

# What's New in the Geodatabase at ArcGIS 10

This session will discuss and demonstrate the key geodatabase functionality in the upcoming ArcGIS 10 release, the next major release of ArcGIS. This session will focus on the new improvements in the geodatabase at 10, including improved data access, replication enhancements, additional topology rules, schema changes, and much more.

<http://video.esri.com/watch/65/whats-new-in-the-geodatabase-at-arcgis-10>

---

## Video Transcription

**00:01** Okay. Anyway. So what we're going to do...we're all impressed that you didn't all go out and have a cheeseburger...

**00:05** ...or something for lunch on a beautiful day like today.

**00:08** You must be really interested in this stuff.

**00:10** So we'll try and do it justice for you, make it worth your while.

**00:14** Listen, things to pay attention to, or what have we been doing on the geodatabase team for the past...

**00:18** ...oh, I don't know, five plus years or so?

**00:21** What we've been doing, you know, sort of starting at ArcGIS 9.1...

**00:25** ...is we've been really trying to finish up the initial information model of the geodatabase.

**00:33** At 9.1, remember we were kind of focused on the transportation domain.

**00:36** That's where the network dataset came into being.

**00:39** That's for transportation.

**00:41** At 9.2, it was actually a very significant release for the geodatabase team and ArcGIS, you know, with all the server technology.

**00:48** But then we also brought out replication.

**00:51** That's what we use and that's the term we use for sharing data.

**00:55** While archiving if you have an enterprise geodatabase so you can see what happened in the past.

**01:01** The spatial types came into being, supported a new terrain model called the terrain.

**01:08** Also brought out some functionality that's used for maintaining your parcel fabrics.

**01:13** The cadastral fabric dataset came out at that time.

**01:15** That's now been rolled into core at 10.

**01:18** And then also remember we brought out high-precision geometry.

**01:22** Okay. So there were a lot of changes that happened at 9.2.

**01:24** At 9.3, we were a little bit more focused on trying to support better distributed workflows...

**01:31** ...as well as improve the versioning experience, okay? with better conflict and change detection management tools.

**01:40** Okay. So, what are we doing now?

**01:43** What are the key themes now that we are focusing on, on basically completing the model?

**01:49** Where again, we continue to evolve and enhance our information model.

**01:54** We weren't done.

**01:55** That's sort of like the first big iteration has been completed.

**01:59** We continue to work on our models, our information model...

**02:02** ...to make it more flexible and powerful for you guys so you can do your work.

**02:07** One of the things we want to do is also help our data sharing technology, replication...

**02:12** ...to make it more flexible, make it easier to use.

**02:16** These were based upon comments and feedbacks that we have gotten from you in the user community.

**02:21** How would you like us to be improving things?

**02:24** Also a key thing at 10 is we want to make the system even more open and allow people to access the data more directly.

**02:35** Say you don't want to use ArcObjects.

**02:37** You don't want to use our desktop applications.

**02:40** How can you get at your data?

**02:43** Okay, so we'll talk about some of the technology we have there to support those sorts of user environments.

**02:49** And then also the never-ending quest to improve our performance and the quality inside the geodatabase.

**02:55** So, let's start hopping through this stuff.

**02:59** The big things inside the information model that we addressed, the main topics were our geometric network model.

**03:05** That's what we use for utilities and the utilities environment, natural resources environment.

**03:12** The network dataset, I mentioned that a little bit before.

**03:14** That's our network model for the transportation domain.

**03:18** Then also topology.

**03:21** Now, more specifically, with the geometric networks, one of the things that we wanted to do was make it easier for people to create...

**03:34** ...design and create their geometric networks as well as add lots and lots of data to existing geometric networks.

**03:42** Now, when we initially released the geometric network at 8.0, remember with the first release of ArcGIS?

**03:49** We had a limit, an internal limit to only being able to build networks that had somewhere on the order of 15 million features.

**03:57** That was a limitation.

**03:58** At that time, nobody had networks that big.

**04:00** You know, the big utility companies, gosh, they had maybe one or two million features in their geometric networks.

**04:06** So we didn't think it was big deal.

**04:08** Well, as time has passed, people have gotten bigger and bigger.

**04:11** Data's just coming at us, you know, this huge pipe.

**04:14** Just more and more data flowing at us.

**04:16** So we now need to be able to handle end users that have just monstrously large networks. Okay?

**04:22** So we've gone and made it such that there is now no limitation on the size of the network that you can initially create.

**04:31** Let's see, also, we have the...we went and reengineered our wizard for going and defining, creating your geometric network.

**04:40** That's been simplified and streamlined, sort of made more modern.

**04:44** That was a nice thing to do it in the 10 time frame.

**04:50** Now, one of the other interesting things that we've noticed in the user community is, there are a lot of people that with time...

**04:57** ...you know, over the past 10 years, they've been adding more and more data into their existing networks. Okay?

**05:04** Now, you know, say, for example, maybe you're a utility company and maybe you go and buy the adjacent utility company...

**05:11** ...and you need to get all their millions of features into your network.

**05:15** Well, previously, we had some...some ways of doing that but it could be a slow process.

**05:21** So one of the things that we focused on during the 10 development cycle was to create a new geoprocessing tool...

**05:29** ...that will allow you to pour data into existing geometric networks very, very rapidly.

**05:36** By that we're meaning we're focused on people that need to add millions upon millions of features to an existing geometric network.

**05:49** Network datasets.

**05:50** This is basically our transportation-focused model.

**05:55** Now, everyone's seen all the flurry of activity that networks, routing networks, and whatnot have gotten on the Internet and...

**06:03** ...you know, with Google and Microsoft, and others.

**06:07** You know, this is a very hot and fertile domain right now.

**06:10** Anyway, what we want to do is take our existing transportation model...

**06:15** ...and continue to improve it and make it even more flexible for our end users.

**06:20** First thing we did was we made the network dataset model itself, you can now version it.

**06:26** Okay. It can be versioned.

**06:28** And much like topology, when you go and edit the features inside your network dataset, inside your transportation network model...

**06:36** ...we'll track the changes and then on a periodic basis, you can then go and have that portion of the network...

