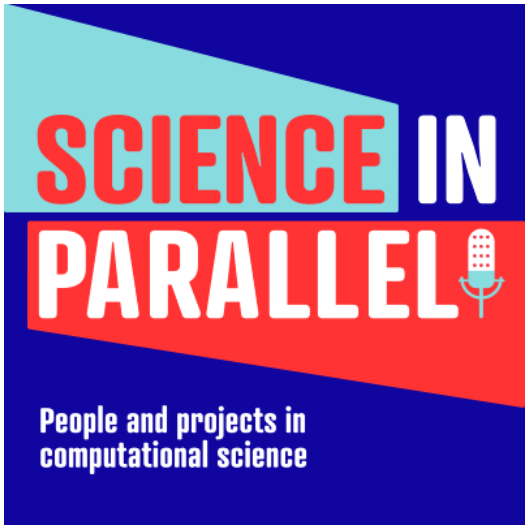

Season Two, Episode 6: Pushing Limits in Computing and Biology

SPEAKERS
Sarah Webb, Amanda Randles, Anda Trifan

Sarah Webb 00:00
Welcome back to Science in Parallel, a podcast about people and projects in computational science. I'm your host, Sarah Webb, and in this episode you'll hear from Anda Trifan and Amanda Randles, computational scientists from different fields, who both work on big questions in biology and medicine using leadership-class supercomputers. Their research projects have attracted some of the highest honors in high performance computing. Between them they've been finalists five times for the annual ACM Gordon Bell Prize or the recent Special Prize for COVID-19 Research.

Sarah Webb 00:40
Their common ground doesn't stop there. Both received the Department of Energy Computational Science Graduate Fellowship, and they worked on urgent pandemic research projects while raising very young children. We discuss the scientific questions that drive them, the adrenaline rush of around-the-clock COVID-19 research, and their words of advice, particularly for other women in science.

Sarah Webb 01:06
Anda Trifan is a graduate student at the University of Illinois Urbana-Champaign, who uses molecular dynamics simulations to study the interactions between proteins and membranes in cancer and other diseases. Through her Argonne National Laboratory practicum, she joined the teams that model the SARS-CoV-2 spike protein and virus aerosols. And Anda was first author on a third paper that simulated part of the SARS-CoV-2 replication process. All three projects were finalist for the Gordon Bell Special Prize for COVID-19 Research, and the spike protein paper won the 2020 award.

Sarah Webb 01:48
Amanda Randles is an assistant professor of biomedical engineering at Duke University. Much of her work has focused on using medical data to build a three-dimensional image of an individual's unique geometry of veins and arteries. That allows them to simulate large scale blood flow toward building tools that could help doctors diagnose and treat disease. Some of that research for their blood flow simulator known as hardy was a finalist for the 2015 ACM Gordon Bell Prize. She has received numerous other honors, including the ACM Grace Murray Hopper Award in 2017.

Sarah Webb 02:31
Welcome, Anda. Welcome, Amanda, it's great to have you here.

Anda Trifan 02:35
Hello.

Amanda Randles 02:36
Yeah, thanks for having us.

Sarah Webb 02:37
One thing, looking at your CVs, for example, is that, Amanda, I know that, say, your undergraduate training was physics and computer science and, Anda, yours is in chemistry. But in many ways, it seems like you've ended up working on very similar types of problems. So I want us get from each of you the development of your interests and how you landed in this computational biology space.

Sarah Webb 03:06
You'll hear from Anda first.

Anda Trifan 03:09
During my undergraduate degree, I studied chemistry, specialty in biochemistry. And I really enjoyed that. I love doing experiments in wet lab. But after my bachelor's degree, I worked in industry for a while for Dow Chemical Company. And I didn't feel like that's what I wanted to do for the rest of my life. When I looked at Ph.D. programs, I looked at very interdisciplinary options, because I wanted to test out what else is out there, and what I would enjoy doing. So computational work has always been really fascinating to me because you're able to immediately visualize the things that you're working on. So especially molecular dynamics, you get to do science at a resolution where you see in much detail what's happening in a biophysical system.

Anda Trifan 04:02
So at University of Illinois Champaign-Urbana, of course, they have one of the best programs in molecular dynamics where they develop NAMD and VAMD. And those are used across the country. I ended up working with Professor Emad Tajkhorshid. In his lab, we work on proteins and specifically, protein-membrane interactions and things like that. So I really enjoy being able to do things at that level and see atoms interact with each other and how that affects a certain system.

