



<http://www.dotnetrocks.com>



Carl Franklin

Carl Franklin and Richard Campbell interview experts to bring you insights into .NET technology and the state of software development. More than just a dry interview show, we have fun! Original Music! Prizes! Check out what you've been missing!



Richard Campbell

*Text Transcript of Show # 264*  
(Transcription services provided by [PWOP Productions](#))



**Donald Farmer on Data Mining**  
**August 16, 2007**  
*Our Sponsors*





**Lawrence Ryan:** The opinions and viewpoints expressed in .NET Rocks! are not necessarily those of its sponsors, or of Microsoft Corporation, its partners, or employees. .NET Rocks! is a production of Franklins.Net, which is solely responsible for its content. Franklins.Net - Training Developers to Work Smarter.

[Music]

**Lawrence Ryan:** Hey, Rock heads! Quit surfing for gender changers and listen up! It's time for another stellar episode of .NET Rocks! the Internet audio talk show for .NET developers, with Carl Franklin and Richard Campbell. This is Lawrence Ryan announcing show #264, with guest Donald Farmer, recorded live, Tuesday, July 31, 2007. .NET Rocks! is brought to you by Franklins.Net - Training Developers to Work Smarter and now bringing world class .NET and SharePoint training onsite to your development team, online at [www.franklins.net](http://www.franklins.net), and by Telerik, combining the best in Windows Forms and ASP.NET controls with first class customer service, online at [www.telerik.com](http://www.telerik.com). Support is also provided by CoDe Magazine, the leading independent magazine for .NET developers, online at [www.code-magazine.com](http://www.code-magazine.com). And now, the man who's attempted to make up for 40 years of no sleep in one weekend, Carl Franklin.

**Carl Franklin:** Thank you very much and welcome back to another episode of .NET Rocks. This is Carl Franklin on the east coast of the United States of America. Richard is not with me for this introduction, but he will be here shortly with our guest, but before we get to that let's start right off with Better-Know-a Framework.

[Music]

**Carl Franklin:** So, today I want to talk about the serializable attribute class and this is an attribute that you can put at the top of any class. Basically, it just says serializable, all right? What this does, it indicates that the class can be serialized and if you don't know what serialization is, fear not, I'll give you a little primer. Serialization is the act of taking an object in memory and turning it into data that can be written to disk or sent across the network or anything, just pure data, and deserialization is the opposite. Deserialization takes that data and turns it into an instance of an object. So, any class that you have as long as there are no non-serializable objects as

properties anywhere in the class, it can be serialized, but you do have to mark it with a serializable attribute. Now, there are different ways that you can serialize an object, but if you want to just create a stream of bytes, you are going to use the binary formatter. The binary formatter is in `System.Runtime.Serialization.Formatters.Binary.BinaryFormatter` and that's what you can use to turn your serializable object into an array of bytes, essentially write it to a stream and you can also deserialize with that. And that is your Better-Know-a-Framework for today. All right, I also want to mention before we get started here that the Sleepless in New York entry deadline has come and gone and so thank you to all who signed up for that and you will be notified shortly whether you made the cut, very shortly from what I understand. So, that's gonna be a whole lot of fun. We'll be doing a show around that event in New York City and good luck to everyone who participates. Also, Infusion is still looking for great developers. They have been hiring .NET Developers for months now and they still don't have enough, big projects going on in New York City. So, if you're a hotshot developer and you want to move to Manhattan and live rent-free in an apartment in Manhattan for a year and work in a really exciting environment, check out [shrinkster.com/kh6](http://shrinkster.com/kh6). Now, through the magic of audio editing Richard Campbell is going to join me for the interview, which was previously recorded.

All right, Richard, let's go ahead and introduce Donald Farmer. Donald has worked in the Microsoft Business Intelligence team for six years and he has worked on both the Analysis Services and the Integration Services product teams. Donald is now the Principal Program Manager for SQL Server Data Mining, working to build a seamless integration of predictive and exploratory analytics with the Microsoft business intelligence offering. Mr. Farmer is a popular speaker at international events for both business and technical audiences with a wide range of interests including data integration, information quality, metadata intelligence, master data management and *predictive analytics*. He is the author of a number of books and articles. Prior to joining Microsoft, Donald worked not only in business intelligence, but in fields as varied as medieval archeology and fish-farming. Medieval archeology, tell me were you in the Society for Creative Anachronisms?

**Donald Farmer:** It sounds like it doesn't cut me out. The team I work for was called the Scottish



Urban Archaeological Trust, which is a mouthful in itself.

**Carl Franklin:** Was that SCART?

**Donald Farmer:** SUAT.

**Carl Franklin:** SUAT. Wow.

**Richard Campbell:** I got to imagine you're talking about a place like Scotland, like there's stuff in the ground.

**Donald Farmer:** Yeah. There's stuff in the ground.

**Carl Franklin:** Sure.

**Donald Farmer:** And we dug it up.

**Carl Franklin:** And fish-farming. Let's just get the fish-farming thing out of the way. So, what lead you to that field?

**Donald Farmer:** I was working in software for salmon farms. They have some really interesting kind of some software requirements to measure fish growth and fish feeds and all sorts of things. It's kind of interesting. I'd be out there with water dripping off my waterproofs working on debugging programs a couple miles out at sea.

**Carl Franklin:** Wow!

**Donald Farmer:** It was good fun.

**Carl Franklin:** Now, we're obviously not talking Scotland here, you're probably in the Pacific northwest by now...

**Donald Farmer:** No, that was in Scotland.

**Carl Franklin:** That was in Scotland?!?

**Donald Farmer:** I was in the Pacific Northwest until I came to Seattle.

**Carl Franklin:** So, they have salmon in Scotland?

**Donald Farmer:** They have a lot of salmon in Scotland. Yeah.

**Carl Franklin:** Wow!

**Richard Campbell:** That would be called Atlantic salmon.

**Donald Farmer:** That was Atlantic salmon, yeah.

**Carl Franklin:** Yeah. I've been brainwashed by the Microsoft people who tell me that you haven't had salmon unless you've had it in Seattle, so...

**Donald Farmer:** I hate to say this, but the salmon in Seattle are delicious compared to the salmon in Scotland. The wild salmon here are fantastic.

**Carl Franklin:** Okay.

**Richard Campbell:** There is only one species of Atlantic salmon and there is half a dozen in the northern Pacific area and they're amazing. Some are better than others.

