Transcript

[music]

Derek Bruff: Welcome to "Leading Lines," podcast for Vanderbilt University. I'm your host, Derek Bruff, Director of the Vanderbilt Centre for Teaching. In this podcast, we explore creative, intentional, and effective uses of technology to enhance student learning, uses that point the way to the future of educational technology in college and university settings.

In this episode, I speak with Katy Borner, Professor of Information Science at Indiana University, Bloomington. Dr. Borner is the curator of a traveling exhibit called Places & Spaces -- Mapping Science. The exhibit, now in its 12th year, features assassinating print and interactive visualizations capturing science and how science is done.

Vanderbilt's hosting the exhibit this spring and due to my own interest in data visualization, I couldn't pass up the opportunity to talk with Dr. Borner while she was on campus.

One quick note. Dr. Borner mentions an online course she teaches called "Information Visualization." She refers to the course as a "MOOC." In case you're not familiar with that term, I wanted to say that it stands for "Massive Open Online Course." The course is free and open to anyone in the world. As a result, it typically enrolls thousands of participants.

Derek: I started my conversation with Dr. Borner by asking her about this Places & Spaces traveling exhibit.

Katy, thank you for being here today. We're happy to have you on campus and sharing some of your work with us. Can you start by telling us a bit about the Places & Spaces exhibit? Where did it come from? What are its goals? What does it look like if people were to see it?

Dr. Katy Borner: The Mapping Science exhibit is now a 20-year effort to bring large scale

maps of our collective scholarly knowledge to general audiences.

The exhibit has been traveling to six continents and more than 200 cities. 240 map-makers have contributed to the contents that you can now see here at Vanderbilt University.

Among others, you will see 100 large-scale maps of science. You'll see some illuminated globes. You'll see interactive illuminated diagrams, and you'll see the Macroscope Kiosk. It's actually wonderful that Vanderbilt University is still viewing the 12th generation of this Macroscope exhibit.

You might wonder what macroscopes are. I'm sure you are familiar with microscopes and telescopes which help you to see things that are too small, or too far away to been seen with your naked eyes. Macroscopes here help you to see structure, patterns and trends in large amounts of data.

If you go and look for the Macroscope Kiosk, then you will be able to pick different data sets to see them in an interactive mode and to explore them on your own time and with your own interests. And to also communicate what you have learned to others.

Derek: Can you give an example of a kind of data set that I could find in the macroscope?

Dr. Borner: One of the macroscopes shows Earth and it overlays very many different real-time data sets, including weather data, including ocean streams, including particles so you can see the dust over the deserts, but you can also see a lot of dust over China.

You can zoom in to different areas. If you are an avid wind surfer then you might like to zoom in to that area you are trying to do your sport in, and you will see what the wind patterns are looking like.

Also, macroscopes show you traffic data, such as shipping data, which is global today. Other macroscopes help you to navigate all the reports of the National Academy of Science, and again other macroscopes help you to get a more detailed understanding of news activity or Twitter activity.

Derek: Can you say more about one of my those last two, because you mentioned maps certainly, and I think some people have a very concrete idea of what a map looks like, there is

geography, there is places, the physical geography of the world, but when it comes to the activity of the National Academy of Science or activity on Twitter, there's not necessarily geography to that, but you still use the term "map." Can you explain that a little bit?

Dr. Borner: The map of National Academy reports uses a topical map where the continents are not Europe, Asia and Americas but instead, say a mathematics, and physics and biology and arts and humanities.

Just like on a geospatial map, if you zoom in, you get to see more details, and so you get to see subdisciplines, and you might get to see subdisciplines of biology if you zoom into Biology. Just like a geographic map, you can now use these maps of science which have topical continents, to overlay other types of data.

You can overlay your career trajectory, you can overlay expertise profiles, you can overlay emerging research areas or bursts of activity. Twitter map in the Macroscope Kiosk actually happens to have a geospatial map. It shows London, and it overlays so called smells over this map, and that is why it is called the "smelly map."

Among those smells are smells that are pretty good. They smell like green parks and forest. Whereas others smell like really yummy food. Others again smell not so nice.

You can actually hover over different areas on the map of London and get to see what kind of smells are typically tweeted about in those areas. Using Twitter's stream data, massive amounts of it, you can then extract certain key terms and use those to visually encode what kind of smells exist in what areas.

Derek: That's fascinating. I can imagine that by mining a lot of Twitter data you can find out all kinds of patterns, whether they're smell-related or geospatial-related or not. In that case the visualization helps you to navigate and see connections between things.

