

ArcGIS Data Reviewer: An Introduction

Michelle Johnson and Jay Cary give an overview of ArcGIS Data Reviewer and its quality control tools that simplify the process of reviewing and maintaining data quality.

<http://video.esri.com/watch/661/arcgis-data-reviewer-an-introduction>

Video Transcription

00:01 Welcome. My name is Jay Cary, and with me today is Michelle Johnson.

00:03 We're members of the mapping and charting solutions team in Redlands, California.

00:08 Today we've got kind of an introduction to ArcGIS Data Reviewer.

00:14 By show of hands, how many folks have heard and/or are using Data Reviewer? Oh, good. Okay. Good.

00:22 So what we're going to do today is we're going to start out by first talking a little bit about data quality...

00:28 ...what it is, how it's being defined, and also why it's important.

00:33 We'll then get into a discussion of Data Reviewer.

00:36 We'll start out by kind of identifying the major moving components of the extension...

00:40 ...we'll talk about different forms of validation that we do against our data to validate quality...

00:46 ...and then we'll talk a little bit later about how, once we've found all these errors, what do we do with them?

00:50 How do we leverage them, how do we make our jobs easier when it comes to maintaining our data?

00:55 And finally, at the end, we'll talk a little bit about resources available to you to help you learn a little bit more about Data Reviewer.

01:05 So this first section, we want to talk about quality, what it is, define its different aspects as well as how do we...

01:11 ...you know, where you might find that technical guidance that tells you the quality and what is good quality.

01:15 We'll then talk about Reviewer in a little more detail.

01:18 At the end of each section, we'll be having a demo, and we'll take questions after the demo.

01:24 If you do have a question that's just burning, you have to ask it right away, put your hand up;

we'll try to get to you. Okay?

01:32 So the first bit, before we even get into the technology piece is, you know, what is good quality, right?

01:39 So when you look...like when you look into what is quality, there's lots of different aspects to data quality.

01:45 How well are my features in the GIS reflective of features in the real world? Are my attributes complete and accurate?

01:52 So all these kind of different elements to data quality conceptually tell us, describe its relationship to the real world...

01:59 ...a little bit about how it's structured in relation to other data layers...

02:03 ...an acknowledgment of where it came from, its source, as well as...

02:07 ...sort of an understanding that this data may have been manipulated through time and lineage that brings it to its current place.

02:17 Given those elements, though, what is good quality data?

02:20 So the question is different for every user; it may be different on a layer-by-layer basis.

02:25 A lot of times, this is based on how are we using the data?

02:29 So if I'm developing a transportation dataset that's going to be used for a national-scale map...

02:34 ...my accuracy requirements, my quality requirements are going to be much different...

02:37 ...than if I'm doing a large-scale project where I'm routing vehicles inside a city.

02:42 So really, we need to think about how the data's going to be used.

02:46 There might be some guidance if we're working for agencies of our contractors.

02:49 It may be some technical specifications put out that kind of define what quality is and what's acceptable and what's not.

02:57 Or client requirements may articulate that as well.

03:01 Now this...we only talk about this in one slide, but I really want to make sure that everyone understands this is not easy to do.

03:08 Defining quality, if you do it right, it's actually hard to do because it takes a lot of consensus...

03:13 ...a lot of understanding of, well, how's this data going to be used today and also what does the future hold for this data.

03:20 'Cause there's a lot of, you know, impact to, you know...

03:22 ...gosh, I didn't realize we were going to use the data this way so we didn't build it that way.

03:26 It's kind of a balance between spending too much money for quality data that you don't need necessarily...

03:31 ...but then later realizing, oh, I really did need that.

03:36 So once these quality requirements are defined, we typically see them incorporated as a series of tasks and methods...

03:44 ...and things we do to meet that quality requirement.

03:47 So quality control, just those things that you do in your production environment to make sure that you produce quality...

03:52 ...to a known standard as well as a methodology to assess it at some point.

03:57 Now, a lot of folks have quality assurance that they do, which is the things you do in a production environment to ensure quality...

04:04 ...but in this mode, we're talking about quality control, which is those distinct things to check data.

04:10 Check it before it goes out the door to a client, check it before we publish the data.

04:16 Normally you'll see these kinds of requirements defined within a quality assurance plan.

04:20 So in a lot of...if you're a contractor, you may have a quality assurance plan for project delivery...

04:25 ...that might also stipulate certain data requirements for you.

04:28 If you're a government agency, you may have quality assurance plans that define...

04:32 ...on a layer-by-layer basis what your quality requirements are.

04:39 So one of the things, and this sometimes is surprising to leadership in some cases...

04:44 ...but, you know, poor-quality data is expensive in a lot of different ways.

04:49 It's expensive to fix after the fact.

04:51 Having technicians go back and correct something that should've been found during production, that's time-consuming.

04:58 It impacts users and decision making.

05:00 If folks are using your data to produce cartographic products or do analysis...

05:05 ...all that becomes suspect if you've got bad-quality data or they weren't checking it.

05:09 So really, regardless of whether you're a data producer or a consumer, you know, we all need to kind of work...

05:15 ...to have an understanding of some...you know, what is acceptable quality for the task I have at hand.

05:20 If we're data producers, of course your reputation is at stake.

05:25 Putting out bad data, not meeting your customer requirements, that could affect funding lines, reputation, things like that.

05:32 So things to consider. You know, we all run spell checker on our Word docs before we shoot them out to have folks review them.

05:38 The IT guys install patches and virus scanners to ensure a quality experience with desktops.

05:44 You know, the same could be thought of with GIS data.

05:51 So ArcGIS Data Reviewer is an extension to desktop ArcGIS...

05:55 ...and it provides a complete quality control environment for managing data quality through a life cycle.

06:01 So we have, you know, rule-based workflows for the automated validation of data.

06:06 We have interactive tools to help facilitate manual quality control review of data...

06:11 ...as well as a framework for managing those errors through what we call the sort of error life cycle.

06:19 Because a lot of these processes are automated, it definitely saves a lot of time in terms of validating your data.

06:26 In the long run, if you implement quality controls at really more as an ongoing part of how you do business...

06:32 ...in the long run, it may actually reduce the amount of work you do through time.

06:41 So a little bit about the framework for how Data Reviewer manages its error life cycle.

06:47 During the review phase, you'll take and use Data Reviewer's tools, automated checks, interactive tools...

06:54 ...to find and discover errors, anomalies in your database. And you'll discover those.

07:03 As they're being found, they're written to a consolidated location and archived.

07:09 During the correction phase, whether it's you or a contractor working on your behalf, they'll use the core ArcEditor tools...

07:15 ...to fix those errors, and Data Reviewer enables that person correcting the errors to annotate in the database...

07:22 ...Was it fixed; is it really an error? If it is an error, how did I fix it and who fixed it, and when it was fixed.

07:30 This is especially helpful when you outsource data production to other shops where you might want to have a little dialog about...

07:36 ...well, I found these errors; you fix them, and back and forth.

07:41 And finally, the verification stage.

07:42 A lot of folks do this optionally, but it's a good idea to, once they've been corrected...

07:46 ...the person who found those errors has the ability to go back, verify the corrections were applied appropriately...

07:52 ...but also maybe rerun some of those validations...

07:54 ...and make sure that no new errors were introduced as a by-product of that correction going on.

08:03 So there's a lot of moving parts in Data Reviewer.

08:06 So essentially what, when we talk about writing errors and managing error life cycle, we do that through the Reviewer workspace.

08:12 This is a geodatabase--could be file, personal, enterprise geodatabase...

