



SDS PODCAST EPISODE 749: DATA SCIENCE FOR CLEAN ENERGY, WITH EMILY PASTEWKA

Show Notes: <http://www.superdatascience.com/749>



- Jon Krohn: 00:00:00 This is episode number 749 with Emily Pastewka, head of Data at Palmetto. Today's episode is brought to you by Prophets of AI, the leading agency for AI experts.
- 00:00:14 Welcome to the Super Data Science Podcast, the most listened-to podcast in the data science industry. Each week, we bring you inspiring people and ideas to help you build a successful career in data science. I'm your host, Jon Krohn. Thanks for joining me today. And now, let's make the complex simple.
- 00:00:44 Welcome back to the Super Data Science Podcast. Today's guest is the exceptional communicator of technical concepts, Emily Pastewka. Emily leads the data function at Palmetto, a cleantech startup focused on home electrification. Prior to Palmetto, she spent more than years building consumer data products and solving marketplace problems as a data science and machine learning leader at huge, fast-growing tech companies like Uber and Rent the Runway. She holds a master's in computer science from Columbia University and undergrad degrees in economics and environmental policy from Duke.
- 00:01:13 Today's episode should be accessible to technical and non-technical folks alike because whenever Emily got technical, she did an exquisite job of explaining the concepts. In this episode, Emily details how data science and AI can make the world greener by shifting us to clean energy. She talks about the team of people needed to bring clean tech data solutions to life, how econometrics plays a key role in nudging consumers toward greener decisions, her top tips for excelling as a data leader, and what she looks for in the scientists and engineers she hires. All right. Are you ready for this marvelous episode? Let's go.



- 00:01:50 Emily, welcome to the Super Data Science Podcast. It's so nice to have you here. Where in the world are you calling in from today?
- Emily Pastewka: 00:01:57 Thank you so much, Jon, for having me. I'm calling in from Boulder, Colorado.
- Jon Krohn: 00:02:03 Nice. We have met in person, or at least I remember seeing you in person. I don't know if we actually talked. But previously, when you lived in New York and you were still working at Rent a Runway, Rent the Runway, not Rent a Runway. That'd be quite a commercial model as well. It's a much smaller market, but Rent the Runway-
- Emily Pastewka: 00:02:21 Quite a different concept.
- Jon Krohn: 00:02:23 Yeah. When you were working at Rent the Runway, you did an amazing talk at MLconf. It must have been MLconf 2021 or 2022. And, yeah, it was a home run. I was completely mesmerized the whole time you were on stage. I took note of your name, and I think I might not have even had a chance to talk to you afterwards. I do remember, and I guess this is not surprising given where you worked, but you also had a fabulous dress. I was like, "Wow." It was the kind of thing that before you went on stage, before I knew that you worked at Rent the Runway, you're like, "No." This person really knows what they're doing in terms of fashion.
- 00:03:08 And then when I found out you worked there, I was like, "Ah, yeah." But yeah, amazing talk and connected with you on LinkedIn, and send you a message pretty much right away. This is two years ago. I sent a message. I'm like, "Emily, you're an amazing speaker. I'd love to have you on the Super Data Science Podcast, the most listened-to podcast in the data science industry And then it was crickets for two years, and then you just reached



out and were like, "Sorry, I was busy, and I'd love to do it." And now, you're here.

- Emily Pastewka: 00:03:38 I'm so sorry for the radio silence there. We did meet. It was 2022. And the crickets reason is I ended up taking a sabbatical that summer, and we'll talk about this, I hope, later in the episode. But I did a part-time master's starting in 2018, that I wasn't sure if I wanted to turn into a PhD or not, and so really threw myself into research that summer, also took the opportunity to spend some time living out of the country, working on language skills. And it would have been tough to balance that with continuing on my normal job and everything that came with it. So great to be here. Thank you for responding when I reached back out.
- Jon Krohn: 00:04:18 Oh man. I was super delighted. It's one of those things that actually happens more often than you'd think, where I ask somebody that I'm really excited to speak to on the show, and I'm just like, "Well, I gave it a shot." And then, yeah. I don't know if two years later is the record, but whenever people reach back out and want to do it, I absolutely love it. So let's talk a tiny little bit about the transition that you just said before we talk about what you're doing now, because I didn't know that. That didn't come out in the research. I guess it's probably not something that we see online about you, but that's sabbatical. So I guess was this while you're doing the master's in computer science in Columbia?
- Emily Pastewka: 00:05:03 That's right. So I finished my coursework in a year and a half, and that would have been in the summer of 2020. And I kept chugging along on the research with this question in mind, "Would I like to do research full-time for a while?" The machine learning space was moving so fast. And then in my time at Uber, I had this opportunity to work with so many people who had doctorate degrees, and the depth of their knowledge was always such a

Show Notes: <http://www.superdatascience.com/749>

phenomenal resource for me. And I think working at a large company where the problems have such a huge impact, the scale just means that a relative change in the accuracy of a model or the success of a system being miniscule can still mean millions of dollars, or it can touch millions of users.

00:05:51 That means I did get the chance to go very deep on problems in my time at Uber, and that was an indicator that I might enjoy deep research for a few years, and just wanted to give myself a chance to really try it out and say, "90% of my waking hours, I'm going to be thinking about one problem instead of spreading myself over everything that comes with a real job." And, yeah, so I spent about three months doing that. And actually, it was probably only two months of it were the research, and then I spent another month hiking the Lost Coast Trail, which is only a couple days.

Jon Krohn: 00:06:23 Where's that? Where's the Lost Coast?

Emily Pastewka: 00:06:27 So that is in Northern California, just north of Mendocino. It's 25 miles of untouched coastline. So you're actually hiking on sand for the most part.

