



## **Interview with Thomas Deely**

**For podcast release  
Monday, March 29, 2021**

**KENNEALLY:** In the hype cycle methodology created by the Gartner research company, an innovation trigger sends a new technology climbing up the peak of inflated expectations. On Gartner's most recent hype cycle for artificial intelligence, knowledge graphs have just reached the summit. Ahead will lie disillusionment for some, but for others, enlightenment and productivity await.

Welcome to Velocity of Content. I'm Christopher Kenneally for CCC. On Wednesday, March 31st, at 11:00 AM Eastern, CCC presents a special town hall program, "Solve for Success: The Transformative Power of Data Visualization." I look forward to hosting the program, and I hope you will join me by registering for a complimentary ticket at [copyright.com](http://copyright.com).

As Gartner has noted, knowledge graphs are increasingly popular tools to make sense of our digital world. They display information in ways that help researchers identify important relationships, relationships all too often submerged in an overwhelming sea of data.

In enterprise settings, knowledge graphs as well as other visualization techniques display customer, partner, and third-party data to yield insights in context that can drive business innovations. The Knowledge Graph Conferences showcase leading innovations and topics in the field, focused on elevating social good applications of knowledge graphs and related technologies. Conference co-founder Thomas Deely joins me today with a podcast primer on knowledge graphs. Welcome to Velocity of Content, Thomas Deely.

**DEELY:** Thanks, Chris. Thanks for having me, and great to be here.

**KENNEALLY:** We're looking forward to chatting with you. As we say, knowledge graphs are climbing up that peak that Gartner describes, the one of expectations. You, though, and your colleagues at the Knowledge Graph Conference are working to sort all that out and to make it really something that people can understand and work with in their own jobs or in their own research. I think it helps for the listeners to hear a little bit about the history of knowledge graphs. They're new on the scene, it seems, in 2021, but really they have been around at least conceptually for many years before that. Indeed, you can go back at least as far as 1945.

**DEELY:** Yes, and potentially even further to 1882. There was a concept of existential graphs. I think what's interesting about knowledge graphs is they're building on layers and ideas and



technologies which have compounded over many, many decades. Today, where we have more data than ever before, we have more compute power, we can link that data together more. We're now asking questions of computers in the same ways we'd ask questions of humans.

Some of our listeners probably have an Amazon Alexa or a Siri, where you're asking a question, and our expectations around the answer we're getting back are becoming higher and higher. We're asking more sophisticated questions. We're hoping for more sophisticated answers. And the engine behind that – there are many technologies. There's natural-language processing. There's machine learning, artificial intelligence. But the plumbing around how those models access the data more and more is being driven by a special type of database called a graph database, and specifically a knowledge graph.

So a long history going back to Vannevar Bush and the Memex machine, Tim Berners-Lee in the early '90s talking about the World Wide Web, the concepts of hypertext and you can link from one web page to another, to Google coming up with the term knowledge graphs about 10 years ago, to today, where that AI and NLP revolution which is continuing to grow at an accelerating pace, with knowledge graphs being the infrastructure providing the data to those models. So there's a lot there, but a very exciting time for the technology.

KENNEALLY: Indeed, it is a very exciting time, but I don't want to let you get too fast to the present day, because I want to just go back to that 1945 article – Vannevar Bush and his conception of the way we were going to file information. It was an automatic personal filing system, I understand, that he called the Memex. This was something that was the beginning of the thinking that did lead, as you say, for Tim Berners-Lee – to his conception of the web, and in 2009 to this notion of progressing from just documents on the web to putting data on the web, and therefore getting to a point where we had these relationships between data that are called linked data.

DEELY: Yeah, the concept has been around for some time. I guess the concept of the Memex machine almost like being – as you mentioned before, it was called a personal filing system, almost like in the same way in our brain, we have our own filing systems. We have long-term memory. We have short-term memory. There's different parts of the brain based on neurons and networks, and ultimately, the way we're going now in this kind of continuous evolution of technologies beginning to mimic more and more of that technology in our brain in machines, if you will, with knowledge graphs being kind of the underpinning around that.