08:16 ...a series of tables for storing spatial as well as attribute information about the errors we discovered in our database.

08:24 We interact through this workspace through what we consider a Reviewer session.

08:29 It's very similar to an edit session with ArcMap, and the purpose of an edit...

08:34 ...a QC session is that we're opening that workspace to interact with those records...

08:39 ...write new records, move errors through their life cycle.

08:44 Once we've got to that part, we're ready to go.

08:47 We have a workspace, we have a session, and now we can start exercising the various tools to discover errors in our database.

08:54 This might mean running automated checks initially, might mean going through and doing it in a batch mode.

08:59 We'll talk a lot about the automated data capabilities, automated data validation that Reviewer has as well as visual review.

09:08 We'll talk quite a bit about visual review; very time-consuming process, but we try to help facilitate that with some tools.

09:17 Once we've written those errors to the workspace, we use the Reviewer table...

09:21 ...to actually interact and visualize those error records, sort them, summarize them.

09:26 Technicians use this tool to navigate to the errors and to update the status of that error as we correct them and verify those records.

09:39 So once we've done our quality control review, and depending on your mode of how you're operating...

09:44 ...if you fix it or someone else fixes it, there's a whole host of reporting tools at 10 that have been introduced to help summarize...

09:52 ...and kind of make it more easily readable formatting of our errors in multiple different ways.

09:58 So this might be a management report that's produced, and it's very quickly generated...

10:04 ...to give other folks an idea of the scope and extent of our errors.

10:09 We'll do a demo of that, a couple different variations, show you how that new reporting capability is.

10:16 Alright. So for our first demo, Michelle's going to be going through...

10:19 ...and we're going to do kind of a high-level pass over of Data Reviewer...

10:23 ...identify the major moving components and run a couple batch jobs.

10:26 Later on we'll get into a little more detail about how that works, okay?

10:34 Actually at this demo, I won't be running any batch jobs, but I will be running an automated check.

10:41 So with Data Reviewer, what I wanted to mention is that it is a separate install to ArcGIS Desktop...

10:50 ...so if you go through, if you go and look at the toolbars, and you don't see Data Reviewer in here...

10:55 ...it means you do not have it installed on your computer.

10:59 So it is a separate install; however, it is a standard extension to Desktop, ArcGIS Desktop.

11:06 So let me add the Reviewer toolbar. Here we have some...

11:11 ...basically all the tools that you will need to perform your quality control on your data.

11:15 So we have some tools here that is used not every single time, but it is used quite often, so it's down in this menu.

11:24 We have our Session Manager button here, and I'll be showing that in just a moment...

11:29 ...and then we have some tools to manually inspect the data...

11:33 ...and then here we have a suite of checks, automated checks, that you can run against your data.

11:38 Then at the end of the toolbar here, these are helpful for performing visual review of your data.

11:46 So let me show you an example of an automated check.

11:49 Here I have a database with address points, building footprints, and some parcels...

11:56 ...and what I want to do is a spatial type of check where I'm validating my address points against my building footprints.

12:02 So I have a rule that my address points must fall within the building footprint...

12:07 ...so I want to find all the points against my building footprints that are not within the footprint.

12:14 So I can set my spatial relation to within, I can use the NOT option, then I can run this check on my current extent.

12:28 And it found two records.

12:31 Because this is a point feature, it will pan to that location, so let me just change my scale.

12:40 And here I have a point that is next to my building but not quite within the building footprint.

12:46 And, again, if I navigate to the next error, similar type of situation. So these are errors.

12:54 Now, we looked at these errors, but we want to be able to store these errors in the Reviewer table...

13:00 ...so in order to do that, you have to start a session.

13:04 So here is the Session Manager dialog, you browse to the workspace you want to store your errors in...

13:10 In this case, I have browsed to a file geodatabase. ...and then you can select your session.

13:18 I'm going to go ahead and select this one; I already have some records in this session.

13:25 To access the records, you open up the table, and here stores all the errors that you have found through your QC...

13:33 ...whether it's automated or visual.

13:36 And by double-clicking on a record in the Reviewer table, it will navigate you to that feature.

13:46 And this is really helpful when you're going in through the correction process and having to make all the fixes to this data.

13:55 So the table has information on what feature class, you know, has the errors, the type of error that it was, and the error status.

14:05 So if we group the column, we can see that these are all the different types of errors that we have.

14:13 So we have some invalid domain values and we have some points that are not within a poly...

14:18 ...some polys not within our polys.

14:20 So you can get a quick summary of the Reviewer table by grouping the columns.

14:26 So that's just a quick overview of Data Reviewer.

14:30 You get the toolbar, you've got the table, you have the automated checks, and then there's some tools of visual QC...

14:36 ...which I'll be showing in the next demo, or the next couple of demos.

14:42 Any questions at this point? Yes, sir.

14:46 [Audience question] How do you get the...how does it pick up the reviewer technician?

14:50 Is that something that when you log in it recognizes who you are?

14:55 Yes. So this is based on your Windows login, and then you can modify this too.

14:59 You can just go ahead and edit it to whatever name you would like if you wanted to. Yes, sir.

15:07 [Audience question] The check you did there was real similar or looked similar to a regular topology check.

15:13 What's the difference, or is it the same or different or better or...

15:18 So the question was that this check that I performed, the geometry on geometry check, is very similar to database topology.

15:25 And yes, it is. However, it's not the exact same. It is different.

15:30 Database topology resides in your geodatabase. There's some rules that you set up.

15:35 There's more rules, I think, that you can use with database topology than what geometry on geometry may offer.

15:41 However, the advantage is, of geometry on geometry, is that your features don't have to reside in the same feature dataset.

15:48 They could reside in different geodatabases.

15:51 Also you can define your rules to the attribute level, where you can apply a SQL WHERE clause.

15:55 And I believe database topology, you can only go down to the subtype level, so...

16:01 Okay, so we'll go ahead and go back to Jay.

16:07 Alright, thanks, Michelle. So now we'll dig a little bit deeper into this automated data validation.

16:16 So we'll talk a little bit more about some of the checks we have out of the box...

16:20 ...and talk about how we can take those and run them en masse.

16:25 So this is really the strong suite of Data Reviewer, its ability to take the...

16:30 So we talked about the quality assurance plans that these business rules are defined within there.

16:35 If you think about it, those business rules are sort of kind of a logical sort of discussion about quality...

16:40 ...but we can take those business rules and implement them in an automated way.

16:45 The value of being able to automate it is, number one, it can be done, requires really little human interaction...

16:51 ...once those rules have been built.

16:53 Also, it, because it's automated, you can oftentimes validate 100 percent of your database.

16:59 So it's very efficient from that point of view.

17:02 It's also this idea that we can persist these through time, so just like Michelle configured this check...

17:08 ...you can also configure these checks and then distribute them throughout an organization.

17:12 We'll talk a little bit about how we do the detail of that later.

17:18 So this is a picture of the Data Reviewer check poster, which can be downloaded from our product page on esri.com.

17:24 And what it does is it outlines all the different checks that we have with Data Reviewer.

17:29 At release 10, we have around 42 different checks available, and they're organized into these 11 functional areas.

17:36 You might have noticed that when Michelle was pulling down that Check dialog...

17:39 ...you saw sort of that functional breakdown of checks.

17:45 Some of the real popular check categories are things like table checks...

17:48 ...where we have different methods for evaluating the attribution of our database.

17:53 One of the real powerful checks there, though, used quite a bit, is the execute SQL check.