**06:44** ...have the connectivity rebuilt. Okay?

**06:46** This is much like topology with how we track areas that have been edited through dirty area management...

**06:52** ...and we support an incremental validate process. Okay?

**06:57** So, want to be able to support in the versioned environment partial rebuilds, that's great.

**07:02** One of the big things that you see these days is being able to support time-dependent traffic data.

**07:09** By that, you know, say you're trying to do a routing application or doing a...

**07:15** ...figure out the best time to drive from, say, here to L.A. starting at eight o'clock in the morning.

**07:20** The driving directions that you get or the traffic you encounter at eight o'clock in the morning...

**07:24** ...is very different from the traffic that you'll see, like, at either noon or at five p.m.

**07:30** You know, we've got the commutes and reverse commutes and whatnot.

**07:33** Previously, we could only handle, you know, just one, what is that free-flow speed?

**07:38** And we would give you directions based upon a free-flow speed.

**07:41** It would not vary by time of day, but now by integrating better with the big data vendors such as NAVTEQ and Tele Atlas...

**07:47** ...we extended the network dataset to support time-of-day-based routing.

**07:53** So the route that we will provide to you will be different depending upon historical traffic conditions.

**08:00** You saw on the big stage Matt Crowder, he gave a really nice demo.

**08:04** He was the one that was talking about LeBron James and Cleveland and health care and all that.

**08:11** That's the location-allocation solver.

**08:13** This is a very sophisticated analysis routine that if you have a bunch of customers or end users or places that you have to deliver stuff to...

**08:21** ...where do you optimally allocate the places to distribute from so that you can minimize drive times and be as efficient as possible.

**08:29** Also Matt showed us the enhanced barriers.

**08:32** Remember that was the snowstorm sort of on the northeast half of Cleveland.

**08:37** Now you can have polygons to define barriers rather than just little point barriers.

**08:41** It's much more flexible.

**08:43** Also, we enhanced the information model of the network dataset to support something we term in...

**08:51** ...we use, internally was termed a hyperedge, that allows you to do very rapid solve operations...

**08:59** ...network analysis operations that can work on a national or continent-wide basis.

**09:05** Think about it for a second.

**09:06** If you start here in San Diego, you want to get to New York City, but you know you're going to be getting on interstate freeways...

**09:12** ...there's still going to be somewhere between 3- and 5,000 street segments, interstate

highway segments...

**09:18** ...that you'll have to traverse to get there.

**09:20** Hyperedge is, what we're doing is, we're taking large collections of these edges or these features...

**09:25** ...linear features that share similar, you know, similar attributes or properties, and we make one big edge.

**09:33** So then to get from say here to New York City, maybe it's only a hundred hops rather than 3- or 5,000 hops.

**09:40** That's how we can make things much faster to do these analysis operations. Okay?

**09:45** That's a very important issue for our user community.

**09:50** Okay, so we've already talked about this.

**09:55** And I guess I got ahead of myself a little bit.

**09:59** Well, here are the further details that you can read at your own leisure.

**10:03** But, yeah, we're basically talking about the new solvers and the enhanced barriers and the long-distance routing with hyperedges.

**10:11** So, that's enough PowerPoint chattiness.

**10:14** We'll see how this stuff really works.

**10:19** Yep. Okay, thanks, Eric.

**10:22** So I'm going to...you okay in the back? Good?

**10:25** I'm going to take you through about three demos that are going to highlight some of the topics that Eric just talked about...

**10:31** ...and show you some of the functionality that the Network Analyst team has been working on and is available in 10.

**10:37** The first demo that I'd like to show is going to highlight the new editing functionality that we have...

**10:42** ...and the ability to track edits with your network dataset.

**10:45** So what I have here is a map of Paris, France, and I have a number of locations that I'd like to find the route for.

**10:52** These might be locations that somebody would like to visit and they want to get the optimal route while they're in Paris.

**10:57** So I can go and solve on that information.

**11:01** And you'll see that it returns me back the route from all the locations starting at my hotel.

**11:07** If we zoom in to my hotel though, I notice that I'm actually missing some roads here.

**11:12** I don't have the most up-to-date data in my network dataset.

**11:15** Now previously I'd have to update the data and then rebuild the entire network dataset, and in this case...

**11:21** ...that could take anywhere from 10 to 15 minutes.

**11:24** But now at 10, like Eric mentioned, we have dirty area tracking with the network dataset, so that makes it a lot easier.

**11:30** So what I'm going to do is turn on another layer which highlights the roads that I'd like to bring into my network.

**11:38** I'll select those roads and then I'll copy and paste them into my street feature class.

**11:47** And when I do that you'll notice that I get these hatched polygons.

**11:51** Anyone who's familiar with our topology model, these will be very, very familiar to you.

**11:55** These are dirty areas.

**11:57** You can see their extent over the features that I imported.

**12:01** So network datasets are now tracking those edits that you're performing.

**12:05** So now that I've done this, I can rebuild my network.

**12:09** And you'll notice that it took just a matter of a couple seconds to rebuild, not 10 or 15 minutes like it would for the entire network.

**12:15** Again, because it's only highlighting those features under the dirty areas.

**12:19** And now when I re-solve, you'll notice that instead of heading down south and taking a road by the river Seine...

**12:27** ...it's actually just heading straight north and using those new features immediately...

**12:30** ...in the exact same edit session into which I added them.

**12:33** So for anybody who's actually editing and maintaining their network data in network datasets...

**12:38** ...this makes it a lot easier to get new edits into your networks and consume them extremely quickly.

**12:46** So the second demo that I'd like to show, we're going to head back...we're going to head over the Atlantic...

**12:50** ...and we're going to go to Upland, California.

**12:52** And we're looking at a shot of downtown Upland here.

**12:54** I'm going to highlight some of the work the network team has done to enhance barriers within Network Analyst.

**13:01** So I have two stops that I'm going to create a route between, one in the north which is my warehouse depot...

**13:05** ...and one in the south which is a restaurant.

**13:09** I can go and create that route and then zoom in to downtown Upland where that is.

**13:16** Now, this route's fine, but I know for example that north of this restaurant on this street there's actually construction going on.

**13:24** So I actually can't take this route as it exists because they've closed the north side of the street to traffic.

