Daniel Boley Oral History

Conducted by Jeffrey R. Yost

Minneapolis, MN (via Zoom)

October 6, 2022
Abstract: This interview was conducted by CBI for CS&E, a multi-year project extending from the 50th Anniversary of the University of Minnesota Computer Science Department (now Computer Science and Engineering, CS&E). The oral history begins with Boley’s early interests, undergraduate work at Cornell, and completing a doctorate at Stanford University. It explores the Computer Science Department environment in the 1980s, its administration, Boley’s teaching, and research in various areas of numerical analysis, data science, and machine learning. This includes his work, often allowing graduate students to follow their interests, in applications such as health/medicine, navigation, etc. He discusses this work with Vipin Kumar, collaborations across departments in the College of Science and Engineering, and with other colleges such as the College of Liberal Arts, and the discussions and debates, and launch of the immediately popular and fast-growing Data Science Program.

Keywords: University of Minnesota, Computer Science, Computer Science Department, Data Science, Numerical Control, Machine Learning, SIAM, Mathematics, Data Science Program, Stanford University, Cornell University.
Yost: My name is Jeffrey Yost. I’m here today in Minneapolis on Zoom with professor in the Department of Computer Science and Engineering, Daniel Boley. This is an oral history for the Computer Science and Engineering Department and for the Charles Babbage Institute for Computing, Information and Culture. It is October 6, 2022. Is it alright if I call you Dan?

Boley: That’s fine.

Yost: Thank you. Dan, can you tell me when and where you were born?

Boley: I was born in 1953, in New York City.

Yost: And did you grow up there as well?

Boley: I grew up there more or less until 10th or 11th grade, and the last two years of high school were in Ithaca, New York.

Yost: And as a student in your pre-college years, what were your interests, both in terms of school and outside of school?

Boley: I was always interested in science and math, and I’m trying to think, I have this vague memory of attending a science camp of some sort for junior high students. And my parents were both engineers, academic, in engineering department. So, it went with the family, though my sister was—I had an older sister who definitely was not science-orientated.
Yost: And were one or both of your parents on the faculty at Cornell? Is that why you moved to Ithaca?

Boley: Yes. Originally my father was on the faculty at Columbia, then he moved to Cornell. My mother had gone to graduate school but didn't have a—I think she finished with a master's degree and never had an academic appointment at least as far as I know. Women in those days were, you know, housewife was more typical, in those days.

Yost: So, you attended Cornell for your undergrad. Is that correct?

Boley: Ah, yes.

Yost: And what year did you start at Cornell?

Boley: I have to calculate back. I think it was in '71. The previous year, my senior year in high school, I was actually spending the morning taking classes at Cornell. Then at lunchtime, I'd walk by the house, grab a bite, and then walk down to the high school. So, I entered in '71 with almost, it was like I think 22 semester credits and so that was enough to shorten my stay at Cornell as an undergraduate. It took three years, so I graduated in 1974.

Yost: And I understand you majored in mathematics. Did you know you wanted to major in math from the start or were there other potential majors you considered as well?

Boley: I don't recall looking around a lot. I was doing a lot of computing with the computers available in those days so. I think I always had a bent between mathematics and computers. But they did not have a computer science major in those days so as far as declaring a formal major, it was in math. I was also interested in, fascinated by the constructions like in pure mathematics. I remember I was fascinated by
the construction starting with something as simple as Peano’s postulates for the natural numbers and you can define objects that end up finding the integers, the rationales and the reals that fall out, and I was kind of fascinated by those kinds of constructions—

Yost: Right.

Boley: The computing facilities were all done with punch cards in those days.

Yost: So, was it handing your punch cards over to a computer operator or was there a, in the 70s, was there a timesharing system that you used at Cornell?

Boley: Most of what I did was to take cards that I could hand it to an operator who would feed it into a card reader. They had, later on, they had some teletype, typewriter type terminals for direct access but those were very special, and if I recall, they were also more expensive to use. Somehow, they limited access to those. So, most students did their computing by doing them on punch cards. They had these rooms full of key punch machines where people would type in their programs.

Yost: And did you have any computer programming courses at Cornell, even though there wasn’t a major?