**Donald Farmer:** Absolutely. I'm kind of eating my way through them.

**Carl Franklin:** Wow! Oh well, you learn something everyday on .NET Rocks! don't you?

**Donald Farmer:** Absolutely.

**Carl Franklin:** So, let's talk business intelligence. SQL Server Data Mining, I saw the website which is at [shrinkster.com/rel](http://shrinkster.com/rel), the [sqlserverdatamining.com](http://sqlserverdatamining.com) data site. This is an effort of the SQL server team?

**Donald Farmer:** It is. Yeah. The SQL Server Data Mining site is really just a place where we can pool all the various pieces of information together. Typically in Microsoft, we have a lot of different websites. We have MSDN websites and we have TechNet websites. There are some materials, which can't fit into any of those categories. There's a ton of articles out there in magazines and on the web. So, we just provided one kind of portal where we can pool all that together and we called it [sqlserverdatamining.com](http://sqlserverdatamining.com) for want of a better name.

**Richard Campbell:** It's a great site name, but you know maybe we got to back up a couple of notches here because data mining, it really sits on top of



analysis services which has its reach all the way back in the OLAP server, SQL Server 7.0 era.

**Donald Farmer:** That's true. Yeah.

**Richard Campbell:** I'm a data guy so this is certainly something that I have dealt with for quite some time now, but I think an awful lot of folks aren't aware of just how much has happened over there. I remember back in 2000 when it was announced that OLAP Server can be part of SQL Server, I was working in data analysis with other products like Arbor Essbase and folks said, "Hey, you're familiar with Microsoft. Is this the end of the world for OLAP for us because if Microsoft is coming to the party, this is it, right?"

**Donald Farmer:** Right. Yeah.

**Richard Campbell:** Well, it's not quite that grim, but it's going to be interesting because it's going to bring it to them. Back then, OLAP and Analytics, they were very expensive technologies. Now, it has become much more commoditized, much easier to approach.

**Donald Farmer:** I think what's really interesting is that it wasn't really the end of the party for anyone. It was the end of a particular kind of party and the start of a new one. If you'll look at the growth of SQL Server Analysis Services over the years, in fact they started in SQL Server 7.0, we had OLAP services in 7.0 and you look at the growth from there to today where we're now leading OLAP application. So, see that we've got that growth primarily by growing the market.

**Richard Campbell:** Right.

**Donald Farmer:** Not by necessarily kind of taking over traditional OLAP shops and perhaps even kind of competing with Essbase. We've actually built a new type of market which simply didn't exist before and when we introduced data mining in 2000 with one or two algorithms that was just kind of "toe-in-the-water" in the data mining space, but now we've got a full range of algorithms. We're full on as a high quality data mining product and our intention is to do the same with data mining, is to get that same kind of growth where we introduce data mining to people who would never have considered it before and then get that kind of growth that we saw in the OLAP market.

**Richard Campbell:** Most people, I think, think of data mining as just writing SQL queries. Maybe we got to get your definition or what will you say the real definition of data mining.

**Donald Farmer:** Yeah. If you read the newspapers just now you'll see a lot of talk about data mining and talk about FBI Data Mining and the CIA Data Mining. In that sense, they are really talking about just large scale database search, that we've got a ton of information on people, we've got all our -- I don't know what records they have, all the phone records, all the library records. We're searching through all them in some kind of fishing expedition and that's a kind of modern usage of data mining, but which is a little bit less precise of what we are talking about. We're talking about I guess what sometimes is called knowledge discovery in databases and it's really a set of statistical learning algorithms which will discover patterns of various sorts within your data sets and then having discovered those patterns, build a model that represents the patterns and then use that model to predict what might occur with future data. It's really about finding these rules and patterns and applying the rules and patterns to the future. It's not about just trolling through large volumes of data, searching for records.

**Carl Franklin:** When I was in the business and doing a lot of work in SQL Server, admittedly this is back in SQL Server 2000 days, oh way back in the day, all of a sudden for what we called storage and search became data warehousing and data mining. I guess the scale of it and the scope it is just much bigger and that's why we use big words like warehouse because in a warehouse you have lots of things that are categorized and tagged and they're just put in big places, but you don't really do anything with them. Mining is going and researching and finding maybe corollaries between pieces that you didn't really know ahead of time were going to be linked.

**Donald Farmer:** Yeah. Maybe in that sense but data mining probably relates back to my background in archaeology. You kind of go through all the debris of the past looking for the kind of nuggets of knowledge that are in there. I think data mining is quite often used in that sense. We do tend to talk about it. There are a couple of phrases which are becoming common. We talk about knowledge discovery or machine learning. We can talk about



data mining. A more modern phrase I feel like or a more popular phrase is *predictive analytics*.

**Carl Franklin:** Wow!

**Donald Farmer:** *Predictive analytics* is where people are, I think they're really kind of capture that sense of "I build rules" or "I discover rules and patterns" not just for the sake of discovering rules and patterns, but for the purpose of applying them to new data in the future. So, I kind of like that phrase *predictive analytics*. It's quite useful.

**Richard Campbell:** The real concept here is I want to put my energy into the customers that will make me the most money, put the most energy to the most results. I can't give the same amount of energy to everybody because then I won't get the optimal results. One of my personal data mining examples is when I was working with a company that was building residences, *townhomes*, and one of our analytic pieces spat out that we had too many nurses that are buying residences in this one particular building. We realized that it wasn't nurses per se, but it was single professional women. It was an upscale place, but it was very secure and well organized and close to the things that women cared about. It wasn't us that found that data, it was the data mining model actually in the original SQL 7.0 version that that worked on where it said there's a relationship between these chunks of data that you haven't identified, but I see it. So, we were able to shape our advertising campaigns to cultivate that customer type. We focused on the things that single professional women will care about and we advertised it in the places that single professional women would look at.

**Donald Farmer:** Right. So, having identified that kind of market segment, you are then able to direct your energies into that and optimize your returns from it.

**Richard Campbell:** And as a result -- part of my reaction to that, it was a self-fulfilling prophecy because we had too many of them, we cultivated more and so we had way too many of them.

**Donald Farmer:** I see.

**Richard Campbell:** But the reality was, yeah, we sold out a building by --- we didn't realize the elements of that building that were important to those

people and so the data helped us find that fact out and then take advantage of it.