**13:29** So I'm going to go and place a new barrier on the north side and re-solve.

**13:34** This is a really nice enhancement that we made to point barriers at 10.0.

**13:38** Previously, it would restrict traversing across that entire edge.

**13:42** We're now tracking the percent along individual edges...

**13:45** ...which allows you to more accurately model real-world situations when you're solving.

**13:52** The second part of this demo highlights line barriers.

**13:56** Upland, like a lot of cities in California, has a farmer's market where a lot of growers and farmers...

**14:02** ...from around the area bring their vegetables and their produce down to downtown Upland so people can get organic produce.

**14:10** If I'm traveling to this restaurant on the day that the actual farmer's market is happening...

**14:14** ...I have to take that into consideration.

**14:16** So if we go and turn on that layer that contains the information of where that farmer's market's occurring...

**14:22** ...you can see that there's a number of streets in red, they're going to be closed to traffic...

**14:26** ...because usually what they'll do is they'll set up shop directly on the street.

**14:31** So I'd like to actually consider that in my route.

**14:33** Now previously what I would have to do is actually create a number of line barrier...

**14:37** ...a number of point barriers around the exterior to try to make sure that I could somehow cordon off this area.

**14:44** But at 10, what I can do is actually load in those line features as barriers to my network and then re-solve on it considering them.

**14:57** And you'll notice what it does is it actually avoids those streets entirely and comes at my restaurant from the east.

**15:03** That's avoiding the closed streets and the farmer's network [sic].

**15:07** The third type of enhancement to the barriers that I'd like to show is the scaled cost.

**15:13** Matt also showed this in the plenary on Monday.

**15:17** There's something about when events happen in the downtown core, the slowdown and the traffic congestion that you see...

**15:23** ...isn't just limited to the streets where the event is happening.

**15:27** It also has an effect that buffers around those streets.

**15:30** So it's not just enough to take into account that I can't travel on these streets.

**15:34** I also need to consider that it's going to be slower for me to actually get to this restaurant because of the traffic congestion...

**15:40** ...and just people walking on the street.

**15:42** So I'm going to go and create a new polygon barrier around this street.

**15:51** And I'm going to go and set its cost to be two.

**15:55** And essentially what this means is that any trip that I take within this barrier is going to take twice as long as streets outside the barrier.

**16:04** And when I re-solve it, hmm, I don't get the answer I'm looking for.

**16:13** What normally would happen is the actual route would miss the barrier completely...

**16:17** ...but it wouldn't prohibit it from actually traveling inside of it.

**16:20** It would take the north side of the barrier, avoid it where possible...

**16:24** ...but it would still allow the route to actually travel inside of it and get me to the restaurant.

**16:27** So what it's doing there is just simply considering the fact that this area represents a slowdown...

**16:32** ...but it doesn't restrict me from traveling in that area.

**16:37** The last demo I'd like to show is how we've integrated time into the network dataset.

**16:42** Like Eric mentioned, we're taking account of historic traffic data.

**16:46** A lot of the data providers that you get your network data from have started to provide historic time traffic data...

**16:52** ...that models what the speeds are on the different roads at different times of day and on different days of the year.

**17:00** We all know that the route that you're going to take isn't just dependent on the network itself...

**17:05** ...but also the time of day that you're traveling and the day of the week.

**17:09** So what I have here is a map of downtown San Francisco, and I've got a number of locations that I'd like to solve on.

**17:17** So when I go and create my route, you'll notice that it's traveling eastward and it gets on Interstate 80 as soon as it can.

**17:24** Because it's decided that the sooner you can get on the interstate, the faster you'll get to the routes that you want to travel on.

**17:31** But this is showing me the ideal route.

**17:33** What if I was actually going to travel on this network and take this route at five p.m.

**17:39** Well before I do that, let's actually go and take a look at what the data looks like at five p.m.

**17:46** I'm going to go and use the new time slider window that's available in ArcMap and turn on my network.

**17:52** When I do that you'll see a number of lines that are highlighted in both green, yellow, and some orange and red.

**17:58** And if you look down at the bottom, you'll see in my TOC I'm highlighting what those different colors mean...

**18:03** ...where green is free-flowing traffic and red in the extreme is stop and go.

**18:09** Now what this is showing me is historically what the traffic is like in San Francisco around the time right now...

**18:13** ...just around twelve p.m.

**18:15** But let's go and take a look at what it's going to look like at five p.m.

**18:22** When I change that, you'll notice that quite a bit more red showed up.

**18:25** Some of the green moved away, and I got much more yellow.

**18:28** What this is showing that on the roads at five p.m. on Tuesday, it's traditionally slower on these roads.

**18:37** So let's go and turn our route layer back on and actually go and ask the network dataset to route considering that information.

**18:49** Now when I do that, the two things that hopefully you'll notice is that...

**18:52** ...I'm no longer traveling through the surface streets near the downtown core to get to the interstate.

**18:57** Instead what it's doing is routing me on O'Shaughnessy Boulevard to get to the second stop.

**19:02** It's also exiting Interstate 80 much sooner than it did before...

**19:06** ...before it encountered the stop-and-go traffic on the way outside of San Francisco.

**19:11** So these are just a few examples of some of the work that the Network Analyst team has done to augment the network dataset at 10.0.

**19:17** Eric?

**19:18** Okay. Okay, now those were some pretty interesting demos to really show you some of the greatly enhanced functionality...

**19:29** ...that's been brought forward with the network dataset, our transportation model at 8.0.

**19:34** I think you'll agree that that's much more useful and interesting.

**19:37** I like seeing the time-of-day routing and whatnot.

**19:40** Now, finishing up the information model portion discussion, with topology, we ended up...

**19:47** ...based upon a lot of feedback from you guys, the users, we ended up implementing six additional topology rules.

**19:56** You remember, the topology rules are how you essentially define in a...

**20:02** ...essentially define what constitutes valid topology and good data for your organization.

**20:08** Remember, unlike coverages, it's a very flexible model that lets you specify whatever collection of rules you want across your data.