Sarah Webb 04:39
I think it's your turn, Amanda.

Amanda Randles 04:40
Mine is somewhat pretty similar if when I when I started out I was, my main degree in undergrad was physics. So it was taking the really trying to understand the fundamental side of like, you know, what are the equations and how do we capture it. And then I was also doing computer science, but like similarly, I went to industry before going back to graduate school. I worked at IBM for a while, and when

I was there, I got to work on the Blue Gene system, which is where I first really got exposed to parallel computing. And I got really excited about, you know, not just building the supercomputers that we were doing more at IBM side, but how do we actually use them? And what kind of questions can we solve that you wouldn't otherwise be able to? So that was what prompted me to go back to graduate school.

Amanda Randles 05:20
I ended up getting my Ph.D. in physics. So I've always been kind of on the physics/computer science side, but then taking that to solve biomedical problems. So I really like I've always been passionate about the medical questions and the biomedical application but take a physics-based approach to how do we develop these models. That was kind of what brought me into like the project I'm working on. I started in graduate school of, you know, we're doing physics-based modeling to capture fluid dynamics using the large-scale computing skills, and kind of bring all that together to work with the clinicians and answer clinical questions. And so it's been that that merging of different fields, I think, is really interesting throughout.

Sarah Webb 05:54
Can you each talk to me a little bit about the biological questions that drive you and how computation fits into those questions.

Amanda Randles 06:04
I started my research lab in 2015. I had been at Livermore as a Lawrence fellow, and while I was there, they were all so incredibly supportive and that, like they let me put together a workshop right before I left where we just invited people from all over the country to come talk about research, and they would just come but we brought together researchers from cancer, from cancer biology, we had interventional cardiologists fly out. So we had people who are experts in visualization. Anyway, we had people from all over and bring them together and really go over. . . we talked about Livermore. We had people from the big computing side and discussed some of the capabilities. And we really just had a three-day brainstorming session of what are the key challenges in our areas? Where could we really, you know, make an impact? And out of that, we ended up coming up with a few key questions in cancer research and in cardiovascular work that ended up really being the major projects I started my lab focused on.

Amanda Randles 06:54
From a physicist's standpoint, it's always hard. There are many questions, I think are really interesting and use, you tell it to the cardiologist and they're like: That's great, but that would not change how I would practice. And like is not necessarily a useful question. So it's always been, it's been really helpful to have the clinician there to say like, how do you actually practice? What are you doing? There are a lot of invasive diagnostic tests. Where could we start? The first step was one of the main projects that came out if it was current practice was using coronary angiography instead of CT to determine if someone needs a stent. And they had just the fundamental question of could we use coronary angiography to create 3D models. Because coronary angiography are pictures of someone's heart. Can we use those images to create a 3D flow simulation? So that instead of having to put someone in the cath lab, insert a wire into their heart to measure the pressure in their heart, could we use a simulation to accurately measure the pressure gradient across a narrowing in your vessel and determine should you or should you not place a stent. We just finished a 200-patient clinical trial where we're able to

show that you can actually do that. You know, we just finished the clinical trial now, and it's been seven years since this question first came up.

Amanda Randles 08:01
It helped kind of identify some of the key questions in that area, and then bring to that using these computational models, we can get much more complex measurements and get more complex phenomarkers, like describing the hemodynamics, we're looking at quantities like vorticity, which is describing with the turbulent nature of the flow. You can't measure that easily directly with a guide wire, but you can get those characteristics or shear stress on the wall or other complex metrics using these simulations. So now we're looking into questions like: you know the vorticity, are you more likely to be able to predict adverse events in two years? Can we find more complex markers you can really only get out of the simulation to then guide treatment and inform diagnosis on that end?

Amanda Randles 08:40
So I think those kinds of questions are really exciting, like, what can we pull out when we couple, you know, an experimental setup or an in vivo measurement with these simulations? And how do we add to and augment that data on the other side? We also were interested in when you're when you're reading a fluid structure-interaction model, you're looking at cells. Simulation gives you that unique capability to change one factor and hold everything else completely the same. And you can run the same experiment, you know, it's different than running something in vitro, where the humidity in the air may change, like, it's very hard to exactly replicate exactly what you had done beforehand. But in a simulation, we can change like just the bending stiffness or just like the size of the nucleus of the cell and see like how, you know, pinpoint how changing this one quantity is going to result in, you know, behavior of that cell and really sort out what is the contribution of that one factor.