Jon Krohn: 00:06:37 Really?

Emily Pastewka: 00:06:39 Other than that, that was pretty brutal. But other than that, I highly recommend it.

Jon Krohn: 00:06:42 It's like sand in your boots.

Emily Pastewka: 00:06:43 Yes.

Jon Krohn: 00:06:43 You're like one mile in, and you're like, "Oh, that's a lot of sand in my boots."



- Emily Pastewka: 00:06:47 It's a high-effort hike, but I got perfect conditions, couldn't have asked for better, and love a good solo trip. So that's great.
- Jon Krohn: 00:06:58 Oh, you did it alone?
- Emily Pastewka: 00:06:58 I did, I did.
- Jon Krohn: 00:06:58 That's like a one night of camping in that kind of 25 miles?
- Emily Pastewka: 00:07:02 It's typically two or three. So [inaudible 00:07:04] two.
- Jon Krohn: 00:07:04 So it's really arduous.
- Emily Pastewka: 00:07:06 It's not very hilly, which I think sets it apart from a lot of multi-day hikes that people think about. If you do the whole PCT or the Appalachian, you're definitely encountering mountain ranges. Being on the coast, the most challenging thing is you're really keeping track of the tide tables. There are a couple points that are so narrow between a cliff and the ocean. But if you don't time it properly, then you're not crossing. So that was fun for me to study environmental science in college. And any chance, I have to nerd out on climate data is fun. I'm taking an Appalachian course right now, prepping for backcountry ski season. And learning about the different types of snow is also super cool for me.
- Jon Krohn: 00:07:50 All right. This is really cool, and we're obviously super off tangent, especially for data science podcasts, but I'm finding this really fascinating. Other than the Lost Coast Trail, do you have any other highly recommended hikes for our listeners?
- Emily Pastewka: 00:08:01 I haven't done many other multi-day hikes since I was a child in summer camp, to be honest, which is why it was such a treat to do that one.

Show Notes: <http://www.superdatascience.com/749>

- Jon Krohn: 00:08:10 Really, multi-day hikes as a child in summer camp?
- Emily Pastewka: 00:08:12 Yeah. A lot of people went to math camp. I probably should have done that. It would have been strategic, but I was a nature nerd.
- Jon Krohn: 00:08:18 Well, okay. So despite not going to math camp, you were at Columbia doing a master's on deep learning. You were applying deep learning computer vision. So again, this was not really where I planned to start this episode. But since we're there, do you want to tell us a little bit about that? And also, it seems like you then made the decision after those few months, doing the hiking on the Lost Coast Trail, and getting really deep into it. It seems like you made the decision to not do the PhD, I guess, and go back into industry.
- Emily Pastewka: 00:08:52 I did. The defense went really well of the master's research and definitely spoke with... I had three advisors from different departments and spoke with a few of them about that option, but I really love seeing the impact of my work on the world. I've known that for a long time. And I'm glad I gave myself the chance to answer this question, but it really helped me remember why shipping code is so fun. And I think going to startups, which is what I did next, was the natural opposite end of the spectrum and a natural move in that way because I realized I hadn't maximized the code shipping side of the work. I'd always been somewhere in the middle. I'd given myself the chance to explore the research side really deeply. The next move was a tiny startup where speed was of the essence.
- 00:09:47 But, yeah, to get into the research, the master's focused on deep learning. And I did it, like I mentioned, or I started it while I was part-time, while I was full-time at Uber, loved my role at Uber, wasn't ready to leave, was learning a ton through working there. And in my time at

Columbia deciding to do research, I looked around for programs that would interest me and professors who had research groups with openings. And this Smart Cities group led by Dr. Kostic in the Electrical Engineering Department-

Jon Krohn: 00:10:21 Yes. This did come up a bit in our research. Yeah. This is awesome. Yeah. Yeah, Keep going. Keep going.

Emily Pastewka: 00:10:27 Yes. So this Smart Cities research group funded by NSF had a component related to self-driving cars, and I thought, "This is amazing." I love learning from my Uber colleagues. We have a whole autonomous group, so I can have this whole extra set of experts to pull in as I'm learning with this Columbia academic community. So joined that research group. And what's funny is a big part of the value I added early on was just good software engineering practices, reducing the chaos. You did a PhD. I'm not sure if this resonates with your experience. But even now, I was at a AMS conference on climate forecasting recently, and talking to somebody from NOAA about technology, and he said, "The research is really hard to wrangle." This Columbia research group was not at the level of NOAA, and it was a mix of graduates and undergraduates and post-grads, and so working on just version control. And sharability and reproducibility was pretty big first focus, but from there, really digging in on the problems.

00:11:41 I mentioned wanting to work on the vehicle aspect of self-driving. And the group had just put cameras on two buildings at the corner of 120th and Amsterdam on Columbia's campus. And so there was new data coming out of that, and what data person doesn't love new data. So we were thinking about the problem to solve there. It's how can you use image data from multiple cameras to detect vehicles and pedestrians moving through an intersection, ideally with the end result of making traffic

flow more efficient, which is climate positive, as well as making people safer, and started with needing to map the images from both cameras into one planes that overlap with one another, so one homography and then, from there, running object detection and tracking models on each image stream individually.

00:12:35 And so I was working with a group on that, and we published some research about whether this quite distant bird's eye view of an intersection is useful image data for these types of models. We weren't building custom models at that point. It was taking the best in class in the late 2010s, so that was the YOLOv3, YOLOv4 times SDS, a few other options. And, unsurprisingly, pedestrians were very challenging to detect and track. Cars are a little easier, tracks a little easier. And one of the biggest problems with pedestrians, they lose their shape from that distance, but they're also easier to block from other objects. And so a truck would go by, and you'd lose visual on a person walking, or a tree would have leaves, and you wouldn't see a person behind that. And that to me was a really big opportunity for the multi-camera setup, where one camera might not have an inclusion blocking important objects that we want to track, and the other camera might.