**20:16** So these are six of the new ones that we brought forward at 10.

**20:21** Now, next interesting topic.

**20:27** Mentioned this a little bit before, you know, trying to make our replication...

**20:32** ...our data-sharing technology even better and more useful to our community.

**20:37** Now, just a quick refresher on replication.

**20:42** Basically, replication is a...it's a pretty cool technology that allows you guys...

**20:47** ...to be able to share data across a heterogeneous collection of geodatabases.

**20:54** By heterogeneous I mean it's sort of like, unlike basic relational database replication where, you know...

**21:02** ...maybe you can go from SQL Server instance to SQL Server instance...

**21:05** ...or from Oracle instance to Oracle instance.

**21:09** This is from geodatabase to geodatabase.

**21:13** One geodatabase could be sitting on SQL Server, another could be on Oracle, another could be on DB2, or Informix, or Postgres.

**21:23** So it's pretty neat, and it also works in an asynchronous manner...

**21:26** ...meaning you determine when you want the synchronization process to occur. Okay?

**21:32** So with this, as of 9.3, there were sort of four main models or use cases or topologies that replication worked really well with.

**21:47** Here's kind of a cheesy diagram but I think it explains to you what the differences are.

**21:52** You know, we often see people with regional offices.

**21:55** We do see the hierarchical level model where you're sharing between, say, federal, state, and local agencies.

**22:02** A very common one actually with, say, a lot of utility customers and whatnot...

**22:07** ...now as we get more ubiquitous computing is the mobile user thing...

**22:11** ...where you have your headquarters geodatabase and you have people replicating or sharing data out to their field units...

**22:18** ...their mobile units.

**22:20** Another one that's very, very common, think about like government organizations.

**22:25** We saw some of this on the big stage yesterday where, you know...

**22:28** ...you'll have your GIS professionals and whatnot working behind the scenes on their own geodatabase...

**22:34** ...but then periodically, that's considered the production system...

**22:38** ...periodically they're going to want to export that out to a publication system...

**22:42** ...such that the end user community can then have access to at least read that data. Okay?

**22:47** And you want to separate the two.

**22:49** You don't want your live hot stuff that's being edited in the back offices to also be edited necessarily by your end user community.

**22:58** So that's the production-publication model.

**23:00** Now, the new one that was brought forward by the replication team inside the geodatabase was to support centralized data centers.

**23:09** Here's another cheesy graphic trying to show that.

**23:12** You know, you do have some organizations where the data is not flowing from...

**23:19** ...sort of the central headquarters down to the end users in the field offices and then back again.

**23:26** Sometimes you only want it to be a one-way flow of information.

**23:30** And in this case, now we have three states, Idaho, Montana, and Colorado, all replicating...

**23:35** ...pushing their data up to the big centralized mountain server or whatnot. Okay?

**23:41** Now, other things that were...so that's what we term one-way replication, child to parent.

**23:48** Some other interesting things, there's some schema mapping across replicas, meaning, you know...

**23:53** ...table names don't have to be the same at both ends of the replication.

**24:00** Workflow also, we want to be able to support more replication or workflows...

**24:06** ...when we're dealing with nonversioned database or the file geodatabases.

**24:10** We don't want to constrain you to just operating on enterprise geodatabases.

**24:19** Another good demo.

**24:26** Those are my cheesy graphics, by the way.

**24:30** Now see here. [Laughter]

**24:31** I don't appreciate that, sir.

**24:33** I certainly don't appreciate your tone.

**24:35** It was just disparaging and insulting.

**24:37** Quite.

**24:39** Okay, so I'm going to go through a quick demo, just highlighting some of the work the replication team has done at 10...

**24:44** ...and highlight what Eric just walked through.

**24:48** The demo that I'm going to show here is highlighting that centralized data center workflow.

**24:52** Like Eric mentioned, a lot of our replication technology is centered on satisfying these user workflows.

**24:57** This was one that we got specifically from several different users out in the field who essentially asked the following from us.

**25:05** During different disasters, both natural and man-made, it's extremely important for the data to be in one central location.

**25:15** Think of a city or a county government, for example, that controls their own data...

**25:21** ...but they would like to push it up to a central authority who can house the data in a read-only way...

**25:28** ...and keep it online for disaster management scenarios.

**25:32** There's been several examples in the last couple years where you can think that...

**25:35** ...that would have been extremely useful to have smaller agencies on the city, county, and even state level governments...

**25:43** ...pushing their data up to a central authority who could manage it and keep it up and online for when a disaster hits.

**25:50** That's really what this workflow is meant to show.

**25:53** What I'm going to show here is really a feature function demo.

**25:56** It's not really showing a solution, but it's showing you the actual technology that the replication team has built...

**26:01** ...and hopefully you can get some ideas of how to apply this in your own workflows.

**26:06** So, what I have here is my central geodatabase that's already been set up.

**26:13** I've got three different counties [sic] here.

**26:15** I've got Oregon, Idaho, and Montana.

**26:17** You can see I have the different county information colored through yellow to blue.

**26:22** I've also got a collection of points on it.

**26:24** I've got hospitals and schools.

**26:26** Typically in disaster scenarios, hospitals, obviously very important.

**26:30** But a lot of times schools are the central location where a lot of the disaster supplies are kept...

**26:36** ...and where a lot of people actually congregate at in order to manage the disaster when it's happened.

**26:43** So, right now what I have is this one central database with all of the information from those three states...

**26:48** ...and I'd like to set up this replication scenario out to the state of Idaho to be able to pull their data as they update it...

**26:56** ...and keep it in this one central location for disaster scenarios.

**27:00** So, what I've already done here is I've got my data loaded in this central geodatabase.

**27:05** Idaho already has their data loaded as well.

**27:07** What I need to do is simply notify the two geodatabases that each other exists and create the relationship between them.

**27:15** So let me walk you through that.

**27:19** I'm going to go and create a one-way replica.

**27:22** And this is the thing that's different at 10.

**27:24** Previously, the one-way replica like Eric mentioned was satisfying the production-publication workflow.

**27:30** This one is a little different though, because the edits are actually going from the child up to the parent.

**27:37** So it's a slight twist on that previous workflow.

**27:40** In this example, the child geodatabases are the different state geodatabases, and the parent is this larger geodatabase...