Often when I work with faculty to add more visual thinking tools into their teaching, I will make the point that for a lot of these, the main idea to remember is that special relationships convey meaning. Whether it's geospatial map where adjacency has the meaning, or concept map where connecting two nodes implies a relationship between those. That's often a piece of it.

I want to talk about the curriculum piece as well. With the exhibit being a public exhibit, what are you trying to convey? What goals do you have for the exhibit?

Dr. Borner: Many of these maps, and also some of the interactive data visualizations, were originally created for experts. It's actually quite a translation process to make them understandable to a general audience, which you might meet in a public library or at a conference, or in other public places.

What the exhibit curatorial stuff does, is to help researchers and experts with that translation step. Because we believe that in today's time and age you have to be able to not only read and write text, but also to read and write data. Data visualizations oftentimes have a very important role to play when communicating key facts from one scientific discipline to another or from experts to novice users or from one generation to another.

That could be a child explaining to his mom and dad how Snapchat works. Or the other way around where somebody might explain to his or her child how a certain theoretical concept works. Oftentimes drawings are made, and oftentimes there's some data behind what is communicated.

But really empowering anyone to embrace large-scale data sets, and to use some of those new tools and approaches to make sense of the data, and to generate actionable visualizations, that's also a main mandate behind the exhibit. Which not only shows maps and macroscopes, but also points people to online resources such as the information visualization MOOC, for instance, that helps people learn how to empower themselves and source new tools.

Derek: Say more about that. Why is this kind of visual literacy important? To what extent should we be building it into our normal curriculum?

Dr. Borner: Most scientists today, they always use statistical techniques to analyze data. However, there are also temporal, topical, geospatial, and other types of analysis which are more and more relevant for answering when, where, what, and with whom questions.

The latest question refers to network analysis and visualization. Oftentimes you have to go to different sciences to learn about how to, for instance, make a geographic map. You would do this in a cartographic geographic department. You would take a course there. For statistical

analysis you would go over to math or to statistics. If you wanted to do lexical analysis you would go to a linguist and learn about topic analysis and mapping.

Some of the network science approaches you can find in physics courses, but also in biology and social sciences, and many other sciences now. By having a set of tools that actually does all these different types of analysis to help you answer these complex questions about how networks evolve, or where bursts of activity exists in the landscape of geography or science is really helpful for just getting your foot in the door and learning to have fun with large-scale data sets.

Derek: I teach a statics course occasionally and we do a little bit of data visualization. I actually a few years ago decided the course needed more of that. I was seeing some of these trends and the availability of large-scale data sets which used to be harder to create and access and now that's getting increasingly easy.

I show one of Hans Rosling's TED Talks where he's giving some global health data and he's got this really great visualization that he uses to tell stories really, is how I see it. It's that communication piece. But I know you can also use these tools to find the stories in the data as well. I guess I'm excited by bringing some of these technologies to students to help them grapple with big data sets. Or to find stories that we don't know that are out there yet.

Dr. Borner: Everybody who hasn't watched Gapminder and Hans Rosling's several TED Talks, just go online and do it. [laughs] It's really amazing the abilities he had to communicate data clearly. And the Gapminder tool is also available online. You can feed in your own data.

Derek: How can faculty go about teaching some of these skills? There are different ways to go on this but as I think about working with faculty, talking to faculty here and elsewhere, who might want to incorporate more of these learning objectives into their teaching, what are some pedagogies or practices they may adopt to help teach students these types of skills?

Dr. Borner: I think there are many different ways now to get that expertise first, and then to communicate to students and maybe also parents and children at the same time [laughs] . It's not just relevant for the academic setting where we have students for four or maybe for eight years. I think we need to also teach to older and younger generations and start younger.

Derek: I'd like to unpack a little bit kind of what makes this hard for students. For instance, in my stats class we had an assignment where they had to create an infographic. It was an application project. They had to find some real data, ask some questions on it, do some statistical analysis, and then share that visually. I really wanted them to practice that communication piece of that.

During one of the activities leading up to that, I had them look at some infographics and analyze and see what makes them tick. I wanted them to start to move to that visual piece. I found that some students just kept thinking of the visualizations as just aesthetic. That it's just making it look nice.

I'm wondering how you would work with students to help them understand there's more to it than that. It's not just making something pretty. There's more there. Have you encountered that challenge when working with students?

Dr. Borner: Definitely. I think part of it is that data visualizations are oftentimes sold as eye candy. And then of course students buy it as eye candy. [laughs] If you see a visualization, let's say a big network, and there's no legend that would tell you what the notes represent, or what the linkages represent, then it might be a wonderful piece of art but it is not communicating any kind of insight.