00:13:41 So being able to select the right camera for your detection at the time was the space I zoomed in on, so building a set of meta learners to determine which camera's images you want to use for a prediction at any given point in time, and a sub-algorithm that split the image up into a bunch of different grids, and choosing for each component of each image grid, which camera is best suited, and then stitching those images back together. And if you could stitch that whole image back together, and then only run inference once, you've still got a real-time model that has a chance at being part of a real-time system in the city. So that was a lot of fun for me, and

didn't end up doing as much with the self-driving team at Uber as I'd hoped. They work a lot with LiDAR data which, in my research, anytime you're reading about multiple cameras, you're reading about stereography, and LiDAR is so much better than image data in that space.

00:14:37 And so if I could have convinced the city of New York to go LiDAR with all of these devices, that would have been worth it, but I'm such a pragmatist, I knew that cameras are a lot cheaper. And so I wanted to know what we could do with image data.

Jon Krohn: 00:14:52 Empower your business with Prophets of AI, the leading agency for AI and robotics experts. Whether you seek a captivating keynote speaker, a company workshop host, or even guidance in implementing AI seamlessly into your organization, Prophets of AI connects you directly with the luminaries shaping the future of AI, such as Ben Goertzel and Nell Watson, both of whom have been phenomenal guests on this very podcast. Whether you are a large global enterprise or just beginning your AI journey, Prophets of AI have a solution for you. Their speakers have graced the most prestigious stages around the world, and now you can head to ProphetsOfAI.com yourself to see their full roster or to the show notes where we've got their contact link.

00:15:34 Well, that's super fascinating, and your ability to explain all of this relatively complex research in such easy-to-understand terms where I'm following along so effortlessly, it flashes me back to when I saw you on stage. And I was like, "I've got to get this person on the show." So thank you. You did have a few acronyms in there that we should break down for the audience. So you talked about late 2010s, the big object detection algorithms at that time being the YOLO algorithm. So You Only Look Once. And so I'll have a link to that kind of series of models in the show notes. But there were a



couple of abbreviations that you used that I didn't know what they were at all. So you talked about AMS and NOAA. What are those things?

- Emily Pastewka: 00:16:23 So AMS is the American Meteorological Society. I mentioned I live in Colorado now, and it's really a hub for energy and environment research groups. And so I saw that AMS was holding their annual conference in Denver last year, or I guess it was early this year. And I thought, "Well, that sounds fascinating," and turned out to be very deep research in spaces that I had very limited knowledge of and I learned a lot. NOAA is the National Oceanic and Atmospheric Administration, and they release a lot of great datasets in addition to providing forecasts that have to do with environment in the North American area.
- Jon Krohn: 00:17:03 Very cool. So we're hearing that girl who loved to do the outdoor stuff, the nature stuff coming through, still in a lot of your interest today that you're applying data science to. All right. So that brings me pretty nicely to our next topic, which is what you're doing right now. So you are now the head of data at Palmetto, which is a company on a mission to increase access to clean energy. So here, we see the environment again. How does data science contribute to the adoption of clean energy solutions at Palmetto? And what are some of the data challenges you face in the industry?
- Emily Pastewka: 00:17:38 Yes. I am working at Palmetto, which is focused on the solar industry today. And I think like many tech companies in the energy space, there's a very big vision for the role data can play in terms of optimizing the grid and making our energy systems more efficient. But I'll zoom in on the solar problems first.
- 00:18:02 So this process of going solar as a residential consumer is extremely fragmented. In all likelihood, if you as a homeowner are thinking about going solar, somebody

knocked on your door and you opened the door, which are already bizarre events, and from there you said, "Sure, come on in, take a look at my roof, I'll give you access to my utility bill, I'll tell you some information that you can get financing with." And throughout all of that process, you're giving all of this information up front and then, eventually, the salesperson has to go through and determine with your utility whether you're going to be able to sell energy back to the grid. That's really important for making the finances work out.

00:18:43 They're running a credit check or reaching out to a financier to ensure that you can get a loan. They're validating whatever design that they've created for the solar panels for your roof with a design firm that confirms that your roof is up to scratch and can support solar panels, determining if you need a re-roof, if you need to take any trees down. So all of the steps along the way include a large number of different players. And so part of making this industry more widespread or more impactful is an efficient marketplace process. And so I've worked at marketplace businesses before Palmetto, and it's been really interesting to see how those concepts translate, looking to use data to understand timelines between different steps of a process and optimize the balance of different players in the marketplace. In this case, many cited, you have different types of supply interacting with this demand of the homeowner. And so that's a big part of data.

00:19:48 There are supply chain problems. We're working with different types of hardware, solar panels, inverters, batteries. There are new types of hardware coming out from the US, which are preferred for a lot of reasons, financial, and otherwise. And well, we might want to strike. That's a little bit political, but constantly there's supply chain problems for solar hardware, panels, inverters, and batteries, and a lot of new development in



that space, and you want to work with what customers want.

00:20:23 And then we're also looking at how we can automate some of those friction points more directly. One piece of automation Palmetto's building is a solar design tool. So I mentioned having a physical design of the solar panels that go on your roof and using AI technology to understand what the contours of your roof are, how you can optimally place solar panels to generate maximum kilowatt hours given where you are on the earth, the shape of your roof, and the needs of your home, which we can do using building data to understand your energy footprint. We're also doing a lot on the financing side. So right now, with rising interest rates and inflation, Palmetto saw viability of loans decreasing. And so we launched our own power purchase agreement and hardware leases, which make it possible for-

Jon Krohn: 00:21:15 Wow.