**27:49** ...the central one where all the information is collected.

**27:53** I'm going to go out and navigate to the geodatabase that I'd like to connect to.

**27:56** In this case, it'll be the Idaho geodatabase.

**28:01** I'm going to choose to register the existing data only because, again, my data's already loaded on both geodatabases.

**28:07** I'm just simply notifying them that the other exists.

**28:14** I'm also going to use another piece of functionality the replication team has added at 10.

**28:18** And that's the ability to use archiving to support one-way replication.

**28:22** What this does is it minimizes the footprint of replication on the server.

**28:26** It means less versions get created when you're using replication.

**28:31** So it can mean a smaller footprint running on your ArcSDE server for using replication in this example.

**28:40** Now I can go and actually pair up what I want to associate.

**28:44** Now the important thing here that I'd like to point out is a nice change that the team has made at 10.

**28:48** You can see that in my central geodatabase, my feature classes are named things like hospitals, schools, and counties.

**28:55** But in my target geodatabase, they actually have very different names, Idaho hospitals, and so forth.

**29:01** This is a new change at 10 where you can actually create replicas...

**29:04** ...between geodatabases where the feature classes have different names.

**29:08** Previously that was a constraint or a rule that we imposed that the names had to be the same.

**29:12** We're now giving this flexibility to better model your data and add the names that are meaningful within the geodatabase...

**29:19** ...but still set up the scenario.

**29:24** And lastly, I'll click Finish.

**29:26** And this will go through pretty quickly because, again, all it's doing is just notifying the two geodatabases of each other's existence.

**29:32** So now I can go out and actually open up that Idaho map document.

**29:45** And if we take a look at the version manager, you'll see that there's a replica out to my geodatabase, the central one that I just created.

**29:55** So now I'm going to start editing.

**30:03** I'm going to start editing my geodatabase and I'm going to make some very simple changes here...

**30:06** ...just to demonstrate in a very visible way that we're pushing these edits up.

**30:10** So I'm going to go and select this one county and actually change its population to be 17,000...

**30:19** ...and you'll see that when I do that the color of it changes.

**30:23** I'll also insert a new hospital and a couple new schools.

**30:31** Now obviously that's not how Idaho and the state of Idaho manages their hospitals and schools.

**30:38** Makes for good demo, though.

**30:39** That's how we do it in parts of Canada, eh?

**30:42** I'm going to put a school there.

**30:45** Really what I meant to show here is, this data would have been input over a series of time, right?

**30:50** Decisions would have been made, analyses run to determine where it is.

**30:53** But the key thing is that now I have the data in my Idaho geodatabase, I want to push it up to that central geodatabase...

**30:59** ...so that it's stored in this centralized data center, backed up, if you will, at a different location.

**31:05** So now that that's done, I can go and synchronize my changes, and I'll go from the Idaho geodatabase that I have...

**31:11** ...to the central geodatabase and push those changes up.

**31:15** And again, this is going to go pretty quickly 'cause there weren't a lot of changes.

**31:19** And as I go back to my central geodatabase, I want you to focus on this center area here that

I was editing.

[31:24](#) I'll go and refresh the changes.

[31:33](#) I'll open up my Versioning toolbar actually to refresh the changes.

[31:41](#) Sometimes this is what happens when you practice the demo too many times.

[31:46](#) I'll just pull up the Refresh Version command quickly here.

[31:54](#) Then I'll go refresh my changes and you can see that the new color for that county appeared because of the population change...

[31:59](#) ...and also the hospital and schools that I input.

[32:03](#) So that's just a quick example of one of the big things that the replication team did at 10.

[32:07](#) Really this demo's meant to highlight those three things, the new centralized data center's workflow...

[32:13](#) ...the ability to use archiving to manage certain types of replicas, and the ability to have different feature class names across replicas.

[32:21](#) Eric.

[32:25](#) Very smoothly done. That was classic.

[32:33](#) Let's see. Good demo. Let's see.

[32:37](#) Let's get on to the next of the four major areas.

[32:41](#) We're actually going to spend a good amount of time talking about this topic.

[32:46](#) Let's see. It's all about more open and direct access to the data stored in your geodatabases.

[32:53](#) Now, one of the interesting things when you look back at the geodatabase when we released it at 8.0.

[33:03](#) It did have an ArcObjects API.

[33:07](#) Very nice ArcObjects API, but you couldn't really get at your data unless you were going through ArcObjects for our desktop application.

[33:17](#) The schema information, basically the metadata that we maintained to organize your data, was stored using binary representations...

[33:26](#) ...and the data itself, the geometry was stored in binary representations.

[33:30](#) So basically the story at 8.0 was you had to go through ArcObjects.

[33:35](#) At 9.2, we started opening up the geodatabase a little bit more.

[33:41](#) Remember at 9.2 one of the big things was the new SQL types.

[33:45](#) What this allowed you to do was, again, you could use ArcObjects to get at either the schema

or data inside your geodatabase.

**33:53** But then, if you actually want to get at the data itself, if it was stored in a SQL type, you could get at it through raw SQL. Okay?

**34:03** So that opened up the system a little bit and sort of like the first big step.

**34:08** Now, at 10.0, we're opening up the schema to try and complete the opening of the geodatabase.

**34:18** Now the schema is no longer stored in a binary representation.

**34:23** I don't know how many of you have gone and opened up a geodatabase and looked at the GDB underscore tables.

**34:28** You know that simple collection of 35 normalized, crazy tables and stuff with binary content.

**34:35** At 10, we changed that, okay?

**34:38** Now the schema essentially all the schema information about your datasets is stored using geodatabase XML. Okay?

**34:49** Now, talk a little bit more about this.

**34:52** So instead of having a constellation of 35 or so GDB system tables, you're going to have four tables.

**34:59** And the interesting aspect here is that in one of those tables, the items table, for each dataset there'll be one row in this table...