17:59 So an example of the execute SQL check from the water utilities domain is that there's a simple business rule.

18:06 "Water mains installed after January 1, 2000, should have a material attribute of either iron, ductile iron, or PVC."

18:15 Using the execute SQL check, we can build a SQL query that essentially validates our features...

18:22 ...based on a SQL WHERE clause, and then when it tests the data, it just pulls up records that don't meet that requirement.

18:29 Now, a lot of this can be done through the geodatabase; we talked about topologies.

18:35 The value here is that we're actually validating multiple attributes, not a single value like you

would with a range or a list domain.

18:42 So we can build a query that says if it's this, then it should be that...

18:45 ...so you can look at attribute combinations and make sure that they're valid.

18:51 Another category of checks is our feature on feature checks. We saw a demo of the geometry on geometry just now.

18:57 This is another ability to validate sort of the spatial relationships between our features...

19:02 ...and, in some cases, have a much more granular way of doing that.

19:07 So again, the geometry on geometry check, very popular, very powerful 'cause it has lots of different capabilities.

19:13 So another example from the water utilities industry is the idea that hydrants need to be connected to a hydrant lateral.

19:20 So that business rule can be implemented using the geometry on geometry check.

19:23 So we have hydrants as feature class 1, hydrant laterals as feature class 2...

19:29 ...but in this case, like Michelle showed, we have the ability to select a NOT condition.

19:34 So here we're looking for ones that don't intersect one another.

19:38 So the results of that kind of check would be hydrants that are completely disconnected from the network...

19:42 ...as well as hydrants that are connected but connected through the incorrect subtype.

19:46 In this case, it's hooked up through a service lateral. Okay?

19:51 And as Michelle talked to you a little earlier, you can apply also SQL WHERE clauses...

19:55 ...so if you really need to get down to specific features...

19:58 ...you can build a definition query basically to filter out things that you don't want to have queried.

20:05 So the advanced check category, lots of new checks were introduced at 10.

20:09 These are advanced checks that are...don't really fit conveniently in any of the other functional areas of the check poster.

20:18 One of those is really an important one for the utilities industry is the valency check.

20:25 In this example, we're looking for reducers that have to connect...

20:28 Reducers are required, so you've got two pipes coming together, and they're different sizes.

20:33 We need to have a reducer to take it from size X to size Y.

20:36 So using the valency check, we can check the relationship between lines and point features...

20:41 ...and when they don't meet that requirement, we want to flag it as an error.

20:44 So in this case it's a little bit different than all of our other checks, 'cause it's inferring and finding missing data.

20:50 Normally, automated checks only run against, you know, features that exist in our database.

20:54 But in this case, it's looking at a relationship between two existing features and inferring some missing data.

20:59 In this case, we have an 8-inch diameter pipe coming to a 12-inch diameter pipe; there's no reducer in there.

21:06 So it's finding actually missing data for us, so it's kind of an interesting check from that perspective.

21:13 Other new checks introduced at 10. The metadata check.

21:16 If your organization authors metadata, this is [unintelligible]-level metadata, you can validate in an automated way...

21:22 ...your metadata against the standard, a metadata standard...

21:25 ...or against specific content standards that you might have for your organization.

21:29 We talked about the valency check.

21:31 The custom check that enables you to extend the Data Reviewer framework by developing custom code.

21:38 So if out of those 42 checks, none of them meet your requirements, you can get a developer involved...

21:43 ...to build you a custom check that'll still operate within the framework of Reviewer.

21:47 The results are written to the Reviewer workspace as would any other check result...

21:51 ...so you can have a similar place for your errors to go.

21:55 Same, same with the topology rules check.

21:57 If you're using database topologies, instead of dealing with those errors in a separate way, you can import those errors...

22:03 ...into the Reviewer workspace and have a common place to work and manage those errors...

22:09 ...without having to go to two different places. Question?

22:12 [Audience question] How is the metadata check different from the USGS MP Translator tool for all the [inaudible] in ArcGIS?

22:19 So the question was, How does the metadata check differ between the Metadata Parser, right?

22:25 So one of the value-- So that works great with FGDC metadata, but a lot of folks are moving to ISO metadata.

22:32 So the metadata check supports both schema validation, which is, I think, what the Metadata Parser does, correct?

22:39 So you can validate based on the schema, but it could also be FGDC schema or ISO 19115 standard metadata...

22:46 ...or the North American Profile or INSPIRE.

22:49 We go a little bit beyond schema validation with the metadata tool because you can build specific element value rules.

22:57 Point of contact must be John Doe; phone number must be properly formatted US phone number.

23:03 So it goes beyond--for metadata, it's quite extensive in terms of validating your metadata...

23:09 ...and like any Reviewer check, it just whips through your whole database if you need it to.

23:13 Does that answer your question? Okay. Sure.

23:19 So we have these automated checks that we can configure and run.

23:22 Michelle showed you running it sort of in what I call an ad hoc mode; just interactively configure it and run it.

23:29 Now that works fine, but what you'll end up finding is you'll want to configure it once and run it over and over and over.

23:35 So we had this idea that you might want to configure these checks...

23:38 ...and the input from these checks might come from a lot of different sources.

23:41 You might get business rules from...just from your own experience.

23:44 You might have technical requirements coming down from your customers.

23:47 You might have a quality assurance plan.

23:49 So the idea is you take all those business rules that are maybe through disparate places, and you implement them in Data Reviewer...

23:56 ..you configure...you identify one or more checks that answer that business rule; you configure them...

24:01 ...but then you persist them on disk and you save them.

24:05 They're saved down to disk as a small XML file.

24:08 If you're a one-person shop, you just can rerun that over and over and over.

24:12 And we'll talk about ways of running it in a few minutes.

24:16 What you see here in the bottom right corner is sort of that idea that, you know...

24:19 ...we've got all these business rules that we've implemented with Reviewer checks.

24:22 We store them all in one place, and we can share them amongst users.

24:26 So not everybody needs to be a QC expert; they just need to know how to run a batch job.

24:34 So batch validation, that's a tool that we'll see in a few minutes that takes these configured groupings of checks...

24:41 ...that we've defined for our data and enables us to run them all at once.

24:47 And we'll get into some detail about the extent and how those different things work...

24:50 ...but the idea is to create it once, run it multiple times.

24:58 So you can always run these interactively through ArcMap, which is a way a lot of folks do it...

25:02 ...but it does tie up your computer while it's kind of grinding through all these checks.

25:06 So we've introduced a whole bunch of different ways of running these automated validations.

25:11 The Reviewer service, that's a Windows service.

25:13 You can schedule to have these automated validations happen maybe during off hours.

25:18 Let's say we check--everyone goes home at five o'clock; let's set up a schedule for it to run at night while we're not there.

25:23 First thing in the morning, we come back in, there are our error results.

25:28 Also from a command line, if you're into running a command line, or as a part of a Python script.

25:33 A really neat thing at 10, though, was if you're using Workflow Manager to manage your production workflows...

25:40 ...within your organization, a number of custom step types have been provided at 10 so that QC, automated QC...

25:48 ...could just be a step in a workflow, and it just runs; you don't even have to run it from ArcMap.

25:55 Alright. So we'll go ahead and do a quick demo and talk a little bit more about automated data review. Michelle?

26:00 Thank you, Jay. Alright, so I showed you how to run a check by itself through the toolbar...

26:11 ...but what we really want to do is create what we call a batch job to store all of our checks in.

26:15 So through the Batch Job Manager, this is where you can create your batch job.

26:21 First, you need to create a group, and then within that group you can add a check.

