

3M Podcast Transcript

Mary Ellen Bates:

Welcome to the Springer Nature Podcast, bringing information professionals to text and data mining in which we explore the roles that corporate information professionals can play in text and data analytics initiatives.

Mary Ellen Bates:

Today, we're talking with Jonnie Hauswirth, Rachel Wangerin and Jennifer Nelson of 3M, a diversified manufacturing company where scientists, researchers, and marketers collaboratively apply science to improve lives. Jonnie and Rachel are co-managers of 3M Knowledge Discovery & Analytics. And Jennifer is a senior technical information analyst at 3M Knowledge Discovery & Analytics. All three are working on defining the intersection of information science and data science.

I'm Mary Ellen Bates, principal at Bates Information Services, and I'm working with Jinfo Limited, a global industry analyst group on this podcast project.

I started the conversation by asking Jonnie about how the 3M Knowledge Discovery & Analytics staff became involved in text and data mining initiatives.

Jonnie Hauswirth:

When we think of text and data mining - really lump that into kind of information and data analytics as a whole. And over time, this has just been more from services that we've been providing. We do a lot of custom research with the technical community at 3M, and it just became an extension, partly our own staff interest in where the field was going, what more could we do? And then a little bit of pull from our clients, some demand. They were seeing new tools become available and thought, "Well, how could this apply to what I'm doing?" So, I'm thinking this goes as far back as theme scape and looking at topographies of a technical area, we've been doing it for a long time.

Mary Ellen Bates:

And was this as a result of your promoting or marketing your particular skills in this area or were your clients proactively coming to you, asking for your expertise and advice?

Jennifer Nelson:

I'd say it was both. I've certainly been involved in projects where the client interacts with me and they think they're maybe going to get 20 articles to read. And at that point I stepped in and I was like, "There are some tools we can use to create our own hierarchies of information, to really drill down to the content and figure out some trends without you having to read every article." But then there have also been people who have come to us and said, "I saw this article in Nature, where they talked about an algorithm they used to read the articles. Can I do that?" So, it's both.

Mary Ellen Bates:

What do you think are some of the unique skills and insights that information professionals bring to a TDM project?

Rachel Wangerin:

I think there's a lot of different components that go into it. A lot of different skills that information professionals have. One of them really is at the beginning of the conversation with our internal clients at 3M. Helping them understand what exactly they're looking for. So, if they come to us and want to know everything that ever happened with a specific polymer, that might be great, but if it's a polymer that is in everything, then we're talking about millions of patents. So, helping them figure out exactly what their question is and focus on that then helps us figure out what types of tools would help answer that question.

On the flip side, if they're looking to use data from one of our publishers, we can help them understand the copyright implications, figure out what type of costs we might be talking about, some sort of ballpark, or at least prepare them to have a conversation about how much it might cost to buy data for the project they're looking at. One of the rules of thumb that I've heard before is you don't want to spend a \$100,000 trying to answer a \$10,000 question. So also thinking about that as they try and figure out how they want to proceed.

I think another component of that as well is the fact that information professionals have training and understand how data's indexed. And so, in data science, in text and data mining, you often talk about clean versus unclean data. And what does that mean? And working with our publishers to try and take advantage of already indexed data, you're talking about a cleaner, potentially cleaner sets of data and helping our end users understand what that means and how it's indexed and how to take advantage of that indexing with maybe algorithms they're looking at.

Mary Ellen Bates:

Have you seen any surprises or unexpected benefits from being involved in these initiatives?

Jonnie Hauswirth:

Well, I don't know if it was surprising, but a benefit that we talked about is really raising the awareness of our group. With each interaction that we have with our clients on a project, it's showing the capabilities of the team and how we understand enough of their technology to be able to talk to them about that, but then enough of the information landscape and our own industry, to be able to bring that to bear and the more interactions we have, the higher that awareness goes and the higher recognition that we have then across the company to be able to be at the table, bring that expertise to bear on a project and say that consultative role, but being recognized that we have the expertise to help them get that next step to what they need to find the solution, to get the answer, to get something done.

Mary Ellen Bates:

When you're thinking back on the TDM projects that you've been involved with over time, is there anything that you wish you'd known ahead of time? Any advice that you'd give to someone who's considering getting involved in TDM projects in their organization?

Jennifer Nelson:

When I look back to the projects I was working on four or five years ago, at the time, I really was thinking of the traditional librarian mindset. We don't want to be opinionated. We want to give good data and our users can do what they will with that data. They can come to their own conclusions, but I

have found the text and data mining projects I've been really invested in and really involved in are more successful when I bring my opinions to the table. And that's really what our end users are looking for. Particularly in that consultation role. They don't want us to just say, "Here you go, go have some fun with it." They want us to say, "Here you go, go have some fun with it. Here are my top-level findings. We can talk about it after you look at things." So, I wish I had done more of that earlier.

Mary Ellen Bates:

There are just so many exciting new developments in the field of data science itself. As you look forward, what's next either in general or there at 3M?

Rachel Wangerin:

Right now, our team is really trying to figure out that answer to that question. I don't know that we have it exactly yet, but part of it is upskilling, which I think is talked about quite a bit in information centers, especially corporate information centers, but probably also in academia. Figuring out who on our staff is interested in upskilling, who would like to take that journey and then upskilling in the ways that are going to help us figure out better ways to serve our clients, to find those more efficient ways to look at information, to look at data. So, I think that's a big part of where we are right now is, is just trying to figure that out.

A lot of vendors are starting to incorporate different types of AI, data science, text and data mining. They call it all different types of things, but we're seeing it more and more in our vendor products. But part of what we're exploring too, is tools that are more fundamental, like RStudio, Python, things like that, to figure out what we can do with that as well.

Jennifer Nelson:

Going back to my comment about being more ... wishing I had been more opinionated. I think this is a place to be brave and try things and fail. And frankly tell people when you failed and what you've learned, because that is, that is part of the practice. We can't only do the same things and tell people about the same things we're doing when we're trying stuff in the background, and maybe it's working, maybe it's not.

But if you really try to maybe automate usage statistics and draw some conclusions about topic trends from those, and it doesn't work, that is a good finding because that tells you that you don't need to go off and figure out how to automate all of your statistics collection. You can do what you've done and redirect that energy elsewhere.

So be brave, tell people what you've done. Tell people what works, what doesn't work. Sometimes I think information professionals, maybe this is just me telling you guys about my brain a little bit, are afraid to be wrong. And we can't. We have to be willing to experiment, to fail, to tell people what we've been doing. Hopefully we'll succeed more than we failed, but you can't succeed if you don't try.

Mary Ellen Bates:

What a great note to end on. Jonnie, Rachel and Jennifer, thank you. That was Jonnie Hauswirth, Rachel Wangerin and Jennifer Nelson of the 3M Knowledge Discovery & Analytics team, speaking about the roles that corporate information professionals can play in text and data analytics initiatives. Listeners,

thank you for joining us for Springer Nature's Podcast on key topics on information strategy and management. Learn more about this project by visiting Springer Nature's resources for information professionals.