Boley: They existed—they had computer programming courses. I don’t recall exactly the languages they were teaching in those days. I think I actually learned the programming that I was using—as I recall, I did most of that on my own. So, I took advanced computing classes. I took an algorithms class and a theory class, so these were all advanced classes, and they were trying to discourage me that they, you couldn’t take these advanced classes without taking beginning classes, but I took them anyway and I got As. I seem to remember they used some version of FORTRAN. They had a PL1 and a 360 assembler. I remember spending a fair amount of time using the 360 assembler.
Yost: And were there any impactful mentors to you at Cornell that influenced you in any significant ways?

Boley: Let’s see... there were interactions with various professors for different things. I don’t remember, I don’t recall doing a particular project, anybody in particular. It was mostly advice. The names that stick into my mind—or maybe, I remember taking classes from—I took one class from Hopcroft and one from Hartmanis. And I remember interacting with a professor in the math department, Rinehart, who unfortunately died in an auto accident while I was there.

Yost: That’s too bad.

Boley: So, that was, yes, that was kind of upsetting. Apparently, he was making a left turn from the main highway out of Ithaca, and somebody passed him on the left and hit him. Other than that, I don’t recall, I don’t recall the specifics of my interaction with these faculty members.

Yost: Can you tell me about your decision process for grad school in selecting Stanford?

Boley: , I don’t really remember, honestly, I don’t remember where I applied other than Stanford. So, I—yes, that’s a good question. I seem to remember—I’m actually drawing a blank. I have no—yeah, that’s something that I’m not sure how I ended up at Stanford.

Yost: Ok. So, you completed a masters in ’76, I see, and a doctorate in 1981, and your dissertation was entitled, “Computing: the controllability, observability of the decomposition of the linear time and variant dynamic system and numerical approach.” Can you briefly describe this dissertation research to me?
Boley: So the linear dynamic, dynamic linear system is basically an ODE system of ordinary differential equations with a linear, that’s linear, constantly coefficient and has a linear forcing term so it’s of the form \( \dot{X} = AX + BU \), where \( X \) is the state, the \( A \) is a matrix, so it’s a linear, constant coefficient system, and the \( B \) is a constant matrix saying how the input, the forcing function \( U \) affects the system. And then there’s another set of vectors, \( Y \) that stands for the output. You may not be able to directly observe the state’s \( X \), but you can observe a linear function of the state’s \( X \) which is denoted \( Y \). And the question is if you manipulate the inputs \( U \), can you drive the system to any particular state. And in particular, for example, can you drive it to zero? And the other half of the question is can you observe, if you’ll just read the \( Y \)s basically, can you back track what the \( X \)s are. That’s the observability condition. The first one is the controllability part. And the main part of the thesis was to relate that problem which is something in control theory to a bunch of standard linear algebra methods that were being developed during the previous two decades up to that point. So, there’s a method called Lanczos method in which you generate a crowd of subspace, you have a starting vector, \( X \), and you multiply by powers of \( A \), and that generates a subspace and you want to compute an orthogonal basis for this subspace on the fly, and that gives rise to this special Lanczos method. And it turns out that this is very closely related to the controllability problem. In fact, the dimension of that subspace generated by the powers of \( A \) tells you whether the system is controllable or not. And so, I was looking at numerical algorithms for doing that. The Lanczos method by itself without any modification is not numerically stable but one can stabilize it by doing some extra work. And so, I was looking at numerical methods and linear transformations to the system that would expose the—if it was not controllable, then it would expose the controllable subspace. So, anyway, that’s what the essence of that thesis is.

Yost: Right. Who was your primary advisor at Stanford, and can you describe this professor as a mentor?

Boley: So, the advisor was Gene H. Golub. He’s famous for his work in numerical linear algebra algorithms, like the workhorse algorithms for computing a singular value decomposition. He was very gregarious guy. He really gave me a lot of help when I was feeling lost in my thesis work, and I probably gave him more headaches than many other students. The environment there was very special because
in those days, the computer science department was split up into lots of little locations with the numerical analysis group in the house—let’s see, the main department office was in one building and the offices of some of the people who did more algorithms and complexity theory. The numerical analysis people were in a house a couple blocks away which used to be an old residence for the university president. And the artificial intelligence people were five miles off campus up in the hills. And so, there was, it was a small, close-knit community of people, and there were maybe six to twelve students and two to three faculty. Most of the time, there were two regular faculty and—or three , and then they had visitors constantly coming through. Gene always used to invite us over to his house, which was right on the edge of campus, for outdoor barbeques and things. So, it was all a very nurturing atmosphere.