**Donald Farmer:** There are two things I really hear when I hear that story. One, it reminds me very much of a very common problem with data mining, is how do you actually *align* this discovery of information with your business? It's just not enough to discover these things. We've also got to be able to *act on it*. The second thing that reminds me on this is that how difficult it can be to actually act on some of that information? For example, you have later found that this building was attractive to a group of people who were less able to afford it than others and in that case, you may have actually had real difficulty acting on that information.

**Richard Campbell:** Right.

**Donald Farmer:** This building is most attractive to people who can't afford it.

**Richard Campbell:** That's a problem.

**Donald Farmer:** It is not a great business model.

**Richard Campbell:** Yeah.

**Donald Farmer:** Data mining is actually used a lot in cases where we try to find the best market demographics. One of the largest consumers of data mining is the marketing analysts who are looking first of all to understand what demographics are actually out there and what demographics are responding to the business and the other one is of course trying to optimize the business towards a certain demographic, doing it the other way around, if you like.

**Richard Campbell:** So, we've got to tie this back to development somehow. I mean this is almost a very much a data-related topic. It is one of my favorites, but it seems to me Microsoft just recently has really reached the point where this whole concept of data mining and analytics is able to be incorporated into regular applications. You see the terminology and the behavior much more often now. It's not sort of off in its ivory tower anymore.

**Donald Farmer:** You do, that's for sure. Actually to tie it back to development, I think one of the things that we do that is unique in Microsoft is we are very focused on data mining for our developers.



This is something that we do that's quite different and is something that is a very important part of it for us. What I mean by enabling data mining for developers is what we have been talking about so far, a lot of business cases, a lot of data analyses, a lot of business analyses, but in the long run how do you actually deliver these results to people?

**Richard Campbell:** Right.

**Donald Farmer:** Now traditionally, data mining has been used by people who had specialized data mining tools, which created specialized data mining reports with a lot of statistics, a lot of analysis in there and then they would do a kind of ad hoc analysis. They might run the model, see what results came out, analyze perhaps some complex 3-D visualizations of the demographic space or whatever and then make some decisions on it.

**Richard Campbell:** So, they were analysts and they had a skill in analyzing.

**Donald Farmer:** That was their business, it was their job. Of course, what we're trying to see, what we're trying to develop is data mining for the end user, the typical end user. Organizations very often just can't afford to have these specialized analysts. They may have to buy into an agency service of some kind. So, what we're trying to do is enable data mining for the end user and there are really two parts for that. One is to enable data mining for the office worker, what I mean by that is the Microsoft office worker, the information worker who lives in MS Office or Excel. We've done some great work in enabling that in Excel, but the other thing is to really target the application developer who can then embed data mining into operational application.

**Carl Franklin:** Wow! Now, that's a n interesting idea. So, rather than hiring analysts, you give the developers the tools and say, "We don't want you to write the actual queries and things, but write us an application where we can access this, this, this and this" and the end user takes that application goes to town.

**Donald Farmer:** The great thing about that scenario is when the end user is working in that application. They don't know there is data mining, they don't know that the complex rules engine behind there. All they see is the results that they need to see.

**Richard Campbell:** Any more than a regular user understands that there are SQL queries involved. They should just be invisible.

**Donald Farmer:** Absolutely, yeah. That's really the same with data mining. The world is actually full of these applications. The one that everybody is very familiar with is when they go on to an e-commerce site and they get -- I guess in the case of a kind of online bookstores like Amazon, you'll get an offer that says, "Customers who bought this book also bought this book. Are you interested?"

**Carl Franklin:** Right.

**Donald Farmer:** That uses data mining to discover those patterns of what people are interested in. Now, that's actually very easy, almost trivially easy to build such an application, such a recommendation engine into an ASP.NET application using SQL Server Data Mining. It's actually one of the simplest data mining things you can do is to build a model that captures those association rules as they call them and exposes that to the end user as a recommendation engine.

**Carl Franklin:** So, I'm trying to think as a developer what my tool set would look like when I'm writing a data mining application and I suppose there has got to be some specifics there about the kinds of things that the end user is going to be doing, but you are seeing a lot of analytical tools like controls both for ASP.NET and Windows Forms coming out on the third party market. Is this really what we're talking about? You're basically just saying if they're looking at the relationship between, I don't know, up to five or six tables, they put in the table names and they get graphs for a line for each one and then they can start messing around and bending things and plugging in variables and... Is that the kind of stuff we're looking at or...?

**Donald Farmer:** I don't see us doing those necessarily as a set of controls. Also, there are some web client controls that can be used for visualizing the models. Visualization of data models can be quite useful for even fairly inexperienced users because seeing some of these complex relationships visually really helps. The way this kind of works for data mining and let me just describe the moving parts for those here. There is a data mining model built over source data and that is normally built using SQL

Server, using the business intelligence, development environment for SQL server. They can be built in Excel using an Excel client, but most people will build this in the Business Intelligence Development Studio. Now, that model is then queried by client application using a very simple language called DMX. It has to be distinguished from MDX. MDX is an analysis services query language. DMX is a data mining query language. It is actually very, very SQL-like. DMX stands for Data Mining Extensions for SQL. DMX is a really straightforward query language that can connect to the data mining server using ADOMD.NET and the queries are just very like SQL queries. For example, to do a prediction, you have a model which contains information about past data and now you want to do a prediction based on that model to predict something from the future. The query just looks like a *join*. What you're really saying is "Join my input data with some models and return to me a result set, which actually represents the predictions of future information." That join is very, very straightforward to write and it's just like writing a SQL joint and is exposed just as an ADOMD.NET data set. So, it is very easy for the developer then to get that into their application.

**Richard Campbell:** All right. I want to walk through this one more time because I'm just barely keeping up. I am going to start it from my existing application. I have got my SQL server database there. It has got customers and orders and all that good stuff in it and then I pull out my business intelligence workshop, which is a studio-like component as I recall.

**Donald Farmer:** That's right.

**Richard Campbell:** And I got to build that data mining model. Is that the next step?

**Donald Farmer:** That's the first step, yeah.

**Richard Campbell:** What does that involve? What am I really doing when I build a data mining model?

**Donald Farmer:** That's a great question. To build the data mining model, you really start with some existing data. The first step for me in building that model would actually be partitioning my existing data into two sets, a training set and the testing set. The training set, I'd probably do it 75% of the total data set and I'd hold out a little bit of data such as 25% for testing on.

