GMT20200428-130803_Digital-Pr_1920x1080-converted.mp3

H [00:00:03] OK. And so continue where we left off. So we were looking at question 30. I think we just about wrapped up question 30 discussing about regular intervals.

H [00:00:17] Was there any other? So what then? [name] and [name], where did we get to with that, was it that you have to have logged that you've done a check and that sort of fixity check?

E13 [00:00:30] No, the discussion was about the contrast between implicit fixity checking, ~~which is~~ which is offered in file systems, such as ZFS and explicit fixity checking which we decided that that's what the question was actually pointing at. And the the idea that an archive should maintain an independent record of the file, of the fixities of the files under its care.

D [00:01:08] ~~We've got it.~~ It's got to actually know the checksums and not just rely purely on, on, on. The sort of automated process that's happening in the background that they can't actually proves that the checksums is the same, over time.

H [00:01:24] Okay, great. Thank you. That's great. So, next question, Question 31.

Question 31, 32 and 33

H [00:01:35] If I can change, I can actually change the slide. There we go. So still on the theme of checksums, you've got some definitions there. So the question is, out of a thousand files where a checksum was generated instantly on deposit by the archive at an archive where there is sufficient information management and good system security, how many files can you expect to provide the assurance that the bit stream is identical to when it was added to the archive? This is quite a long question. Take a minute to read through it. And there are three questions in the minute that are similar. It's the underlying bits where the questions that going to vary.

E07 [00:02:27] My first thought was around the time frame you're talking about here. So are we talking about kind of next week or are we talking about in 100 years time? Because your answers to that might differ depending on the timeframe.

H [00:02:48] Thanks, [name]. So I think we want to talk about now as much as possible so can we provide assurance today. Do you think you'd be able to provide assurance now that it is still identical to when it was added? Under these conditions.

E07 [00:03:06] Great thanks

E13 [00:03:08] A can I come here.

H [00:03:09] Yeah.

E13 [00:03:10] A moment. The problem that I have with the question is that it is. It is conflating purpose and result. In that what you can do is expect to be able to confirm that a bitstream is identical or is not identical.

H [00:03:33] Yes.

E13 [00:03:34] What you what you can't do with fixity management is guarantee that ~~a stored~~ a stored bitstream is, of it of necessity, going to remain the same. The point is, you can tell if it's changed.

H [00:03:53] Yes.

E13 [00:03:56] So I think the wording of the question is not quite right.

H [00:04:01] So I think what we're all trying to be after here is firstly, can you tell if it's changed? And then. And do you expect, what proportion do you think you will see that it has remained the same? So you need to be able to provide assurance, you need to be up to be able to tell if it has changed or not. And then how many would you expect it to have changed, given you've got good system security?

D [00:04:32] Yeah, because we're always trying to get a bit of the fact, the there, that there's no risk that someone could actually have changed the recorded checksum as well.

E13 [00:04:42] Yes. Which is why I spoke earlier about an independent record.

D [00:04:47] Yeah. But if if your checksum is part of your system, there is still a you know, there is still a risk that someone could accidentally or maliciously actually change that checksum in order ~~to to~~ to be able to change the file as well.

E13 [00:05:03] Yes. So I'm still I'm still slightly unsure about the notion that you're providing the you're providing assurance that the bitstream is identical. What ~~what what~~ fixity does for you is provide the assurance that you can tell whether it's identical or not.

D [00:05:26] Yeah. But given, so we've you've received files, you've, they came with a checksum or you've generated it soon as they came into it. You have stored that checksum somewhere independently of the files, within your system. Over time, you can verify that the checksum is still if you regenerate the checksum of the file it's still matching what was recorded for it when you received the file.

E13 [00:05:53] Indeed so, that's that's precisely what fixity management does but that's not what the question is referring to. The question is asking you, in addition to comment on the long term reliability of the storage.

H [00:06:12] Yes.

D [00:06:13] It's not yeah, it's not just the storage is the wide, because it is the wider information management and security. So it might not just be the storage that's gone wrong. It's the it's some of those other things.

E13 [00:06:23] But that's not, but that surely isn't is not what question 32 is all about. It if it is, then it's asking an impossibly wide, wide, wide question. My case is that we should be asking here, how many files can you expect to provide an insurance of telling whether it's been altered or not?

H [00:07:01] So I think the way we approach doing the model, we are going to assume that. Oh, hang on, maybe I'm going to eat my words. If you've got good information management and good system, if you've got good information management, you'll be able to, and you have the checksum instant generations on deposit you'll have the checksums,

and you will be able to do that check and this question is ask asking. So you've got to be able to do that check, and then the checked results are that everything's fine. What is the probability or out of the thousand files where you you can make that check and you get the answer that the outcome's fine.