Emily Pastewka: 00:21:16 ... you to put solar panels on your roof and not have a big loan against them. So you're paying either a fixed monthly fee or just a fee for the energy you use. That means for us, we have pretty traditional financial data science work to do. So we're underwriting loans, and we are predicting the risk against those loans. We also need to understand, more than your credit, the climate risk of your home. And we're taking again into account the viability of the hardware and whether this solar system will produce the energy that we've promised you. And that gets into one of my favorite sets of problems we're working on today, which is this predictive maintenance space. So in the past, we helped install solar systems. And then we said, "Congrats, you have a solar system." Now, we own panels on a lot of roofs. And we need to maintain those because it's an asset on our books, as well as the service we're providing a customer.

Show Notes: <http://www.superdatascience.com/749>

- 00:22:15 And so looking at the production curves relative to those designs we've built and looking for both anomalies as well as comparing these forecasts and being able to preemptively determine when maintenance is needed, as well as understanding the different error codes coming out of this diverse set of hardwares to give our technicians a recommendation of the next best action they can take, it's really expensive to roll a truck to a home. So things that we can solve remotely are ideal. And once you've gone solar, you not only want your panels to keep working, but you have this source of free electricity, which really changes your decision-making with regards to what you're doing at your home, the appliances you put in it. So suddenly, gas stoves and fireplaces and water heaters aren't looking as good. So Palmetto's working on providing guidance to customers on how to further reduce bills and increasing energy independence by recommending whether you should buy a certain device or make a certain upgrade.
- 00:23:19 And so a really fun data problem here that takes me back to some of the data sets I worked with 10 years ago, early in my career, is using 15-minute interval data about the energy you're producing as well as what you're consuming, which we get from solar panel inverters, as well as we have the option to get from utilities directly, is to understand your energy footprint of your home and use those really granular readings of kilowatt-hour consumption to understand what devices might be running. And so bigger devices are obviously easier to pick out of something even 10 years ago in the early dawn of machine learning we were pulling out in my time at Opower, things like hot tubs, EVs, but even air conditioning. And if you have room-by-room readings, you can really start to pull out more insights there and make really useful recommendations to people.

00:24:13 And where I start to get really excited beyond solar is when you have access to a large number of these distributed energy resources, that becomes really valuable to the grid at large. You can start selling energy back as a unit instead of household by household. You can use this as an option during demand events to work with the utility and say instead of turning on a coal-fired plant, which you're really hoping to take offline permanently, you now have this network of homes that can sell you energy they're producing. And so part of that is getting consumers to reduce their usage during these demand events and having that trusting consumer relationship. And then part of it is the grid optimization work, and there are some really cool startups in this space right now, approaching it with all sorts of different devices, appliances. And I think solar is a really good entry point.

Jon Krohn: 00:25:09 There was a parallel that came to mind for me there when you're talking about these energy flows from consumers out into the grid when there's a demand event. It sounds a lot like when I order an Uber and I get surge pricing. And so there's this incentive in this automated system that is dis-incentivizing me from using that service or requiring me to spend more. It sounds like a super similar kind of problem.

Emily Pastewka: 00:25:39 Yeah. I'm not sure how much you look at your utility bills or have clicked around your website, but you probably have options for different rates that you could sign up for. Standard rates are typically like you'll pay the same amount for a kilowatt hour of electricity at any hour of the day, but you can also sign up for rates that will let you pay a lot more if you use energy during a time when your utility forgoes high demand and a lot less in the off hours. And that incentivizes people to use the timer settings on their dishwasher and laundry, for instance. Run them at 10 AM, get home, you still have to turn your



lights on, but maybe don't run those other high power consuming appliances.

00:26:20 And that's a version of surge pricing basically. And this idea of using distributed energy resources to further optimize those times is, I guess, you would think of as... So surge pricing is the Uber rider financial impact control of demand. This distributed energy resource idea is the supply side where you're saying like the way Uber offers incentives to drivers to drive at certain times of day. This is offering consumers the option to be a part of the grid.

Jon Krohn: 00:27:00 Awesome. So you have gone through a broad range of different kinds of problems that data science can be used to tackle. So from things like financing consumers to just doing financial data science to this smart grid energy flows, this surge pricing stuff that we were just talking about, understanding hardware codes, detecting maintenance issues, just helping people make better clean energy purchase decisions, having a more efficient marketplace for those decisions to be made in, tons of different things. What are the different kinds of roles in data science or adjacent positions that are needed to fill all those different kinds of demands?

Emily Pastewka: 00:27:45 I think you can hear in that discussion of grid optimization, a need for pretty classic systems engineering and optimization knowledge. That is a huge area of research with such meaningful impact on so many people's lives. There's a lot of excitement about these technologies and also a lot of caution that utilities exercise and their decisioning around who they'll work with and what success looks like. So the quality of that research and the depth of it is certainly a fun challenge and a place that it is fun to think about. But as with any other data problem, sitting underneath that is getting the data. And the data engineering practice at Palmetto is so



important. We are working with, like I mentioned, so many different players in this multi-sided marketplace.

00:28:45 And to do a good job of making forecasts of demand and supply, we really want to understand the details of what's going on at every step in that residential solar sale and install process. That means integrating with or pulling data from local utilities, local jurisdictions, and the requirements that your city may have versus my city on going solar, document specifics, and as well as integrating with a lot of publicly available data sets that we can make intelligent predictions about the energy your home will use at any point of time in the year, image data to help us place panels. So that data engineering and data platform maintenance component is really key for enabling all these different problems. And the model we're pursuing at Palmetto is very federated beyond that. So there's, I'd say, a centralized ML research function and this centralized data engineering and platform development.