Yost: So, you were in the numerical analysis group?

Boley: Yes.

Yost: And did you have much interaction with those in other groups, either faculty or students?

Boley: I had some interaction because, of course, I had to take classes from different faculty outside of the numerical analysis group. And one thing that helped was that the first quarter of Stanford, it was on a quarter system, and I think they still are. So, in the fall quarter, first quarter, all the PhD students who were entering in that class, which was only about 20 or so, had to take a research class. It was more a problem-solving class in which we’d meet and there would be a fair amount of class discussion. And so, as a cohort, we got to know each other, the students, we all got to know each other fairly well. After that class, we would run into them, not very often but it was a way of making sure, it had the effect of making sure we recognized some people on campus and were friendly and what not, and so it was a nice cohort. I remember some interaction, I remember I had some interaction with Donald Knuth, mainly because I took those algorithms class using his volume one, I believe. He was also very approachable to the students.
Yost: Yeah, there were some tremendous stars on the faculty.

Boley: Yes.

Yost: I’ve interviewed Martin Hellman and Ed Feigenbaum.

Boley: Yes. It was kind of interesting. So, there were some faculty that didn’t really interact with, like Feigenbaum and McCarthy, who was active in those days, McCarthy spent most of his time up in the AI lab that was five miles in the hills. It was kind of interesting because in Serra House, the numerical analysis group had the ground floor, and there were a couple of rooms upstairs, which was more of an attic, and Bob Engelmore had his office up there and he was one of the people in the heuristic programming group, as they called it, that worked with Ed Feigenbaum. So occasionally we’d run into Bob Engelmore on the stairs.

Yost: And early in this project I’m doing for both us and the Computer Science and Engineering Department, I interviewed your wife, Maria Gini.

Boley: Right.

Yost: And I understand you met her at Stanford, correct?

Boley: Yes, that’s correct. I used to go up to the AI lab as an interloper because I had no technical reason to go there, but they had more fun computers to play with. For instance, in those days, there was, you didn’t have to worry so much about security because as long as you had an account, you could login on a local machine without a password and it was also a good excuse to get in a nice bike ride. So,
I met Maria there because my father’s family is originally from Italy and I was always looking for people from Italy if they were there, and there she was.

Yost: And you spoke, and speak, Italian, correct?

Boley: Yes, it was probably a bit rusty because I hadn’t spoken it in a long time. I spent 5th grade in Italy, so whatever Italian I knew, I learned as a smaller kid.

Yost: Best time to learn a language.

Boley: Right. Yup.

Yost: It’s a whole lot easier when you are young. So, upon completing your degree, can you tell me about your thoughts with career plans and continuing into the academic realm—

Boley: Well-

Yost: —did you also consider industry or not?

Boley: Well, I remember interviewing at various companies since I’d worked on this control problem. In those days, control was a big thing. The two big users were the military, the defense department and manufacturing, and there’s also the oil industry. I remember interviewing at some of these industry labs, but it always seemed like they were too constrained for my taste. I interviewed at three places, academic places, and one was Rice, one was Lehigh University in Pennsylvania, and one was the U of M. And I decided the U of M had larger, more well-rounded, I mean, faculty in more areas which I thought would be a greater opportunity for collaboration and just having a bigger university seemed like a better
fit. Also, having grown up in New York, I was leaning toward a bigger metropolitan area. Rice actually never got around to making an offer, so that actually never panned out. So, then I came here and was in Lind Hall and I’ve been at the U ever since.

Yost: And when you came here, Minnesota was one of the central locations for the industry with Control Data and Cray and Sperry Univac, Honeywell and others.

Boley: Right.

Yost: Was that influential at all to you?