26:27 So I can create a domain check, and the reason why I would want to create a domain check...

26:31 ...is this data's recently been migrated from shapefile into the geodatabase...

26:35 ...so I want to ensure that all of my features are adhering to the domain constraints that we have defined in our geodatabase.

26:42 So let me select my feature class, Address Point...

26:48 ...and I want to create a domain check for all of my feature classes in my table of contents.

26:55 So let me duplicate that check...

26:58 ...and I can select all my feature classes, and then it'll create the checks for those feature classes.

27:04 Now I can also move around my checks from group to group...

27:10 ...and I can also rename my group to something a little bit more meaningful.

27:16 Now having your groups is a great way of organizing your checks.

27:21 So I'm going to go ahead and add the check that--another geometry on geometry check...

27:28 ...where I'm looking for my building footprints that are not within my parcels.

27:36 Again, set my spatial relation; use the NOT option, and let me rename this group.

27:48 So now that I have my checks configured, I can save this out into a batch job file.

28:03 Let me just navigate to my folder.

28:15 Now that I have this batch job file, I actually have some other checks that I have configured in another batch job file...

28:21 ...that I want to bring in to this one.

28:23 So I can insert that by going to the Insert button, selecting my other batch job...

28:28 ...and now I have all the checks from the batch job I just created and this other batch job I inserted.

28:37 So I kind of wanted to highlight some of the checks that I configured, so here are attribute checks.

28:43 Now, these are basic execute SQL checks.

28:45 Notice that I was able to give it a title, select a feature class, and pass a simple WHERE clause.

28:54 I also have this duplicate geometry check.

28:59 Here I'm looking for duplicate address points. I'm comparing features within the same feature class.

29:04 I have the ability to compare attributes, so if they have the exact same attributes then I know it's an exact duplicate.

29:11 And I also have the option to ignore feature-level metadata attributes, like the last editor and last update fields.

29:21 And then also I have, in my first demo, shown you how to configure the geometry on geometry check...

29:30 ...looking for the address points not within the building footprint.

29:32 Now what I didn't tell you is that my Address Points actually has a field, Address Type, that identifies that point...

29:40 ...whether it should be within a building or within a parcel.

29:43 So this is where I can define at the attribute level what I want to validate.

29:47 So here I have my address points where the address type is building, and I'm comparing it to building footprint.

29:53 And then this one here, I have my address point where my address type is parcel and comparing it to my Parcel feature class.

30:05 So now that I've combined both batch jobs, I can save that out to a new batch job.

30:14 And now that I have my batch job, I want to be able to run this batch job on my data...

30:17 ...and I want to be able to store these errors in a session.

30:20 So let me go ahead and start a new session; I'm going to use my Batch Validate session.

30:26 Sessions are a great way of organizing your errors.

30:29 Now, what I like to do is have a session for automated QC then have a separate session for visual QC.

30:36 And if I have multiple people performing visual QC, then I'll have a session for each one of those technicians doing the visual QC.

30:45 To run that batch job, I'll go to Batch Validate.

30:49 I'll add my batch job; it will load up all the checks. I'll validate my batch job.

31:01 What validate does is to ensure that all of the feature classes that I'm validating are available to me in the table of contents.

31:08 And I have the option to validate on a selection set of feature; the current extent; a definition query...

31:13 ...[unintelligible] definition query in ArcMap table of contents; the full database.

31:17 And if I'm in SDE, I have the option to validate on the features that have been edited.

31:22 It'll compare the current version to the parent version and identify those edited features.

31:28 So I'm going to go ahead and run this on the current extent, and it's running all of those checks.

31:33 And when you run it in a batch job, it'll automatically write those errors to the Reviewer table.

31:39 So it has found 273 errors within this extent.

31:46 I open up the Reviewer table, and you can see that I have a table full of records.

31:52 So I can group by that check title...

31:55 ...and I can see that I have majority of my errors are with my domains and within the Road Centerline feature class.

32:06 So that's a way that you would probably want to run your automated checks is through batch.

32:13 So through Batch Job Manager, you can create your batch job...

32:17 ...and then through Batch Validate over here, you can run your batch job.

32:21 So now, do we have any questions on automated QC? Yes, sir.

32:30 [Audience question] Can you set it up so that it automatically...that does a batch on a certain extent...

32:40 ...or do you have to manually [inaudible] out locations for...if somebody's going to do editing?

32:47 I mean, can you just bring in...

32:50 Yeah. So the question is, Can you set up Reviewer to automatically validate an extent?

32:55 Now I think maybe working with Workflow Manager, with Workflow Manager you have your area of interest.

33:01 You can run the... We integrate very well with Workflow Manager...

33:06 ...so we have custom steps that you can configure in Workflow Manager to execute your batch job.

33:12 So you can, you know, using Workflow Manager, go to the area of interest and then...

33:17 ...run that batch job on that area of interest. Yes, ma'am?

33:22 [Audience question] Do you run the same batch job on multiple feature classes at once, or would you...

33:27 Can you run the batch job in batch, or do you have to do it on separate, each one separately?

33:34 So the question is, Can you run the batch job in batch? I'm a little...

33:40 [Audience question] If you have 14 separate feature classes that all need to be validated at the same time...

33:44 ...like, you know, maybe midnight one day, can you run it all at once, or do you have to do it...

33:49 So what you would do is configure a batch job with all of the checks for all those feature classes...

33:54 ...and then you run that batch job once. You can create multiple batch jobs.

33:58 Maybe you want a batch job for each feature class; it's a different way of organizing your checks.

34:03 And you can run, if you use the Reviewer service, you can select all of those batch jobs...

34:08 ...and, you know, schedule it to run at midnight. Yes, sir.

34:16 [Audience question] I'll talk on that screen where you can check Current Extent or Full Extent.

34:24 Do you have the choice to like do it within a polygon feature class?

34:32 So if...so the question is, Do you have the option to validate features within a specific polygon? And...

34:41 [Audience question] Just like we, you know, we're water, just like these examples you've been showing.

34:47 And like we put a project in, we put what we call a progress polygon around it...

34:53 ...and so you can identify everything related to that project.

34:56 So you could then, after you've drawn that polygon, do your data check within that as opposed to the whole extent because...

35:05 Yeah. [Audience question, cont.] ...within your extent, there could be features that are ranked, you know, 1950...

35:11 ...that aren't going to have those attributes, but you don't know.

35:13 So the question is, you know, if you have a specific extent that you only want to validate and not the features outside of that...

35:21 Well, not extent, but the polygon.

35:23 ...with ArcGIS Desktop, running Reviewer in ArcGIS Desktop, you do not have that ability.

35:32 What you would have to do is select your features within that polygon first and then validate on your selection set.

35:39 I think with Workflow Manager, you can pass that area of interest, and it'll just validate those features within that polygon.

35:46 And then at 10.1, we have--you know, we currently do have a GP tool where you can run your batch job...

35:54 ...but at 10.1, we'll be able to pass that polygon and just validate those features within that polygon.

36:04 Alright, Jay? Oh.

36:05 [Audience question] Can't you do it against a version as well?

36:08 Yes. You...whatever data is-- So the question is, Can you run your checks against a version? Yes.

36:16 So if you're in an SDE database and you have a version, whatever version is loaded in your ArcMap table of contents...

36:22 ...it'll validate those features. Alright. Thank you. Jay?

36:29 Alrighty. So we've talked about automated data review...

36:35 ...and you've seen how quickly we can kind of rip through a database in finding errors...

36:39 ...a lot more errors than your technicians will be able to fix probably in one day.