Amanda Randles 09:27
So that was that I think that's a really interesting and exciting if we can, you know, it's a unique way of looking at data that simulate simulations kind of brain that you can't get get in other ways. And so we were trying to figure out, you know, thinking of how do we really use that and how do we learn more about, you know, what, what is it about certain cells that cause them to stick to different regions in the body? And how can we, you know, how do we actually learn more about cancer cells? Why is it clusters? You know, what do we like, Why do clusters behave differently than single cells and really trying to get into those questions were the things we were initially you know, trying to think of And of course, that also brings into, like, brings into play all of the computational questions of, if you want to look at a large number of red blood cells, like, you need large-scale-- it's very computationally intractable to model a large number of cells on a, you know, a few nodes, you need leadership-class systems. And that brings out all the, you know, all the interesting, high-performance computing questions that kind of come into play. How do we use that data? We've also had a few, like we had just started at that time looking into, like, how do you actually use virtual reality to couple in with your physics based simulations to try to like, have more intuitive interaction when you're doing treatment planning. How do you drive simulations with that? There were, there were a lot of like that, just those kinds of questions we were dealing with ahead of time.

Sarah Webb 10:43

So, Anda, talk about your work kind of early on.

Anda Trifan 10:47
Actually, as soon as I joined the Ph.D. program, I said that the one thing I really don't want to work on is cancer, because everybody worked on cancer. And it's just absolutely so complex that I don't think studying very small systems could, you know, elucidate anything new in any way. So, but of course, that's what I ended up working on. I worked on a very small protein called Ras, and it has three different isoforms. And it specifically worked on K-Ras. This is one of this is a signaling protein very important in our bodies, and mutations in its active site are responsible for about 90% of pancreatic cancers, and many others like lung and other cancers. So one very important aspect of its function is that it functions at the surface of the plasma membrane. And so its interaction with the membrane is best characterized using molecular dynamics simulations, because many times experiments kind of don't give you enough resolution to really see what's happening there.

Anda Trifan 11:59
So when I started working on it, there were no treatments and no really hope in sight. But since then, Amgen actually came out with a drug to target one of the mutants, where one of the amino acids is mutated to a cysteine. And so they developed the drug to bind covalently to that and to help kind of keep this protein in its active or state, or deactivate. And so my simulations were of the wild-type protein, which means the regular, no changes to it. And then I also studied the mutants, which are the ones that cause problems and help this protein stay in an active state, which means it keeps signaling and then cells keep growing. And that's what causes the cancer.

Sarah Webb 12:53
I want to ask you both how COVID shaped your research. And, Anda, I'll start with you, because it sounds like that was very much a turning point for you and a very critical time in your Ph.D. Talk to me about how you got involved with COVID, and what it was like to work on the biggest problem of our time.

Anda Trifan 13:14
I did my first practicum with Arvind Ramanathan at Argonne National Lab in 2020. And I was supposed to finish in April and in March the pandemic hit. And so, of course, as with everybody else, everything just kind of dropped, and everybody started working on coronavirus-related things. And this was a week before my practicum ended. So technically, I could have just left, but I was so curious about how the virus worked and why it's so easily spreadable and why it's so dangerous and how it affects all of our organs. And so none of that was known at the time. We know of SARS-CoV, the original one, and MERS, but this one has such high infectivity rates. And I was really curious about how this is different than why.

Anda Trifan 14:05
And so I asked Arvind to join his coronavirus projects right away, and he was like: Yes, of course, we can use all the hands that we can get. So I started working on smaller systems first, with COVID. There are two main proteases; there's MPro and PLPro. I worked on the smaller one. This one kind of cuts up the viral RNA and then it translates it into the subsequent proteins that need to be used in the viral

replication cycle. But aside from that, we got involved with Rommie Amaro's lab, who built the first virus system, and we also collaborated with Lillian Chong's lab from Pittsburgh and they do this enhanced molecular dynamics method called WESTPA, weighted ensemble simulations. Lillian's lab studied the spike protein and specifically the open-to-closed transition, and that's very important because that is the first point of contact that the virus has with our ACE2 human receptors. And so this open-to-closed transition was very important in trying to figure out, you know, how exactly the the virus binds first to our cells. And so they did that part, and then Rommie's lab did the full simulations of the spike, and then also did the simulations of the spike binding to the ACE2 human receptor.

