

The eLearning Coach Podcast #15

Don't Fear The Data

with Ellen Wagner, Ph.D.

Show Resource Links: <http://theelearningcoach.com/podcasts/15>

Connie: Hello everyone and welcome to episode 15 of The eLearning Coach podcast. Collecting data about our behaviors and preferences has pretty much become a part of our culture by now for better or for worse. In the world of training and education, organizations collect a lot of data about their learners.

How can we take advantage of that data to make us smarter and to make better decisions? The best person I know to answer this question is Ellen Wagner, Ph.D. Ellen is Partner and Senior Analyst for Sage Road Solutions. She is responsible for learning industry intelligence, briefings, enablement, and executive solutions and services. In this capacity, Ellen represents the interests of the Western Interstate Commission for Higher Education.

She is the former senior director of Worldwide eLearning at Adobe Systems and was a tenured professor and chair of the Educational Technology program at the University of Northern Colorado.

For some context, let me tell you something important that Ellen said before the interview, and I quote “there are certain inflection points in industries where things change in profound and fundamental ways. As eLearning did this for learning and development, now data is going to do this for learning and development.” I hope you get a lot out of this interview, enjoy.

Connie: Hi Ellen, welcome to the podcast.

Ellen: Well, Hi Connie, it is great to be here with you today.

Connie: I am so excited about exploring this topic with you. There is just so much I want to learn about it, and I have a feeling the listeners do too. Let us start out with some general context. What is big data and how is it being used today?

Ellen: Well, I have to imagine that there are not too many people, whether they are learning professionals or not, that have heard people talking about big data. It has all been in the news lately. Some of it coming from things like what the NSA is tracking about, our phone calls, our emails, and our various transactions. So big data is pretty much on everyone's mind, but big data is not necessarily all data, and I think big data is starting to sharpen our awareness of the fact that every single time we have an online transaction, every time we engage with somebody over some type of medium, we leave little digital breadcrumbs of everything we are doing.

So the promise of being able to keep an eye on all that stuff is pretty compelling, and we are finding in some places it is taking place. In the learning world, we are really not to that point. In fact, I would be really surprised to find much going on in learning and development that is going to require true big data. Here is a way to think about it. Big data is typically not something that you track in columns and rows, so when you think about most of our learning organizations, we do keep track of a lot of our transactions with one another, but more than anything, we are keeping track of people and places and things, all of which is pretty easy to track in a spreadsheet.

Big data is typically so big that you cannot track it using traditional means. Hopefully, that will help to let people know that when you are dealing with things like Amazon or when you are dealing with things like AT&T keeping track of their phone networks or thinking about weather predictions, you are dealing with modeling that is taking in so much information that you have got to come up with really very, very different ways of finding patterns and finding insights within all that information.

But in learning and development, I know for years, many of us have, we roll our eyes a little bit around the so called smile sheet data, the satisfaction data. In truth, if you are dealing with online learning or some type of digital experience design, most of us do not have a lot of patience if we are not satisfied, do we? So how much time does it take before we are unsatisfied and click off. Well, that is the type of data that we do have at our fingertips and that actually is not that hard for us to get hold off if we decide that we really do want to get more intelligence

around the actual behaviors of the people with whom we engage all the time.

Connie: So getting intelligence about the behaviors, is that why the learning industry should be interested in big data?

Ellen: We should pay attention what is going on in big data because it is going to be impossible for us to live outside the expectations of the rest of the world. Education as a sector is 1 of 17 sectors in the US economy. It is the only one that does not use big data as a regular part of its practice when you get outside of some of the basic research in science or medicine or some of the true applied research areas where you are diving into disciplines. Financial modeling, for example, that is an example of where big data really does step in; retail is a place, but for most of what we do in learning and development, unless we have ways of being able to tie what we are doing into some of those bigger transactional parts of the business, it is smart to pay attention to what the buzz is and it is also smart for us to not get tripped up in the fact that there is a whole lot of little data we already have really readily access to that can help us make better decisions about how we can help guide our practice.