**35:06** ...and its definition will be stored in XML that you can freely and readily get access to. Okay?

**35:13** Now we don't suggest that you go about, you know, modifying XML...

**35:17** ...unless you really know what you're doing, but it's really, what's in my geodatabase?

**35:21** Now it's opened up.

**35:24** So, this is providing more open access.

**35:28** Also one of the things that we wanted to do was make it faster when you're running your applications...

**35:33** ...to browse and search through the collection of datasets inside your information models, your data models.

**35:40** The geodatabase has always been scalable in terms of data volume. Okay?

**35:46** We have users out there with tens if not hundreds of millions of features in a table or a feature class.

**35:53** I mean we can handle that very well, always have.

**35:57** The difficulty was with end users that had very complex schemas...

**36:02** ...very complex data models where maybe they had a thousand different tables inside their geodatabase...

**36:10** ...2,000 tables inside their geodatabase.

**36:13** This new way, this new representation that we have for storing schema is much simpler...

**36:18** ...and it allows us to be much more scalable in terms of being performant [sic] with large complex schemas that you may have. Okay?

**36:27** And this also is serving as a foundation for us moving forward...

**36:30** ...such that we can introduce possibly new abstractions into the geodatabase in the future.

**36:36** So a big simplification effort there.

**36:38** Now, one thing to keep in mind, you will only get that new schema, simplified schema, if you upgrade your geodatabase. Okay?

**36:49** This is a geodatabase upgrade.

**36:52** You can upgrade either your enterprise geodatabases, your file geodatabases, or your personal geodatabases...

**37:00** ...using a GP tool or a command that is fired off of the workspace property sheet.

**37:07** So all you have to do, go to either the connection or go to the workspace itself, right-click on it, brings open the property sheet...

**37:17** ...and then hit the Upgrade button; it brings up a GP upgrade tool.

**37:20** You can also call it through Python if you want because it is a GP tool.

**37:25** In any event, it'll take care of everything.

**37:28** Now first, back up your data before you do that. Okay?

**37:33** Good idea, back up that data.

**37:35** Also, keep in mind, we are not touching your data. Okay?

**37:40** We're not going into your tables and your features and rows and things like that and mucking with those at all.

**37:46** No. We do not touch your data.

**37:49** What we are touching, or what we are changing, is how we internally organize the geodatabase.

**37:56** Those GDB system tables are turning from a collection of 35 down to 4...

**38:01** ...but your data is not going to be touched at all in this process. Okay?

**38:05** But do back up your data.

**38:09** Now, one of the cool things here, and this was talked about on the big stage during the plenary yesterday...

**38:16** ...is that because we've been able to simplify the schema of the geodatabase, how we organize the information...

**38:23** ...it's now made it possible for us to create an open API to file geodatabases. Okay?

**38:32** This is basically it's going to be a C++ API, basically a DLL that you'll be able to download freely.

**38:41** It doesn't require any licensing.

**38:43** And it will allow you to get access to your file geodatabases, basically the simple feature model.

**38:49** It's not intended for updating topologies or network datasets or the more complex models...

**38:56** ...but for your simple feature data, this would be a way that you would be able to both read it as well as update it. Okay?

**39:03** Through a public API. Okay?

**39:06** That does not require any licensing.

**39:09** Freely downloadable.

**39:10** No ArcObjects whatsoever.

**39:12** This is just a single DLL.

**39:15** Now, of course you can read all the data in a geodatabase including your more sophisticated dataset types...

**39:22** ...but you'll only be able to update simple feature data. Okay?

**39:26** So here's a little pictorial representation where if you have a C or C++ client application...

**39:33** ...using this new open API in conjunction with knowledge of the geodatabase XML spec...

**39:39** ...which is another thing you can just download from our Web site.

**39:43** It's a standard thing, just go grab it.

**39:45** You'll be able to write applications that'll be able to read everything in a geodatabase...

**39:49** ...or actually update simple feature inside your file geodatabases.

**39:55** Now, this is actually one of the neatest things, I think, that's happened with the 10 release in terms of the geodatabase functionality.

**40:04** Oh, also, I should go back and let you know the open API, the file geodatabase API, is not in the 10 box.

**40:15** We're still finishing it up. Okay?

**40:17** This will not be going out with the next major release of the software.

**40:22** It's a single DLL.

**40:23** It will be able to go out off cycle, what we term off cycle.

**40:26** That means basically when it's ready, it will go out. Okay?

**40:30** It's a single DLL. It's not a big process to install it.

**40:34** So expect that this fall. Okay?

**40:36** We're trying to finish up that work right now.

**40:39** We expect it to be out this fall.

**40:41** So stay tuned on that one.

**40:44** Now, getting back to accessing data, opening up the system.

**40:48** Well, you know, we have all kinds of different ways to get at your geodatabase spatial data.

**40:53** But then, as we're seeing these days with all these spatial types and whatnot, you know, look at what Microsoft's doing with SQL Server...

**40:59** ...and their geometry storage type or their geographic storage types...

**41:04** ...there's a lot more spatial data out there that's actually not inside a geodatabase.

**41:11** We have users that don't want to, you know, do what we've done in the past where we said...

**41:15** ...well, if you want to be able to access that data, you have to register that data with the geodatabase.

**41:20** We have to be able to lay down our system tables and whatnot. No.

**41:24** This is basically a technology called query layers that will allow you to define layers that have full layer functionality...

**41:32** ...to be able to access data stored in spatial databases. Okay?

**41:39** Now, this is pretty interesting stuff here.

**41:42** It's really opening up your modeling possibilities.

**41:47** Everything in your application, everything in your ArcGIS Desktop...

**41:51** ...that you're manipulating does not have to exist inside a geodatabase. Okay?

**41:56** Now Brent's put together a pretty cool and completely new demo that he's going to take us through. Brent?

**42:05** It's a lot better than looking at PowerPoints.

**42:08** Or would you rather see PowerPoints and a demo?

**42:13** More demos always good.

**42:17** More demos, more cowbell, yeah.

**42:20** So in ArcMap here what I have is I'm consuming a basemap from ArcGIS Online, so just the ArcGIS imagery basemap...

**42:31** ...and I've already gone ahead and added two query layers to map...

**42:34** ...and these query layers are coming directly out of a SQL Server 2008 database instance that's on my laptop here.

**42:39** So really, as Eric mentioned, a query layer is really just a feature layer in ArcGIS that's defined by a SQL query.

**42:46** So what it lets you do is get access to data natively in a spatial database.

**42:50** No geodatabase involved at all.

**42:52** So these really act just like any other layer in ArcGIS.

**42:56** There's one subtle difference.

**42:57** When I open the layer properties for my oil spill layer, I can actually look at the query that was used to define that layer...