It cannot because there's no explanation of what it even represents. Similarly, if you have a geospatial map and you don't really know what area of the world it actually represents, it wouldn't be very valuable for you.

Ultimately, I tell my students that the data visualization they generate should have both. They should be aesthetically pleasing so that they capture people's attention. Even if they are on a two-page spread in the Wired Magazine, let's say. Where on one side you have a gorgeous looking woman with some perfume advertisement, and on the other side is your data visualization.

How do you make people even look at yours? They have to attract attention. Then immediately they have to funnel that initial attention into something that draws people into the data and lets them explore that data visualization and understand it, ultimately.

That's not trivial to really make it easy for people to have a low floor and a high ceiling,

ultimately. But also helps them maybe even go beyond what they see in the data visualization.

Have sufficient background information so that they can go back and even get to the raw data. Could potentially replicate that data visualization because that's the hallmark of science that you could potentially replicate. And ideally even feed in their own data if it's for instance Fitbit data or some kind of bank account statement data.

Imagine you would actually have that entire work floor also easily available and now you could look at your own new data in novel ways. To get there I think it takes scientific rigor. You have to help students understand that they have to provide information on data provenance. Where that data came from. If the data is not good then the ultimate result will not be good either. It's trash in, trash out pretty much.

At the same time they also need to reveal what kind of data preprocessing, cleaning, analysis, and visualization work floor they used. What kind of algorithms, what kind of parameter values so that others can replicate. Oftentimes they will also need to provide initial description of what is actually seen in that data visualizations so that others have an initial description. A little story which helps them understand what they can take away from the visualization.

They might go farther and discover other additional elements, but having a little bit of explanation of what data was used, how it was treated, and what key insights there are, is very helpful and should come with every single data visualization.

Again, if you just see a picture like that over there on the wall, it's a piece of art, it's not a data visualization.

Derek: Sure. I'm hearing a couple of things that I like. One is that having students create the data visualization, but then also create a document where they imitate it, or explain it, or unpack their own process a little bit.

In my course with my assignment, I got this from a colleague at another institution, Sidney Eve Matrix. She had her students do infographics but they had to turn in a designer statement with it. When I'm evaluating them I can see what they did, but then I can see the kind of behind the scenes decisions that may not be evident in the final product but are

important in understanding the student's process. And what process they followed. I like that.

I think you have to grade all that. It has to be built into the assessment, data provenance things like that, has to be built into the whole assignment. I think otherwise students do tend to focus on the final product and not on the whole process and the pieces that really do need to go into it.

I also like what you said about the aesthetics. My students sometimes push back and said it's just looking pretty. But actually designing something so that it engages the viewer, it captures their attention, particularly if it does that as it helps them understand what they're looking at, that's really good design. That's not easy.

Dr. Borner: Ideally it even works on multiple levels. For instance, if you're 10 meters away, it should still be interesting so that people actually come towards it. Then let's say, a meter away, they see new elements and they get more interested and maybe they zoom in farther. Then 10 centimeters away…

[laughter]

Dr. Borner: …of course, it's easier. They can move out pieces because it's printed in 300 or 600 DPIs or you actually need to get close to see all the details. Then, of course, online you can just zoom and zoom again and zoom yet again to explore data. As you have already noted, there's always see aspect of data exploration and data communication.

Exploration would be to look at a data set from many different angles, look at answering these, when, where, what, was, whom, questions. Then, as soon as you have a good answer to a key question, trying to keep that particular visualization in a format -- and it might be a printout -- so that other people can benefit from that insight. That's then meant for communication. That's where storytelling also comes in.

Derek: I was with a friend a few weeks ago at an art museum. We were looking at some impressionist paintings. She's an art teacher. It was fun to watch her look at the painting from afar, and then walk up and look at it very close. There were different things she was looking for as she encountered it that way. I love the idea of doing that with scientific data visualizations as well.

Let's talk a little bit about technology. This is an educational technology podcast. Are there ways in which technology helps teach this domain? Are there challenges associated with using technology here?

Dr. Borner: What I see many teachers do and also students do is to go online and look for online resources, which are either parts of MOOCs or part of online sites where somebody just has a genius ingenious explanation of a certain mathematical concept or a great description of a novel by a really famous author.

They just have their students or themselves watch these examples of how to communicate that concept very, very clearly.