Connie: Ellen, that makes perfect sense to me, I am really glad you explained that. What kind of problems can a learning and development community solve or prevent, and what kind of decisions can we make when we are looking at whatever data we can get around the training and education industries?

Ellen: One of the most exciting reasons for this growing interest in data across all the sectors has been that when you can take a look at the evidence of what is going on and you can start anticipating where you can go in a particular direction is the ability to make better decisions. So for anyone who is thinking about these data as being a lot of research reporting, that is not necessarily where organizations are going to get the biggest value. In fact, I think for many of us, we have collected data in the past,

I can think of just about every learning organization that starts making investments in technology finds themselves looking at some type of learning management or content management system, because it is hard to keep track of all the things that we do without having ways of building automated spreadsheets, and in many cases, what we have been building in LMS's can be viewed really as spreadsheets on steroids, if you will. There are different things that get collected, but for the most part, we have been really trying to do reporting with this information. It is just now that we are starting to get a sense of our own abilities to anticipate a particular trend by looking at where we have been in the past, that the idea of making inferences from data is becoming clear and actually more readily available to the mere mortals among us who do not work in huge organizations with the data staff. Remember that those of us in learning and development typically do not find ourselves in really big organizations. So I know one of the questions we will get to in a bit is what is it that an instructional designer can actually do or expect to get from using data. In some cases, what we are trying to do is simply look at where we have been to see if there are trends we can anticipate to get ahead of the trend a little bit or if we can anticipate points where there might be problems, then we can get ahead of the problem and get a solution in place before the problem even happens. I am going to use an example from real life that might illustrate this.

Connie: That will be great.

Ellen: I am sure that many of us this last holiday season found ourselves presented with some type of a digital device that we are wearing on our bodies to keep track of our steps, our activity, our calories, our sleep, whatever. There are many brands of them, so I am not going to point out any particular brand, but the idea that we as people can now not make excuses for how much activity we think we are or are not having. As a human who was wearing one of these devices, I will tell you that I thought I was actually more active than I am. No big surprise, but when you actually look at the true number of steps one takes in a day and you can say "oh, that is not as many as I thought."

This is not rocket science. What it tells me is that if I believe that I should be more active, I have evidence that shows that I need to step it up a little bit, and it is really that simple. If there are indicators that people can find, that either suggests that they are on a path to success or conversely that they are on a path to rack and ruin, what you want to do is to get ahead of the fact that you should not ever have to be surprised by that.

So giving some type of an assessment where there are certain questions that just simply do not work, you might want to take a look at those, it probably puts the entire validity of your test at risk, the idea of scratching ones head and going “wow, something is not working here.”

Most of us who are experience practitioners will always default to our experience. I mention this because even when we have the very best data in our hands, we will almost always not pay attention to those data if they are counterintuitive to what we believe as experience practitioners. It is one of the reasons why I think people in our discipline and our profession, we as instructional designers, we have really, really good processes, but a lot of us push back a little bit when they get to the measurement and evaluation part of our work.

We say things like “well, our customers won’t pay for it.” Well, okay, but do we want to continue down a path where we actually get to a point where we produce something that does not work? No, of course not. We have all got pride in our own ability to keep track of where we are headed. So I think sometimes this idea of big data or being data savvy, there are points of sophistication and I do not want to sidestep that, there are people known as data scientists who are different than people like you and I perhaps, who really do think about relationships with big patterns in very different ways. We want to encourage that, and frankly, we do want to have those types of specialized skills on our teams as we move forward, but that should not excuse any of us who are working IDs to sidestep the reality that, as a professional we kind of like to pretend that we do not have to measure it, because we are artists, so we know in our hearts it is going to work, and I say that with love.

Connie: I understand.

Ellen: This is not a criticism, and you are laughing because of course I expect that you probably know a person or two who is like that, right?

Connie: One of my best friends but not me.