Boley: I think it played some role, I wasn’t actually doing a lot of work on those platforms, so it didn’t really affect me directly. It was more what computing facilities were available to me as a faculty member. And they had a, I can’t remember if it was a CDC 6600 or 7600 on campus in the computer center, in a building called Experimental Engineering, which used to sit on the site where Keller Hall is now. And the department had just gotten a [DEC] VAX computer the year before I came. By the time I graduated from Stanford—so at Cornell I was using punch cards. When I went to Stanford, they were using punch cards but by the time I finished, Unix time-sharing, had really taken over in the computer science department. And so, I was used to using a Unix-type platform. And so, when they got the VAX, so I was very comfortable with the VAX at the U.

Yost: And were there faculty members that you worked especially closely with in the early years at the U?

Boley: I’m trying to recall, I don’t—in the early years I think I was working a lot with myself and Ben Rosen. I don’t think I actually wrote any papers with Ben Rosen, but he was the other numerical analyst in the department, so he was acting as my mentor. But the other numerical analyst were—occasionally
I would go to for advice, it was Mitch Luskin in the math department and Don Truhlar who were always giving me advice. I’ll have to go back and look at my list but that’s what I—sorry. I do remember that the department involved me in several big group proposals to NSF, which, as far as I recall, I don’t recall any of them being funded, maybe one was funded. But anyway, I’d have to go back and look to see which ones might have been funded. I knew there were several that were not successful.

Yost: In those early years, were you putting in proposals as well?

Boley: Yes, in those years, I was also doing a lot of individual proposals. I seem to recall that, I don’t know, maybe it was my imagination, but I think it was easier to get funding in those days than it is now. The success rate, it wasn’t as cutthroat as it is now. I seem to recall most of the proposals I submitted were actually funded. And think NSF, the impression I also have is NSF, basically their bread and butter were the individual proposals that people submit. And now there’s a big push on these large group projects and expedition programs and whatnot, and the individual proposals are getting buried to some extent.

Yost: Right. Do you recall any of the undergraduate courses that you were teaching in the early years?

Boley: So, what did I teach? Well, I’ve been teaching numerical analysis course ever since my first semester here. I remember the first semester I was teaching a course and at the end of the course, the students applauded me, they gave me applause, and it’s never happened since, so it’s been downhill, my performance has been going downhill ever since. Other than that, I know I’ve taught, I taught the course on assembly language that existed then. I think that was the 68000-assembly language, which is particularly clean and simple, which they don’t teach anymore. And I taught, a couple times, I taught the beginning programming course, which was the Scheme language. That’s a dialect of Lisp. So, I taught that a few times. And I’m trying to think if I—so that was, yes, well, that came later. More recently. I’ve taught them, once or twice, I’ve taught the undergraduate linear algebra class but that did not exist under the quarter system.
Yost: What was your early teaching philosophy with undergrads, and can you discuss how your teaching philosophy has evolved over the years?

Boley: Yes, I don’t recall specifically my teaching philosophy. At the beginning, I think the general rules were to give the students some respect, be open for questions. As I recall, I was mostly, most of the courses I had were lecture format. For the programming classes, they had these discussion sections, so a fair amount was helping with the TAs. I tried to explain things in an intuitive way. Over the years, it’s changed in that students expect more interaction. So, over the years I’ve instituted a few changes, so now I give periodic, weekly in-class exercises which are not graded. They just have to—they just get credit for doing them, for attempting them and, just so they get a chance to work with the material that was lectured on because if they don’t actually use that material, it doesn’t register as well. And I try to get the students to interact some more, that’s probably the biggest change over the years. So, at the beginning, it seemed to—the students were more amenable to—they were more used to the lecture format.

Yost: Right. So, you, of course, were active in SIAM and one of your papers in the early 80s was “Computing and the Kalman decomposition of numerical method.” Can you—this I assume this extended from your dissertation. Can you describe that and the reception to this early research you were publishing?