Anda Trifan 15:34
And what we in Arvind's lab did was take all of that data and use machine learning to extract very relevant and important information from them. Because typically, when you have hundreds of nanoseconds or microseconds of simulation, it's very difficult to kind of sift through it visually, or with any kind of analysis tools, and extract what might be interesting or relevant at a biophysical level. So everybody kind of played a different part in this project. And this led to our Gordon Bell submission. And eventually, we won with that project in 2020. Yeah.

Anda Trifan 16:18
And then last year, the pandemic was kind of tamed by having the vaccine, but people were still getting infected. So it for me, it was very interesting to see what happens post infection. And so I started studying the replication-transcription complex. This is responsible for replicating the viral RNA and making more viruses within ourselves. And we did some work on trying to simulate how RNA is being transcribed within this complex. And also, it has a mechanism to backtrack, which means it finds errors within the RNA and corrects them so the virus can keep replicating. So we simulated the backtracking mechanism. And currently, I'm studying remdesivir, which is one of the only drugs that have been approved so far for post infection. And I want to know how it works at an atomic level because it did work for SARS- CoV, and that's why it was repurposed for SARS-CoV-2. But nobody really knows why it stops this replication mechanism and how it aids in symptom relief sooner than, you know, placebo. So currently, I'm doing simulations of remdesivir within the RTC, and it's such a complicated name, RNA-dependent RNA polymerase, and how it functions. So it's been quite the ride since 2020, and, really, there's so many things to be studied about coronavirus, and so many proteins that are crucial to its function and how they work together. So I've kind of had a taste of working on everything from the outside of the virus with the spike, and then its replicating mechanism as well.

Amanda Randles 18:16
So did you have experience with the machine learning side of that before the coronavirus projects or did that was that brand new?

Anda Trifan 18:23
No, that was all brand new for me. People from NVIDIA and they, they're the experts in actually writing the code that we use, but I was able to apply it and extract the biophysical meaningful things that we were able to do. So in our project, everybody kind of had their own role. The people that did the coding and that side of things were not experts in biophysics, so they were able to write the machine learning part of it and pipe the HPC. They were able to optimize that, but then they weren't really able to

interpret the results. And that's where, you know, my expertise came in. And we looked at the different confirmations of the spike. And we looked at, you know, how, for example, one of the things that we extracted from these simulations was the stalk of the spike bending, which was seen in experimental results, but we would have to have milliseconds or, you know, microseconds of the simulations going in order to observe that also computationally, but with machine learning, we were able to speed that up by like 10 or 11 times, because we were able to pick out some of these rare states and then restart simulations from there instead of, you know, having one very long trajectory, where we would need to bring it to that experimental observed point. So it was extremely interesting and such a honor for me to be able to work with all these incredibly intelligent people in our field. So that was probably the most exciting time in my career so far.

Amanda Randles 20:02
So I've one more follow up just so I think this is really interesting of like my experience with the pandemic was like it brought out, it brought people together to work in a way that like you actually collaborated quickly and effectively, How did you do the handoff between the computer scientists and the biophysicists? How did you guys actually do that and bring together people from NVIDIA with, you know, the lab, academic researchers?

Anda Trifan 20:22
Yeah, so I was in actually like, pretty unique position because I was closely collaborating with both the CS [computer science] people and the biophysics people. So for me, it was really exciting because all of these people were located geographically in very different parts of the world. And so, Zoom enabled us to really work together because any time we had any questions, you would just jump on a call. And it was it was very effective in communicating our issues. So I don't remember much pre-pandemic how it was, but I feel like it wasn't as easy to communicate with people across the country. Yes, you would make appointments, but I feel like the pandemic really, really just enabled scientists to work together faster and more effectively. We learned how to communicate much more effectively and much quicker. But I did feel that immediately after we started talking about the Gordon Bell submission, and how things just really progressed very fast. So I think we decided to do the submission in July or August, and then that much of the work was done before individually. So like the steps were taken in Rommie's lab, they had the simulations and Lillian's lab. And then we had kind of a backbone for what we wanted to do with the machine learning part. But really, this collaboration started very quickly and rapidly before we had to submit in November.