36:42 But let's talk a little about, you know, the things we can't do with automated checks.

36:47 You know, even though we have 42 different checks, there's still going to be some things that have to be done visually.

36:51 So we'll talk about how we can help facilitate that process, and we'll do a demo showing some of those tools as well.

36:58 So really, you know, visual data review, a lot of folks probably have traditionally used that for their entire QC process.

37:05 And, you know, in general, you know, it's a decision-making process; it needs to have folks who are understanding of the data...

37:11 ...can make consistent judgment calls today, tomorrow, the next week...

37:14 ...however long this process is for visually interrogating our data.

37:20 It is a time-consuming process, though, right? Takes a lot of time to do.

37:25 So we'll talk about some ways a little bit later about how we can kind of help facilitate a little bit there.

37:31 Data Reviewer does help, though, kind of organize and provide a structure...

37:34 ...for how you might go about doing visual QC of your data.

37:38 Has some tools for helping managing that process, as well, along.

37:42 And again, because you're using a single workspace to log all your errors...

37:47 ...errors you would get through visual QC are then stored along in a single repository so that when you--

37:53 ...you have a complete record of the errors from your database.

37:58 So one of the tools we have to help kind of manage this process is our overview window.

38:03 This is just a simple window in ArcMap that allows you to visually look and see the extent of your study area...

38:10 ...divided into smaller pieces using a grid, and then some simple navigation tools to move a technician who's doing visual QC...

38:18 ...from one cell to another cell to another so they don't miss anything as they're doing their QC review.

38:25 We also have the ability to toggle these cells as having been reviewed or not reviewed.

38:29 So you can...you know, obviously you may not finish visual review in a single day...

38:34 ...so the next morning you come back in, you'll know where to pick up again.

38:39 Now because it is so labor-intensive to do visual QC, a lot of organizations don't do 100 percent review of their data.

38:47 So at 10 we introduced a new check, the random sampling check, which allows you to do a random sample on a...

38:53 ...using a number of different methods to arrive at a random sample of features, the idea being...

38:58 ...I'm going to take a sample; some organizations may say, hey, I need a, you know...

39:02 ...I have a, you know, take 500 features. Other ones might be a percentage of the features.

39:08 We'll also have the ability to autocalculate that random sample size based on a confidence level and a margin of error.

39:14 So if you don't know how many features you need to, what would be a statistically valid sample...

39:19 ...you can use the autocalculate to figure it out for you based on your confidence.

39:25 The random sample results are written to the Reviewer workspace like an error.

39:29 That way you can use the Reviewer table to navigate to those randomly selected features...

39:34 ...visually review them, and mark them pass/fail.

39:39 Now later on, we'll show you some of the reports that can leverage that information...

39:42 ...to help you have a concise report of whether the data's passed or not from your visual review.

39:47 And finally, also, by polygon grid. So a lot of organizations don't do a random sample of features.

39:52 Maybe they have a grid that they do a random sample of the grid and then 100 percent visual QC of the features within that grid.

39:59 So we also support that type of sampling as well.

40:04 In an enterprise environment, folks may want to be able to compare maybe different edit versions to one another.

40:10 At 10, we introduced a Version Differences tool...

40:14 ...similar to the core tool where you can check the differences between parent and child versions.

40:20 This one actually allows you to compare sibling versions.

40:23 So if I have multiple editors working, and as the QC manager, I want to see what they're working on...

40:27 ...to make sure they don't step on each other's toes, I can use the Version Differences tool...

40:31 ...to compare those two sibling versions before they post and reconcile.

40:36 The results are written to the Reviewer workspace so I can go back and review them as I would any other error.

40:44 Other ways of writing errors to the Reviewer workspace include the Commit To Reviewer Table tool.

40:50 So essentially, any process that produces a selected set of records--maybe it's a network trace, maybe it's a model.

40:57 Whatever that selected set of records are, you can commit those as errors to the Reviewer workspace.

41:03 Also interactively you can commit them, as well.

41:09 The Capture Missing Features tool is for those times you know there's just data missing...

41:13 ...and I really quickly need to note it, provide a little bit of descriptive information about what's missing...

41:19 ...and how severe is it of an issue so that my technicians can go back later and add that feature.

41:25 If I'm particularly energetic, I have some sketching tools that I can sketch point, line, poly geometries.

41:32 And if I'm a good enough sketcher, maybe the technician just takes that, copies it into the database, and they're down the road.

41:41 Alright, so let's go ahead and we'll take a quick demo of the visual QC process. Michelle?

41:48 Thanks, Jay. So as Jay mentioned, automated QC is very useful, but it cannot catch everything.

41:57 So I've been assigned to review map sheet 2217...

42:02 ...and my boss has asked me to check my building footprints against the image data that I

have.

42:09 So what I'm going to do is create a grid over my map sheet area, and this will help me perform a systematic visual QC of my data.

42:19 So I've just opened up my Create Polygon Grid wizard...

42:24 ...and so what I need to do is browse to the location where I want to store my grid, give my grid a feature...

42:31 ...give my grid a name, then I'm just going to use this first option to drag a bounding box over the area that I want to create my grid.

42:44 So I'm just creating that box over my map sheet 2217.

42:49 And I want to create a 4 x 4 grid, and then I would click Finish.

42:53 Now, I've already created a grid for this demo, so I'll just go ahead and cancel this and turn on the grid that I've created.

43:01 This is a 4 x 4 grid and then I can use this grid in my overview window that Jay mentioned.

43:08 And if I double-click, if I use the Select tool and double-click on my grid cell, it'll navigate me to that location.

43:16 Now if you notice over here, this grid cell's already shaded, so I've already reviewed that area...

43:21 ...so I know that I need to move on to the next cell.

43:25 So I'm reviewing this area, and I can already see that I've a missing building footprint.

43:31 So then I can use the Flag Missing Feature button to identify this missing footprint.

43:36 I select my feature class and then enter a status of Add, and then that record will be written to the Reviewer table.

43:45 Once I've finished reviewing this area, then I can mark this cell as being reviewed by just clicking on the Change Cell Status.

43:55 And then I can use the navigation tools to move on to the next cell.

44:01 And in this area, I can see that I have some building footprints that are not aligned to the buildings in my image.

44:08 Now, the automated QC would not have caught this, right?

44:11 So the checks that I configured in my automated QC is looking for my points within my building footprints; that's okay.

44:19 And I also had a check looking for my building footprints within my parcels, and that's okay.

44:25 So the automated check would not have caught this, so this is why visual QC is important.

44:31 So I can select those features that are misaligned and commit them to the Reviewer table, saying that they need to be moved.

44:55 So now that the rest of...I've looked at the rest of this data, and it looks okay, so I can mark that as being reviewed.

45:01 So you kind of get the idea of how the systematic review works with Data Reviewer.

45:07 So now that I've done my visual QC, I wanted to show you the records that were written to the Reviewer table.

45:17 So if I double-click on the record, it will navigate me to that location, and this first one here is a missing feature.

45:23 So here's my missing feature, but you don't see any indication of that missing feature...

45:27 ...so what you can do is click on the Symbolize Reviewer Feature Records button...

45:33 ...and it'll symbolize all the records in the Reviewer table in your map display.

45:37 So there's my missing feature.

45:42 So as Jay had mentioned, performing a visual QC is very time-consuming...

45:47 ...and if you have lots of data and maybe not enough resources...

45:51 ...what you may want to do is perform a sample of your data and then perform the visual QC on that sample.

45:59 So I would like to show you that sample, sampling check that we have.