00:29:46 And then because there are so many business problems related to any different data sets and components of the marketplace and other problems we're working on, we have analysts embedded in many different parts of the company trying to marry that data with all of the business outcomes we're looking for and the challenging problems. And that role is the hardest to describe. It can take so many different shapes. And I think that's pretty normal if anywhere I've worked, and maybe highlighted in the energy space just with how much is going on right now.

Jon Krohn: 00:30:24 Large Language Models are revolutionizing how we interact with technology. With companies rapidly adopting models like the GPT and Llama series architectures, the demand for skilled LLM engineers is soaring. That's why Kirill & Hadelin, who have taught Machine Learning to millions of professionals, have

Show Notes: <http://www.superdatascience.com/749>



created the 'Large Language Models A-Z' course. Packed with deep insights on Tokenization, Input Embedding, Transformers, Self-Attention and LLM Tuning this course will help you gain hands-on experience with LLMs and stay competitive in today's job market. Enroll at SuperDataScience.com/llmcourse for your free 14-day trial. This course is exclusively available in the SuperDataScience Community, you won't find it anywhere else. Once again, the link is SuperDataScience.com/llmcourse

00:31:11 Yeah. Getting smart people to be able to handle all of those diverse types of data flows that the data engineers are bringing in and doing various kinds of analyses that are needed to allow decision makers in, I guess, any part of the business to be making better decisions, whether it's across the marketplace or the financial decisions or the energy flows. So, yeah. So sounds like the main areas that I jotted down there, as you were speaking, were systems engineering, data engineering, machine learning research, and data analysts.

Emily Pastewka: 00:31:42 [inaudible 00:31:44].

Jon Krohn: 00:31:44 And I imagine, maybe as part of that platform team, there's probably software engineers acting as like a glue between all these things or maybe that's mostly the data engineers themselves.

Emily Pastewka: 00:31:55 I'd say we have pretty strong data engineers by really close partnerships with software engineers. And something that's been really fun in this space is how collaborative the entire tech ecosystem feels right now. We all see this end goal of a more efficient grid and the big barriers to getting there. And there is a world where one company wins big at the cost of others losing, but it's not clear who that's going to be yet.



00:32:26 And it is clear that nobody's doing it alone. So I'm talking with other startups constantly who are, for instance, building an API aggregator so that I don't need to manually integrate with every single solar panel inverter. Instead, I might be able to have a standard format of data that is cleaned and validated through some other company that then provides that same service to people trying a lot of other things in this field. So speaking with people who are working on different subpieces of the tech stack has been really, really fun coming back to this space. So yeah, to your point about partnering with the engineering talent required, I don't think about just the engineers within a data team, but the broader ecosystem of engineering feels really collaborative right now in this space.

Jon Krohn: 00:33:15 Sweet. And are you hiring for any of those roles? I imagine we have a lot of listeners who would like to be making a big social impact with their existing skills in this space. It could be data engineering, data science, data analysis, software engineering. I'm sure we have tons of listeners in those areas who are already established as professionals who would like to be making a big positive social impact like you can be at Palmetto, or we could have listeners who are maybe getting into the field and are looking to be inspired by these kinds of possibilities for these application areas. So, yeah, are there particular areas that you're hiring for at this time?

Emily Pastewka: 00:33:58 Yes. It's mid-December when we're speaking right now. So as with many companies, not as many roles posted today as we close out the year. But one that is up on the site right now is an analyst role who would be focused on that financial underwriting area specifically. It's one of our newer areas. And so a lot of room to run in that role. Then, hopefully, next year, you would come to the Palmetto careers page and see a lot more posted. But I'd also recommend there's a group called Work on Climate.

Show Notes: <http://www.superdatascience.com/749>



They have a Slack channel. And I think it's a great resource for learning about roles available as well as people's path to finding jobs in the climate tech industry. It's a challenging space to learn right now, just there's so much going on. And so much of it is pretty new. There's been a ton of investment in cleantech over the last few years.

00:35:02 The administration has put so much investment into climate-friendly infrastructure and incentives that make VCs want to invest in this space. And so if you want to work on climate, there's a lot of options, but they're not always the really big companies that are easy to find. So I really appreciate the existence of the Work on Climate group.

Jon Krohn: 00:35:29 Yeah. I'll be sure to include that Work on Climate link in the show notes. And now, we didn't discuss this before we started recording, Emily, but although it is mid-December at the time of recording, by the time people hear this, it will be mid-January. So there may be more roles already available on the Palmetto site. And something that we were talking about before we hit the record button was that you do anticipate quite a bit of hiring in the future for particular kinds of roles. So I don't know if you want to get into that more forward-looking stuff, whether those roles are available immediately or not.

Emily Pastewka: 00:36:01 Yeah. I'll get into it generally. I think curiosity about the types of problems I just described is really key. Those are the things we're going to be working on. They have to be exciting to you. And I would also say, in general, I always hire for humility. You have to be willing to learn from your colleagues, and we're all learning about this industry as it develops in real time. And thinking you've got the answers doesn't help. We're looking for people who can bring expertise from different fields and know that they have that expertise and yet are willing to have a

beginner's mind in a parallel topic or an adjacent topic. And so that's my method for building teams always, as people who are really willing to accept that there's this concept of being T-shaped, where you're maybe a very deep expert in one thing.