Obviously, there are a lot of commercial providers out there which have a million dollar to really teach mathematics 101 in a perfect way. Normally a classroom teacher doesn't have a million dollar to optimize his or her course in that way. By really creating these bridges to, especially, free online materials, I think we all benefit because we have ultimately better materials for us, and for our students.

Then teachers become more, and more curators, which stringed together good bits and pieces of how to teach a certain concept well in a hands-on or in a very theoretical, or anything in between way, and brings these online resources into the classroom, because guess what, when student go outside they'll have to do lifelong learning.

Science and technology is moving so fast that after two, three years of not taking additional schooling or additional online courses, students' knowledge and skills will be close to obsolete. You cannot expect, in a very fast-moving area of science, for instance, to just do four courses, of course work, and then have that knowledge in place.

Derek: Sure. We've got to teach our students how to learn. I think it gets back to one thing we said earlier about helping students understand a set of principles, a framework, a way of thinking about data, that then they call apply as they learn new tools.

As new techniques emerge, new different types of data emerge, some of those ways of thinking about data will certainly carry through. When I think about some of the more connotative -- of course this has been a conversation in the mathematics school for some time -- do we let students use calculators?

Most of my colleagues here at Vanderbilt do not use calculators when they teach mathematics, and they have their reasons for doing so. I found it helpful to give students access to a lot of tools, because it allows me to ask harder questions, and allows me to focus more on the concepts actually, by offloading some of the computational work to a device.

I imagine with data visualization there's a bit of that too. There's the figuring out what good data is, and what's the right kind of visual heuristic to understand, or see stories. There's a tool that will help you do that, and you have to learn how to use that tool.

It's that conceptual work that's often harder to teach, but it's the kind of conceptual work that students will take with them, and learn from in their future learning.

Dr. Borner: Yeah. Many of the students which we have in the information visualization MOOC, they already live a future where new data sets become available every single day. When they come into their office in the morning, and these are corporate offices often times, there's a new data set, and it's highly relevant for their biological experiment, or for their chemistry profession, or for their teaching job.

In addition to these, there's always a new tools, there are new approaches, there are new methods. It can feel very overwhelming to come to your office every single morning [laughs] and to see this influx of new data tools, approaches, and of course also new ways to tell stories.

If you are a journalist, I think you also get to see new methods and new ways to defeat the beast of foolishness [laughs] of this, hopefully very well-researched stories. How do you empower today's students to live in such an environment and I think it has to do with having certain theoretical frameworks they can hold on to in this very flexible, very dynamic world.

But, also having the means to use social networks, and to use all kinds of mechanisms from Google to Twitter to other social media to be on the pulse of time to really understand what kind of new techniques are going to make it. You should invest time learning them, and which ones might be soon obsolete because you have only 24 hours each day, and many of them are very much booked up already.

You have to still stay on the cutting edge, because otherwise, bad things happen along, people don't get course they deserve, or students don't get the kind of expertise they would

like to gain from a university.

Derek: Couple of other things I've heard as you've been describing all of this, one is that I'm guessing with your MOOC, the tools that your student use are probably free or open access tools. In some cases at least, right? I mean that's the way technology helps is that we have greater access to tools, greater access to data sets, greater ways to share or collaborate.

You mention the team projects in your course and I imagine that they're able to actually all work on the same data visualizations, collaboratively which used to be harder than it was, and also to pull on their own data as you said. Your own Fitbit data, your own Keeper data. I find that doing that, the collaboration piece and having students generate their own data gets them more interested and invested in the process and more willing to try something a little bit weird to them.

Dr. Borner: Definitely. Students are strongly, strongly encouraged to, not only do their homework we prescribed, but to start using these tools in order to analyze their way on data. Many of these students are actually already in the industry so they have seen large-scale data sets, and they now get to have the means to analyze, and look at that data in new ways.

And yes you are right, open data, open code, open education go beautifully together because oftentimes a new data set needs a little bit of hand-holding. A little bit of explanation of when to use it and when not to use it. Similarly, tools need to have descriptions of how to run a certain workflow and again, what parameter values to use when and when not to use that tool.

At the same time, I would also argue that if you now have projects that students do together, they actually all have a focus and they have a goal and that also aligns different cultures, different types of expertise.

Oftentimes, in these teams, which we try to have formed in the second part of the course, we advise that they should try to have a data wrangler, somebody who can deal with big data. They should have a data analyst. A designer, if possible, not always possible, but often possible. And then you have to have a very strong report with your client.

You have to have somebody who is a little bit of more extroverted so that he or she can talk and set up meetings with the main stakeholders to really design his or her dream to make

sure that his or her needs and priorities are met.