E13 [00:07:49] Well, that is what I would would claim is an impossibly wide, wide question. Because you're basically asking what the probability that a bitstream remains remains constant. The point about checksums of fixity and fixity management and systems security and all the other things is that you can tell whether it's remain the same or not. It doesn't help you keep it the same. But it helps you identify whether it is the same. That's my point.

H [00:08:46] So I think the risk. I think what we are trying to answer here is both of those in one. It is because you've got to know if it's changed, you've got to know whether it has been changed and then you have got to see well has it been changed, has there been a bit flip, and everything else which is what the system security aspect is trying to protect against.

H [00:09:09] But do you think, so you're saying that's quite broad to be asking both? Do you think it's?

E13 [00:09:15] Yes, I'm I'm saying that to suggest that fixity management provides an assurance that the bitstream is preserved is a wrongheaded question.

H [00:09:31] Oh, no, I agree, yes.

E13 [00:09:33] I'm I'm saying that fixity management is all about being able to identify and demonstrate that the bitstream has remained intact.

D [00:09:49] I'm not quite seeing why you think that's different to what we're asking you, asking the bitstream. We're saying the bitstream is identical to when it arrived. How is that different to what you're saying?

E13 [00:09:59] I'm not I'm not talking about the the way in which fixity management is performed. I'm talking about what fixity management as a technology achieves and it achieves being aware of the long term fixed fixity of a bitstream. It provides information.

A [00:10:32] I'm having to think about the question now.

D [00:10:40] Yeah. Perhaps there's a slight, you're trying to say how many files, for how many files will the bitstream be identical? So you have the assurance that it's the same as when it was added to the archive?

E13 [00:10:52] No. That is what I'm saying is an impossibly broad question. I'm I'm making the point that fixity management doesn't assure that the repository, that the storage has worked perfectly. What fixity management does is to tell you whether it has or not.

H [00:11:14] Yes. I guess a question we are asking in a minute, so this is under a situation with good systems security. And then we do ask it under poor system security to highlight that, ok it's a broad question but if you know you have good system security maybe you might feel that you you're more likely to be able to provide that, that it will be identical than if it bad system security.

E13 [00:11:41] The wording in question 33 is similarly flawed. That, the fixity management is not assuring that the bitstream is identical. It's assuring that you can tell whether the bitstream is identical.

A [00:12:04] How would you change the question [name] to, would you narrow it down to just one aspect?

E13 [00:12:13] Yes.

A [00:12:13] ok

D [00:12:16] I think it's, so you're objecting to the word assurance as much as anything.

E13 [00:12:20] No, I'm concerned about the word assurance being being directed at the whole of the bitstream. My problem is that assuring that the bitstream is identical. It requires consideration over the whole of the repository process. The reason the reason that you do fixity management is in order to be able to identify and take management action in respect to those bitstreams that are no longer identical. It doesn't help you retain them being.

D [00:13:12] Okay. Yes, I see what you mean. So, your effectively saying we've missed out. So if if, if. Over that time, something had been detected where that would have gone wrong. Unless you actually take some action about that.

E13 [00:13:27] Yes.

D [00:13:28] Yes. So I think. Sorry. So that. Yeah. So that that that that needs to be in there as well. We are assuming that if you have detected any issues along the way, you have acted on those and corrected them. But the other key thing is having also having that good system security so that you can be sure that the checksum itself hasn't been altered.

E13 [00:13:47] Yes. Yes. Yes,.

D [00:13:48] Yes. Okay. Yes. So we we we've we've kind of elided over that step that we are assuming that corrective action has been taken if issues have been found along the way.

E13 [00:14:00] Right. OK. I see the note from [name] coming through.

A [00:14:08] OK. So it sounds as though we need to clarify that. by adding that thing that David said that was elided and that. And that's fine, we can do that, we can do that for 31, 32 and 33 are the same, are based on the same premise. Is that right? yes.

H [00:14:32] So are we saying that we need to add on assume that if there had been I assume that if any errors had been found, they would have been corrected? No.

D [00:14:43] Yes. Yeah. So checksum issues have been encountered on the way, you've taken action to get back a good version of the file. Because we're not considering here the risk that the file storage itself has, you know, something's gone wrong with that. And we're not looking at whether you've got enough copies and so on. We're just saying, given everything else.

H [00:15:11] Any found previous issues have since been rectified. Have been, yes.