00:36:55 But if you build a vertical bar chart of all of your skills, some of them are low, and some of them go deep. And if you're going to work on some of those skills that you're not as deep in, being able to work with your colleagues and ask them questions. And then I would also say generalists are really valuable at startups overall, but also in the energy space. It may sound like I'm talking about a lot of hyper-specific terminology, and my background has muddled through this space in a couple different ways, but I think a lot of the value that I've brought to Palmetto is my background on marketplaces and consumer data. And it's not energy-specific necessarily. So there are a lot of different skill sets that can move the needle. So don't think that you don't have an in.

Jon Krohn: 00:37:49 Very cool. And, yeah, so I hadn't heard of this T-shaped people idea before. But being broad across many skills and deep in at least one, I totally get that. It makes a lot of sense of the analogy. And, yeah, so T-shaped people, they could come for any of these kinds of roles, like systems engineering, data engineering, ML research, data analysis, software engineering across the board. It's probably conceivable that Palmetto will be looking for those kinds of folks in the future.

Emily Pastewka: 00:38:19 Absolutely. And I stole that concept from a book called *Analyzing the Analyzers*, which actually my boss, when I worked at Opower, had us read at a team offsite 12 years ago. And, obviously, it landed. So I recommend that. It's a little dated now. But the overall concept just really resonates with me. And it's a good read, short rehirible ad.

- Jon Krohn: 00:38:47 Nice. So you've alluded a bit, as we've gone on this episode so far, about your past. And one thing from your past that I find pretty fascinating, and we might be able to get a bit of great conversation here, a bit of deep conversation here, is around econometrics. So it's not something that I know a ton about, which is partly why I want to ask about it, but I know that it's really fascinating. So to mitigate climate change, aligning incentives and desired consumer behaviors to reduce carbon footprints is one of the big road blocks.
- 00:39:21 For example, getting people off of fossil fuels and onto electric vehicles would make a difference, or convincing people to fly less, for example, would make a difference. Actually, and just the kinds of things that you mentioned like being thoughtful about having a timer set so that you're using high-drain devices at low-demand times of day. So, yeah, these kinds of behavioral changes, behavioral economics, econometrics are powerful tools to study these kinds of effects. You've published research in the past on these topics, like social proof and norming, and the ability of these concepts to influence people's environmental behaviors.
- 00:40:08 You said at least once in this episode, coming back to the environment, so I guess this is what you're saying. So you've done this kind of research in the past. And after Uber, Rent the Runway, Columbia University, now, you're back doing this kind of climate stuff. So, yeah. So that was a very long question. I had all kinds of different things in there, but one, just confirming that this is what you mean by coming back to climate, and then letting us know what social proof and norming mean as terms, and how this impacts what you're doing today.
- Emily Pastewka: 00:40:42 Yes. So I started my career actually doing environmental consulting, and very quickly realized I'd rather build things and tell people what to build. So it was cool

working with DOE and EPA data, Department of Energy and Environmental Protection Agency data, but I moved into tech by taking a job at Opower. And Opower's fundamental learning that led the founders of Opower to create the company was about how behavioral science could be used to drive people to save energy in their homes.

- 00:41:24 And the study was done in the mid-2000s, where researchers put one of five door hangers on the doorknobs of a couple hundred households in Southern California, and they partnered with the local utility to take the consumption data from all of those homes and measure pre and post the intervention. And the five-door hangers had different messages. They all asked the consumer to save energy. But one of them was, do this for the environment. One of them said, "Do this for your community." One of them was for financial reasons. One was season dependent. It's summer. Be more efficient. And the last one said, "Your neighbors are doing it. Why can't you?" And that fifth one had the highest impact on people's energy savings rates.
- 00:42:18 And so the Opower concept was really running with that and saying, "If we can design large scale experiments where we send monthly or at least semi-regular emails and physical pieces of mail to people's homes with insights about their energy usage and critically a comparison to their neighbors, can we get them to save energy?" And the reason we had to run it as a large scale experiment was to be able to convince regulators that those energy savings were real and that the utilities that we worked with could claim those savings as part of state regulatory programs.
- 00:43:00 That was what made it a business and not just a nonprofit or an annoying mailer. So as part of running this business, we were gathering data from every utility

we worked with. And in addition to using that data to measure the impact of our interventions, it was what we used to define what your similar home's neighbors group was. So that was an unsupervised learning problem. And we were using not just how much energy you use, but also any data we have about the home. So the age of the home, size of the home, any known appliances, pools, solar panels, things that really impact the energy footprint, and that helped us derive a meaningful set of neighbors that we could compare you to because if you're too high or too low, it doesn't become helpful.

00:43:52 So that was one really interesting data problem at Opower. But the other one was that for a utility to want to pay for this program, they needed some level of predictability. Changing the demand in a service area too much without being able to plan for it is high risk. And so the forecasting component was where I spent a lot of time. And building that model, we pulled in all that same data about how we derived your similar neighbors, but we also pulled in other factors that we thought might help us understand your propensity to make changes to your behavior.

00:44:24 And part of that was the content that we put on the reports. So it's how do you compare to your neighbors as well as any tips we could provide or specific comms from that utility. But it was also more broadly spending being done at the state or local level, and it was demographics. And I think, finally, weather, of course, is really important just to know how much you could save. And as I was building those models, some of the features that were really important in predicting energy savings are pretty easy to guess. If you're a very high energy user, you have more space to bring it down. Similarly, if you live somewhere where you use both electric heat and you have an air conditioner, those are discreet things that you can change your behavior on. But some of the factors that

we found were a little harder to understand and probably the most interesting. So one of those was the state and local government efficiency spending. And we found that even controlling for demographics and location and other factors, this was a meaningful predictor.

