraas: It's probably not very readable. Wanted to point out one thing that should stand out like a sore thumb and kind of did to me then, and I would do things differently now. The very last line there, you see on overall collection size, like

333

00:52:44.780 --> 00:52:51.960

Mitch Fraas: 3 or 4 columns over. You know, we've got things like really well granularly described, like John Briscoe, medical notes: 200 images.

334

00:52:52.200 --> 00:52:56.960

Mitch Fraas: Forever medical notes: 200 images. Lecture notes of physic: 2,500. You know,

335

00:52:57.300 --> 00:53:10.249

Mitch Fraas: You've got individual description. Great. The very last line. What is that? Overall collection? Says 22,300 images. But what are those Pen Medical School theses 1805 to 1828. That's that's not. That's not like the other stuff.

336

00:53:10.360 --> 00:53:16.969

Mitch Fraas: In this case, this is part of a very large project. We made a very convincing case in our narrative,

337

00:53:17.070 --> 00:53:25.209

Mitch Fraas: but in hindsight, not getting better control of this last category. That giant we hold.

338

00:53:25.320 --> 00:53:34.660

Mitch Fraas: Penn was the first medical school in the country. And until the 1860s/1870s. All MD's had to write a thesis or dissertation.

339

00:53:34.680 --> 00:53:40.800

Mitch Fraas: And we decided to digitize all of them up to 1828, in many volumes, all handwritten.

340

00:53:40.910 --> 00:53:42.350

Mitch Fraas: And

341

00:53:42.980 --> 00:53:58.909

Mitch Fraas: it should have been a wake up call, seeing that number compared to these that I needed to do a little bit more work. There's me with the medical dissertations and those black bindings there. And what I did was I

342

00:53:58.940 --> 00:54:05.930

Mitch Fraas: went through, and there were 44 volumes of those pre 1828 medical theses.

343

00:54:05.940 --> 00:54:15.719

Mitch Fraas: 44 volumes great. I took a look at a couple of them and said, you know, these are about 500 pages each. I'm not gonna look at all 44.

344

00:54:16.140 --> 00:54:21.740

Mitch Fraas: I'm just gonna extrapolate and multiply 44 by 500. And I got that 22,000

345

00:54:21.750 --> 00:54:25.030

Mitch Fraas: number. I don't know where the 300 is. But,

346

00:54:25.350 --> 00:54:41.299

Mitch Fraas: you know, I thought that was perfectly reasonable. I think this is a perfectly fair way of estimating especially smaller size, you know, if it was a run of 10 volumes, or 15 volumes, or 20 volumes doing this kind of brute force, sampling and multiplication, I think, is totally justifiable. I'm gonna talk in a second about margins of error, and why

347

00:54:41.590 --> 00:54:44.349

Mitch Fraas: things may or may not be more or less justifiable.

348

00:54:44.480 --> 00:54:58.330

Mitch Fraas: In this case, I also felt good because I knew, even if I was off by some that we had in house digitization capacity. Even if we didn't have funding, if I screwed up by a little. So I felt

349

00:54:58.500 --> 00:55:08.689

Mitch Fraas: better about this than spending a lot of time individually counting leaves of like 44 volumes of medical theses. Say, another key consideration here

350

00:55:08.820 --> 00:55:22.309

Mitch Fraas: that makes an enormous difference... I was like, you know, I don't know if I should even mention this. Maybe it's too obvious, but different places, and I think both are justifiable, have attitudes to blank pages.

351

00:55:23.010 --> 00:55:28.389

Mitch Fraas: The medical theses are written entirely on single sides of paper.

352

00:55:29.020 --> 00:55:36.900

Mitch Fraas: With maybe a couple exceptions. But as a policy here, our library is a very particular

353

00:55:37.160 --> 00:55:40.360

Mitch Fraas: thing. One of our academic focuses is the history of the book.

354

00:55:40.420 --> 00:55:56.920

Mitch Fraas: We've done digitization since the late nineties, and our policy has always been, we do not change the physical object. So it means taking pictures of the entire physical object, regardless of whether there is text. So we take pictures of the binding, the spine, the front, the back, and every single blank page.

355

00:55:57.170 --> 00:55:59.929

Mitch Fraas: So that doubles, if you think about

356

00:56:00.160 --> 00:56:21.859

Mitch Fraas: scale here. Doubles the number of pages, doubles the cost in that way, which, internally for us, was justifiable. Because we have this, we have a universal policy. It applies to every single thing we do. So, scholars know that anything we digitize meets these standards so they don't have to guess like was the back of this not blank? Did they forget to shoot it? I think, in other instances

357

00:56:21.980 --> 00:56:32.650

Mitch Fraas: and for other purposes, at other institutions, not shooting blank pages can be a fantastic idea and a way to not boost your costs to the stratosphere.