17 [00:15:17] So what are we trying to count? Are we trying to count the number of files that don't remain the same?

E13 [00:15:24] Its measuring bit rots. 32 and 33 are about bit rot.

H [00:15:31] yeah in essence, Yeah. We're talking about a bit about bit rot, a bit preservation of be able to assure, I know that it's identical and that that hasn't happened yet.

A [00:15:46] OK, so do we need to amend 31 with the correct, the assumption that the errors have been corrected and that a proper version has been recovered and that's what we're talking about here.

E13 [00:16:01] Well the the as I infer it, watch what you're trying to detect, trying to give yourself assurance of is that you can detect bit rot.

H [00:16:18] Yes, yes, yes. How much and how frequent you expect that to be.

E13 [00:16:24] Yes. Yes, so why don't you just ask that?

D [00:16:30] I think yeah. OK. OK.

H [00:16:33] So maybe it's how many files...

E13 [00:16:38] Well, for example in in in question 33. It's how many files can you expect to provide assurance that 'bit rot' can be detected.

H [00:16:51] And so, But there is bit rot can be detected and then we also want the frequency of how many will it be, do expect it to be detected in.

E13 [00:17:03] Well, isn't that what that question would then be? How many files can you expect to provide the assurance that bit rot can be detected?

H [00:17:11] Well, that that that that rot is detected.

E13 [00:17:15] Yes. Oh, no. Can be.

H [00:17:20] But we want to ask not only, you have the ability to detect it. Then what's the probability you do detect it as well.

E13 [00:17:28] In which case you need to break that up into two questions.

E07 [00:17:33] I'm not sure this is just about bit rot, is it? Because that wasn't my reading of it. So even with good system security, there's still a chance that something will go wrong, that someone who has access to the files with good reason accidentally changes one of them. In which case the bitstream won't be intact. So is it really just about bit rot?

H [00:17:59] I think it's not just about bit, yeah it is about security. Yeah you have good security systems so people can't hack in or it's less likely people will hack in and damage it. Yeah, incorporating a range of things be it accidental, malicious and bit rot or or

anything else. So maybe there is two questions. Maybe one is, can you provide assurance?

E13 [00:18:24] And the fixity management would apply to both. So perhaps bit rot is the wrong term to use. Perhaps one should talk about file corruption.

E07 [00:18:36] But it's not just about corruption, is it? So file corruption is a particular thing. Someone going in and altering the file, either maliciously or accidentally, is just as much of a threat. Perhaps more so. But that's not corruption.

E13 [00:18:55] Oh, I would argue that, that it is, because the file is not the same as it was when it started out. The files file has been corrupted.

D [00:19:04] The file has been altered in some way. Yeah. And the thing that's what we were trying to drive at, but um.

A [00:19:17] I think we need to I think what I say now is that we need to move on and I think we need to come back because I think I've got this quite a few questions and suggestions around this as well. So I think lets, and as this is 31, 32 and 33, there were all variations of the same thing.

H [00:19:39] So, the only difference is good and bad system security. And there is a slight difference in that question 33 is talking about files where they were deposited with a checksum. So you might feel like there was a slightly different risk for files where you've had to generate the checksum instantly on deposit or which ones you were given a checksum with initially in terms of then being able to later on, provide assurance, but you'd be able to tell if the bitstream has changed.

A [00:20:09] Yeah, it's slightly different, but I think the the fundamental of the system security, corruption threat, is still underlying those ones. So shall we leave them out now and move on to 34?

H [00:20:27] Yeah. Yeah. We'll come back.

A [00:20:30] We'll come back to them and have to think amongst ourselves. Yep. Yep. The numbers were revised once we. If we get back to definitions.

H [00:20:39] Yeah.

A [00:20:40] Yeah.

Question 34

H [00:20:41] OK, 34 is still along the lines of, um, bit preservation but let's see how we go. Our of 1,000 files with their bitstream stored and accessible, but where you cannot guarantee their integrity, how many would you expect to be bit preserved? so this is in a situation where you cannot guarantee that they haven't been changed.

E07 [00:21:27] That makes sense. I think. Yeah, so that's if you've got nothing in place to monitor it and fix it.

D [00:21:36] Yes. And you don't have a checksum for it. So they may coincidentally stay the same but you can't, you haven't been doing anything actively to try to ensure that they are.

E07 [00:21:50] Yeah, that makes sense.