**43:04** ...and I can change it, I can make modifications to this query.

**43:07** So I'm doing a very simple query here.

**43:08** For an example, I'm doing select star, so selecting all the rows in the trajectory oil table.

**43:14** And what this table is, the data it contains is just the daily snapshot of the oil spill information.

**43:22** And because that information has a time component to it, I can use it with the time slider and make it time aware in ArcGIS.

**43:29** So I've done that, and I've added the time slider to map.

**43:32** So it means I can just jump through and see what the oil spill looked like each day.

**43:37** This is really powerful functionality, and this is pulling this directly...

**43:40** ...querying my database tables directly, and integrating it with ArcGIS.

**43:44** So you could use this to integrate with your feature classes, other layers and map, and work with this data natively in the database.

**43:53** So I'll go ahead and create a couple more query layers.

**43:55** I'll first create a very simple query layer.

**43:57** And how you do that is just go to the Add Data pull-right and you open the New Query Layer

dialog.

**44:02** The first thing I'll do here is make a connection to my spatial database.

**44:07** It's the database instance I have running on my machine.

**44:09** It's the same database that I took data from the two previous query layers.

**44:14** And when I make the connection, I get to see all the tables and views that I have access to...

**44:19** ...as the user I've connected from in this table list here.

**44:22** So I can use this to build up my query and as I go through, when I select any of these tables, I get to see the columns in those tables.

**44:30** I could see this is the date value that I was using to use with the time slider and map.

**44:35** And I can use this column information and pull this into the query layer which eventually helps me building up that query.

**44:41** But if I just want to do a very simple query and see all the rows in the table, I can just double-click on the table...

**44:46** ...and it adds a simple select star query to the query, and I can go and modify this if I wanted to constrain it.

**44:53** But I really just want to see the total path, everywhere the oil has went over time...

**44:59** ...so that total cumulative polygon, and this has been captured in a table called the uncertainty zone.

**45:06** So I'm just going to do a simple select star query from that.

**45:08** I'm going to paste in the uncertainty zone name to name the query.

**45:12** And then the next step, I have to validate the results of that query.

**45:16** And what this does is just check to make sure that the results that returned from the query conform to ArcGIS's data modeling standards.

**45:22** So it checks stuff like, is there only one shape field in this result set? Does it only have one spatial reference?

**45:28** And if it doesn't, the dialog gives you an opportunity to change that query so it does conform.

**45:33** So since it validated okay, I can select Finish, and I can add that layer to map.

**45:38** So here I could see the cumulative polygon of everywhere the oil might have touched over time.

**45:46** Now this is an example of a very simple query but I can do more complex queries...

**45:49** ...and I can really do anything that is supported through SQL in the database.

**45:55** So I'll follow the same workflow, connecting my SQL Server instance...

**45:59** ...but I've actually gone ahead and created a query here which is a little bit more complicated.

**46:03** So what I want to do is find the intersection of that uncertainty zone with oyster lease polygons I have in another table...

**46:10** ...a SQL Server table in my database, finding all the oyster beds or oyster leases that have possibly come in contact with the oil spill.

**46:18** So here I'm using the SQL Server ST Intersects function, and this really does a spatial intersection down in the database...

**46:25** ...to give me the result of the intersection between uncertainty polygons with the oyster leases.

**46:31** So I can go ahead and copy and paste this directly into the query window, and I will just name this Oyster Leases.

**46:43** Here, following the same workflow, I'll validate the query and then selecting Finish, it'll add those results to map.

**46:50** So you could see here anything that's supported...anything I could do via SQL in the database I could also do through query layers...

**46:57** ...as long as the result set conforms to ArcGIS's data modeling standards.

**47:02** So I'm going to go ahead and zoom to that layer, ArcGIS, and you can see this is the result as it pulls up and renders here.

**47:11** This is a result of all the thousands upon thousands of oyster leases and oyster beds that have been affected to date by the oil spill.

**47:20** So say I want to share this information out with other stakeholders, other users within my organization.

**47:25** Well, at ArcGIS 10 through ArcGIS Online, I can quickly and easily share that information out...

**47:31** ...with other people who might be very concerned about this information.

**47:35** So the first thing I'm going to do is open the properties on the layer and just give it a quick description.

**47:40** So I'll call this Affected Oyster Beds.

**47:48** Just to give it a little bit of metadata so people understand what they're working with and then right-clicking on the layer...

**47:53** ...I can create a layer package and upload that layer package to ArcGIS Online.

**47:58** So I'm already logged in to ArcGIS Online, so I'll call this layer package Oyster Leases.

**48:05** I'll first validate it and then I'll share it out with ArcGIS.

**48:10** To save a little bit of time, I'll just copy and paste the description into the summary here.

**48:16** Great metadata.

**48:17** That's true, I'm big on metadata here.

**48:19** And then I'll just add a couple tags to this so it can be searched on.

**48:24** And I'll call it Oil Spill and maybe Oysters.

**48:30** So I've added some tags, and now do I want to share this with everyone who has access to ArcGIS Online...

**48:35** ...or do I want to do it just for specific members of groups?

**48:38** So I've actually gone ahead, and I'm a member of the Gulf Oyster Fisheries Special Interest Group...

**48:43** ...so I really just want to share it with members of that group.

**48:45** So I can do so easily through this and selecting Okay, we'll go ahead and package up this layer.

**48:51** So not only the layer properties, how I've visualized it and symbolized it in ArcGIS, but also the data that went into creating that layer.

**48:59** So what this means is I'm now uploading that layer package to ArcGIS Online...

**49:04** ...so anyone that goes ahead and consumes this layer doesn't really need to know any SQL.

**49:08** They don't really need to know anything about databases at all.

**49:10** All they need to know is what the layer is and then they can bring that down into their map documents and work with that information.

**49:17** So it's gone ahead and uploaded that map package or that layer package rather to ArcGIS Online.

**49:21** So I'm going to close down the map doc...

**49:23** ... 'cause I'm done with it, and I'm already connected to ArcGIS here in the Web browser.

**49:30** So I'm signed in as myself.

**49:32** But say I was signed in as a different user in my organization who also was a member of that group.

**49:37** They can go to the Groups tab here and they can see all groups they're a member of.

**49:42** So they'll go to the Gulf Oyster Fisheries Special Interest Group and they'll now see that layer that I've uploaded.