Amanda Randles 22:01
Yeah, I just it's a phenomenal having so many different groups coming together to work on one project from like: You couldn't have started earlier than March. And then you had a finished product by November. It's a very unique setup,

Anda Trifan 22:13
I really, really enjoyed it, I have to say it was like adrenaline rush all the time, because something was due at any point, right? As anybody working in science knows, it's not a given that you get any meaningful results from any kind of experiments. So it was all just really exciting to be able to contribute

to science, and at the same time to kind of merge all of these different methodologies together and come up with a project that was so relevant at the time,

Amanda Randles 22:47
Especially as a graduate student, that's, it's a very exciting opportunity. But I feel like that's very brave, you don't know how that's going to impact your graduation. You have the path to the PhD. And this is a very, like it could pay off, or it could just delay things.

Anda Trifan 23:01
So I honestly, I didn't even think about any of that at the time. Like to me this project was so interesting. And the problem that we were trying to solve was just so fascinating that I didn't even consider whether or not would be part of my thesis, or if I'm even allowed for it to be a part of my thesis. But I was completely willing to work on it, you know, separate from everything that I had to do in grad school, because I was learning new things and whether or not they were directly related to what's going to end up in my thesis, it didn't really matter, because these are skills that I can apply at any time during my career. And the collaboration with so many really smart people, for me was probably the best aspect of it because I learned so much from so many different fields. So even if this meant delaying my Ph.D. by six months, I was willing to take that.

Anda Trifan 23:52
But for me, specifically, it was even more intense because at the time, I had a five-month-old baby. So a lot of a lot of the meetings that I was in, I was either like with her on my lap or she was like screaming in the background. So it added to kind of like the chaos and the adrenaline of getting everything to work. Also, that I had to feed her like throughout the night because I was working at hours that other people weren't.

Sarah Webb 24:23
One of my questions was what working on this problem at this moment where obviously life was jumbled to was like for both of you. And, Anda, I think you've spoken to that. But Amanda, I'd like to sort of get your perspective on that. And I also want to hear about your COVID work too, but I think we're in the life jumbling of COVID. What was that like?

Amanda Randles 24:43
No, I like mine was very similar with a COVID project we jumped into, but I have twins that are two years old now as well. So like I had like six month old twins. We worked on trying to like how to split ventilators and trying to address the lack-of-ventilator issue that was really hitting in like March and April 2020. And like the first meeting we had with Microsoft to work on that, I had like both twins sitting with me. Like it was very much like one would start crying. I'd pick her up, but like, you know, they'd see the one crying. When she'd stop crying, and I'd put her down and the boy would start crying. And it just like rotated the entire meeting. And it was just like, yeah, nice to meet y'all. This, this is how it is.

Amanda Randles 25:19
I did not run all the simulations. It was the students that were running more, but I was like, up on Slack, you know, feeding in the morning, like, okay, I can see how the jobs are doing and like check in. With

me like it was weirdly helped, like not like it was like a little helpful at times of like you were up at all hours to be able to check on things. And it's like, I'm here around the clock, because that's just what it is right now. Yeah, it was, it was definitely more interesting.

Anda Trifan 25:42
I do think that's another unintended consequence of the pandemic, this realization that people have personal lives and our personal lives have become much more intertwined with our professional lives. Because when you're at home, you're not necessarily able to separate those very well-- especially, you know, with your kids at home. So I feel like as a graduate student, that's not really known well, among my peers, because I'm literally the only one in our lab who has kids. But I feel like as you advance in your career, PIs and postdocs, researchers have kids, because they're at that point in their lives. And so it became a lot more acceptable to have your kid in a meeting or, you know, have to stop for five minutes to whatever, give them a snack. I don't know. So,

Amanda Randles 26:35
Yeah, it was really interesting. Even like the Microsoft team, we worked with one of the leads there actually had twins. So like this never would have come up by normal working environment, he ended up like giving me advice.

Anda Trifan 26:45
Do you also feel like this social aspect of our lives have been a lot more pronounced since the pandemic?