Question 35

H [00:22:03] I think in that case, maybe we'll move on, that seems to have gone past quickly. Question 35: Out of the thousand files, which are bit preserved, at an archive that has full access to rendering tools, so it can access all the tools it needs, but the files have insufficient technical metadata, how many files would you expect to be able to render? So you have everything you need, but you don't have that sufficient technical metadata we were talking earlier about droid reports, or validation tools. So we don't have that. Insufficient detail. How many would you still expect to be able to render? Is that clear?

E07 [00:23:00] I think so, to me, it is.

H [00:23:05] Without creating, yes, without its without having done a droid report or something else, so, the way I thought this is me as a non digital archivist, I will have a word doc and I will have word and I will open it in word and I will see it and think, yeah, that looks fine. So I haven't used any technical metadata report or done a droid report. So I. I'm still rendering. I've still rendered it. But have I actually rendered it correctly? So sometimes you'll spot whether or not you have but sometimes you might be missing, and sometimes you will, you won't need that extra report. And it will be rendered fine. It will be rendered properly.

E13 [00:23:51] Can I ask if this is the analogue to question four and five? Were we agreed the tools to render meant sufficient rendering? Not, not extravagant rendering?

D [00:24:08] Yes, I think so, yeah. Yes. OK.

H [00:24:12] Over what timeframe? Again, now, how would you expect to be able to render now, I think is all we've got.

H [00:24:24] If you hadn't done a droid report, it would be difficult to know if you have, yeah, you could argue impossible to know if you've actually rendered it correctly. Maybe that is something you think. But it's not impossible to still render it correctly, even if you don't have that technical validation. There is still a chance that if I'd have had that droid report or that validation report, I wouldn't have done any better or rendered it any differently. Or less correctly than I have done otherwise.

H [00:25:08] Any more questions? Does that clarify things for people?

H [00:25:20] OK. I think I'm going to move on.

H [00:25:24] But I reiterate as well during when you answer these questions we will be online. Martine will be online and if anything else isn't clear, do do ask to come online and say these are all the first round questions as well. And which we will then discuss and maybe discuss why people answered differently. And then the proper ones, which we'll use as the data for the model will be the results we do tomorrow. So if anything isn't clear, hopefully things will be clear because we're going through the questions but don't feel like,

this doesn't have to be perfect either. We will have discussion. This is just the first step. So.

Question 36, 37 and 38

H [00:25:58] Question 36. Again, we've got some repetitive ones now. Out of a thousand born-digital files, for how many would you expect an archive to know their conditions of use. So here we've got a new term that we've derived, which is knowing the conditions of use and any restrictions on the digital material, including legal status, copyright, intellectual property, freedom of information, etc.. So sort of it's not the same as content metedata. It's more about, but it's metadata with information to know what appropriate use of the files are.

E07 [00:26:43] So my my question on this one is around what you mean by the word expect. So you could say you would expect the archive to know the conditions of use for 100 percent of their files, because that's that's good practise. But it's, but I think what you're actually getting at is in actuality, how many would you anticipate that people do have the conditions of use?

D [00:27:10] Yeah.

H [00:27:16] And [name] says, is this any archive or our own archive? As Martine said earlier answer from your experience. So let's say you archive now or one you've recently worked if you don't and how many in general you would expect.

E06 [00:27:38] And this is possibly slightly pedantic. But would this be pre or post sort of analysis? Because, you know, a lot of the time you don't know until you open a file up. And have a look or read through it. So are we expecting that to be, are we expecting this knowledge to be about the conditions of use to be something that is, something that comes with the information on accessioning or something we then have to find out?

A [00:28:10] No, I think it's something you've ingested and you have, you know, full intellectual control of all of those records or as much as you would have, you know, initially, you might do some more later, obviously at different stages. But yeah, I think you've had a chance to establish something at this point.

H [00:28:38] OK. And question 37, 38 are identical. But there, its all broken down by born digital, your digitised and your surrogate. So digitised being the maybe quite bespoke TNA case of you, a digitised file from an analogue record where you don't, well, the record is the digitised file, you don't have the analogue. It's the same question. So is there mean is there a difference when you answer this between born digital, digitised and surrogate files as an archive about knowing their conditions of use?

H [00:29:30] Fab, I'm going gonna move on.

Question 39

H [00:29:34] Question 39, 40. Nearly there! 39: Out of a thousand U.K. archivists, how many would you expect to say that their digital collections were protected with access restrictions or permissions?

E13 [00:29:56] Can I seek some clarification here over the phraseology, access restrictions and permissions. At first glance, this would relate to file access restrictions or permissions. Is that is that what you mean only? Or do you, or are you allowing more general access restrictions, permissions that may be identified, for example, in catalogue?