**Richard Campbell:** Okay.

**Donald Farmer:** I'd probably do that randomly. In the next version of SQL Server will actually do that for you automatically, but in the current version, you might do something like use integration services or a query to split these sets. Once I've actually got my training data, which is a sample of data which represents my business cases, if you like, I then connect to that data from the development studio using the Data Mining Wizard I think. Give me a new mining model and the next step is really to choose the algorithm that you are going to use. The algorithm is just statistical algorithm which will go and find patterns in that data. Typically, what you'll be doing is saying, "There's a column here that I want to be able to predict." So, let's say I've got some customer data and I've got information on customers' ages and addresses and zip codes and occupations and annual income and so on. Now, I may want to say, "Okay. Based on the patterns, can I predict which of these customers is going to be..." well, a good example we often use is, "Are they going to be eligible for a gold or silver credit card or a buyer's card?"

**Richard Campbell:** Right.

**Donald Farmer:** So, if we had all that information then we already had columns which said these existing customers with all these attributes are gold or silver customers. Then we can say, "Based on this incoming data, I'd like to predict in the future who would be a gold or silver customer." So, I select the card type as being a predictable column and I select all the other columns as the input columns. So now, when I run the algorithm against that at the end of the Wizard, what that does is it searches for the patterns in my existing data that most strongly represent customers who are, say, gold or silver card members. Those patterns are then captured in the mining model and deployed to the server and at that point, I can now query the server to reuse those patterns again, new data. So, does that make sense?

**Carl Franklin:** It's complex.

**Richard Campbell:** It's not a simple thing. Maybe, let's go through one more example, say, the *you might want to buy this book too* kind of model.

**Donald Farmer:** Yeah. That's a great example. You've already bought this book and other people



who bought it have bought many other books. Building that model is really pretty straightforward. You have a set of data that's probably your transaction data which has the idea of customers and the idea of books that they bought.

**Richard Campbell:** Right.

**Donald Farmer:** Training the model against that is really just a question of telling it, "I want to predict the sets of books that people are likely to buy given a particular input book." So, customers who bought, whatever, the latest Harry Potter also bought six other Harry Potter books and I want to be able to take that pattern. It's really all you need to build that model is a very simple piece of information, the idea of a customer and the idea of the books that they bought.

**Richard Campbell:** Then it learns again, say, 75% of the data and then the 25% is where you then say, "Okay. Now, based on what you know about this guy and what book this guy bought, what other books did he buy?"

**Donald Farmer:** That's right. Yeah. The reason that you keep this 25% out is that you do want to be able to test not just how well the model is trained against the existing data, but how well it performs against new data that it hasn't seen yet. The reason for that is that essentially you are going to use this model for predictions because you really want to test this predictive capability.

**Richard Campbell:** Make sure that it is right. I can never imagine especially we're talking about like book buying, it's gonna be 100%.

**Donald Farmer:** That's right.

**Richard Campbell:** It's going to be better.

**Donald Farmer:** The quality of the model and the quality of the predictions that you are willing to tolerate at the low end or happy to see at the high end again really depends on your business model.

**Richard Campbell:** Right.

**Donald Farmer:** It's just a very low cost to offering somebody and making an offer to someone then you can probably have a model that is reasonably predictive, but if it's such a very high cost

acting on the prediction then of course you want it to be as good as possible.

**Richard Campbell:** I could imagine in the gold and silver credit card model, you're talking about taking on some significant liability to offer the higher end card.

**Donald Farmer:** If you get that wrong, yeah.

**Richard Campbell:** Yeah.

**Carl Franklin:** Well then there is also the sort of real time services like these jukebox services like personalized radio stations. Have you heard of this?

**Donald Farmer:** Yes, absolutely. Yeah.

**Carl Franklin:** That's got to be using the same kind of thing.

**Donald Farmer:** They absolutely do. Yeah.

**Carl Franklin:** Yeah. You basically say, "These are some songs that I like. Now, entertain me." It figures out based on what other people listen to.

**Carl Franklin:** Hey, do you find that the horizontal scroll bar is the most annoying thing when you're trying to read that impossibly long line of code? Well, maybe a 19-inch LCD monitor would help. Telerik challenges you to explore their new reporting product and have a chance to give your workstation a facelift. A 19-inch NEC monitor could be yours if you answer a few easy multiple choice questions about Telerik Reporting. Just spend a few minutes and see how easily you can generate Windows, web and PDF reports. Play with the drag and drop data binding. Experiment with Telerik's acclaimed CSS-like customization of reporting items. The Reporting Tool is fast, compact, and very easy to deploy with a mere *Xcopy*. Even if you don't get top marks in the quiz, you can still be a winner. The modest score of 7 correct answers out of 11 questions secures you a complementary Telerik Reporting Developer license that you can use in your personal and professional projects. So, go to [telerik.com](http://telerik.com) and give it a try. It's fun, it's interesting, and it can get you a free license or a new monitor.

**Donald Farmer:** For developers, there are some interesting scenarios, which pop out which are not just about e-commerce applications and not just about



recommendations, but there are some interesting scenarios around things like data validation. So, on the [sqlserverdatamining.com](http://sqlserverdatamining.com) site, there are some examples which are actually hosted. On the right-hand side, there's a link that says, "Live samples" and these hosted applications give you an idea of what data mining can do for scenarios which are, if you like, noncommercial. One of them, a data validation sample, works like this. Given some customer data that has already been input to a system, we can find the patterns within that customer data. So, if you input somebody's age and their education and their salary and their name as you're taking down that data perhaps in a call center, then once you've got a set of those records of existing customers, you can train the model and discover the patterns of that and then as somebody enters a new record, you can validate it by saying "Does this look anything like the patterns that we have already discovered?" And if it doesn't look like the patterns that we've already discovered, then you can raise a warning to the user, in this case call center operator and say, "That looks like that data might be wrong." So, somebody has been entering data that says somebody is buying a mountain bike and they're 120 years old and we can say, "You know, that looks like bad data."

**Richard Campbell:** *Betcha not.*

**Carl Franklin:** Yeah.

**Donald Farmer:** That looks wrong.

**Richard Campbell:** Just in terms of just straight data validation, it's a more advanced way to validate. I was immediately thinking the fraud angle that this doesn't look like a real person. They're trying to fish or they're trying to come up with some scheme to steal from you.

**Donald Farmer:** Absolutely. That's something that we're all probably familiar with and we've had credit cards stopped for transaction queries or so on. They're looking for patterns of activity that may be fraudulent and that's something that data mining is extensively used for. It doesn't need to be that kind of high-end application. It can just be very, very useful and very simple web applications can take advantage of this.