And our ability – the early stages of the web, it was more around defining a web page, getting documents up there, linking to web pages. Google had their famous search ranking algorithm which drove their search engine, which made them the company they are today. They figured out that the more entities that linked to a given web page, the more relevant that web



page was – so that concept of the relevance of a piece of knowledge or information forming part of the algorithm.

Then they evolved that towards their concept of a knowledge graph. Let's say you do a query on who is Leonardo da Vinci? Having a database that understands Leonardo da Vinci is a person. It's not a person or not an object. And then having that meaning and context around the search, being able to link that to other aspects of data and provide a more intelligent answer to the question that the person on the other end may be asking.

KENNEALLY: Right. And for businesses, for research institutions who begin to use knowledge graphs, what are the kinds of assets that they are working with?

DEELY: Yeah, great question. There's a wide variety of use cases, and it's one of the reasons we founded this conference, because the technology is beginning to get a lot of adoption across every industry. It's a little similar to, let's say, the cloud or blockchain, where more and more, these technologies are cropping up whatever the sector you're in, whether it's finance, health care, media, e-commerce, etc.

Also, there's lots of public, open domain datasets, and then there's private datasets. I'll give an example of one of each. In our first conference, we had Airbnb talk about how they were building this feature, Airbnb Experiences. So if you're booking a vacation to, let's say, Italy, and if Airbnb can figure out a little bit about you in terms of – is it a family trip? Are you going with others? What are your interests? Are you an outdoors person, into activities? Are you more like an indoors person? Are you into food or culture? And blending that into, let's say, a set of experiences that they could potentially offer you around the fact that you've actually booked an Airbnb apartment for a week in Rome or some nice place in Italy.

Obviously, that also has to be linked to what's available in Rome. What are the offerings for the outdoor person? What are the offerings for the person who's more, let's say, interested in culture or based on a certain age? So linking together data around the Airbnb consumer which the location that that person has booked their Airbnb rental and the time that they're there. That's a really interesting broadly applicable use case, where Airbnb is consuming information they have, but also perhaps some public information that's available.

Then there are also large-scale datasets – open data is the term. One of the workshops we have during our conference is the United Nations have set out these grand goals around the Sustainable Development Goals. I think there's like 17 or 19 goals. These are grand goals which range all the way from eliminating poverty to making sure everyone has access to water to improving education, etc.



Now, there are tons and tons of publicly accessible datasets around economies or patterns or behavior or farming patterns, etc. that the UN has access to. What they're looking to do with knowledge graphs is figure out, if I can move the needle on education, which is one of the SDGs, then that will also have a trickle-down effect in terms of economic growth. So I'll be able to impact multiple SDGs by focusing on the education. But building those linkages between the goals and the metrics to achieve them and the interaction between each of the SDGs is something where you need lots of datasets, and you need an intelligent database to be able to build those inferences. On the flip side, if, let's say, one of the SDGs might be around economic growth, if you're doing a great job there, then perhaps the climate or environmental score might be at risk. So lots of interrelationships, lots of public data.

Those are two examples. One is more of a private data, but combining private and public. And then the UN one, they're partnering with Accenture, who are our lead sponsor, on accessing lots of publicly available datasets which the UN has access to.

**KENNEALLY:** Right. That conversation that the whole world is having about data and research is, of course, focused in 2021 on the COVID-19 pandemic. And just another example of the ways that knowledge graphs is helping to yield insights is a project that you have cited as well, the covidgraph.net (sic) project. We will have as a guest on our town hall Dr. Alexander Jarasch, who's the head of data and knowledge management at the German Center for Diabetes Research. What can you tell us about covidgraph.net? (sic)

**DEELY:** Again, this is a great example of a community-driven project, where you have multiple contributors supplying different datasets, and you have many members participating in that.

I think the interesting thing about COVID – at one of our talks last year, Konstantin Todorov, who's focused on fake news, used the term that not only are we in a pandemic, we're also in an infodemic. So we have now access to more information than ever before. Particularly if we look at COVID, all these kind of sources on, let's say, something as simple as measuring the rate of spread, whether it's people who are ill, mortalities, tests – all this kind of confusion around the state a given country is at in terms of mitigating that – multiple different views.