Ellen: Precisely, and I have of course have always been data driven, so it is the art of design, and I do not want to have anyone think that we are going to have to give that part of it up, but I do want to share a story, and I hope it is illustrative. I have used the expression “living under the sword of data” for a while now, and it does refer to the sort of Damocles where sometimes you make decisions where you have to operate and making decisions where you are not comfortable because you do have the sword of the decision-making hanging over your head. Leadership can be tough and making decisions? You want to find ways of being able to have those decisions work in your favor.

But this is actually a reference to something more explicit. A few years ago now, a guy by the name of Douglas Bowman, who was the lead designer, visual designer, at Google, resigned from his position and actively blogged about the fact that he was doing so because he was distraught. He was distraught because as a designer, he had found it impossible to live within the engineering culture at Google when it came to feedback on his visual designs, and he went on to point out that it was difficult for him to sit in meetings where engineers would debate the value of a 3 or a 4 or a 5 pixel line of a visual design that was not being driven by data, in his expression.

So his resignation came with this pronouncement that I can simply no longer live under the sword of data where my designs are concerned, and it really resonated with me at the time because it was just at the point when learning and development people were starting to get the word that data was going to start driving the ways in which we approached putting our learning experiences together. The way that we might select content from an array of content that had been voted up through some type of crowd-sourced input.

This is one of those things where it is easy to talk about it in the abstract, but I know a lot of designers, instructional designers, who come to this practice believing that they are going to be architects of big learning ideas and sometimes we are a little surprised to discover that maybe we are not the architects, maybe we are the general contractors or maybe we are the carpenter, and these are all really cool things but it suggests to me that we take a look at what the skills are there that help us do the

best we can, and it comes down to measurement. Connie, a lot of it comes to measurement, and most of us do not feel comfortable in that space.

So it is just a long story to say that I think designers everywhere are finding themselves as really having to think differently about their practice and that how nice to be amongst designers of all stripes that have to figure out how we can be both qualitative and quantitative in how we approach our practice.

Connie: So you are saying that we need some of both?

Ellen: Oh, absolutely. Who among us wants to paint by numbers all the time, really?

Connie: But in the real world, right now at this point, what kinds of patterns related to learning are people finding?

Ellen: Let me tell you about a specific project that I have actually been working on, because I know that most designers like big ideas but we also like to know how it actually works when you touch it. About 3 years ago, while I have been working as one of the executives for a higher education collaborative group, some of the members, universities and colleges, were talking about how cool it would be if we as a bunch of practitioners could aggregate our student's data, de-identified student data, to form a single federated database that would allow us to look for new patterns that could inform us about student loss and student momentum, which is basically technical speak for what makes people drop out and what makes them persist over time, so that they can move toward actually completing a degree or a high value certificate.

Anybody who pays attention to education these days knows that there are all kinds of calls for reform in education and increasing America's college graduation rate and being more economically competitive as we move into the true information age, and unless anybody thinks the colleges and universities are not keenly involved in this and concerned about this, they are, but like most organizations thinking differently about data, it can be really hard.

There is a lot of accountability in higher education, and many of it comes from how an individual school competes with others. So it is not easy to think about sharing all your data to do something proactive, but we found 6 universities and colleges that agreed to this, we did a pilot program, we

got some money from the Bill and Melinda Gates Foundation to actually support the work, and what we discovered is that we could take all of these records from the students, put them into a single data set, run predictive analytic on these numbers and come up with patterns that actually helped us figure out how to create models that, now that we continue to exam it, we can tell literally, if I am working with a specific institution, we can tell them what 17 students are most likely to fail Math 101.

Connie: Wow!

Ellen: It is wonderful to be able to use these patterns, and it is a very sophisticated way of doing predictive modeling, but using consistent reliable variables for building the model, so that in our case, it is actually going to customize to every school that uses it, and that is pretty important because the things working in data that we are all finding is that it is really easy to talk about how everybody has to work together to raise the bar, but everybody competes with each other.

So how do you find common ground? Well, you figure out how to have the big overarching conversation but you never lose sight of the fact that each organization really does have to take care of their own constituencies. Again, I do not want to get too deep in the weeds on this.