**Carl Franklin:** So, we are talking about SQL Server 2005 analysis services as being the center of the universe in Microsoft for data mining?

**Donald Farmer:** Yeah. It's the center of the universe in Microsoft for analysis and *predictive analytics* being, if you like a kind of advanced form of analysis absolutely lives in there and there's quite a lot of advantages to, for example, using OLAP and Data Mining together or using them separately for that matter. This is where it lives and you do need to have analysis services, that means SQL Server Standard or Enterprise Edition, in order to take advantage of Data Mining. That's true for the developers, true for the data analysts as well.

**Carl Franklin:** You don't need to have reporting services, but if you want to do reports that work against your models, that's probably the easiest way to do it, huh?

**Donald Farmer:** Absolutely, yeah. You can use the DMX Query Language to build reports.

**Carl Franklin:** Right.

**Donald Farmer:** The other thing is even if you're using the Excel add-ons and the Excel add-ons are really an awesome feature. It's freely downloadable feature. If you have SQL Server and you have the new version of Office 2007, you can download these add-ons, which enable data mining directly inside the office environment.

**Carl Franklin:** I just felt about 10,000 developers just go, "Oh, no! Now, my boss is going to ask me for this in Excel."

**Donald Farmer:** Well, it's pretty controllable.

**Carl Franklin:** They love their Excel, don't they?

**Donald Farmer:** The nice thing for developers in Excel is that we actually provide a trace button so that they can see what their boss is doing as well.

**Richard Campbell:** That's pretty cool.

**Donald Farmer:** Actually, that is not the intention of a trace button. A trace button is of course to enable you to track what calls are being made to the server.



**Carl Franklin:** Sure.

**Donald Farmer:** So you can find out what's gone wrong if anything has gone wrong or what models were built and so on, but I can see other uses for it, you know?

**Richard Campbell:** Jumping back, we talked about the sequence from the database, now you built the model, and then once that model works and you can see that it's predicting successfully with their existing data, now you're just making DMX queries through ADO.NET to, say, "Given this customer, what's the likely booklist?"

**Donald Farmer:** That's right. Yeah.

**Richard Campbell:** The challenge here to me sounds like the model is the challenge. Once you get the model right, querying against is pretty painless.

**Donald Farmer:** I'm going to say that's right actually. Querying is painless and the model is perhaps conceptually the most difficult thing to get your mind around. That's actually one of the advantages of the Excel add-ons is that they simplify those models greatly and they allow people to, if you like, learn by example. Otherwise, if you're not going to learn by example, you're going to have to learn from scratch and that includes understanding a bit more about what the different algorithms do and what circumstances they can be used in -- because different algorithms are good for different scenarios. So, you can start from that site and learn, but another thing we're looking at within Microsoft is whether we can provide some packaged models, some kind of get-you-started models for well-known schemas, so for example we could create a customer schema perhaps like a view and say, "Well, look as long as you can get your data into that view, here's a model which you can build over and train over that data."

**Richard Campbell:** Right. So, then you sort of eliminate those things. Every time I play with data mining, I was always asked the question, "Which algorithm do you want to use?" and I never really knew which one to use. There wasn't that many, but you sort of guess, "I'll just try this one and see what the results are like and if it wasn't any good, I'd do it again with a different one."

**Donald Farmer:** Yeah. Professional data miners do that too. They very often run every model, every algorithm and then having created the whole set of models for each algorithm, they'll then use a validation just to find out which it is, but there are some heuristics of course as to which algorithms are based for which circumstances.

**Richard Campbell:** Right.

**Carl Franklin:** I'm looking at the list of algorithms that are available in a SQL Server, decision trees, clustering, time series, association rules, sequence clustering, naïve days. Interesting.

**Donald Farmer:** You've got to love that one. Yeah.

**Carl Franklin:** How about neural network? Is that what I think it is?

**Donald Farmer:** Yeah.

**Carl Franklin:** Wow! I had no idea.

**Richard Campbell:** Interesting you can say that too because when you were describing the sort of learning process, right away I thought, "Fuzzy logic modeling."

**Carl Franklin:** Yeah. This is how the brain works.

**Donald Farmer:** Yeah, exactly. Neural nets are actually very powerful and useful for a quite a lot of scenarios such as classification and estimating future values and so on.

**Carl Franklin:** The other two are linear regression and logistic regression. Yeah, I know what a couple of those are, but that's pretty awesome. I have no idea when you would use what.

**Donald Farmer:** Yeah. So, we provide some advice in the books online about that. Again, one of the nice things that we've done, say, in the Excel interface so people can start to learn what is that rather than thinking about the algorithms, the Excel interfaces talk about the tasks that you do with them. So, for example rather than saying, "Do you want to use decision trees or clustering?" They say, "Go and find customer clusters" or "Go and find categories within this data" and then under that they will use



different algorithms. So, you can take this very task-oriented interface for categorized, find exceptions, associations and from that you can learn which algorithms are used and what the models look like and how they perform. That's probably a much more friendly interface than expecting people on page one of the Wizard to just choose between linear and logistic regression.

**Carl Franklin:** Yeah.

**Richard Campbell:** Now, I'm looking at the live samples and there's the data validation sample which to me is fascinating because it doesn't really sound database-driven at all. How does this work where I'm mentoring these items and it's making suggestions code wise or you're actually making trips to the server to get the data together?

**Donald Farmer:** What happens is all the customer data in that database has already been scanned and the data mining model has been trained against it.

**Richard Campbell:** Right.

**Donald Farmer:** So there's a set of rules and that set of rules will be much smaller than data. The data could be terabytes of data and the set of rules that comes out will be quite small and lives in memory on the server. That set of rules is being queried, but not all the data is being queried. So, just imagine. I mean I don't think there's much data behind the model online, but it could easily build over terabytes of data, but the query at each time is just a query to the rules engines saying which rule is most appropriate for this data and if we don't find the rule that is appropriate then we can return that information to you and say, "We couldn't find any rules and here is how your data fell outside those rules." So, that's a very lightweight interface. There is a server in the background. There is a server and it's running SQL Server Analysis Services, but it's actually serving up a fairly lightweight rules engine.

**Richard Campbell:** Yeah. I guess that makes sense. The big thing here is that you are getting all the data together and it almost looks like you would send this off to the validator before you would send it to the database.