A [00:30:36] I think this is around us saying whether things are, no, it's not about whether the, I don't think it's around whether its password protected, for example. I'm thinking of looking to David and Hannah but i am pretty show that wasn't our definition was it.

D [00:30:55] Yeah. I don't think we're talking about technical restrictions now. So the fact that they're on the file system that only certain people can can access. I don't, that wasn't what we were after this was about, because then we go onto sort of ask about redactions. So it's about whether you would actually be able to make it available to people.

E13 [00:31:19] Thank you.

D [00:31:23] Sorry I realised when we came back onto this, it wasn't actually very clear. I think it's one of those we were when we were writing it, we were already in that frame of mind as to what context, we were in.

A [00:31:35] so, what we're saying, because [name] has said about [can't hear]. And [name] has mentioned does this relate to digital rights? I am wondering if there isn't an overlap now between our conditions of use, which is around legal status and it does that cover, as you know, and things like closure, because in our 36 you've got conditions of use, which is freedom information restrictions, which records are closed or open and intellectual property around copyright and things like that.

H [00:32:12] I think this is implementing that, sorry carry on.

A [00:32:14] I just wasn't sure how well we were thinking of access restrictions differently from, was this access restrictions as in you can't provide access to them because, you know, you physically can't, you don't have a way of providing access to the digital records. So it's not that is it I don't think because we talked about access was something different when we decided earlier on take them out because it was too broad. I would want to do more of a investigation of that.

E12 [00:32:43] I must admit. I've kind of read this question and I'm kind of I kind of read it from kind of user access permissions into your digital asset management system. But maybe that's maybe that's why I'm getting it wrong. So I'm kind of saying and, you know,.

A [00:33:02] So I see, so access to where it's held.

E12 [00:33:05] Yeah. So. So in our digital asset management system, you know, we have different user groups who can have different access to, say, a proxy or a master file. And within that, they can either view it and or they can download it depending on which user group they are within. So that's how I would kind of read that question, from my kind of knowledge.

E13 [00:33:31] But is that not exactly what David said? The wording was not intended to be?

E12 [00:33:40] Yes. That's probably just, how I.

E13 [00:33:44] Oh, indeed. Yeah.

D [00:33:48] I can't, I am trying to remember now. Sorry. Yeah.

E13 [00:33:54] There is some ambiguity here.

D [00:33:57] Yeah.

A [00:34:01] I think you might, because that one would kind of be covered. Wouldn't it, really? Because it sounds like a thousand UK archivists and we're also talking about, ah so I suppose the previous questions have been about the number of files and their conditions of use. This is about archivist saying that their collections as a whole have the same kind of conditions of use in place that people know whether they can access them or not?

H [00:34:28] I think this is about I think as [name] asks. I see this is about implementation of the conditions of use. So you might know that this file is only supposed to be accessible by Alex or by one group of people. You could know that. But that doesn't necessarily mean you're implementing that and you have those restrictions and permissions properly set up and you're managing that.

A [00:34:53] Right. Yes. So that makes more sense.

D [00:34:55] Yes, I was wrong then. So in parse, it is those technical restrictions. Yes, yes. I've confuse the issue.

A [00:35:02] Yes. We're not talking about individual files and knowing if its open or closed or copyright restricted, it's about knowing whether you can actually access them from the archive, in the correct way.

H [00:35:19] Yeah so its implementing those conditions, of use, be it these are closed for 30 years or, yes but certain people can't see it and be that internally or even with X conditions on internal use and external use.

E13 [00:35:37] So does that mean that we're now agreed that that...

D [00:35:42] Yeah I was wrong and have confused the issue, so it does include technical kind of countermeasures.

E13 [00:35:47] Yeah, but is it solely that? It that that that is referring to just technical access control?

D [00:36:02] So just so it might not just be at the file system level. It could be a wider system controls as well. But yes, it does actually actually include that.

E13 [00:36:12] Lovely. Thank you very much.

H [00:36:14] OK. Anything else to clarify on question 39. OK. Thanks all.

Question 40

H [00:36:24] Question 40 out of a thousand U.K. archivists, how many would you expect to say that they had at least some knowledge to be able to digitally redact part of a document for Web publication? It's going to, at least some knowledge, so, they know perhaps where to start or how to start to be able to do this.

A [00:37:03] The only question that we don't have any questions about.

H [00:37:07] OK. Oh. Oh,.

A [00:37:09] We have.

A [00:37:13] It's more specific, I suppose it specifically about redaction and digitally redacting something.