And then on the other side is the innovation, like coming out with the vaccines, having all of the research. Think of how rapidly we have been able to deploy today – just a year into the pandemic, we have at least half a dozen different vaccines, which maybe a few years ago might have taken decades or many years to build and distribute. Now, they're beginning to be rolled out across the world. So the quality of research, having a knowledge graph validating the quality of research, patent applications linking to prior groups, measuring the progress of clinical trials, and then applying it to progress and rollout across different geographies.



That's a little bit of background on the graph. Interestingly, at our conference, we will have three pharmaceutical companies present. We'll have AstraZeneca, we will have Novartis, and we'll have Hive, who are working with Bristol Myers Squibb. We'll also have a workshop on personal health records. Think of all the data about us that is now available through our behaviors, whether you have an iPhone or a Fitbit or an Apple Watch, and how that can potentially feed into better medical processes – more on the preventative side rather than on the curative side for us as individuals.

So broadly, health care, life sciences, is a particularly rich area of opportunity for knowledge graphs given the volume of data that exists in the space and then just the complexity of understanding that data and finding that source of truth. Because with health care, it's one area where you can't really mess up. If you do something wrong, it can have devastating effects.

KENNEALLY: Well, the other piece of the Knowledge Graph Conference that I hope you will tell us about, Thomas Deely, isn't so much the science of this, but the community that is coming together around knowledge graphs.

DEELY: Yes. When we launched the conference in 2019, we weren't quite sure how it would go. It went really well. We had 200 people. We did very little marketing. We've never done much marketing, so it's been through word of mouth. And then we noticed the energy at the conference. There was so much excitement, so much exchange of ideas, so much fascination around learning this new technology which frankly most people have still not heard of. So we see a rich opportunity for us to be a force in accelerating adoption, increasing awareness, and making it more accessible.

A fortunate outcome of the pandemic for us is this time last year, we were all set to do an on-site conference at The Forum at Columbia University. We decided around this time – it was mid-March last year – we're going to pivot. We're not going to cancel. We're going to keep the event. We were a believer in, I guess, the new domain we're moving into of digitization and virtual events. So we dived headfirst into the virtual events. We spent a lot of blood, sweat, and tears to create a very curated experience whereby our attendees could interact with speakers on a Slack channel – so asynchronous communication, which is actually much stronger in a virtual space than in a conference, where you might have to wait in a line of 10 people to get to the speaker. So we were able to provide our attendees more access to speakers, and then we were able to record all the talks.

Because a virtual event is a great equalizer, we were also able to reduce the price. We didn't have to pay for food or venue, so lower ticket price. And then the person who's joining our conference from Brazil has the same level of access as someone who's joining our conference from New York, because they're both watching it from their living rooms.



So that was a huge positive for us. We grew to 600 people who joined. And then some of our attendees said, this conference is great, but what about after the conference? I'd like to continue the conversation. So over the summer, we opened up our Slack channel to the public, and we were pleasantly surprised over the summer, as every week, there were new people coming on there from all over the world, to the extent that today, we have 1,300 people on our Slack channel. We have a member directory. We have a job board. We have a Q&A platform. We have networking. And people are coming on there, they're asking questions, they're looking for introductions, they're sharing the projects they're working on. We now run events in that community every week. It's exciting that the conversation can happen before, during, and after the conference in this new online community that the pandemic has accelerated. It's a signal of the level of interest in the space that we're able to tap into.

KENNEALLY: Thomas Deely, co-founder of the Knowledge Graph Conference, thank you for joining me today on CCC's Velocity of Content.

DEELY: Thank you, Chris, for having us, and hopefully you can join us in May at the conference.

KENNEALLY: I will look forward to that. For more information, everyone, about the Knowledge Graph Conference coming May 3-6, visit [knowledgegraph.tech](https://knowledgegraph.tech). To register for the Wednesday CCC town hall, "Solve for Success," go to [copyright.com](https://copyright.com).

Our producer is Jeremy Brieske of Burst Marketing. You can follow Velocity of Content on Twitter and Facebook. I'm Christopher Kenneally. Goodbye for now.

END OF FILE