358

00:56:32.740 --> 00:56:43.500

Mitch Fraas: Same thing with other kinds of questions around blanks. So that'd be, you know, blank sides of audio cassettes. Whatever it is, I think that's considered,

359

00:56:43.540 --> 00:56:45.940

Mitch Fraas: justifiable thing that you can

360

00:56:46.000 --> 00:56:59.480

Mitch Fraas: make sort of a call on your own. In this case, we also didn't know, to be fair, since no one had looked at all 44 volumes. We were not confident there wasn't stuff written on the backs of pages. Happened occasionally, but

361

00:56:59.700 --> 00:57:01.469

Mitch Fraas: I think it was not

362

00:57:01.860 --> 00:57:05.230

Mitch Fraas: very common. So what happened here?

363

00:57:06.220 --> 00:57:07.830

Mitch Fraas: And what about

364

00:57:07.900 --> 00:57:10.789

Mitch Fraas: errors. So whoops.

365

00:57:11.200 --> 00:57:12.849

Mitch Fraas: I estimated 22,000

366

00:57:13.100 --> 00:57:20.419

Mitch Fraas: distribution page of this, the final report to CLIR recorded that we shot 49,863 images

367

00:57:20.980 --> 00:57:22.839

Mitch Fraas: from those dissertations, which is

368

00:57:23.440 --> 00:57:30.939

Mitch Fraas: more than double. Couple of reasons turned out to be more volumes on catalogs that I didn't know about, so that's fair.

369

00:57:30.970 --> 00:57:41.609

Mitch Fraas: But the biggest difference was that the average of the volumes clocked at 831 pages by the time we got through all of them. So the volumes I had looked at

370

00:57:41.700 --> 00:57:44.469

Mitch Fraas: we're short of representative.

371

00:57:44.620 --> 00:57:49.529

Mitch Fraas: The good news on this is twofold, I think. One,

372

00:57:49.790 --> 00:58:02.590

Mitch Fraas: I mean, in our particular case we had capacity in the grants as different institutions that were partners did more or less, and the pandemic also played into this. So we were able to do all of it. For various reasons. Thanks, CLIR.

373

00:58:03.030 --> 00:58:06.829

Mitch Fraas: The other reason is that in this case,

374

00:58:07.300 --> 00:58:15.240

Mitch Fraas: we knew ahead of time, I knew ahead of time, that the margin of error here would be okay, even if way off, like this is not a good margin of error. I'm not not really proud of this.

375

00:58:15.300 --> 00:58:24.279

Mitch Fraas: But I knew, given our digitization capacity here, and that we were doing it in-house, and we were internally charging, that even a factor of 2

376

00:58:24.510 --> 00:58:28.859

Mitch Fraas: would be okay. I knew that going into this that we could cover

377

00:58:29.150 --> 00:58:35.200

Mitch Fraas: probably, and have the capacity, probably, for a pretty big margin of error. And so,

378

00:58:35.340 --> 00:58:38.039

Mitch Fraas: in retrospect, my sampling was not good,

379

00:58:38.430 --> 00:58:41.110

Mitch Fraas: but I still think defensible

380

00:58:41.180 --> 00:58:48.230

Mitch Fraas: in terms of knowing that it would not truly break the bank. I think if if you've got a much tighter, if you're sending stuff out to a vendor,

381

00:58:48.570 --> 00:59:03.780

Mitch Fraas: having much tighter sampling, and knowing that your margin of error cannot be this big, is really important. I have the audio cassette example, there, of how much you know estimations can change, based on the number of minutes on the audio tape. So

382

00:59:03.950 --> 00:59:11.689

Mitch Fraas: knowing kind of where you stand with this, if you're gonna have any wiggle room of more or less costs, I think, makes a big difference.

383

00:59:11.820 --> 00:59:15.299

Mitch Fraas: And a contingency plan if you start

384

00:59:15.330 --> 00:59:16.790

Mitch Fraas: to run into

385

00:59:16.880 --> 00:59:26.459

Mitch Fraas: big discrepancies in amounts of material halfway through. In our case, we thought about it. Once we hit the 22,000,

386

00:59:26.710 --> 00:59:32.549

Mitch Fraas: we had to decide. Do we keep going? And for various reasons, including the pandemic internal capacity,

387

00:59:32.620 --> 00:59:42.799

Mitch Fraas: changes in funding models, we decided to go for it, but our contingency plan would have been to choose a different cut off year because the volumes were arranged in a certain way.

388

00:59:42.950 --> 00:59:52.169

Mitch Fraas: So it would not have been a disaster to say we digitized all of the volumes up to 1824 instead of 1828. So you know, that's intellectually defensible. It's not great

389

00:59:52.180 --> 01:00:01.060

Mitch Fraas: from a consistency standpoint of doing what you said you would do, but it at least makes a lot of sense. Doesn't stand out as clearly wrong,

390

01:00:01.140 --> 01:00:05.360

Mitch Fraas: In house of the estimation. I think if I had a super tight

391

01:00:06.130 --> 01:00:10.880

Mitch Fraas: budget capacity and digitization capacity, I would have sampled

392

01:00:11.110 --> 01:00:18.169

Mitch Fraas: 3 or 4 times as many, so maybe sampling as many as a you know, a quarter of the items, just to begin to realize that the page counts were much,

393

01:00:19.200 --> 01:00:29.169

Mitch Fraas: I'm trying to think of that 500 versus 800, like a pretty significant range above what I thought they were. Not a proud story. I hope this doesn't go in the the CLIR report about

394

01:00:29.260 --> 01:00:38.279

Mitch Fraas: delinquent former applications. But it was very instructive to everyone involved, including our Budget office.