**Donald Farmer:** That's right. Yeah.

**Richard Campbell:** Then it would say, "This doesn't look right." I actually put in some bad data and it came back and said, "Yeah. This doesn't look like this is an appropriate age."

**Donald Farmer:** That's right. It doesn't smell right.

**Richard Campbell:** Yeah. It doesn't smell right.

**Donald Farmer:** Some of those rules can be very difficult for human beings to infer. So, for example, the normal way of doing this if you were writing data validation there is you would in fact be saying, "The maximum age of a customer is 110."

**Richard Campbell:** Right.

**Donald Farmer:** That's a hard rule and if there's an entry which is 112, that's not valid. The kind of rules that we find in data mining are, "Yeah. We've got no customers that are 110 who are buying a mountain bike." It may well be that one day a customer comes along who is 112 and buys one for their great, great, great, great grandson or whatever.

**Richard Campbell:** Right.

**Donald Farmer:** In that case then of course your hard-coded business rule would reject it. The data mining rule might learn that, in fact, there was a legitimate customer.

**Richard Campbell:** Ultimately, the rules are modified. Going back all the way back to that whole suggesting books thing, what's challenging about that is new books coming all the time...

**Donald Farmer:** That's it.

**Richard Campbell:** And new customers coming all the time.

**Donald Farmer:** That's true.

**Richard Campbell:** Purchase behavior changing over time.

**Donald Farmer:** And this is why I said that now and then the data mining model may learn that there is a valid customer. One of the very important things you should do with data mining models is you don't



just build them once. You in fact train them continuously.

**Richard Campbell:** Right.

**Donald Farmer:** Not necessarily continuously, but at least regularly. We have data mining within Microsoft that we use for our own marketing and they run a system where they in fact use a data mining model to do the predictions, but at the same time they have two or three candidate models sitting in the background being tested and as soon as one of them starts to outperform the model that is in production, that will be promoted into production, the reason being that things do change over time. At this point, we're in the height of summer as we're talking and it's a beautiful day out there and there's probably people out there buying beer and barbecue fuel and that's a fairly typical kind of summer sale.

**Richard Campbell:** Right.

**Donald Farmer:** In the winter, they're not going to be buying those things. They may probably still be buying beer, but they're not buying barbecue fuel along with it so in the middle of December if somebody buys some beer and you suggest, "Hey, do you want some barbecue fuel?" They're going to think you're crazy.

**Carl Franklin:** Well, you've never been to Richard's house for a barbecue.

**Richard Campbell:** I remember a story like this when we were explaining the whole data warehouse OLAP analysis concepts back in the late 1990s. There was the beer and diaper story.

**Donald Farmer:** Right. Yeah.

**Richard Campbell:** Do you remember that story?

**Donald Farmer:** Oh yes! Absolutely.

**Richard Campbell:** It's a funny story. I'll tell it or you could tell it, but it's just that this whole idea of predictive analysis.

**Donald Farmer:** I believe the story is an urban legend.

**Richard Campbell:** Yes.

**Donald Farmer:** It was used as a potential example, but the idea was that a retail customer supermarket mines all its data about who is buying what and it discovers that there's a correlation between people buying beer and people buying diapers. I think particularly on a Friday evening was the idea. There's a number of potential explanations depending on your sociological bent on this one, but one of the explanations was that people, men in particular, stopping on the way back from work on a Friday evening, ready to watch sports for the weekend or movies with their buddies, they'll pick up some bottles of beer and then thought, "Hmm... Hang on. I'm feeling pretty selfish here just buying beer and I'm going to get into trouble when I get home so why not pick up some diapers and show that I'm a bit of a new man as well." So, they find this correlation that nobody else would have found if they hadn't done data mining.

**Carl Franklin:** Well, it's still a theory. It's not validated or anything.

**Donald Farmer:** The correlation could have been validated.

**Carl Franklin:** Right.

**Donald Farmer:** This one is an urban legend, but I think what's useful about it is it does tell you something that's pretty significant about data mining and the way in which people use data mining. It's because people are looking for some sort of, I could say some sort of magic, but it's more like some sort of innovative insight. They're looking for an advantage. They're looking to find something...

**Carl Franklin:** Yes.

**Donald Farmer:** That no one else has found. So, data mining results to be successful, it's not just enough that they are accurate. It's not just enough that they are actionable in the way we're talking about earlier that you can do actually do something with the information. People also want them to be innovative. If data mining just gives you the same results as you would get from reporting or analysis, reporting or OLAP, then it wouldn't be so exciting. It's the idea that you might get results that you couldn't discover just by exploring your data, but it's actually going to discover something new that you would've overlooked when you were exploring that data.



**Richard Campbell:** Of course, the actual element on that beer and diapers story is to make sure you put the diapers by the beer.

**Donald Farmer:** Absolutely.

**Richard Campbell:** Because the guys going in there for the beer, and some of them are smart enough to go, "Oh, I better get diapers," but most of them aren't and so if you put the diapers nearby, they're more likely to buy them.

**Carl Franklin:** I think if you drink enough beer, you'll need a diaper.

**Donald Farmer:** Exactly, yeah. I think nowadays, you might put the Depends by the beer and see, you know?

**Carl Franklin:** That's right. Diapers are for me.

**Richard Campbell:** Oh, man.

**Carl Franklin:** We'll edit that one out. *No we won't.* So, let's get back to the algorithms because this is the fascinating thing for me. What if I want to use an algorithm that's not in the list, that isn't supported by SQL Server? Am I out of luck?

**Donald Farmer:** No, you're not out of luck. There are nine algorithms in that list and if you want one that's not there, there are a couple of options. One is to ask to provide it or look around for someone who has already written it, but there's also an option to write your own algorithms.

**Carl Franklin:** How interesting.

**Donald Farmer:** You can write your own algorithms and manage codes and install them into SQL Server Data Mining and there are quite a few people who've done that. I wouldn't say there's a big aftermarket of algorithms yet, but there are certainly people out there who are writing them and they can be pretty useful. People can sometimes write specialized versions of algorithms or they may write some entirely new algorithm or some algorithm that they learned in college that they're rewriting again for the SQL Server platform. We certainly encourage and provide a lot of help and assistance to people who want to do that.

**Carl Franklin:** It seems to be though there's a lot in the box just with the, what is it, 11 or 12 that you have.

**Donald Farmer:** I think it's nine and then there are some variations on them. Yeah.

**Carl Franklin:** Yeah.

**Donald Farmer:** I mean there is a lot in the box and frankly, they meet a lot of the business needs and the request that we've had for other algorithms have tended to be from people who've got some really highly specialized requirements.