H [00:37:21] Yes. And this is more on the access side as well. OK, I'll move on. All right,chat is flashing, OK. Good.

Question 41 and 42

H [00:37:43] So out of a thousand UK archivists, how many would you expect to say their organisations born digital materials were made available to users? And then questioned 42, it's their organisations digitised materials were made available to users? Now I think I've got a, I might put in a clarification here. I think in 42, I think it should probably be digitised as including digitised and surrogate.

A [00:38:19] I did wonder if there was a surrogate one.

H [00:38:22] Yeah. So I think 42 probably should actually be corrected.

A [00:38:26] And say digitised then surrogate.

H [00:38:30] Yeah. It's saying digitised materials, not digitised records.

A [00:38:38] Yeah

E12 [00:38:41] Can I just ask and made available to users on site or publicly available at home?

A [00:38:53] I think either, as long as access is made is possible.

E07 [00:39:00] OK? Even if the access is very ad hoc, so they haven't got them available to download online. But if someone was to make an enquiry, they may be able to make them available in the search room.

A [00:39:14] Yes.

E13 [00:39:16] Can I just seek clarification over here over the definition of users, are we including, for example, subject access to closed material?

A [00:39:34] We weren't thinking of those, no. And we'd have to make them available. That's that's the law. So, yes, we would more of a...

E13 [00:39:43] That's the the point that I'm making here is that the born digital material that's referred to might be might be closed for quite some considerable time, so it's not available to users in the general sense, but it is available to a user.

A [00:40:07] Yes, if it was their information, it would be made available to them. Yes.

D [00:40:12] This is more about the technical availability rather rather than the restrictions on use I think.

E13 [00:40:18] So for users, one can read user, to a user.

D [00:40:26] If they are the only ones who are legally allowed to access it then yes. But this is about having the capability to put, technically to provide access.

E13 [00:40:35] Excellent, thank you

H [00:40:36] That's probably a good way of wording it, there. Having that capability to technically provide access.

E13 [00:40:42] Yes.

H [00:40:45] And we have the distinction here. So 41 is about born digital materials and 42 is about any digitised materials that do have an analogue version. So that is either digitised records or surrogates.

H [00:41:02] OK. Erm, commercial. Would material be being available on third party commercial sites be included?

D [00:41:16] I think if that's the choice you've made as to how to give access, then yeah.

H [00:41:22] Yeah.

A [00:41:26] I think [name] your immediate access or on request, it could be either.

E15 [00:41:33] Just come back to the wording where it talks about their organisations born digital. So if you're a city council that only includes institutional records, not necessarily other records held by that institution. Or is it born digital records held by that organisation?

D [00:41:51] I think we should, held by I think it was really what we were after rather than.

H [00:41:55] Yes.

E15 [00:41:56] It could be them having a diction between.

D [00:41:58] Yeah, where we're coming from TNA mindset, where where there is isn't anything very much that that's not a government records, there is some but.

E15 [00:42:11] OK. It is again,just checking that.

D [00:42:15] I see now that it's it is ambiguous in that sense, but that which, we're just showing our TNA group think when we were writing the question.

A [00:42:30] Has that answered everything? There are some questions in the chat. Yes. OK.

Question 43 and 44

H [00:42:39] And final slide for now, so out of a thousand U.K. non-cyber security incidents, how many would you expect to be due to data being posted, e-mailed or faxed to an incorrect recipient? Not sure how common faxing is nowadays, but we thought we'd put it in there. And question 44 out of a thousand U.K. non-cyber security incidents, how many would you expect to be due to loss or theft of paperwork or data left an insecure location?

H [00:43:23] So just to clarify on these, we are saying that, say a thousand non-cyber events have happened, what proportion of those would be due to it being sent to the wrong person? And what proportion would be due to loss or theft of paperwork or data left in an insecure place. Human error.

D [00:43:47] They're both human error.

H [00:43:50] Yeah. The first one definitely more than, could argue, the second one. Might be.

H [00:44:06] OK.

A [00:44:14] ah, [name], we don't want to. We want your your experience answer, rather than for you to check facts, because we can check facts too, but if we did do that, we wouldn't need to do this exercise.

E01 [00:44:31] I understand. But I'm just checking that we are putting those figures into the actual results as well.

A [00:44:37] When we have actual figures, we will be putting them into the model. Yes,.

E01 [00:44:41] good that's all I was doing I wasn't planning to check.

A [00:44:46] I'm sure.

H [00:44:51] So in in regards to those questions 31, 32 and 33 about the checksums and assurance, Alex and David, do you think maybe if we come up with a suggested rewording and let everybody else know as soon as possible.