395

01:00:38.360 --> 01:00:42.519

Mitch Fraas: But also an instructive point here, that

396

01:00:42.640 --> 01:00:54.759

Mitch Fraas: at least at Penn, when we do audio cassette digitization, we send it out to a vendor, and we have to pay actual money versus internal capacity, a photo there of one of our great camera operators who are on salary, so

397

01:00:55.000 --> 01:01:01.180

Mitch Fraas: we know we have a margin of error there, because we're not, we're sort of charging ourselves. So important institutional considerations.

398

01:01:01.200 --> 01:01:04.390

Mitch Fraas: I'm running low in time. So I'm gonna

399

01:01:04.690 --> 01:01:06.099

Mitch Fraas: skip ahead to

400

01:01:06.300 --> 01:01:14.379

Mitch Fraas: the representative samples part. This is always a fun part of the application for me. Once you've done the work of doing some estimation and

401

01:01:14.460 --> 01:01:21.369

Mitch Fraas: selection from what you want to do, and making your argument, getting to show that in a couple really key

402

01:01:21.520 --> 01:01:23.360

Mitch Fraas: ways.

403

01:01:23.830 --> 01:01:31.309

Mitch Fraas: You know, this is this is where you you don't have to show why all a thousand pages of those financial reports are

404

01:01:31.410 --> 01:01:37.849

Mitch Fraas: really gonna enhance access for communities, but where you pick the page of the financial report that talks about

405

01:01:37.940 --> 01:01:49.749

Mitch Fraas: changing representation on the city council, and you know, battles over funding free clinics in the city that you've got like a really visual and compelling

406

01:01:49.770 --> 01:01:57.019

Mitch Fraas: example. Here's a couple from our Manuscripts of the Muslim World grant. I would say,

407

01:01:57.070 --> 01:02:07.100

Mitch Fraas: picking things that have real representative value. I clearly went here for pretty. That's not always possible, but with very pithy

408

01:02:07.420 --> 01:02:20.870

Mitch Fraas: sort of comments or justifications of why these are emblematic of the larger mass of material that are going to be digitized. It can also sort of help focus, if there's multiple partners,

409

01:02:21.070 --> 01:02:31.380

Mitch Fraas: focus on what really are the highlights and strengths of your different collections, if you're all collaborating together. And it's sort of a nice challenge to say, you know, like, if we were to pick the highlights,

410

01:02:32.050 --> 01:02:34.130

Mitch Fraas: you know, what would it be? And why?

411

01:02:34.700 --> 01:02:42.099

Mitch Fraas: You know, in this case, you know, we've got both an example that shows art, historical interest. And then another example, that shows textual interest.

412

01:02:42.120 --> 01:02:52.390

Mitch Fraas: And have sort of different audiences of scholars. I think that was a nice mix for the Health of a New Nation Grant of medical theses and stuff like that.

413

01:02:52.450 --> 01:03:02.470

Mitch Fraas: I think I picked some Benjamin Rush notebooks talking directly about, I think I maybe I picked the ones dealing with yellow fever, I can't exactly remember, but that had nice illustrations of

414

01:03:02.860 --> 01:03:12.990

Mitch Fraas: sort of medical apparatus, and that showed the real engagement in these student notes and in professor's lecture notes with the material and I think,

415

01:03:13.370 --> 01:03:17.469

Mitch Fraas: I can't remember whether I we did these representative examples before or after

416

01:03:17.480 --> 01:03:22.609

Mitch Fraas: writing the main narrative, but it certainly helped me at least thinking about our own collection of like

417

01:03:22.670 --> 01:03:24.609

Mitch Fraas: what do I want to say about

418

01:03:24.960 --> 01:03:30.220

Mitch Fraas: what it is we have, and why we're doing it. So I think that

419

01:03:30.760 --> 01:03:36.829

Mitch Fraas: is a sort of helpful framework to go with. So you know, I think,

420

01:03:36.910 --> 01:03:46.630

Mitch Fraas: if you were reading these applications, and you look at the representative samples, and you look at the narrative, they should sort of go together, and you can see why it meets the amplifying voices

421

01:03:46.700 --> 01:03:55.129

Mitch Fraas: rubric. So I'd want to make sure the represented samples are not totally out of left field, you know, or like the complete outliers and what you're gonna digitize,

422

01:03:55.170 --> 01:04:03.569

Mitch Fraas: But really speak to the main point of your digitization project, which is probably self-evident, but I think it can be tempting,

423

01:04:03.930 --> 01:04:13.129

Mitch Fraas: in this case, for instance, it would have been tempting to only choose the prettiest of manuscripts for everything, even if those manuscripts were not

424

01:04:13.470 --> 01:04:26.789

Mitch Fraas: that intellectually or textually interesting, and we made some decisions not to do that. We included some, some snippets of text manuscripts to show that this wasn't just a project about digitizing gold leaf. And I think you know, those decisions do matter.