46:04 So let me end this session and start a new session for sampling, and I'm going to configure my sampling check.

46:17 And I want to sample my building footprints.

46:19 I'm going to use this autocalculate method and then use the default parameters for the confidence level margin of error.

46:29 Then I'm going to run this on my entire database...

46:33 ...and it'll randomly select the features within this building footprint feature class.

46:40 And then it'll write all the records, if I so choose, to the Reviewer table.

46:47 So there's over a thousand records written to the Reviewer table.

46:51 Now if I open that Reviewer table, I'll have all of my records in here.

46:58 I want to point out this review status. The review status says, "Needs review."

47:02 So once you've done your sampling, all the records in there will have a status of Needs review.

47:08 And then you would double-click on that feature, inspect it, and then, if it looks okay, then you

can update the sampling status.

[47:17](#) So if it's acceptable, you can pass it, and then the information in the Verification column will get updated...

[47:28](#) ...and also the review status will be updated.

[47:31](#) So then you can move on to the next one, and if that's acceptable, you can update that status.

[47:41](#) So you get the idea of how the sampling works.

[47:44](#) So if this one fails, it fails, but note that the Verification column did not get populated.

[47:51](#) So this way, when the person goes in and makes a correction to these sampled features...

[47:55](#) ...they can populate the correction information and then the verification information separately.

[48:03](#) So what this sampling method that we used, autocalculate, it tells...

[48:08](#) ...based on the number of features you have in your feature class, it determines your sample size.

[48:12](#) And also, this acceptable failure threshold identifies the number of features that you're allowed to have...

[48:22](#) ...before this dataset's considered failed.

[48:25](#) And we have a report that we can generate, and I'll be showing that in the next demo, so it'll tell you whether it passes or fails.

[48:35](#) So those are two ways that you can use Reviewer for performing visual review of your data...

[48:40](#) ...using the overview window to systematically visually QC your data and then also using the sampling check...

[48:46](#) ...to get a statistical sample of your data and review those features individually that way through the Reviewer table.

[48:56](#) Any questions on the visual QC?

[49:05](#) Okay. Alright. So now we've been talk--

[49:10](#) We first started talking about automated data review, logging lots of errors to our Reviewer workspace.

[49:15](#) We've logged a bunch of errors to our Reviewer workspace based on our visual QC review.

[49:21](#) So let's go into a little more detail about how we work with those records, and more importantly...

[49:25](#) ...how do we summarize and report those findings out, you know, to other folks, maybe contractors or to management.

49:33 So as you've seen in the demos, the Reviewer table's really an interactive table.

49:37 It allows you to...there's a relationship between those errors and features in your map...

49:43 ...so double-clicking on them navigates you to the error.

49:45 The value there is that as technicians work through the data, they can very quickly get to the problem that's at hand.

49:53 We also talked about the verification statuses, the ability to--sorry.

49:59 ...to keep the life cycle of that error through its--sorry. ...to keep track of its error through its life cycle.

50:06 So we've changed, you know, as it's being corrected, we'll note that in the database and then we'll have the verification as well.

50:13 And it's all being tracked in a table.

50:15 You might have noticed we were also grabbing the user name and a time-date stamp.

50:18 So we can kind of keep track of our errors actually being fixed in the database or not.

50:24 You know, we can very quickly look at the Reviewer workspace to see we've got a bunch of errors; where are they at in the process?

50:30 Are they getting fixed or not?

50:37 We've already seen an example of some of the summarizing work we can do with Reviewer, with the Reviewer table...

50:43 ...so simply dragging and dropping, dragging a column to the gray title bar there allows us to group our errors.

50:49 And it's really helpful to kind of get a first-blush look at the trends in our data.

50:52 Are we seeing lots of errors in certain feature classes?

50:56 And by further subgroupings, we can kind of look at it and think, well, are these systematic errors in my data.

51:03 We had an earlier example of that domain check problem. That might be a real simple fix.

51:08 Even though we may have hundreds of records in error, that might be something we can fix quickly.

51:13 Whereas when we don't see systematic errors, that might mean for a manager to sort of think about...

51:17 ...well, this might take a little while to fix.

51:23 One of the really neat things we brought out at 10 are a whole host of new reports that are generated for you...

51:28 ...out of the Reviewer workspace.

51:30 We'll demo this tool, but essentially what it does is allows you to create an automated report in Excel...

51:36 ...and break it down by different categories, by feature class or origin table, by subtype.

51:41 We have the sampling report as well.

51:43 And the value here is that it's literally, click, click; there's your report.

51:46 It's in Excel so you can then go and customize it in some way if you want to do further summarization or filtering of those results.

51:54 It's a presentable report, though, right out of the box.

51:59 Alright, so in this demo, we're going to work with the Reviewer workspace; work with the table a little bit...

52:04 ...work some errors through their life cycles, as well as run a couple reports to show you how that works. Okay? Michelle?

52:13 Thank you very much, Jay.

52:15 Alright, so here I have my Reviewer table open, and this is where we store all of our errors...

52:20 ...and it's kind of the heart of Data Reviewer.

52:23 This is where you keep track of your errors; the information about your errors...

52:26 ...so the title, the feature class, the check that was performed...

52:31 ...and if you use severity, you can log the severity of the error, whether it's a critical error and needs to be fixed immediately...

52:40 ...or you know, if it's maybe not so critical.

52:43 It keeps track of who found the errors and also keeps track of the correction status and the verification status...

52:51 ...so that's the error life cycle that Jay was talking about; who made the corrections and then who verified those corrections.

52:59 So with the Reviewer table, you can double-click on a record; it will navigate you to that feature...

53:04 ...and then you can go in and make your correction to that feature using the core tools.

53:10 Now, once you've made your correction, then you can keep track of that correction...

53:17 ...by right-clicking on the record and enter your correction status.

53:22 So I've resolved this, and then it will update the correction status information.

53:28 So I briefly showed you how you can group records in the Reviewer table.

53:34 One thing that I like to do is group my records by feature class and then I can sort them by the object ID.

53:42 And this helps me focus my correction on each feature, if that makes any sense.

53:48 So say, if you notice here, my object ID here is 467 and 467, so this one feature has two errors...

53:56 ...so I can fix both errors at the same time.

53:59 And if I fix both errors at the same time, I can select both records and then enter in the correction status at the same time.

54:10 So by grouping the record, it helps you focus your correction process.

54:14 You may want to focus based on feature class, which I have done here, or maybe you want to focus it on the type of checks.

54:20 Maybe you want to fix all your domain errors first or your, you know, attribute errors.

54:26 If you group multiple feature classes, or multiple records, so say by feature class and the check...

54:40 ...you can get a summary of the errors that you have and then you can export this into a simple text file by going Generate Statistics.

54:58 And it'll create a text file of basically that summary that you have in your Reviewer table.

55:05 And then, if you want to, you can e-mail this to your manager so that he can get an idea of what those errors are.

55:11 But if you want a little bit more detailed report, we do have that ability, new at 10, to generate reports.

55:19 So here we can generate a report by--here I would like to generate a report by feature class, which is origin table.

55:32 And it'll create this report in Microsoft Excel.

55:38 And here I have my feature class, Address Point; my checks that I ran against that feature class; and the title.

55:46 And then it keeps track of the features that I validated, the number of errors that were found...

55:50 ...and then it'll give me a percent accuracy of that check. And I can get the percent accuracy at the feature class level too.

55:59 Here it's 99 percent, but if you notice that my domain check was at 97 percent.

56:07 So this is a great way of identifying kind of the accuracy or the quality of your data...