Boley: I think there was—I never know how, a lot of times how these papers are treated but I remember there was a series of papers on that topic that grew out of the thesis work. The one that caused the biggest stink was the short paper I published early on even before I graduated from Stanford on the Kalman composition because I found a bug in the Kalman. The Kalman paper that actually introduced it. And so, I published a short paper on the fact there was this big mistake. And a later paper published a numerical method that was basically a corrected version of it, showed that with a certain small modification, you can correct the thing. I don’t remember the exact details at this point, I would have to go back and look at the paper. What I do remember actually meeting Kalman, who was not a
nice guy, and he really took it badly. In retrospect, I probably should have waited to publish until I had the whole correction thing settled instead of doing it piecemeal.

Yost: So, in your early years on the faculty you were submitting individual projects to NSF. Were there other agencies and did you seek funds from industry as well?

Boley: I don’t recall ever talking to the industry. I remember making some attempts at some of the military funding agencies like AFOSR (Air Force Office of Scientific Research), if I recall. I think the part that I was never, never really succeeded was in connecting the work that I was doing, which was interesting in an abstract way, connecting them to particularly anything connected to the particular mission of the agency because all these other agencies had these specific missions. And so, I—and also, I think I didn’t have a lot of patience to go through the kind of selling job that the agency demanded. So anyway, none of those panned out, let’s put it that way. NSF was more amenable because they were more interested in the intrinsic merit of the science without a particular end mission in mind. And they also took a more hands-off approach in regarding periodic reports and things and deliverables.

Yost: Do you recall any of the strategic debates in the department and its direction in say the 1980s or early 90s?

Boley: So, during the time that I was an assistant professor, I don’t recall being involved in much in the discussions or at least contributing to the strategic directions. I never really thought about the strangeness of the fact that this department head was only an associate professor. I think the year, either the year that I was up for tenure or just before, we had a big falling out, and I’m sure you’ve heard this from several people. The net affect was that the department head, Kurt Maly, ended up leaving entirely and they appointed— we went through a series of department heads, and at this point, I don’t even remember the order in which that happened. I think at one point we had David Fox come in, and then he left, and then there was a period—I mean, the department fell apart administratively. You probably have the history down of who was running the department when because we had the three associate deans for a while. Which we call it (triumvirate except that was a misnomer because they were
not all three men. And then we had Ahmed Sameh come in as a department head for a year, then he left and then he came back. So, we had a big turnover, and there was a lot of nasty things going on, and the one thing I do remember is that, even though there was a lot of nasty messages—and it was actually a bit of a scary time—as far as the promotion of the untenured faculty, it seems to have gone through fairly smoothly. I mean it didn’t really, amazingly enough, it didn’t really affect the promotion process as far as I know. Since I was involved for two years running because if I remember right, my promotion went in before all this happened. But Maria’s promotion happened during this whole thing. So, I was privy to this thing for more than the usual one year.

Yost: Right, so some administrative challenges for a bit. At the end of September, I interviewed your colleague Vipin Kumar, who you’ve published with and in looking over your research and my discussion with him, one thing that really struck me was the partnering in various areas of science and both of your research evolving to contribute to and collaborate with various scientific projects and also publish with scientific publications, journals outside of just computer science publications. Can you tell me about—you mentioned you had an early interest in science, but can you give me a sense about how that evolved from when you first started really collaborating in verticals, in the sciences with the biological sciences and medicine?

Boley: I know that I got a lot of inspiration by working with other faculty on individual problems. I mean, all those collaborations lasted for maybe two, three years and when it finished that, it ended. So, with Vipin, it was a matter of looking at these algorithms for clustering, as a way of dissecting an unstructured data set into its component parts. The biological interest… Vipin made the connection, but that was working with a professor in the Biotechnology Institute. It turned out that they had a problem that looked like it was basically linear algebra in nature. It was actually more associated with linear programming. So, if you have this issue of metabolic networks, we have a lot of chemicals, metabolites that are generated by all this network of reactions and their constraints because the metabolites in the system can’t vary—the concentration has to stay in a steady state, and then a lot of the reactions are one way so there’s a mix of inequality constraints and equality constraints. Just like you’d have in a big linear program, and the one thing missing is an objective function. But they were trying to look at ways of manipulating this network to come up with certain tasks. The methods were linear algebra in nature and so they connected through Vipin, they connected to me, and I ended up getting a student involved and it was about the time they were creating the bioinformatics and computational biology program.
and so I joined that program as a charter member. And so that got me interested in various other issues with the, within bioinformatics. And so, since then I’ve been on and off doing some work in bioinformatics, it really depends on students, if they’re interested. So, I’ve had some students that are interested in bioinformatics and some that are not. That’s probably a fair description.