**Richard Campbell:** You're getting to analytics that are specific to the given enterprise there and I'm thinking archaeology is a candidate actually. There's a lot of interesting analytics I've seen about archaeology for predicting this is likely where something will be.

**Donald Farmer:** Yeah. I mean some of the areas in which you might see some developments in the future might be around stuff like spatial analytics, but one of the questions we have to kind of look at there is, again, are the algorithms that we have interesting enough or good enough for that. I think over the years you'll see us adding more algorithms, but I also want to kind of encourage this aftermarket algorithm whether that aftermarket is commercial or shared community source, I'm not too worried, but certainly it would be nice to see more algorithms out there.

**Richard Campbell:** You've mentioned spatial. Right away, I'm jumping on geospatial data, the idea that we could have predictive models looking at the addresses of customers and saying, "You are likely to find more customers here."

**Carl Franklin:** Big Brother.

**Donald Farmer:** It's the equivalent of what people do. One of the nice things about data mining is that it simply works at a scale with a speed and with an efficiency...

**Carl Franklin:** Right.

**Donald Farmer:** That allows you to discover these innovative insights and make innovative predictions.



**Carl Franklin:** Absolutely.

**Donald Farmer:** But the process is pretty much what people would do even without predictive analysis. The beer and diapers is an example of somebody finding a pattern that would simply have taken a huge amount of work to find any other way and would probably have been overlooked, but there are plenty of scenarios where people are doing that work already.

**Richard Campbell:** Right.

**Donald Farmer:** I always say that we all do *predictive analytics*. If I'm a shelf keeper and in January I'm talking to the sales rep and I'm working at how many boxes of Valentine cards I want to stock for February 14th then I'm doing predictive analytics.

**Richard Campbell:** Yeah and you're doing it out of your gut rather than out of your data.

**Donald Farmer:** Absolutely. Yeah. I'm doing it from my experience rather than an analysis of the data. I'm kind of doing it from my own gut feeling analysis.

**Carl Franklin:** So, getting back to developer experience for a little bit. I'm reading about the viewer controls from the Business Intelligence Development Studio and SQL Management Studio, are redistributable.

**Donald Farmer:** That's right, yeah.

**Carl Franklin:** If you want to get that same kind of look in your application, you can download and plug them in.

**Donald Farmer:** That's right. They're reasonably easy to integrate as well and they do give you those fairly complex visualization of data mining that can be pretty useful for users because very often the rules set, something we discover are pretty complex and when you're using them in an application of course it's nice to be able to present them to the user in a very simple way where you just give the user a number or you give the user a suggestion, but if you're somewhere in between the developer of these rules or somewhere between an analyst and a simple end user and you actually want to see some of these

rules and understand how the rules interact then the visualizer is great for that.

**Carl Franklin:** Are these controls managed controls? Are they COM controls? What are they?

**Donald Farmer:** These are managed controls. Yeah.

**Carl Franklin:** Very good. That's so good to know.

**Richard Campbell:** So, everything is .NET now.

**Donald Farmer:** Yeah. Everything is .NET now. Yeah. That's the future and of course one of the reasons being is it's very, very simple to develop against these things.

**Carl Franklin:** Yeah.

**Richard Campbell:** I don't think we could probably do a data mining discussion without saying the word *dashboard* at some point.

**Carl Franklin:** Oh yes.

**Donald Farmer:** Dashboard.

**Richard Campbell:** Yeah.

**Carl Franklin:** All right. Next.

**Richard Campbell:** Isn't that like the normal...?

**Donald Farmer:** Next question. Yeah.

**Carl Franklin:** Yeah.

**Richard Campbell:** You're just going to move away from it.

**Carl Franklin:** When I think business intelligence, I think little googley gadgets and gauges and dashboards and things.

**Richard Campbell:** Arrows are pointing up. That's good. Arrows are pointing down then that's bad. That's all I know about my business.

**Donald Farmer:** I hope you don't mean googley gadgets in the sense I think you mean.



**Carl Franklin:** No, no, no, no.

**Richard Campbell:** Don't say the G word.

**Carl Franklin:** I didn't mean that at all.

**Donald Farmer:** No. I think...

**Carl Franklin:** Blinky lights.

**Donald Farmer:** People absolutely do think in those terms about how do I make this visualization kind of straightforward and easy and just point towards the trends that're going. I've been doing a lot of work recently on how do you build predictive dashboards. The classic dashboard that people see when they are building these scorecards and dashboards is "here's what happened last month. Here's the way your sales have been going and look, they're going badly," or "here's the number of customers who turned away from your service last month." There's a problem with that, which is it only tells you what happened and you've got to do the interpretation of what's going to happen.

**Richard Campbell:** Right.

**Donald Farmer:** Sometimes it can be straightforward, sometimes it can't be. What I like about data mining is that with data mining you can build predictive dashboards, dashboards per se, not just "this is what happened," but "this is what's likely to happen." Give us a prediction of how many customers will turn next month, at the very least giving you some time to do something about it.

**Carl Franklin:** Right.

**Richard Campbell:** Yeah. "Stop telling me how I failed."

**Donald Farmer:** Yeah. Yeah, exactly. If all we're showing you is dashboards that tell you what happened, that doesn't really help you very much. Some of them of course were involved in planning and tell you, "Okay. This is what happened. This is what we'd like to happen," but that additional predictive element of "this is what is likely to happen" on top of that is I think a really, really valuable addition. The great thing about SQL server for doing this is because the language that we use for querying predictions is the same as the language that's an extension to SQL. It can be used very easily. It's

quite possible to build dashboards that are built over very simple predictive queries. You're not having to include additional objects you can use standard reporting services, performance point server, performance management application or ProClarity as a VI tool or Excel for that matter.

**Richard Campbell:** Right.

**Donald Farmer:** To build these.

**Carl Franklin:** I'm thinking out there for the average developers listening to the show looking for a new edge on their career, this is probably a way to make yourself very valuable to a company is to learn the ins and outs of analysis services and data mining.

**Donald Farmer:** I think this is a real advantage for developers to have this under their belt because I talked to a lot of companies who know this stuff exists and think it's beyond them and to have people in-house who actually have those skills in this area would be tremendously valuable for them.

**Richard Campbell:** They probably already own it. If they own SQL Server, they already own all the software necessary to do this.