A [00:45:12] Yes.

H [00:45:12] Based on the discussion we've had and the need for clarity. And so if people hold off perhaps from answering those and we will get that out ~~as soon~~ as soon as possible. And of course, we have got tomorrow to go through the answers and discuss and check that we are clear as well. Does that sound like a good way forward?

A [00:45:32] I think so, yes.

H [00:45:34] OK. And so. Thank you all for that.

H [00:45:41] I think in that case, we've gone through all the questions. Thank you. I know that that took a while. And as I say, we this isn't, this is the first round of estimates, once you've had a go to answering them and we've had another discussion about all our answers tomorrow, we'll hopefully all be a bit more familiar and know a bit more how intend to answer them. I'm going to change slides a sec so I'm going to stop and find the other ones and resume where we were.

H [00:46:12] So.

A [00:46:15] Yes. So we are.

H [00:46:18] Welcome back.

A [00:46:19] Yes that's it, and you should all have copies.

H [00:46:22] But let me, OK. So, screen share. And bear with me.

H [00:46:44] OK. Welcome back. So I think you should all have a copy of elicitation questions and helpful definitions and an answer sheet. There's actually Excel sx since I wrote this PowerPoint, so I know some of you have let me know if you haven't had them. I assume that by now, you all have. A couple of reminders of what to do next. So we'd like you to answer these on your own. You know, just do what you think based on your experience, based on your understanding of the question. We will be discussing them and going through ideas. So it's not about getting anything wrong or right. It's just about ~~your~~ your premonitions and what you think. How you would answer them. So don't be tempted to look anything up or confer with others and say well what did you think for this? We will have time for that to do in a strcutured way tomorrow. Your own genuine beliefs and opinions. So if you complete that in your answer sheet Excel file, which you should all have. So there is a preamble sheet and then an answers sheet and we will have hopefully left it so that you can enter your, for each question, you will enter your fifth percentile, your 95 percentile and your fiftieth percentile, just like we did with watermelon. So we have that answer form for you. So we will ask that you could complete that answer form and email that completed back to Alex, just Alex, by 7 p.m. tonight. So if you manage to answer them all in the next half hour and send them, fab. But I think that's unlikely and don't rush. But by 7 p.m. tonight would be good so that we can collate them so we have chances to discuss tomorrow.

H [00:48:34] And I've just had a question just to confirm we're answering each percentile with numbers out of a thousand. Yes. Every single thing is out of a thousand. So there was also a hint in the in the spreadsheet. If you put a number less than zero or more than a thousand, it should tell you that's incorrect because everything's got to be out of a thousand. And if you complete your name on the answer sheets as well, where it says a slot, that would be helpful. But as you'll emailing Alex, she'll know who they're from. To re-emphasise, we're not going to be sharing these with anyone else. They're just used as agreed to create our range graphs and discussions tomorrow for all results will be anonymized. You won't be identifiable. So please just give your honest answers. We're not going to be naming, naming and shaming you for people who've done different answers or anything. It's purely, it will be confidential.

D [00:49:27] Unless you choose to identify yourselves during the discussion.

H [00:49:30] Yes. Tomorrow there will be anonymized results. But if you choose to say that was that line was me, I answered that because of this, then you're welcome to. We're not going to, us as a project team aren't going to release identities of individuals that would be up to you tomorrow. And I guess the main thing is, even though we're going to sort of say goodbye soon until tomorrow, Martine and Alex and I will be online all afternoon until 7:00 p.m. to answer any questions. So if you go back to something, you think, oh, I can't remember what that meant or how am I supposed to answer? What does the fifth percentile mean again? We are online so do ask us. Come on line on the chat or come on line and physically speak yourself at any point between now until 7 p.m. That's absolutely fine. And again, you've got, you could, if you need a break, if you've got some other things to this afternoon, it is in your own time as soon as it comes to us tonight.

H [00:50:31] And so final reminder on the next slide, which I will leave up. And the helpful hints is you have your lower bound, your best estimate and your upper bound, try and make sure that you have them in the right sizes. So your best estimate is not going to be higher than your upper bound or lower than your lower bound and there are warnings in the spreadsheet as well to help you with this. And the fifth percentile and 95th percentile, are lower and upper plausible bounds, not physical bounds, because for all of them the physical bounds is zero and a thousand. And we've programmed that in there. We want your plausible what would realistically just its unrealistic if it's if the answer is higher or lower than this. They don't need to be symmetrical about the best estimate. I don't think that's really come up but you just because your low is 50 and your middle is one hundred doesn't mean you're upper has got to be 150. It can be whatever, whatever gaps in between then you like. And we need your estimates. Not someone else's. Not what you think you know, yeah, that is not not, what anyone else says, just what you think from your experiences and your opinions. And that's what will make this work and be really valuable.