56:15 ...is by running these checks and then creating these reports.

56:19 Now I did that sampling in my previous demo, and I wanted to show you how you can create a report out of that sample...

56:28 ...that identifies whether or not your data passes or fails.

56:31 So my sampling report is in session 5, so I can select which session that I want to create my report on.

56:40 I'm going to select the sampling report, and because there are four different methods of doing your sampling...

56:51 ...it creates a worksheet for each one of those.

56:54 So I did the autocalculate method, and here I have my building footprint; it tells me the total number of features I have...

57:01 ...the sampling number, and then the number of errors I'm allowed to have for this dataset to be considered acceptable.

57:10 And now it tells me, based on the review that I did, I had 37 errors, so based on that, this passes.

57:20 So this tells me whether or not, as long as my QC has been complete...

57:24 ...tells me whether or not this data is acceptable or not acceptable.

57:27 It kind of gives you that pass/fail information.

57:36 So with Data Reviewer, you're able to perform your automated checks and then use tools to perform a visual QC.

57:44 All of your errors are stored in the Reviewer table, and then you can generate a report that gives you like the quality of your data.

57:54 Do we have any questions at this time? Yes, sir, in the back.

57:59 [Audience question] If somebody finds errors in like a statewide database...

58:05 ...can they send that error to somebody that's working that one [inaudible], working that same data...

58:10 ...and [inaudible]...they zoom to the errors then?

58:15 So the question is if you have a large dataset and maybe somebody else is managing that dataset, and you find an error in it...

58:24 ...and you want to be able to send that information to the user and then they can look, zoom in to that location to see the error.

58:31 You can export the features in the Reviewer table. If they don't have Data Reviewer...

58:37 You'll have to export the features in the Reviewer table.

58:40 If they have Data Reviewer, you can just export those records to another Reviewer workspace...

58:45 ...and send them that Reviewer workspace, and then they can just import those records and then zoom to it.

58:50 But if they don't have Data Reviewer, by symbolizing the records here, you add these feature classes...

58:59 ...and then you can just export these feature classes using just the core functionality, Data Export.

59:05 And then you can share that feature class or shapefile, whatever you export it to, to that user.

59:16 Also, as Jay was reminding me, we do have the ability to--let me ungroup my table.

59:22 ...ability to export the entire Reviewer table into an Excel spreadsheet.

59:30 So it will take all the records in here and just copy the contents into the Excel spreadsheet...

59:34 ...so then you can share that spreadsheet to the user as well.

59:41 Okay.

59:42 [Audience question] Can we set up the Reviewer workspace in SDE so that other users can also see the edits?

59:49 Yeah. The question is, Can we set up the Reviewer workspace in SDE? Yes, you can.

59:53 And actually, we have two white papers that talk about the best practices of setting up your Reviewer workspace...

59:59 ...in a SQL Server SDE and then Oracle SDE. So those can be accessed from our website.

1:00:07 [Audience question] And then can you also set up like the sessions you have to specific group of users to access specific sessions...

1:00:13 ...and [inaudible] specific groups?

1:00:18 The question is, Can you set up your sessions for specific users?

1:00:22 [Audience question] Meaning if you have the QC and then the QA and [inaudible] sessions.

1:00:27 Specific group of users can access only those sessions--

1:00:30 Oh, setting permissions to certain sessions for the users. I don't...

1:00:35 At this time, we do not have that ability to have only specific users access certain sessions.

1:00:44 You may be able to do that through the RDBMS, but that would get pretty messy.

1:00:48 'Cause you're talking role-level security at that point. Yeah.

1:00:52 [Audience question] But can he have like roles who can do a specific [unintelligible]?

1:00:57 If you have like a multiuser, could [unintelligible] QC [unintelligible], for example, [unintelligible].

1:01:08 So you can set privileges in the Reviewer table to people who can only write to the Reviewer table...

1:01:16 ...or who can only read to the Reviewer table, so you can do it at that level.

1:01:24 Question over here?

1:01:25 [Audience question] [Unintelligible] feature class name [unintelligible] before taking...before running a session?

1:01:40 Could you repeat that question?

1:01:41 [Audience question] If you have [unintelligible] client, finds that are the same construction every time.

1:01:52 So you have to run it, but every time with another finding. So feature class name.

1:01:59 So, yeah. So the question is, You have data coming in, but the feature class name is different...

1:02:06 ...but you want to run the same checks.

1:02:08 So with Data Reviewer, when you build your batch job, it is important to have the same schema...

1:02:13 ...so your geodatabase name could be different, but your feature class name will have to be the same...

1:02:22 ...for it to automatically re-source. I believe at 10 we do have the ability to, through Batch Job Manager and batch validate...

1:02:33 ...you can change the workspace of the checks, but you can also change the feature class.

1:02:39 But you can only do that at one feature class at a time.

1:02:42 So it's a little bit of a lengthy process depending on how many checks you have that are pointing to that feature class. So...

1:02:52 [Audience question] Could you reference the alias instead of the actual feature class name?

1:02:56 Actually, Reviewer's based on feature class, so it will...

1:03:00 When you configure your check through the dialog here, it'll list the feature class name.

1:03:10 Question right here?

1:03:11 [Audience question] I see that you created your sessions in ArcInfo, but can you give this to a technician with ArcView...

1:03:18 ...and have the reference session?

1:03:19 Yes. So that was a good question. He noted that I'm using ArcInfo here, but was wondering if a technician...

1:03:26 ...if we were to pass this on to somebody else who's using ArcView, could they use it?

1:03:30 Actually, Data Reviewer recommends ArcEditor license, but...

1:03:35 The reason why is because if you want to go and make the corrections, you need to be able to edit the data.

1:03:41 But if you're just doing quality control and just finding the errors, ArcView license is okay too.

1:03:48 Another question?

1:03:49 [Inaudible audience question]

1:03:52 The question is, Can you search for geometric network errors?

1:03:55 Did you notice that I opened up that connectivity rules check?

1:04:00 Little bit of foreshadowing there for you.

1:04:03 So we have a connectivity rules check that will validate your features against your geometric network rules...

1:04:10 ...if you have your rules defined. Now, if you don't have rules defined, then it will let...

1:04:15 ...then you won't be able to select anything here. [Audience comment] Okay.

1:04:22 And then also, geometry on geometry, some of the examples that he gave on geometry on geometry check...

1:04:27 ...like looking for hydrants that are not connected to a lateral and stuff, you know, you can use some of the other checks that we have here.

1:04:33 [Audience question] I wasn't sure if, like that specific example was actually looking for intersections...

1:04:39 ...or actually the geometric network...

1:04:41 It's just looking for the intersections, so that is a good...yeah.

1:04:46 [Audience comment] It's like...one thing that's hard to spot is where we have a service saddle that's on a...

1:04:53 ...has to be on a main where the main is broken. And sometimes that's a little hard to identify [unintelligible]...

1:05:01 ...seeing if there's a node there on the main.

1:05:06 So, yeah. Any other questions?

1:05:14 Alright, so let's go ahead and we'll wrap this up a little bit.

1:05:17 So kind of in summary, you know, we talked about ArcGIS Data Reviewer. It is a standard extension to ArcGIS Desktop.

1:05:25 And it really, as we've kind of discussed, it really covers the whole quality control process...

1:05:31 ...from the ability to automate your business rules...

1:05:34 ...to make that an automated process rather than a manual process and save you a lot of time.

1:05:39 We also showed you a lot of the interactive tools that we have out of the box...

1:05:42 ...that allow you to facilitate the drudgery of doing manual data reviews.