Yost: And what’s now the College of Science and Engineering was, of course called the Institute of Technology in those years. Did the Institute of Technology promote interdisciplinary collaboration within and beyond the college in the 80s and 90s?

Boley: I think there was a general notion of welcoming such collaborations, and as far as budgeting goes, I think most of the people I was collaborating with were all in the college, so it—I didn’t have any issues with collaborating with any people outside the college. And within the college, it seemed like, certainly I don’t remember any particular roadblocks, so I think it was—they didn’t have a, I don’t recall using any kind of systematic platform or organization or anything in the college that was promoting these things, but they certainly were cooperative when these opportunities came up.

Yost: I’m going to name several categories of your research areas and if you could briefly describe each of them, summarize key work that you did, that would be great. Numerical algorithms for very large-scale sparse dynamical systems?

Boley: Well, that basically was my thesis work and stuff that came afterwards. So, a lot of it was related to problems for matrices. So, that was a continuation of the thesis work, so it took a few papers to fully develop that thread. It turned out—a lot of the techniques that I encountered there turned out to be useful in some of the other research threads that I followed.

Yost: Can you tell me about Web ACE? ACE being automated categorization and exploration?

Boley: Ok, so that was a clustering—my part in that was the clustering algorithm. So that was my first dabbling in anything related to what would now be called machine learning. It turned out that there is a
spectral method for clustering. I think the idea there was to automatically organize—if you visit a pile of web pages to try to organize them by topic and the basic idea that my contribution was to look at an efficient way to cluster the pages based on the words that they used. This translated to a spectral clustering which you basically use adjacent—well, you create a—spectral clustering is usually used on graphs, but this was a way of doing it on tabular data. And it turned out that the method that I had was easy to incorporate and it worked pretty well. It was an opportunity to see how linear algebra could then turn into a text analysis problem.

Yost: And this was the start of a considerable amount of research on machine learning and including and especially scalable methods and machine learning. Can you talk about how your research on machine learning and specifically scalable methods evolved?

Boley: It evolved into different applications. They went in different directions. The clustering method that I developed during that time with the WebACE—well, one was this method of Principal Direction Divisive Partitioning which turned out to be a general-purpose method. We used that in several different applications. So, when I had a later student doing some work in bioinformatics by analyzing the flu virus, we used this kind of clustering method to differentiate among the different flu viruses. And it showed that linear algebra could play a role in analyzing these methods. So later on, later machine learning is more toward optimization—the mathematical method underneath is more toward the area of optimization, and so some of the methods carried over. It turned out—so one of the methods that was popular in optimization was this method called ADMM; alternating direction method of multipliers, and it’s a type of primal/dual method, and there were a bunch of analysis for this method, but it seemed like the performance and practice was usually much better than the theoretical analysis, which happens actually more often than you would think in numerical methods. And so, I discovered you can use the matrix value analysis of the operators to say something about the way ADMM would behave on a model problem, that in a much more detail than just giving in general bounds.

Yost: And what about vehicle navigation and localization using multiple navigational aids?
Boley: Let’s see, so a lot of these methods, so my general practice has been to give the students basically free reign to work on what they want as long as it involves something that I have some expertise in like linear algebra. And in this case, it was more of a single processing thing. So, one of the other threads in numerical linear algebra is recursive algorithms where you compute a decomposition and then you dynamically update it without having to recompute it from scratch. And so, we were using an approximate—it’s not really a singular value decomposition but we called it a rank revealing decomposition where you expose the subspace corresponding to the large singular values and the subspace corresponding to the small ones, which the latter one, that’s called a noise subspace. And you want to try and separate the noise from the signal. And then as you get more signals, you have to update the decomposition, and to do this dynamically. And that, from the vehicle navigation— the part that was interesting to me was that you take the signals that the vehicle is using to figure out where it is and you pass it through this system, recursive updating system to filter out the noise to try to get a better handle on—to get an accurate localization. And that was the idea there. So that was an example where the student really drove the research more than me. And a lot of the work that I’ve done has actually been that mechanism. So, like the whole work on the flu virus as another example, that’s where the student wanted to work on that and involved some linear algebra, so interesting.