425

01:04:28.310 --> 01:04:36.140

Mitch Fraas: So that's my "end slide," it is one of our camera operators working on medical dissertations, during the pandemic.

426

01:04:36.400 --> 01:04:44.990

Mitch Fraas: And with that, I realize, I'm just about at 3:10. We could talk for a long time about a lot of these things.

427

01:04:45.070 --> 01:04:57.740

Mitch Fraas: You know, I sort of talked briefly about file formats. There's a lot of different opinions out there, and professionals in the world of digitization, who have strong feelings about file formats

428

01:04:57.810 --> 01:05:01.190

Mitch Fraas: and types, but I think getting clarity for your

429

01:05:01.550 --> 01:05:06.759

Mitch Fraas: capacity, and whether you're presenting the sort of you know, full

430

01:05:06.890 --> 01:05:15.759

Mitch Fraas: archival quality TIFS and MOV files, or whatever to the public versus access derivatives. You know, these are all important questions that

431

01:05:16.020 --> 01:05:20.770

Mitch Fraas: seem more on the technical side, but also have implications for your

432

01:05:20.960 --> 01:05:24.759

Mitch Fraas: selection process. Same with text,

433

01:05:24.960 --> 01:05:41.290

Mitch Fraas: you know, having searchable PDFs for modern material is what nearly everyone in the research world, and I think many people in the public expect, and that's often not always what libraries deliver. So thinking carefully about your audience there, I think, is another consideration.

434

01:05:41.490 --> 01:05:49.270

Mitch Fraas: Obviously, there's another session on rights issues. But that's a that's another fun one, and can really change the

435

01:05:49.380 --> 01:05:56.750

Mitch Fraas: the way you select manuscripts, or select items, which is one of the reasons I highlighted authorship there with the Anderson papers.

436

01:05:56.800 --> 01:05:59.860

Mitch Fraas: And rather than and keep droning on, I think I will

437

01:05:59.910 --> 01:06:13.829

Mitch Fraas: wrap it up there and say that I'm looking forward to hearing from you, because everyone has different institutions, different experiences, different kinds of material, and it'd be great to get into, you know, some questions and examples and things that are relevant to you.

438

01:06:17.860 --> 01:06:18.650

Mitch Fraas: Thanks.

439

01:06:22.680 --> 01:06:37.750

Alyson Pope: Yeah. So we are prepared to take some questions. Mitch is gonna field questions specific to today's focus on assessing collections for digitization. A reminder, please use the Q&A Box for questions rather than the chat.

440

01:06:40.000 --> 01:06:43.449

Alyson Pope: Sharon, would you like to read him one first?

441

01:06:43.550 --> 01:07:05.960

Sharon Burney: I sure will. The collection we have proposed has been surveyed and inventoried at the box level. We're doing a deeper dive, a deeper survey for this grant application. We also hope to receive funding, to hire someone to further help with processing and selecting for digitization at the item level. Is this a good approach for these types of collections?

442

01:07:06.220 --> 01:07:11.070

Mitch Fraas: That's a tough question. I will,

443

01:07:11.680 --> 01:07:16.529

Mitch Fraas: the part that I can't really answer that I kind of want to defer to the CLIR team after I finish is the

444

01:07:16.760 --> 01:07:17.670

Mitch Fraas: the sort of

445

01:07:17.910 --> 01:07:25.859

Mitch Fraas: programmatic or financial implications of the cost of metadata and how that fits into your overall plan and budget and all that because that's a big thing.

446

01:07:26.200 --> 01:07:27.099

Mitch Fraas: I think,

447

01:07:28.330 --> 01:07:37.839

Mitch Fraas: this is gonna get me in trouble with some archivists, for a lot of people, a lot of use cases, a lot of researchers interacting with objects digitally

448

01:07:38.000 --> 01:07:41.119

Mitch Fraas: often expect item level interaction.

449

01:07:41.440 --> 01:07:42.240

Mitch Fraas: I'm

450

01:07:44.240 --> 01:08:01.280

Mitch Fraas: I cannot say that that's always a good thing, but I think we cannot say that it is not common, and that is not really what people expect, even if we sort of wish people followed archival practice and we're okay with box level description. There are a lot of great archival reasons why item level description is like, not tenable. I mean, it's

451

01:08:01.350 --> 01:08:11.089

Mitch Fraas: so many reasons. But I think, for if you're trying to make an argument about engaging specific audiences, and you can say

452

01:08:11.150 --> 01:08:17.660

Mitch Fraas: at the box level, we don't think we can do that. And here's why. Like we think no one's going to look through

453

01:08:18.090 --> 01:08:27.280

Mitch Fraas: the, it's described as financial minutes, it's got 1,000 images. It's a thousand page PDF, and and we just we think that is just not going to do it

454

01:08:27.420 --> 01:08:35.329

Mitch Fraas: for the purposes of this grant. I think that's justifiable. I think in other other cases you could make a very good case that

455

01:08:35.450 --> 01:08:38.960

Mitch Fraas: you know the box

456

01:08:39.740 --> 01:08:41.590

Mitch Fraas: labeled with a particular, you know.