Amanda Randles 26:54
Yeah, it's I think it's I don't know, you people talk about it more like, we had just little things like one of the articles about us on the ventilator project, like there's a picture of my whole family in the article, which is not normal. Everyone with young kids kind of gets it: You either don't have a daycare slot; your daycare is closed half the time. Like it didn't, you know, it kind of continued. And I think it's really gotten people to be much more understanding and kind of talking about, like, well, how are you doing? Everybody's a well-rounded person, and it is okay to have these other sides.

Anda Trifan 27:22
Yeah, I also think it helps women in science a lot, because we don't want to seem as if our kids are taking time of our lives, and we're all just really focused on our careers. But that's not really true, right? I mean, you do have to split your time. And you do have to cater to all these different aspects of our lives. And so I think it just brings to light to that we do so much on a daily basis that was kind of, you know, ignored or pushed to the bottom before. So we're seeing that other people are also kind of struggling with their work-life balance, and it's okay to do that. And it's part of life. And I'm, I'm really glad that it kind of was brought to surface because it's not easy sometimes.

Sarah Webb 28:10
I think there's so many ways in which this time was hard, and it reshuffled how we think about everything. And it's just really interesting to hear the two of you kind of talk about that. But Amanda, I

want to be sure we talk about your ventilator work. You were here managing twins, and you have an older child, too, right?

Amanda Randles 28:29
Yeah. So we have a four year old now. But yeah.

Sarah Webb 28:32
But you got involved with COVID, too. Tell me a little bit about how you got involved and what you worked on.

Amanda Randles 28:37
It was right in the beginning, when, you know, we were hearing everything about like the ventilator shortages in the US. And we had previously worked with some people on the medical side of the Duke campus, and they really started a project trying to figure out how do we make the ventilators we have go further. So they were trying to find out ways to split ventilators between patients that didn't necessarily have to be matched. They could have different weights and different properties. They were questioning if you could use flow simulation to figure out how to actually tune the ventilator and create the product to split the ventilator. They kind of came to us and asked if we take our blood flow simulation and do an air flow simulation of the ventilator to try to help them tune their model.

Amanda Randles 29:13
It really all happened really quickly. That team pulled together in March, before we even got together. They created a brand new tool that could split the ventilator be 3D printed and be accessible to hospitals around the country. And then I asked my students, you know, if anyone wanted to be involved, who wanted to help, and we ended up creating three different teams that were trying out different models and figure out what we needed to do. From the scientific end of everything, all of our experience modeling blood flow was not necessarily that helpful in converting to airflow. Students created a brand new model. They had to validate that model. They worked with the team to do like benchtop comparisons with mechanical lungs. It was pretty fantastic of just like they created a brand new model within two or three weeks and validated that using tools we had never used before. We got to the point where we had the simulation we had had a validated and we wanted to have all of the data available so we can put it in for FDA emergency use, review and have it available for the doctors.

Amanda Randles 30:07
The idea was the doctor would put in information about the ventilator, that they happen to have the two patients they were trying to match. And it would tell you how to tune the apparatus that this team had created-- the weight of the patient, height of the patient ventilator-- it kind of explodes quickly with a number of parameters. And so we joined in with that HPC Consortium for COVID that came together early March, April 2020. We ended up work with Microsoft. We're working with people from completely different backgrounds, it just came together with no ego, no questions about IP, like it, people just jumped in immediately and wanted to help.

Amanda Randles 30:37

But the part, I think, is really interesting about the story of like, we submitted the application to the consortium on like a Monday. We were paired with Microsoft by Tuesday/Wednesday. We had our first call with Microsoft on Thursday. And then we ran all of the simulations we needed by that following Monday, and we ran like 800,000 hours of compute time in the cloud that weekend

Anda Trifan 30:57
That's incredible.

Amanda Randles 30:59
No one slept that we, like, you know, like, researchers from Microsoft, who had no idea what they were signing onto, like stayed up all weekend to help, you know, monitor this and make sure it worked. The first meeting we had with Microsoft, they brought in fluid dynamics experts in case we needed that help. They got in touch with people from MATLAB to like, try to figure out some issues, we were having to get things to scale on the backend. They really pulled in anyone, everyone to make this happen quickly. You know, we kind of approached them. And we're like, you know, we know we could be smarter, you don't have to do a brute force, let's model everything. But our dream would be let's just model everything and get it done. And they just gave us the compute hours and helped us finish it. That part was really amazing. We didn't end up getting emergency-use approval because, you know, the pandemic moved on. We didn't end up needing it. I think it's good information to have. We have the models out there. And it's helpful to have for the future. And it's, it's a good thing it wasn't used. Yet, it's an interesting side of it.