00:45:38 And so this was a finding that we took to a conference at one of my co-authors, was on the policy side. And so we took these findings to the 2004 ACEEE conference, which is the American Council for Energy Efficient Economy, a heavy hitter in the space in terms of getting smart people all together to talk about some of the learnings. And the hypothesis we took to this group was, so we took our findings, obviously, but paired it with this hypothesis that this type of spending was priming customers to be more receptive to interventions on energy spending or on energy efficiency.

00:46:19 So, essentially, if you have been seeing your utility investing in changes and your local government talking about those investments, you are more willing to take action when you see that you're an inefficient household than somebody who didn't have access to the resources, the education, the social norm, and support to do that. And the data bore that out. And it was a really great finding to be able to talk about with policymakers and utility leaders who really have this question, is all this education spending I'm doing meaningful? It's super hard to measure. Maybe, you can measure email open rates and click-through rates, but what's actually sinking in and how is it changing behavior? And what we were able to show is that spending on education coupled with spending on these explicit energy programs, two plus two equal five. So those are really exciting finding.

Jon Krohn: 00:47:16 Oh, nice. Yeah. I've never thought about how two plus two equals five is always this like 1984 Orwellian concept of a bad thing that we want to avoid, and misinformation.

But in that context of the sum is greater than the whole of its parts, it actually makes a lot of sense. I've never heard that before. Speaking of Orwellian and changing people's behaviors, so something that we talked about a lot in episode number 715 of this show, we had Professor Allen Downey on the show. And he talked us through tons of examples of everyday situations where psychological research is interesting and where it has a real world impact, like take away message that we can use in real life. And one of the things that I found most fascinating in the episode that is cover to cover, absolutely riveting is that he talked about how attitudes shift over time.

00:48:20 So there are different ways that it shifts. So for example, it shifts by generation. So younger people are more likely to have more liberal values on things like, it would have been years ago, things like just inter-race relationships and more recently things like gay rights and transgender people. And so these kinds of liberal ideas are much more accepted by younger generations. There's this like generational effect. But people also change over time. So an individual over their lifetime will shift, although typically not as quickly as a new person who's born. The shift over time is slower. And so there's this interesting thing that I talked about on the show, which is that obviously, in some ways, this is really obvious, what I'm going to say. It's not profound at all. But in other ways, I feel like we don't think about it enough.

00:49:26 It does not something that I talk about in conversation a lot, which is that anything that we think... We have some of our own original thoughts where we tinker around with some data science model for a while and we figure out some new approach that maybe no one has ever thought of before. We could potentially publish a paper on it or it ends up being something viable that adds to our company's intellectual property, whatever. But most ideas

that we have come from the outside. And of course the time that you're born into is going to have a huge bearing and where you're born, when you're born, particular family born into obviously all of these things are going to hugely influence your perception of the world. There's no absolutely correct perspective. It's constantly shifting over time.

00:50:20 And so a lot of what you're talking about is things like having the door hanger that says, "Your neighbors are doing it, why can't you?" The idea here is to nudge people in the direction of a behavior that is, you know, better for society. And so you probably kind of get where I'm going here, which is there a right amount or a wrong amount of nudging people's opinions? To some extent, every conversation or every advertising campaign, and it's interesting how when it's a consumer behavior like buy more hamburgers, there's very little pushback of like, "Oh, you're trying to brainwash me," when obviously that is what's happening. Exposure changes your mind. When you see ads for the hamburger, you're more likely to get a hamburger. Me talking about hamburgers right now is increasing the probability that our listeners are going to buy hamburgers.

00:51:19 This message was not paid for by the hamburger lobby. But I should have been talking about veggie burgers. That would have been better for the environment. Here we are on a climate change episode and I'm talking about this methane-belching animal that we do horrible things to, to get our burgers. I wish I could go back and change all that. But, anyway. So, yeah. So you get where I'm going with this. Maybe, it was something interesting to say. I don't even have like a very specific question, but it's like something around especially when the government is funding it, and this is changing behaviors, yeah. I don't know if you just have any interesting thoughts that you can contradict me on or build upon.



- Emily Pastewka: 00:52:02 Well, I think you made a really important point there, which is so much of what companies are trying to do, especially consumer companies, is change consumer behavior in some way. And as a data scientist, I've worked at exclusively consumer companies. Even if they had a B2B sales motion, it was ultimately a B2B to C. So I was working directly with consumer data, individual humans were my subjects, and I tested my research subjects. And that's really something that I've been thinking about for a long time. And it's a reality that I've chosen to operate within by picking who I work for, and saying, "If I'm going to change someone's behavior with my work, that's what data science is quite effective at helping do, especially now that I've realized I'm not a pure research person. As much as I love R&D problems as well, I want to work somewhere that's moving the needle in a direction that feels right to me.
- 00:53:11 And so I don't think I'm answering your question by putting boundaries on how much we can move consumer behavior, but just recognizing that it is a huge part of our economic system, and the companies that operate within it, and choosing to put your skills to use to change behavior in a way that resonates with your values is a good step to take. And then in terms of government spending, vote. Make sure that people choosing how to spend that money align with your values as well. I'm sorry, I don't know if you can vote being Canadian.
- Jon Krohn: 00:53:51 I guess I'm allowed to vote in Canada.
- Emily Pastewka: 00:53:53 Yeah. It's true. Can you [inaudible 00:53:57].
- Jon Krohn: 00:53:58 I don't know. I haven't done it. I don't know. I haven't been a very good civic... Yeah. I don't know. I haven't lived in Canada since I was 21, I guess. And so now, it's been, what is that, make it 17 years of going to different countries? And I don't know. I guess there is like some

postal voting, but I don't know. I haven't looked into it. I think it varies by jurisdiction how long, what requirements there are for postal voting, and stuff. I don't know. I haven't been doing it. I apologize.