457

01:08:41.710 --> 01:08:46.070

Mitch Fraas: technical report on, you know,

458

01:08:46.270 --> 01:09:01.949

Mitch Fraas: Athletic achievements or whatever. So it's just like a thousand page report about a very specific subject. Someone's gonna go to that and flip through it and skim and read, and they sort of know what it's about. It's a concrete thing. Then, doing some sort of like, item level, like, Oh, yeah. Volume one is about this volume 2 is about that.

459

01:09:02.240 --> 01:09:03.850

Mitch Fraas: Kind of silly, but

460

01:09:04.319 --> 01:09:09.789

Mitch Fraas: the budget implications and the like grant application things are significant. So I'll let the CLIR folks

461

01:09:09.830 --> 01:09:13.469

Mitch Fraas: talk about the sort of balance of paying for metadata digitization.

462

01:09:14.270 --> 01:09:25.020

Sharon Burney: Yeah, I agree. There's a couple of... it's hard for us to determine this one, because we don't understand how large your collection is, and I don't understand,

463

01:09:25.240 --> 01:09:33.749

Sharon Burney: are you talking about finding funding for that now? Or that funding coming from the actual grant if you get awarded?

464

01:09:33.770 --> 01:09:43.560

Sharon Burney: There is some monies allowed for this process of determining what's in assessing the collections, determining, but you're gonna have to

465

01:09:44.490 --> 01:09:59.789

Sharon Burney: come up with that amount in your application before you get awarded. So you're making an assessment based on your understanding. We understand just as Mitch clarified in his presentation that

466

01:10:00.330 --> 01:10:16.599

Sharon Burney: sometimes there'll be a surprise. Sometimes you'll end up funding a lot more, and sometimes you may have duplicates in there that you don't even know is there? And that's expected. We want you to provide the reviewers with an estimation of what you believe to be in these collections

467

01:10:16.600 --> 01:10:30.349

Sharon Burney: and that your project, plan and design and budget all come together cohesively to make an estimation of how much your digitization costs are going to be. So,

468

01:10:30.490 --> 01:10:39.069

Sharon Burney: that's where I'm confused. Is this outside funding coming from the grant? Are you writing that into your budget proposal, or is this outside of that.

469

01:10:39.180 --> 01:10:52.719

Sharon Burney: But you can always email us more specific questions, with your specific situations, to the program box at hiddencollections@clir.org, and we can help you with a more nuanced answer.

470

01:10:54.670 --> 01:11:01.840

Mitch Fraas: The one thing I'd add quickly, the Muslim world grant we did, we were able to make the case for using some of that money

471

01:11:01.920 --> 01:11:06.950

Mitch Fraas: for metadata work, because we felt that Arabic script, metadata work...

472

01:11:07.350 --> 01:11:11.779

Mitch Fraas: Like we could not do the project without describing these things.

473

01:11:12.030 --> 01:11:27.400

Mitch Fraas: But the way we sort of justified that was saying, the institutions would pick up and greatly reduce their internal digitization costs to help offset this. To say, you know, we know that this is a significant expense, so instead of charging \$2 and whatever 47 cents per page.

474

01:11:27.500 --> 01:11:36.370

Mitch Fraas: you know, the universities are gonna say, we're charging a dollar and 43 cents, whatever it is, so that we had to make some big trade offs to make that work possible within the limits. But,

475

01:11:36.390 --> 01:11:39.050

Mitch Fraas: I assume that's what everyone you know, it's it's trade off.

476

01:11:41.880 --> 01:11:57.640

Alyson Pope: We have one more. What advice do you have for those of us who are digitizing materials on a specific subject from across multiple collections. What should we consider when selecting materials for digitization in this circumstance? Yeah, that's a good

477

01:11:58.310 --> 01:11:59.300

Mitch Fraas: a

478

01:11:59.570 --> 01:12:10.669

Mitch Fraas: the medical hub grant, actually, was sort of like that. You know, often people we're picking material relating to Philadelphia medical education, and it was like, you know, one folder in a

479

01:12:10.930 --> 01:12:17.170

Mitch Fraas: collection of someone's papers who was in Philadelphia for a month and wrote about medical education.

480

01:12:17.360 --> 01:12:18.720

Mitch Fraas: I think the

481

01:12:19.730 --> 01:12:21.539

Mitch Fraas: the big questions there are like.

482

01:12:22.330 --> 01:12:28.389

Mitch Fraas: how are you doing the selection? Like, is there, in some cases, it's really obvious.

483

01:12:28.880 --> 01:12:34.239

Mitch Fraas: you know, ways of selecting of like on this highly specific topic. We only have

484

01:12:34.440 --> 01:12:44.289

Mitch Fraas: 23 things in whatever about it. So we're gonna choose those 23. It's not like 10 are in, 3 are out. Whatever. But making sure that if you are making

485

01:12:44.670 --> 01:12:57.310

Mitch Fraas: tough choices like on even things that seem specific, like I'm gonna work on, we're gonna digitize everything related to Vietnamese Americans in South Philadelphia. Seems highly specific. There is a ton of stuff.