What we have been finding though is that you can do a pretty good job predicting which students are at risk, but what it revealed was that simply knowing who is at risk really was not ever going to be enough, so the big ah hah for all of us was having a predictive model to tell you what is going on is super-important, but if you are not prepared to step up and address what you are finding, then you are going to not ever really be able to realize the full value of what your data are telling you. So it is a two- part conversation, and I just wanted to make sure that we got both of those parts in there.

Connie: I understand that. What would be a corollary of that type of pattern to a company, to a very large corporation?

Ellen: What if you were looking for ways to identify the different competencies on two levels, one would be your incoming workforce. If not just finding the best graduate per say because the thing we are finding is that businesses want to bring really skilled people into their organizations for specific types of work but they want to be able to know who is most

likely to be flexible enough to be successful in other places in organization.

So being able to anticipate these successful hire in organization is analogous to the types of things we are seeing in universities, is how do you recruit the students who are most likely to be successful under the different types of programs and conditions.

The other thing you want to be able to do, and here is another corollary, is once have people in your organization, how is it that you can support them in ways that allow them to realize their full potential? So it is not just on recognizing, getting people into the pipeline, as it is once people are in the pipeline, you have got to make sure that if you want them to grow that you can anticipate how to do that in ways that help you realize your investment as an organization but also help get a person to be in a position to give the most value back.

Connie: Okay, so you are saying that you can figure that out possibly by doing learning analytics, you can find those patterns?

Ellen: It is possible if one sets up the queries in the right way. Now, this is not like Magic 8-Ball. So if anybody thinks that asking a question of your data and expecting the answer to emerge fully formed, it does not happen that way. One of the other big learnings that we have had from working with data projects is that, about 80% of your effort in a big data project is going to come with making sure that your data are clean and valid and reliable and that they are going to reveal to you the types of answers you are looking for. If your data do not work together or if it is inconsistent or if it is not complete, you are not going to get anything. We have talked for years in the computer world about garbage in--garbage out, this is no different than that. So making sure that what you are doing is clear, if you do not know what you are looking for to begin with, it is going to be pretty hard to know if you have got the right data for answering those questions. Is it possible to find a way of identifying competencies that people have that map more directly to the competencies required for success in an enterprise? Absolutely.

Connie: Okay. So what are some other ways that that type of data is collected in an enterprise?

Ellen: Oh that varies within the enterprise, and this is one of the places where I think we have some tremendous opportunity for doing a little bit of what I

would call data readiness, because we have never thought about trying to find data that would complement the big question, we have never thought in terms of extracting the data to do insight work.

Different organizations are going to take different amounts of time to get ready for this. In the example I was telling you about, one of the things we had to do was to create common data definitions that every single one of the universities and colleges that we have worked with would be able to use, and that can take a heck of a lot of work when you are looking at the research university over here would be able to talk to the community college over here and know that when they talked about grade point or withdraw, that those words actually meant something.

Connie: Right.

Ellen: So what you are going to find even inside a big enterprise, you are going to have to make sure that what learning and development tracks is relevant to what sales and product support tracks which is the same thing that other parts of the organization might track. I am fairly certain we are going to find that in some cases, it will be initially like comparing apples to zebras, but this is why having the big question in mind makes this whole thing a lot less onerous because it is far easier to imagine finding the data you need to answer the question if you have the question in mind first.

Connie: That makes sense...really fascinating. Are many companies or institutions hiring statisticians or people with the skills to analyze the data, I am curious?

Ellen: There is shocked realization across learning and development, whether it is in education, whether it is in training, there really are not a lot of people with these types of skills in our particular slice of the ecosystem. You are finding that a lot of data scientists are actually coming out of biomedical research, medical research because of the need to do so much of the huge modeling that is required to start thinking about big data and using some of the pattern recognition software.

Here is the funny little thing about working with this for learning and development. Anybody who has ever taken a research class, if you have gotten a master's degree, typically somewhere along the line, you probably had to have a research class of some kind, and one of the first

things that you learn in that class is data snooping is bad, you cannot do data snooping, do not do it.