**49:49** So selecting the layer, they can see more information about that layer, if anyone had comments on what the layer was about...

**49:55** ...what was in that layer, they could see these comments here.

**49:59** And if they want to use it within ArcGIS, all they have to do is select Open and that'll download that layer package locally to the disk.

**50:09** Opening that layer package will then open that layer within ArcGIS.

**50:16** So this is a quick way to share information with other stakeholders and display that information.

**50:21** You could see this user doesn't need to know anything about the SQL, the complex SQL that went into generating this layer.

**50:27** They just need to be a member of a group and have access to ArcGIS Online.

**50:33** So that's a very desktop-focused user story, but I want to show a second part to this demo...

**50:37** ...that's a very quick explanation of how query layers can also work in ArcGIS Server.

**50:42** So if I was to publish any of these map documents as a map service, you could consume these within web applications.

**50:49** And I've gone ahead and done this with some other data...

**50:53** ...and this is the map document that I've been using to seed the map service.

**50:58** So it doesn't really look like much.

**50:59** Really I'm just using it for symbology, but this is also a query layer here.

**51:03** And I'll just open the properties on this layer.

**51:06** And what I'm doing is I'm using some native SQL Server time functions to the DATEDIFF function...

**51:14** ...and actually what this query is doing is just giving me all results within the last 180 seconds.

**51:20** So this is time-aware layer of all the 911 calls in the New York City area.

**51:26** And it's logging this to a database table, and I'm refining this massive amount of information just showing me the last three minutes...

**51:33** ...or a sliding time window of the last three minutes of the most recent calls.

**51:37** So I've gone ahead and already created this map document.

**51:39** I published it to ArcGIS Server, and I've consumed it through a web application...

**51:45** ...and this is a web application that's actually consuming this service which is showing me the sliding time window.

**51:50** And it'll periodically refresh to show me the latest calls.

**51:55** So this really demonstrates how you can consume native database information and quickly

and easily bring it into both...

**52:01** ...ArcGIS Desktop and ArcGIS Server through use of query layers which we added at ArcGIS 10.

**52:15** That's pretty cool stuff, Brent. Thanks. Good demo.

**52:20** I've never seen that one before.

**52:21** That's a nice demo.

**52:23** Also I got very hungry with you talking about steaks and oysters all that time.

**52:28** And I'm sure everyone else in here who's missed their lunch is hungry as well.

**52:34** What? Oh.

**52:38** OK. Well let's see.

**52:41** I think we have enough time for this PowerPoint slide and then we're done.

**52:47** Basically, you've seen some of the functionality we've been releasing with the geodatabase at 10.

**52:55** You know, we were, in the 9 time frame, we were trying to complete the initial information model of the geodatabase.

**53:01** Now we're just trying to refine it and extend it and will be continuing to do so in the future.

**53:07** Now you saw the enhancements we did to the information model for our...

**53:11** ...network technologies that are used for both utilities and natural resources, the geometric network.

**53:18** You saw the work that's being done for our transportation network model.

**53:22** Craig gave a great demo there with time-aware traffic routing and all kinds of interesting things.

**53:27** Very cool and flexible barrier support.

**53:30** And then you saw the new topology rules that we've been introducing based upon all your feedback.

**53:37** Now in the domain of data sharing and whatnot, our replication...

**53:40** ...we're supporting that other key replication scenario that Craig also demoed.

**53:47** Where from those three different states, Montana, Idaho, and I don't know where the other one was.

**53:52** I'm not very good with geography.

**53:55** Yeah. New Jersey?

**53:58** No, this is funny.

**54:00** Basically a statement has been made in our group by one of the other developers.

**54:03** I can't take credit for it.

**54:04** But, a lot of computer science guys like me said, everything I've learned about geography I've learned while fixing bugs.

**54:13** Very true.

**54:15** In any event, so we have that new workflow in the replication domain or data sharing technology.

**54:20** We're trying to open up the geodatabase to make it even more open and flexible for you.

**54:25** You know, first we simplified the schema.

**54:27** Okay. That's cool.

**54:28** It's all now in XML.

**54:30** Joined that bandwagon.

**54:32** The other thing is we've gone and we're finishing up right now a new freely distributable API to file geodatabases...

**54:42** ...that'll give you read/write access to simple data and read access to more complex information models stored in your file geodatabases.

**54:50** And then finally, Brent gave us some really cool demos that brought together a lot of key pieces of technology...

**54:56** ...that we've been introducing at 10 involving query layers at the heart of it all.

**55:01** And then taking it up to ArcGIS Online and sharing this stuff.

**55:05** This is one thing we really look forward to to see what you guys, what you gals will do when you get back to your offices next week...

**55:12** ...and you start using ArcGIS 10 nonstop in everything that you do.

**55:17** We want to see people start sharing data and see what comes up because it is fun.

**55:20** You saw yesterday how you can click on buttons and you can see, oh, what's the most recent stuff people have been throwing up there.

**55:26** I think it should be quite interesting.

**55:28** Plus, you know, some of those downloadable apps for, you know, your smartphones and whatnot, iPods, Androids, flip phones, stuff like that.

**55:40** I love my phone. I actually downgraded my Android last weekend.

**55:44** Quite seriously, the people at Verizon could not understand anyone turning in an Android and wanting an old-school flip phone.

**55:50** But, it was fun. It was a good conversation.

**55:53** You should have gotten a pager.

**55:56** Fax machine.

**55:58** In any event, you've seen all this stuff.

**56:01** I hope you'll find it useful, and we're going to be continuing to move forward very aggressively in the future...

**56:07** ...with enhancing the geodatabase to make it more open, faster, and flexible and useful.

**56:12** So, anyway, thanks a lot for coming.

**56:14** Please fill out those surveys.

**56:16** And if you have questions, we're kind of out of time, but come on up.

**56:19** And other members of the geodatabase team, there are a couple other criminals in here...

**56:23** ...if they'd come forward, too, and help flag questions, that'd be great. Thank you.