In addition to also doing a great job doing the data analysis and visualization, but if you don't know perfectly well what exact data analysis and visualization truly answers some question, then you might design something that nobody needs at the end. That's not very satisfying in life or for the client, to be honest.

Derek: That's fascinating and helpful too, I think. Imagining the roles that need to be there, and having students either enhance their existing skills in one of those roles, or trade off and try a new skill set, I think that's really important.

Dr. Borner: And they learn talk to each other. The designers learn to talk to the programmers, the programmers learn to talk to the project manager, which might be the official title for the person who keeps the contact to the client but then of course, they also have to write a scientific paper about it.

It's not always clear who has those scientific writing skills.

[laughter]

Dr. Borner: It can be any one of them, or they all write it in a very collaboratively, but they oftentimes get a publication out of these projects, and oftentimes they have a great portfolio item.

Several students just come to the IVMOOC just to get that part of the course and they of course take the initial theory and hands-on part, but what they really want to do is to apply their knowledge. As an engineer, I believe that they have to apply what they just learned otherwise they don't know what they just got.

Derek: I think you're right. For a lot of students who take a MOOC like that, they're less interested in the certification piece and more in the product that they help produce because that's something they can share with employers, future employers and colleagues. That's really persuasive.

We have one question we ask all of our guests at the end, and it is this. What is your favorite analog educational technology?

Dr. Borner: I have many. We do a lot of 3D printing. You can actually take a data visualization let's say migration pattern…

Derek: I would argue that's a digital technology.

Dr. Borner: No, it's printed.

Derek: The product is physical. I'll give you that.

Dr. Borner: The product is printed, but if you, for instance, have a map of the US with flows of people, and people arrive oftentimes in New York still. That's a mountain and then they move over their life to different places, and then they die somewhere.

If you look at Florida it's definitely a valley. People don't come back.

Derek: [laughs]

Dr. Borner: That's one of my favorites because it tells a little sad story. What else?

Derek: You'll design this map virtually but then print a physical copy that has that topography on it.

Dr. Borner: Almost every library now has a 3D printer somewhere. You should check that out.

Derek: Thank you very much for speaking with us today. I really enjoyed this conversation, and I'm looking forward to your talk here in just a little bit.

Dr. Borner: Yeah, on maps and macroscopes, and if you're here at Vanderbilt University, check out the hundred maps. If you just spend one minute with each of them it's a hundred minutes. Sorry for that.

Derek: [laughs]

Dr. Borner: The macroscopes are maybe more engaging, so bring your family over. You also have a number of really cool speakers coming in. Really neat new workshops will be taught, and ultimately there is a student competition I heard.

I hope you can point them all to the Web links that comes with the competition because I'm going to look forward to judge some of those. I'm curious to see what you have to show of your work.

Derek: That's fantastic. Thank you very much, and in the show notes of the podcast I'll put links to your website and the exhibits websites so that folks outside of Vanderbilt will know where it's going next.

[background music]

Dr. Borner: Very good. Thank you.

Derek: That was Katy Borner, professor of information science at Indiana University, Bloomington. Places & Spaces is on display at Vanderbilt University through April 23rd. Here on campus look for the exhibit in the Sarratt Student Center, the Central Library and the Wond'ry, Vanderbilt's new center for innovation and entrepreneurship.

Vanderbilt is hosting a number of events this spring on the theme of data visualization including, the student data visualization competition with cash prizes.

We've got a couple of students planning to revise their infographics for the competition. See the show notes for more on data visualization events at Vanderbilt this spring.

I do have some sad news to share. Hans Rosling, the statistician I mentioned in the interview, passed away in February just a couple of weeks after I spoke with Dr. Borner. His TED Talks changed how I thought about data visualization, and I share them with my students, to show how really good data visualization can help us find and tell stories. Rosling's work in global health care was significant, and he will be missed.

Take a few minutes to watch one of his talks today. I have put a few links in the show notes. You can find the show notes on our website leadinglinespod.com. We welcome your comments and questions there and on Twitter where our handle is @leadinglinespod. You can subscribe to our pod cast through iTunes or your other favorite podcast app.

If you like what you hear in the podcast, please leave us a rating and review on iTunes that helps other listeners find the show. Leading Lines is produced by the Center for Teaching, the

Vanderbilt Institute for Digital Learning, The office of Scholarly Communications at the University Library and the Associate Provost for Digital Learning. This episode was edited by Rhett McDaniel, who also designed our logo.

Look for new episodes the first and third Monday of each month. I'm your host Derek Bruff. Thanks for listening.

[music]

Transcription by CastingWords