Because if you are developing a hypothesis, you have got to control the conditions of testing your hypothesis in the scientific method to make sure that what you are doing is tenable at whatever level of significance you said as a statistician to show that you can infer from your data that something is likely to happen.

That is the whole point of using what are known as inferential statistics. You try to control your conditions to the point where you make your very best guess. The thing about predictives is that you are not guessing. You have got so much data that you are working with that you are seeing the answer in front of you, so you are really not testing hypotheses anymore, you are actually looking at the true results of the work because the numbers are that big, and that is counterintuitive and almost anathema for people who are learning educational research methods because we do not think that way, we had to think really differently. So I mention this because there are a number of places where you are starting to see people jumping in and providing certificate programs around predictive analytics.

Connie: Oaky, so you are saying that now people are using predictive statistics to work with big data rather than the inferential. Let me ask you a question about the Experience API, the newer learning and development standard, do you think it has the potential for improving the quality of data for learning analytics?

Ellen: Well, I think it has potential. I think that the work that the teams have been doing and the companies that have been jumping onto testers are doing really important work because we are going to have to all get a heck of lot smarter about what we track, how we track it, particularly as we start moving into technologies like Google Glass or even the types of GPS sensor things we have on our mobile phones right now.

It would be a shame to waste those data if we can pull things from there that do give us insight. I have lived in the software world long enough to know that having a standard toward which people can focus their development energies is important, it is the common definitional framework that you need to get the work done.

What I also know is that standards typically do not get adopted until you tip into a critical mass of people who are already using that standard. I am thinking very specifically about SWF and thinking very specifically about PDF. Those did not become standards because Adobe or Macromedia said they should or that they worked with standards organizations to make sure that it happened.

Those became de facto standards and eventually Adobe released those to the standards communities, so that everyone could work with it, but it was after there had been enough adoption, that there were entire industries that were created around, say the open SWF specification. As you know, so many learning tools that had used flash as foundation for content creation were able to do so because the SWF spec was available for people to work with.

Connie: Right, so now we are really at the early stages, and I do have an interview scheduled for a podcast with someone speaking about the Experience API.

Ellen: I just want to make sure that the idea of exploring where this can go is important, and I applaud what the efforts that are going on. There are probably going to be a number of competing efforts that are going to be looking for some of the same types of harmonizations, so the Experience API is getting some good traction, I think that you might want to ask about some of the other things that are starting to go on, but the point is that if we are finding ways to make all of these data more readily available and help us get smarter about ways in which we can help use those data to make our practice better, why would not we want to explore it? I just do not think we should overpromise on what I can do now, and I say this with great hurt and with great feeling.

There are some amazing things going on with predictive analytics using Hadoop technologies and true big pattern seeking work in semantic analysis and neural networks and other types of places in predictive analytics right now that are going to absolutely change the way we think about learning design. I also know that for most organizations to make a decision about even the simple things that we keep track of in our spreadsheets is like moving icebergs.

So I do not want us to find ourselves getting frantic at the idea that this is going to pass us by. It has taken is a really, really, really long time to

look at this thing called eLearning to tip into broad public consciousness that it is probably an okay thing for us to be talking about now.

Connie: Really, I know.

Ellen: I guess my excitement in all of this is every so often there are moments in design evolution where those of us who are creating the experiences actually have a chance to really do something profoundly different, and I think this is one of those times. So do not fear the data. If there is one thing I would say, do not fear the data and do not be freaked out that “Oh my gosh, I do not know what to do.” Think about ways in which you can get better and think about how you would make the case to get there.

Connie: Ellen, are those your parting words because I was just about to wrap up? Do not fear the data is a great note to end on. Thank you so much Ellen.

Ellen: You are welcome.

Connie: One thing I love about listening to Ellen speak is that she has a mega big picture view of things, and that is helpful for all of us to understand where things are headed, and if you are interested in show notes, go to the <http://theelearningcoach.com/podcasts/15>. I have put a lot of resource links there for you. As always, thanks for listening, and I will catch up with you at the next podcast, thank you.