Yost: It must keep things interesting to have students bring in all these different types of applications.

Boley: Yes.

Yost: So, playing into this a bit, can you describe your style and approach to mentoring graduate students?

Boley: Let’s see—I’ve never formally thought about it, but I encourage students to work on, I mean, I’ve been letting students if they know what they want to work on or have some idea— I encourage them to work on. So, I get—that’s one reason why the topics that I’ve worked on, even recently, there are several threads that seem totally unrelated and it’s mainly because I’ve encouraged the students to follow their interests. And it’s also the case for the master’s students that I’ve mentored through the
Data Science Program. I do expect students to—I give them a lot of, try to give them a lot of autonomy in putting together the papers. I meet with them regularly and so far, it’s been very, I mean, I think it’s been successful because the students have all made it through pretty smoothly, probably more so than my own case.

Yost: You were a founding member of the Data Science Program. Can you tell me about the context for the formation of that program in 2015?

Boley: So, the person who really pushed that, started the ball rolling was Vipin Kumar. He—I think it may have started when he attended a White House event—I’m not sure workshop or get-together, I’m not sure what exactly it was, but an event at the White House to promote data science as the new thing, to maintain the competitiveness of the United States. And so, he decided that we really needed to start some program in data science. And in those days, it seemed like the thing to do would be to—a master’s program would be good for training expertise so we would have a population of people who would be experts in—be able to gain expertise that the industry could look to. So, we had a working group of faculty from various departments. The departments represented were Computer Science, Electrical and Computer Engineering, Statistics and there were some people from Biostatistics. And there were lots of faculty discussions over a whole year, some which got quite heated. There were some industry representatives that were in the mix too, and we finally settled on the structure that we have today with a few minor changes. And I was asked, it seemed like the natural to become the director of graduate studies at the beginning and it’s been pretty successful since. It seemed like something—the right thing at the right time because even without any publicity, the first year we were accepting applications for this program, there was very little publicity because it was very late to be approved by the board of regents in the summertime. So, the application period started in the fall, and with almost no publicity, we already had 250 applications.

Yost: Wow, great interest from the start.
Boley: And since then, it’s grown to about four or five hundred applications a year. Right now, the situation is somewhat different because the buzzword that gets everybody excited now is not—data science is being supplanted by ‘machine learning’. Of course, the Data Science Program covers that. It is similar. The international—a lot of our applications are from abroad, like India and China, especially in China, the situation has changed quite a bit and it’s not clear if we’ll continue to get applications from China right now, but we’re still getting quite a few applications.

Yost: You mentioned earlier in starting those programs, there were some debates and heated discussions. Do you feel comfortable telling me some of the issues that came up and were debated?

Boley: I think the debates were, how many, like the balance between statistics and computing, whether we needed something more applied, like the industry people wanted or something more academic. I think it ended up ignoring the industry wishes to some extent and we divided the topics that people would have to cover, the students would have to cover into three areas: statistics, a topic we called algorithmics, which was basically the fundamentals of data mining, machine learning, and infrastructure, which is basically the computer underpinnings of anything that would be implemented at large scale, like databases and parallel computing, that sort of thing. And one of the big debates was, “How many credits of each?” And at the beginning we were going to say the student could choose to take more credits in one of those three areas and fewer credits in another one of those areas and we ended up making it the same across the board which still leaves a fair amount for electives. But I think little things like that cause a lot of discussion; how to balance—because those statistics people wanted to make sure they got enough enrollment. The other part that needed negotiation was to figure out how to—the MOU that was needed to figure out how it would be governed and how the money would flow. That’s usually a big deal but somehow that—one thing that saved a lot of grief is that the faculty committee was basically talking about the academic requirements, never got into this administration-money business. Once it was agreed that since most of the credits would be taken in either computer science or ECE, that the administration of the program, even though it was supposed to be cross-college, interdisciplinary, the administration would be housed in the college of science and engineering. Once that was decided, the dean’s office started the negotiations on the MOU and the money business and
that saved me a lot of grief because Chris Cramer was the Associate Dean for the college at the time, did a lot of the grief work with the CLA to figure all of that out. It saved us—it saved us a lot of grief.