Sarah Webb 31:51
That's just super cool.

Anda Trifan 31:52
Yeah, it's very, very exciting.

Amanda Randles 31:55
It was ours was much shorter, like so like a month and a half. But like it was very much the same like very high adrenaline, how do we like get like everyone was kind of sitting there, like you want to do something and you want to help. I think it was good to try to feel like you're productive and helping in some way. Because I think everybody probably felt a little helpless at that time.

Sarah Webb 32:13
I wanted to get your perspective on what that's like to kind of be both at the cutting edge of the science and the cutting edge of the computing. For example, do you ever get frustrated by the computers not quite being able to model what you want them to model yet?

Amanda Randles 32:26
I love trying to figure out like, like, how far can we push it with the next biggest system. Like we're part of the Aurora early science program? And it's it's super exciting to sit down and brainstorm of like, what can we do? But then there, there have definitely been a few moments. There was one where like you forgotten exponent, and it was like, Oh, we thought we could go to like a one micron resolution was like

no, the next biggest system, we can go from a nine-micron resolution to an eight-micron resolution. That is much less exciting. There are questions like: It is physically intractable, to do a brute-force simulation of all the red blood cells in the body. That is not going to happen on today's computers tomorrow. Like we need different algorithms; we need different methods. It's not just “let's throw more processors at that, really making sure you're you're accurately bounding your problems and understanding “yes, we have access to the biggest systems, but that doesn't mean you can do anything.”

Amanda Randles 33:12
And lately it's for a long time, we played with questions like in situ visualization. And it was more of a fun to visualize while you have it, but not necessarily necessary. And now we're at the point where the data is a real issue. Even you know, Summit, or the systems today, we're using the entire memory and then we're running a million time steps. So you're creating, let's say, like a petabyte of data. And then you're doing that for a million times steps, like you have to be processing on the fly, that you can't make a video of the entire thing at the resolution you want to. We had a big push for a project last year, we were trying to just to get a picture of the shell of the data of what we were simulating. And that was on Summit. The person making the picture was at Livermore. And it was 450 gigabytes just for the shell. And it was like just trying to figure out how to get how do we get that to Livermore for the picture? We now have the capability to run these simulations, but actually, you know, how do you analyze it and do something productive with it.

Amanda Randles 34:07
And it's not frustrating with like, what, you know what we can't do with the systems now, but it's more of the, you know, we have the capability to do this amazing research, but it's hard to really capitalize on it and make use of it. And we're now we're working on a lot more of like the advanced algorithms to make it you kind of shift the way we're modeling and like what kind of like you had to tie in things like machine learning, how do you accelerate? How do you how do you do those sides? I think a lot of what we've done in the last 10, 15 years has really been pushing-- let's go to the bigger scale. And so if we you know, we did the first model that you can take the entire supercomputer and run the full 3D human body and that's great. Like we can model one heartbeat. If you want to model, you know, a whole week or an hour or someone's time, fundamentally like you're never going to get all of Summit for two months to actually model even an hour of someone's life so that we can hit the scale. How do we extend the time periods we can model and how do we how do we make the most of the time when you do get something like all Summit? How do we make the most of that of that compute resource? And how do we make sure we get everything we need to out of it. I want to play in the space of like, you need the biggest supercomputer, and I think I think that's, that's really exciting and fun. But um, yeah, trying trying to make that tractable and do things with it is where we're at.

Sarah Webb 35:16
So, Anda, for the kinds of things that you're working on, and obviously COVID was pushing Summit and what was available to the limit, I'm sure.

Anda Trifan 35:24

A big problem is, so now we're able to build systems like the aerosol, for example. And it's in the billions of atoms range. But one major issue is the amount of data that comes out of it, for sure. And so struggling with having to even transfer to be able to visualize any of those simulations, because you need, we don't quite work on the petaflop scale, but still, you have terabytes of data that need to be analyzed. And so one issue is, how do you manage transferring even any of that data from like a supercomputer like Summit to a local workstation, so you can literally just visualize five nanoseconds of a trajectory. So that becomes a problem really quick with expanding systems.