486

01:12:57.380 --> 01:13:02.899

Mitch Fraas: And it can't be: we're going to justify everything. So it's like, okay, we're going to do everything

487

01:13:03.190 --> 01:13:19.590

Mitch Fraas: created before 1990. Like what is, what is the criteria? And is it defensible with the people you're working with? Sounds confrontational? Is it justifiable? Does it seem like a good idea? Does it make sense to say, you know, we're gonna focus on first generation only, and

488

01:13:19.660 --> 01:13:32.190

Mitch Fraas: we could apply for another grant for other generations of people's papers or recordings. Like just nailing down those criteria so that everyone feels, even if it's a trade-off, people don't love it. Everyone's like, okay, we've made this decision.

489

01:13:33.170 --> 01:13:37.980

Mitch Fraas: I would have preferred XY or Z. But like, I know why we're doing this. This is exciting.

490

01:13:38.010 --> 01:13:47.270

Mitch Fraas: and I think then, making clear within the digitization project, the provenance or source of these things. So it's not, we have this problem sometimes that

491

01:13:47.290 --> 01:13:54.409

Mitch Fraas: people assume everything is held at Penn because it's been selected from many different places, and we host it.

492

01:13:54.500 --> 01:14:00.020

Mitch Fraas: And making sure it's clear to people that yes, this one folder on X topic came from

493

01:14:00.170 --> 01:14:09.750

Mitch Fraas: you know, Mikveh Israel Synagogue downtown, and we don't own it. It's related to this collection. We digitized it. And I think there's some work we could have done to make that

494

01:14:10.030 --> 01:14:17.459

Mitch Fraas: crystal clear. And some of the metadata, it's there, but you know, making it clear to people. But yeah, that's a tough one.

495

01:14:19.990 --> 01:14:37.469

Alyson Pope: So this one is really probably one to be fielded by us. Our audio collection was donated by the producer of the collection based on his ownership rights, and there were no performance releases collected from the musicians. Our digitization workflow will include metadata work as well as gathering releases from the musicians or their estates.

496

01:14:37.570 --> 01:14:46.070

Alyson Pope: This affects access outcomes. There are 200 plus reels with that number of performances. Will you have a webinar on this scenario?

497

01:14:46.110 --> 01:14:56.980

Alyson Pope: So this will probably be best addressed in our rights ethics, and re-use session. I believe that is Session 4. And

498

01:14:57.360 --> 01:14:58.560

Alyson Pope: that

499

01:14:59.400 --> 01:15:12.150

Alyson Pope: will have more to do with what you are going to be able to address from an access perspective based on what comes out of that session and their explanation of what rights you have.

500

01:15:13.680 --> 01:15:25.279

Alyson Pope: There's a sub comment: we are applying for the grant to pay for this hunt for releases. That can be something that happens during the process. But it would

501

01:15:25.350 --> 01:15:40.080

Alyson Pope: would, it can be rather lengthy. I know that we see that this can be a rather lengthy process. So it may be something that you want to begin before being awarded. And that's definitely something that will be better addressed in Session 4.

502

01:15:41.710 --> 01:15:53.380

Alyson Pope: Mitch, did you leave out materials that were part of the same collections to avoid copyright violations, or did you digitize and set up controlled access?

503

01:15:53.620 --> 01:15:55.060

Mitch Fraas: Yeah.

504

01:15:55.430 --> 01:15:59.540

Mitch Fraas: Oh, this is a great question. In none of these,

505

01:16:01.090 --> 01:16:04.410

Mitch Fraas: none of our CLIR grants did we leave out...

506

01:16:05.140 --> 01:16:17.849

Mitch Fraas: Did we leave out logical material for copyright reasons. Like in the Marian Anderson collection, we chose to just do her home recordings and do, I think, all of them. Because we knew the rights were gonna be okay on that.

507

01:16:17.880 --> 01:16:21.429

Mitch Fraas: We definitely chose not to do some other audio because of that. But

508

01:16:21.600 --> 01:16:28.670

Mitch Fraas: within those logical series, like archival series, we did them all and just avoided an entire archival series.

509

01:16:29.020 --> 01:16:31.469

Mitch Fraas: Controlled access is something

510

01:16:31.530 --> 01:16:38.370

Mitch Fraas: we are currently struggling with. We have sort of controlled access for some of our media types in the reading room,

511

01:16:38.400 --> 01:16:45.500

Mitch Fraas: but our long-term solution 2025/26 on that for controlled access. So

512

01:16:45.840 --> 01:16:58.509

Mitch Fraas: I think in the future that'll be a big part of determining things and allow more flexibility. Instead of just saying we'd love to digitize the series of these recordings in their entirety, but half of them are problematic.

513

01:16:58.550 --> 01:17:09.160

Mitch Fraas: We might be able to say, yes, we'll do it because we have an access solution, as opposed in 2017, I was like, well, we're just gonna avoid that whole series for that reason.

514

01:17:09.600 --> 01:17:10.689

Mitch Fraas: Yeah, that's tough.

515

01:17:18.860 --> 01:17:21.509

Mitch Fraas: Because they're also, this will be covered in the right session, I'm sure, but like

516

01:17:21.640 --> 01:17:25.360

Mitch Fraas: different risk tolerances. I mean, this is not really a CLIR issue, but

517

01:17:26.620 --> 01:17:36.830

Mitch Fraas: I would say a lot of academic university libraries vary wildly on how risk tolerant they are and then smaller organizations, community organizations, also.