H [00:51:47] We will get back to you about questions thirty one to thirty three. And we've also got a definition of the digital asset management, as mentioned earlier, that I will also send when we come back with clarifications about 31 to 33. Other than that, well, any questions now? If not, the plan is to return your answers to Alex and then rejoin the same call again. We'll close after 7 p.m. tonight, but re-login tomorrow morning for 9 a.m. start, where we will review what we've done and then start a discussion about the answers we've given you've all given today.

A [00:52:30] Ok, any other questions, anyone, now or? I'm just seeing if anything comes up in the chat. Otherwise. Yes. Otherwise, can't see anything coming up, so otherwise, thank you very, very much for your attention and your contributions today. We do really, really value this. And I look forward to seeing you here at nine o'clock tomorrow morning to hear all about your what you've come up with. And that's it! But we will we will be here if you want ask up any questions or three of us will anyway.

H [00:53:09] Thank you. Feel free in the meantime to either mute and be off video. So you keep us on in the background or to totally log out. But then if you need to ask us any questions to log back in and we will re-admit you to the meeting. Whatever works for you.

A few minutes later…

[00:00:00] Really with her at the beginning of the project, we are still at the really early stages of digital preservation. So I feel that some of my answers are very, very inexpert.

M [00:00:10] Do feel free to make a note of that in the column.

[00:00:15] The column to the right or where you put the answer?

M [00:00:20] So you put your in numbers, but then you know, using the column to the right to make a note of anything you're uncertain about because you can discuss it tomorrow.

[00:00:28] Yeah.

M [00:00:29] And you can raise that then and say like I feel very inexpert on this one. And also what typically we find is on the questions where a particular expert feels less expert, they just have very wide uncertainty bands. The low will be very low and the high would be very high. As we come to the discussion, we say well, this is has got a very wide uncertainty band for this estimate, this estimate and this estimate, is this because the individuals are uncertain about where the true value should lie? Or is it because they think that the the underlying system is uncertain? And then people typically say, oh, I was one of those. I was uncertain myself.

[00:01:11] Yes.

M [00:01:12] And some guess will say well I think actually the system is really uncertain, it could, genuinely I think it could be that low. We'll say what are your reasons for that? Talk us through on how you think about that. I really think it could be that high, do talk us through the reasons why you think you think that's and then we haev a chit chat around that.

[00:01:30] OK. OK. So it's it's just it's partly about capturing where there isn't knowledge in a way.

M [00:01:38] Yes. Yes. Because we don't want to be overconfident with this. Because we're putting together a model to say if this is this you're going to be fine if actually we haven't captured people's uncertainty about that.

[00:01:50] Yes. Yes. OK. It just does feel as I'm going through with those ones where I feel sometimes it would be helpful for me to say, you know, this is really based on not very much knowledge at all, but purely on kind of, you know, a process of kind of deduction based on what little I do know, which is so little, it feels like it could almost be false information for me to even be speculating. Do you see what I mean?

M [00:02:13] Yeah. Yeah. Which is why in teh IDEA protocol, I love the IDEA protocol that we do have the second attempt tomorrow.

[00:02:21] Yes. Yes. To refine...

M [00:02:28] Seeing what everybody else has put, not neccessarily who has put what but what everybody else has put, you can compare your answer, we can have a discussion about why it could be that high, why it could be that low. (Yeah.) Captures it and this kind of thing. Any then you may feel like you have more information and you can adjust, if you think your answer needs to be adjusted, you can adjust it then. (Yeah.)

M [00:02:47] So yeah. So these answers all to get you thinking everything through, thinking or all the reasons of why it could be as high or low as that. (Yes). And to have a proper discussion tomorrow.

[00:02:56] Yeah. So people tomorrow who do have greater expertise in these areas will share their knowledge and then we kind of we refine based on that. OK. And that's reassuring. Yeah, I agree.

A [00:03:06] Yeah. You can put comments in the column F as well. It says warning messages will appear but you can add comments.

[00:03:14] Yeah. Okay. That's good to know because I do feel that's that would, that would be helpful. I would feel better if I can do that. Just do a few little qualifying statements.

M [00:03:23] Okay. Yeah. Do. That would be helpful.

[00:03:26] All right. Thanks then. Bye.