**Donald Farmer:** This is something that I find all the time in executive briefings. I'm called in to do an executive briefing and people sometimes say, "Well, why are you talking to us about data mining?" I say, "Because you already got it."

**Carl Franklin:** Right.

**Donald Farmer:** This is something you're not using and I see jaws dropping open and people say, "Oh, we've got data mining?" I say, "Yeah. You've got the fastest, most powerful data mining engine already and it's sitting there deployed and a number of servers for your enterprise and you're not using it," and they are stunned to hear this.

**Carl Franklin:** Well, Donald, let's keep the love going maybe with the dnrTV episode where you could actually show us how to, from a developer's perspective, do some of this stuff. What do you think?

**Donald Farmer:** I think that will be great. I've been telling you how visual it is. I think we better show you some of it.



**Richard Campbell:** Yeah. Show some visualizations.

**Carl Franklin:** Excellent. Ah, I can't wait for that. That's going to be fabulous.

**Richard Campbell:** I do think the appeal here is that from a programmer's point of view is these are the tools that influential people in the company tend to use.

**Donald Farmer:** Yeah.

**Richard Campbell:** The sorts of visualization tools. Where are we going? What are we doing next? Those folks, you make them happy, you're going to be happy.

**Carl Franklin:** You become very valuable to the company too.

**Richard Campbell:** Absolutely.

**Donald Farmer:** I think that's true. This is an incredibly useful skill to have. At the TechEd Conference that we had in the US, Rafael Mizrahi did a session for developers specifically about using data mining to build intelligent applications and this case was a beautiful one. You can use data mining to build intelligence into all your applications whatever they are because any of your applications which process data, any application which is recording data can take advantage of this.

**Carl Franklin:** I guess the first thing is just knowing what questions to ask, right? What are those questions that the users are going to be asking when they are visiting our website, for example, and what kind of information can we give them that would just be above and beyond?

**Donald Farmer:** That's right and those questions ultimately drive the choice of the predictable columns that you choose for a model in technical terms. What it is that you want to predict becomes the predictable and how you're going to make that decision becomes the input and at the very simplest level, that's the easiest way to kind of think of that process.

**Richard Campbell:** Suddenly, I have an urge to turn the mining modeler against log files, start predicting when I'm going to have more errors.

**Donald Farmer:** Absolutely, yeah.

**Carl Franklin:** Oh wow!

**Donald Farmer:** Start predicting when you're going to run out of disk space.

**Richard Campbell:** Right. Stop thinking about the customer. Start thinking about your information system.

**Donald Farmer:** Yeah. Yeah.

**Carl Franklin:** Interesting.

**Donald Farmer:** Data mining over things like SQL Server and domain management views becomes pretty interesting.

**Richard Campbell:** Yeah. You've got all this data here. Mining is not magic. All it's doing is finding relationships, stuff that you didn't know about. So, just giving it a set of data and telling it how these things are connected together then it finds the anomalous relationship.

**Carl Franklin:** Is there any relationship that it can't find? What is the big difficult, holy grail now in 2007 for data mining? What are we still waiting for?

**Donald Farmer:** In terms of relationships that it can't find, I would say that the most difficult relationship to find are ones which are offset where you'd have to go through a number of relationships to find it. So, my favorite example of that is let's say I've got some information on customers who have returned goods to me. I know which ones completed their orders and I know which ones were unhappy and returned the items to me and I have a lot of information. One of the pieces of information or two pieces of information I have are their order date and the delivery date, the date they placed the order and the date the goods were delivered to them. Now, with those pieces of information, at present, if I do a data mining model what I'm going to do is I'm going to look at all the order dates and find out which ones are significant and I'm going to look at all the delivery dates and find out which ones are significant. All I'm going to find there are, are there any order dates that

are significant and are there any delivery dates that are going to be significant.

**Carl Franklin:** Yeah.

**Donald Farmer:** What is actually significant is perhaps if the distance between the order date and delivery date is too great. So, I might find out for example that if the delivery date is more than 15 days after the order, customers are more likely to return goods, but at moment, a present data mining can't find that relationship unless I explicitly put in at the right column that calculates that difference.

**Richard Campbell:** Right.

**Donald Farmer:** Here is something that I think is important, go and mine it.

**Richard Campbell:** Right. You have to create the elements of data that that number of days to delivery, you define that.

**Donald Farmer:** Yeah.

**Richard Campbell:** The miner won't create relationships that way.

**Donald Farmer:** So, I have to hypothesize about that and I make that calculation. I think one of the things that we're going to be trying to do in the future more and more is find those relationships for users, make some of those hypotheses for them probably using heuristics and so on to do that.

**Carl Franklin:** Well, is there anything that we haven't covered?

**Donald Farmer:** Wow! We've covered a lot of ground, haven't we?

**Carl Franklin:** Let me ask you this. If and when we do this dnrTV, what would you like to show?

**Donald Farmer:** I think I'd love to show two things in particular. The data mining add-ins for Office which are just beautiful in terms of the design and their ease of use and until you actually see them, it's really difficult to get across just how powerful and easy to use they are and then I'd like to show some of these web applications, some of the interfaces that developers can build. How you can embed data

mining into your own applications. I think that would be powerful to show as well.

**Carl Franklin:** All right. Excellent. Well, Donald Farmer, thanks a lot for coming on the show. It has been enlightening for me. It's always great for me to hear two people who know what they're talking about, talk about a topic intelligently, you and Richard.

**Richard Campbell:** There is always a funny dynamic on this show where sometimes it's a Carl show, sometimes it's a Richard show.

**Carl Franklin:** Yeah. This one's yours, buddy.

**Richard Campbell:** It's so right.

**Donald Farmer:** Well, it has been a great fun for me. I hope you've learned something from it, but certainly for me it has been a blast.

**Carl Franklin:** I certainly have and it has been "a gas" for me as well. Thanks a lot, Donald, and we'll see you next time on .NET Rocks!

[Music]

**Carl Franklin:** .NET Rocks! is recorded and produced by PWOP Productions, providing professional audio, audio mastering, video, post production, and podcasting services, online at [www.pwop.com](http://www.pwop.com). .NET Rocks! is a production of Franklins.Net, training developers to work smarter and offering custom onsite classes in Microsoft development technology with expert developers, online at [www.franklins.net](http://www.franklins.net). For more .NET Rocks! episodes and to subscribe to the podcast feeds, go to our website at [www.dotnetrocks.com](http://www.dotnetrocks.com).