1:05:49 The value being that we're writing all these errors all to a common workspace so that we can track them holistically.

1:05:55 So these might be errors coming from multiple sources, but we have a single place to store them...

1:06:00 ...where we can report out, keep track of how we're doing in progress moving forward and correcting our data.

1:06:07 You know, the value here is that, because these are automated processes, you can save a lot of time...

1:06:11 ...so even if you're a one-person shop, this can save you a lot of time through automation.

1:06:15 If you're a large team, the value is that you can have QC defined one time, those rules distributed across the team...

1:06:22 ...and everybody's doing the same QC. It's not this person doing it one way and that person doing it a different way.

1:06:28 We have a common framework for doing this work.

1:06:33 So lots of resources available for you. There's eval copies obviously available for you to test out.

1:06:39 The Data Reviewer poster that I showed a little bit earlier with all those 42 different checks is available for download as well.

1:06:45 We have both instructor-led training as well as a couple free classes on the Virtual Campus you might want to take a look at.

1:06:51 One is a web course; one is a training seminar that's been recorded.

1:06:55 We have the Reviewer Resource Center as well we released at 10, so you'll find things like...

1:07:00 ...we've got a very active blog that we'll go through in excruciating detail about how you might use Reviewer...

1:07:05 ...for water, wastewater, for metadata checks, things like that.

1:07:08 We also have a number of gallery items you might want to download--video, demos, and all that kind of fun stuff.

1:07:16 Questions? We have an alias, datareviewer@esri.com, where if you have questions...

1:07:20 ...you can send e-mails to us and we'll get back to you and let you know how that works.

1:07:25 And the other thing to note too, talking about the resource center, is if you visit the resource center...

1:07:30 ...and some of the other vertical industries there, we're really trying hard to whenever there's an industry model, for example...

1:07:36 ...like with the water, the Water Utilities Resource Center, we'll build a number of checks...

1:07:41 ...based on our own SME input on that industry model and publish those out.

1:07:46 It's a straw man, a starting point. It's not going to be a hundred percent solution...

1:07:50 ...but it'll at least get you started and start thinking about how you might want to implement it.

1:07:53 So we have the grand plans for, you know, sort of marching forward and also putting QC rules up in other industry areas as well...

1:08:00 ...so kind of keep an eye out for that in the area of your interest.

1:08:05 So we've got a lot of Reviewer sessions going on at UC.

1:08:08 So right now we're kind of halfway through; we have a number of demo theaters.

1:08:13 We didn't really talk about assessing positional accuracy, so we have a demo theater where we'll be talking about...

1:08:19 ...how you can assess positional, both horizontal and vertical, accuracies of our data--of your data in demo theater.

1:08:25 We have a tech workshop happening tomorrow--or actually, I wrote ahead, so tomorrow morning...

1:08:29 ...we'll be talking about what's happening at 10.1 if you're interested, first thing in the morning.

1:08:34 We're going to rerun this guy tomorrow, and then what do we got...oh, the custom check demo theater.

1:08:38 So if you have those 42 checks and they ain't going to work for you, come talk to Shankar about how to develop custom checks...

1:08:46 ...but still implement them within the Reviewer framework so it's a part of your overall QC process.

1:08:54 We also have folks down in the Geodatabase Island; we've got a couple workstations set up.

1:08:58 We're also doing wastewater health checks this week, so if you're a water utility...

1:09:04 ...and you happen to have a sample set of data in your pocket, you can find some time, come down and talk to our guys.

1:09:09 If there's time; I don't know if there's a lot of space left.

1:09:11 But they'll run through your data with you using the industry checks that we publish out on the water utilities site...

1:09:17 ...give you some feedback about how that sample data looks.

1:09:22 Alright, so a little bit of logistics also.

1:09:24 So the session feedback's gone digital this year, no little, little notecards, so when you have a moment...

1:09:30 ...please avail yourself of this website and let us know how we did. We appreciate the feedback.

1:09:36 Any last questions before the end? We good? Got a couple questions right here.

1:09:41 [Inaudible audience question]

1:09:47 No, you can create it with ArcView.

1:09:50 Okay. Any more questions? In back?

1:09:52 [Audience question] Early on, you mentioned it can work with large datasets. How do you define a large dataset...

1:09:57 ...or what do you consider a large dataset?

1:09:59 How big is Hong Kong Lands?

1:10:00 So the question was, you know, we say "large" datasets. You know, let's talk about how large is large.

1:10:06 Michelle, you did some pretty large work with Hong Kong Lands, right?

1:10:10 I don't remember how many features was that, but at least recently we were just validating some features.

1:10:16 That was like 1.7 million. That's large. But probably anything over...You know, if we just have, at a feature...

1:10:24 ...just one feature class, I would say over a couple hundred thousand features I would consider large.

1:10:31 How big's a batch job?

1:10:34 Jay was asking how big are the batch jobs. Batch jobs vary.

1:10:39 They can be really big where you have hundreds of checks, even thousands of checks

depending on your business rules.

1:10:46 Question in back?

1:10:47 [Audience question] Does it work with version 10 and 9.3? Do you have those versions on it?

1:10:53 So the question is, Does this work with version 10 and 9.3? At...

1:10:58 Before 10, Data Reviewer was called PLTS GIS Data ReViewer...

1:11:02 ...so there's a slight name change, and yes, it's available at 9.3; 9.2, 9.3, 9.3.1, and 10.

1:11:09 So at 10, we just have a renaming, ArcGIS Data Reviewer.

1:11:13 Question over here?

1:11:14 [Audience question] So does Data Reviewer have all the cross-functionality now, or is it separated?

1:11:21 So the question is...

1:11:23 [Unintelligible audience question]

1:11:28 Oh, oh.

1:11:29 The other components of PLTS? It's now called Production Mapping.

1:11:36 Right. We had a big name change at 10, so some folks are familiar with Data Reviewer...

1:11:41 ...as a component of Production Line Toolset prior to 10.

1:11:46 So at 10, the whole box got shuffled, and now PLTS Foundation is now known as Production Mapping...

1:11:52 ...and there's some other minor name changes, but Data Reviewer...

1:11:54 [Unintelligible audience question]

1:11:58 That's why we can change our business cards every time, but it's still a component of Production Mapping.

1:12:04 It's also available stand alone. We really didn't talk too much about it in the past about a stand alone...

1:12:08 ...but it's always been available stand alone for folks who don't need all the other bits and pieces of Production Line Toolset. Question?

1:12:16 [Audience question] In one other session, they mentioned in 10.1 that [unintelligible] date and editor stamps.

1:12:26 Is there any way to implement that in 9.3.1 now, or is that...

1:12:32 Yes, you can, using Production Mapping, or PLTS.

1:12:36 [Audience question] So you can?

1:12:37 Well, so PLTS has the ability to keep track of feature-level metadata.

1:12:41 You can identify what fields you want to store the last editor and the last update fields, and by setting up what we call...

1:12:50 Oh, was it the knowledge base tables? Yeah, knowledge base. Yeah.

1:12:53 So if you go to Geodatabase Island, there's a girl there, Amber, she's working...I think she's working there...

1:13:00 ...but she can help you answer that.

1:13:03 And the thing to remember is that it's feature-level meta-..

1:13:06 What PLTS does is way...it does a lot more than just track the last editor.

1:13:11 Essentially, you may have half a dozen different feature-level metadata fields...

1:13:16 ...and you can have those autopopulate with Production Line Toolset or Production Mapping at 10.

1:13:22 Any other questions? Alright, thank you very much.