Anda Trifan 36:12
But another one is, okay, now we have simulated the virus and an aerosol. Where do you go from there? So what are what are the things that you want to answer? Because most biophysical things happen on order of milliseconds or, you know, microseconds. And realistically, right now, we would need, you know, Summit for months to be able to really simulate something of that timescale. So, of course, one of the first steps in any emerging technology is to be able to do it. So proof of concept is always very important. And just being able to put these systems together and simulate them for one nanosecond is an incredible victory for our field. But from here until you are actually able to extract information that you can apply, let's say, tomorrow to a biophysical problem, I think we still have quite a ways to go. But again, we need tools like machine learning, or just smarter weights to analyze these larger trajectories, because you want to be able to, like Amanda said, capitalize on being just the sheer fact that you're able to do these things.

Sarah Webb 37:32
What advice would you pass along about getting started in this field?

Anda Trifan 37:37
One of the things that I'm passionate about is getting more women in science because, in most of my meetings, I am one of the only girls or women in the meeting. I think one of the deterrents of seeing more women in this field is that a lot of us don't think that we can handle our personal lives and our professional lives at the same time in such a demanding field. And I would say from experience, and, you know, Amanda is an amazing example of this, that you absolutely can and there's resources like other ones of us that have been there to rely on. And so I think it's very important to not give up your personal life for your career necessarily, or, you know, wake up 15 years later and regret that you didn't do what you wanted also on a personal plan. So, specifically for women in science, it's very possible to do everything that you wanted to. Of course, it's insane at times, but I think it's worth it, and you don't need to sacrifice either part of your life.

Amanda Randles 38:44
It is very hard at times. But it's also, you know, it's incredibly rewarding, and it is possible. And we're all, you know, found it easier, like throughout the pandemic as you can reach out. It has opened things up where you're not just reaching out to people who live down the street from you, like, it's been helpful to reach out to other female faculty members who have kids who have those and can give advice and are there and like having the support network is like, I think we're, we're all much more open to that. And really, you know, it's okay to ask for advice of how to do it. And also just listen and know, if you need to

say like, it's hard. I just need someone to understand that it's hard. I think it's been helpful to have to know that other people are finding it hard, too, but it is possible.

Amanda Randles 39:25
And I think the big thing I know that kind of resonates with me that's kind of come out in both of our explanations is like finding something that you're truly passionate about. Like the COVID situation is very unique, but if you are truly passionate about it and excited about making an impact, I think that is it is worth finding out a way. You don't want to look back in 20 years and be like there were other things that could have been much more excited about. Try to figure out how to how to work on the type of problems that really get you out of bed in the morning and make it worthwhile to make this trade off. It is difficult to try to balance family and work, but it's worthwhile when you have, you know, amazing family time and then also are genuinely excited about the thing you're doing during the day.

Anda Trifan 40:01
And to build on that one piece of advice that I would have is find a team where you feel supported and where you feel like your ideas or your work is being put forward and not pushed down. Throughout my career, when I worked with people who were supportive, I definitely was a lot more productive. And I mean, obviously, that makes sense, but it's not always the case. So for me, that's a that's a very important step when when choosing a lab or when choosing a workspace.

Sarah Webb 40:35
Well, on that note, I want to thank you both for your time. This was such an amazing conversation, and I learned a lot and I hope you enjoyed it, too.

Anda Trifan 40:43
Yeah.

Amanda Randles 40:44
Thanks for including me.

Sarah Webb 40:46
To learn more about Anda and Amanda, please check out our show notes at scienceinparallel.org, which includes links to an article from the 2022 issue of DEIXIS magazine about Anda and a Microsoft article about Amanda's COVID-19 ventilator work. This episode concludes Season Two and our six-episode series exploring how pandemic events have shaped computational science work and workplaces. Please subscribe and rate us on your favorite platform, and we'd love to hear from you via email or on Twitter.

Sarah Webb 41:22
Science in Parallel is produced by the Krell Institute and highlights computational science with a particular focus on work by fellows and alumni of the DOE Computational Science Graduate Fellowship. Krell manages this program for the U.S. Department of Energy. Our music is by Steve O'Reilly. This episode was produced and edited by me, Sarah Webb.