Yost: Was this program fairly early, in terms of research universities, to have a large Data Science Program and do you have a sense of whether the program here had a different character or definition to it compared to some of the other existing an emerging data science programs around the country?

Boley: Yes, I don’t remember, I know there were other data science programs being already developed around the country, so it wasn’t an original idea to create one here. I think it was already the case that Georgia Tech had their on-line program that was drawing students in the thousands, but it was purely on-line program. And I seem to remember there were other programs like at NYU. There was a big thing going on at Berkeley, and even here at the U, the business school was creating their business analytics program, which called data analytics. Vipin managed to convince them to change the name to business analytics before ours was created so there wouldn’t be any confusion. Now we’ve lost that because there are actually several other programs in CSE with the name Data Science and ‘something’ and it’s—I’m sure it creates confusion in people when they—

Yost: Right.

Boley: —distinguish our program from those that are different programs. But in those days, there were a lot of schools I think were in the process—there were few programs around the country, and there were a lot of schools in the processes of developing them. I think our program was or is distinctive in being more rigorous than a lot of the other programs. So, we actually have people taking a fair number of graduate courses in computer science and in statistics. These are all technical courses of various sorts. So, in that sense, there are a lot of programs that are designed for people who come from a non-math, non-computing background, so they have some computing experience but not a lot. And so, they learn a lot of the tools, whereas we tend to teach people enough so that they could design their own tools or adopt existing tools or modify them internally to make or to handle their application. At least that’s the idea. Hopefully it works...but, anyway, yeah.
Yost: Among your many honors, I see that you were named distinguished member of the ACM in 2020, so, congratulations for that!

Boley: Thank you.

Yost: Can you tell me what that honor meant and obviously you were heavily involved with SIAM, ACM. Can you just broadly talk about professional organizations and what they’ve meant to your career?

Boley: There was a time when I was doing a fair number of activities with SIAM. Early in my career, I was publishing a lot at in the SIAM venues. The conferences, SIAM conferences, generally did not have proceedings. They were mostly occasions to get together and present things. So, there wasn’t anything to referee. I helped in organizing some of these conferences and there were—the control work I did was published in a lot by IEEE conferences that did require refereeing and I helped in refereeing. And it was a good way to meet people and what not. I was the poster chair of one SIAM conference, but other than that, I haven’t taken a big administrative role on any of those.

Yost: What about any ACM-SIGs?

Boley: The ACM; the computational mathematics was more in the—the SIAM was basically more the society that looked at that. So, I’m pretty sure I was a member of some of the SIAM SIGs. But ACM, I have to go back and look. We didn’t play a big role. Let’s put it that way.

Yost: Before we conclude, are there any topics I haven’t brought up, things you’d like to discuss?
Boley: The department, my career here, the department has been pretty supportive over my career, they've been fairly cooperative in my teaching assignments and office space and just helping with research and funding collaborations. I mean, the school in general, finding collaborations over the years. I'm thinking, all and all, it's been a pretty supportive environment. I do have this feeling that the administrative burden has changed over the years, and some things have gotten a little more rigid administratively which is a bit frustrating, and also the character of the students has changed and I'm not sure if it's because they are used to all these electronic gizmos that didn't exist way back when. But that has changed a bit, the character of a teaching environment and I'm not sure if it's for better or worse. I think it's fair to say people's attention span has gotten shorter.

Yost: Yes, MIT Emeritus Sherry Turkle wrote an excellent book on that. Well, it's been a great pleasure and I thank you so much, Dan, for your time and thoughts and insights in doing this interview. And in the not too distant future we will have the transcription done, some are ahead of it in the queue. We do that using humans rather than using voice recognition systems. We've experimented but we find it works better to use a transcriptionist—

Boley: Yes. Right.

Yost: There'll be a month-long window for you to make any edits or changes or deletions you'd care to make to the transcript. Again, really appreciate you doing this and have a great rest of your day.

Boley: Ok, well, thank you. I enjoyed this.