518

01:17:37.380 --> 01:17:40.990

Mitch Fraas: But I think there are some smaller organizations who have a pretty

519

01:17:41.440 --> 01:17:44.110

Mitch Fraas: high tolerance for risk, knowing that

520

01:17:45.080 --> 01:17:48.759

Mitch Fraas: they aren't a, you know, 1 billion dollar institution that's likely to

521

01:17:48.830 --> 01:17:57.220

Mitch Fraas: attract the ire of a recording label. It doesn't mean there's no risk and it's not problematic. But I think within legality there's a lot of risk.

522

01:17:57.440 --> 01:18:09.760

Mitch Fraas: It's not a question, always, of strict like, are we allowed to? It's sometimes a really tough judgment call. So I'm sure that will that will come up, and the kind of institution you're at, and who you work with will affect that

523

01:18:10.160 --> 01:18:34.589

Sharon Burney: Also, as far as the rights, ethics and re-use are concerned, which we'll get into that later and with the new iteration of the program, there are some instances where that is necessary. So as long as you have a plan in the take down due to ethical concerns, or if you want, you have to have some fees associated with it. There is a place for that on the collections form, and explaining the justification of why, too.

524

01:18:34.640 --> 01:18:35.690

Sharon Burney: So

525

01:18:37.460 --> 01:18:57.160

Sharon Burney: We have another question for you: Following up on your answer to one of the last questions about level of metadata. If we are digitizing 63,000 images, we don't necessarily need to create 63,000 metadata records one for negative. We can possibly do one metadata record for each envelope or

526

01:18:57.190 --> 01:19:02.319

Sharon Burney: negative, since the envelopes contain several negatives on the same subject. Is that acceptable?

527

01:19:02.730 --> 01:19:17.940

Mitch Fraas: Oh, I mean I should, I would say the reviewers might have different, like metadata gets people going so I don't know. Whenever, I'm like, yeah, that's acceptable. I don't trust that. But I think largely like, given our internal debates here, I know we have curators, people with different opinions.

528

01:19:17.970 --> 01:19:22.479

Mitch Fraas: We've made those decisions on like, the yeah, the envelope, the folder, the contact sheet level

529

01:19:22.570 --> 01:19:31.639

Mitch Fraas: make a lot of sense. When I was thinking of an example I saw the other day, which is from a particular event. You know, a

530

01:19:31.850 --> 01:19:38.599

Mitch Fraas: let's say a musician visiting a high school, and someone had taken 130 photos. He's a journalist. So it's a lot of photos.

531

01:19:38.760 --> 01:19:56.780

Mitch Fraas: And you know, we're not gonna put 130 photos with individual metadata. What people want to know is, here are the 130 photos of that event on August 3rd, whatever 1985, with this person, and you can scroll through the 130 photos that are slightly different head positions of people. And I think that makes a ton of sense.

532

01:19:59.230 --> 01:20:10.950

Alyson Pope: How did you avoid feeling overwhelmed, coming up with quantity estimations? Lol, did you have someone help you come up with those numbers?

533

01:20:11.120 --> 01:20:14.759

Mitch Fraas: Well, you could clearly see how well I did. I yeah, I think

534

01:20:14.800 --> 01:20:23.749

Mitch Fraas: I had our digitization head. I don't wanna, like, impugn him in the blame for this estimation error. But he was really helpful in

535

01:20:23.920 --> 01:20:29.760

Mitch Fraas: he's given his experience seeing a lot of stuff come in for digitization, could sort of eyeball a lot of stuff and say,

536

01:20:30.030 --> 01:20:36.459

Mitch Fraas: you know, usually he's good at time estimates. Like, that's gonna take so long, or this long, and to get sort of a gut check on

537

01:20:36.540 --> 01:20:38.499

Mitch Fraas: you know what this is going to entail.

538

01:20:38.540 --> 01:20:45.629

Mitch Fraas: In retrospect, this would have been a great case for some student. Again, this is at an institution that we have

539

01:20:45.830 --> 01:20:49.170

Mitch Fraas: you know, students on work study, and not every institution has this.

540

01:20:49.190 --> 01:20:57.799

Mitch Fraas: To have a little more sampling done, I think I would have maybe asked a student to spend a grueling hour, you know, counting pages on a couple of these.

541

01:20:57.960 --> 01:20:59.980

Mitch Fraas: I think that would have been a good use of time.

542

01:21:00.020 --> 01:21:02.970

Mitch Fraas: I also yeah, I think that

543

01:21:03.180 --> 01:21:05.520

Mitch Fraas: having more people involved, I think there was a little

544

01:21:05.900 --> 01:21:18.679

Mitch Fraas: a little bit off the cuff there, that I would do differently knowing instructively what happened. That my estimates, you know, those 2 or 3 volumes I looked at sure, we're 500 pages, but

545

01:21:18.720 --> 01:21:30.940

Mitch Fraas: the average turned out to be 130. So, I should, CLIR is, you know, just to make it clear or apparent that CLIR is not like, you know, seat of your pants funding everything. This was during the pandemic, and

546

01:21:31.050 --> 01:21:35.470

Mitch Fraas: to CLIR's credit, we were in contact with them all the time. Once we hit our image limit,

547

01:21:35.500 --> 01:21:38.170

Mitch Fraas: we said, you know what?

548

01:21:38.380 --> 01:21:47.600

Mitch Fraas: You know, we've still got 20 volumes to go, and in our case the other institutions in our area were still closed and had no digitization going on

549

01:21:47.690 --> 01:21:53.769

Mitch Fraas: for Covid reasons. And we actually were able to employ our folks who are essentially alone in the building

550

01:21:53.940 --> 01:21:57.659

in parts of late 2020 doing this work and everyone

551

01:21:57.670 --> 01:21:58.780

Mitch Fraas: we

552

01:21:59.020 --> 01:22:24.330

Mitch Fraas: there was communication with the PIs and with CLIR all along so the estimation error wasn't just something like discovered 3 years into the project, or something. It was something we realized pretty quickly and were able to come up with a budget plan with CLIR on that it needed a change order, or whatever the approval was, so I shouldn't be so cavalier about it. It's an example of where estimation could matter, and also how you know there are solutions. It's not. It's not a disaster. There's ways to work around it.

553

01:22:24.750 --> 01:22:53.330

Sharon Burney: Yeah, I wanted to just say, if awarded funding we have a grant modification process, and it does happen all the time where they get in there and they find duplicates, or there are things that couldn't be salvaged or saved, and people want to change. Do a Grant modification request to change the project deliverables, which can happen. Or, sometimes they save a lot of money in digitization, and they wanna add on to the collection. As long as it fits the same cohesive project scope,

554

01:22:53.360 --> 01:23:20.509

Sharon Burney: the theme stays cohesive, you don't anticipate any rights ethics, any additional rights ethics and re-use issues with the proposed nominated materials. This is something that is normal that happens. We, there's no need to freak out. We asked for an estimation for a reason, and that's your vendor, and the funding and the price quantity somewhat looks reasonable as far as digitization for idle.

555

01:23:21.000 --> 01:23:26.349

Mitch Fraas: And that's, I think, where the margin of error matters to, like, when I'm on our digitization committee here, and

556

01:23:26.670 --> 01:23:43.940

Mitch Fraas: we always, when people are struggling to come up with numbers and they're like, yeah, I think it's 130 folders. Or maybe it's 126 folders. We all just like, oh, okay, that's enough. Like we, we know, that's like so close that it does not matter in the world of what we're doing. Even, I would say, 22,000 versus 49,000 images

557

01:23:44.700 --> 01:23:48.179

Mitch Fraas: is like a big error. But it's not a million images. Like,

558

01:23:48.250 --> 01:24:02.680

Mitch Fraas: you know, understanding where you're gonna, when the orders of magnitude really scale up that, especially if you're doing like video or audio recording. If you're getting into the, you know, 50 to 100 TB level like it does matter, whereas at these smaller image levels

559

01:24:02.800 --> 01:24:03.850

Mitch Fraas: the errors

560

01:24:04.140 --> 01:24:08.530

Mitch Fraas: do matter for budget reasons. But, like, you know, the orders of magnitude that are going to be

561

01:24:08.680 --> 01:24:09.860

Mitch Fraas: safe.

562

01:24:11.850 --> 01:24:26.029

Alyson Pope: So we just have a few quick reminders before we're done for today. First, we want to do a quick check in with you again about the topic. We're going to launch another poll now, if you could take a quick moment to complete it, I'll drop the screen share.

563

01:24:45.510 --> 01:24:46.590

Alyson Pope: Okay.

564

01:24:50.500 --> 01:24:59.000

Alyson Pope: Looks like that's about complete. Thank you guys for participating in your feedback.

565

01:25:00.650 --> 01:25:14.050

Alyson Pope: As you work on your final application, remember, there are a number of additional sessions planned over the coming months. The next one will be on March 20th and we will address designing an achievable and sustainable digitization plan.

566

01:25:14.560 --> 01:25:21.070

Alyson Pope: You can sign up for these sessions on our [Apply for an Award](#) page. I believe we are dropping a link in the chat.

567

01:25:26.450 --> 01:25:50.439

Jane Larson (she/her): We're so glad that you all were able to join us today. And we wanna extend a huge thanks to Mitch for leading this session. As we continue our assessment of the program, we appreciate your continued feedback. So again, we've shared the links to the [Webinar survey](#) as well, and we look forward to hearing how this session went, what you might like to see in future sessions, and how we might improve.

568

01:25:50.760 --> 01:26:04.560

Jane Larson (she/her): Please don't hesitate to send questions to the CLIR Grants team as they come up between now and the final submission deadline of May 8th and as needed, our team will also be available to consult with Mitch on questions specific to today's topic.

569

01:26:04.580 --> 01:26:08.449

Thank you, and we hope to see you at another session soon.

570

01:26:09.890 --> 01:26:11.699

Alyson Pope: Thanks everyone. Thanks all.