Emily Pastewka: 00:54:31 [inaudible 00:54:33].

Jon Krohn: 00:54:33 I'm sorry, Canada. All right. And now, we need to cut to our ad from the wheel-hunting lobby. Sorry. We won't be doing that. We tried to have very ethical sponsors on the show. Hamburger, hamburger. So, yeah. I don't know. It was just kind of food for thought with that. I guess I'm all for it. I think that the point that you're making there is a big takeaway that I had from it is the work that we choose to do. So people listening to this show, it is more than just a paycheck. You are contributing to these shifts in the way people think, whether it's in a subtle way or a more direct way. It's on the more direct way. If you are in marketing and you're taking on a vegan burger client versus a hamburger client, that's going to have a different impact on the climate.

00:55:33 And so to the extent that you can be trying to shift maybe even in a small way, you could be convincing management to say, "If we take on this vegan burger client, I'll do it for less pay, or we'll figure out some way of making that work better to try to be shifting the world towards something that is aligned with your values." Yeah. So I think that that's a good takeaway message there. You can't understate the impact that every decision you're making as a consumer or as a worker is contributing to one way or another, some kind of change in people's minds, as well as literal physical atoms on the planet that could be more positive or more negative for future generations.

00:56:24 Anyway, a bit of a lecture. So we've talked about things that you've done in the past. A little bit, we've touched on

them. So Rent the Runway that you worked at, Uber that you worked at. And now also at Palmetto in all of these roles, you help build machine learning teams and launch major data initiatives. And so, yeah, do you have anything like high level, I guess it could be low level, whatever level you want, strategies or leadership principles that you'd like to share with the audience that helped you manage these super high performing teams at super high performing companies and drive progress at them?

Emily Pastewka: 00:57:07

I don't think there's a one size fits all leadership style or set of rituals to impose on a team. I learned that, in particular, as many of us did through COVID, where so much had to change in the way we worked together and supported one another. And as we've continued to work from home, what worked in the early days of COVID has really changed. And so, expanding, taking that same concept and extrapolating it to change is constant and accounting for it as you figure out how to lead your organization is a huge part and just being willing to try things and recognize when they're not working and take someone else's idea and run with that is a big part of my style.

00:58:00

I think it's coming back to that humility that I hire for. I try to keep myself held to the same standard of willingness to put aside my ego or my belief about what's right and make decisions based on what's in front of me, no matter whose idea it is. But general principles that are pretty consistently serving me well, one of them is, again, hire the right people. It's not a silver bullet. You have to do a lot to enable those people too. But the wrong hiring decision we all know is really painful. And the right hiring decision is absolutely joyful. And from there, I think it really can differ by team as well as the type of company you're at. I would run things really differently at Uber than I was at a tiny startup. But when you're working with technologists broadly, one thing I find really valuable

is separating the immediate needs and immediate priorities from the long-term vision and making it very clear to your team how every step of a roadmap ladders up to that bigger goal that you're very excited by.

00:59:23 But in exchange, also making it really clear to leaders around the company where they're going to have to wait to see something because it's not an overnight outcome. And in some cases, thinking about Rent the Runway, for instance, there were certain projects that we deemed full research mode. And we were purely experimenting with 20% of our time while we were shipping code on other projects that were already in production. And so, constantly measuring incremental improvements on things like recommender systems that were live on the site. But telling our leadership also that 20% of time was really held aside for demoing new machine learning models as they came out, reading papers and testing how we might incorporate them into our stack. And then sharing those research learnings as well with the ideas for products to build on them was, I think, what kept the exact buy in there.

01:00:17 So separating those ideas of shipping value and discovering has been really good as part of both the motivation and the translation layer to exec teams, especially now where I think every exec team is trying to be very, very lean and cost-conscious. And discovery for discovery's sake is a little harder to finance than it might have been 10 years ago.

Jon Krohn: 01:00:45 Yeah, but interestingly, also simultaneously with the kinds of work that we do and our listeners do, more valuable than ever. We are in the spot where the most innovation and excitement is possible. And so that 20% investment, that seems like a great deal. With smart people who are willing to take some risks and try some new things, the potential for dramatically changing a

product or business operations to be more revenue-generating or more profitable, it's never been a better time.

- Emily Pastewka: 01:01:20 No. And I think finding those product and business leaders to be your champion as you start experimenting is so key to taking something from discovery mode. I'm getting it on the roadmap in the future. At Rent the Runway, at Palmetto, I have the people who are open to thinking about the future. That's the type of company I like to join and just being really narrow in who you choose for each project so that you don't get the whole company spun up around an idea. You really want to be able to incubate it, but still have them ready to run with you when it's time to set your priorities for the next half or the next quarter, however you do that. That's been really successful for me to not be the lone voice saying, "I promise this model is cool. It's going to have an impact." You have to bring the right stakeholders along for the ride.
- Jon Krohn: 01:02:15 I see, a great tip. Related to the diversity of your background, but making it even more diverse, you have a really a collective background. So it isn't just Uber, Rent the Runway, Palmetto. We've touched on it a bit in this show, but your background spans economics, environmental science, computer science, and business, the energy sector, transportation, retail, climate tech. How has this diversity of perspectives shaped the way you approach problems or build teams?
- Emily Pastewka: 01:02:43 Yes, it's definitely an eclectic background. And I think it served me quite well. I think of myself as a real generalist. And at times, I've wished I was more of a specialist. I've worked with people, like I said, who've done really deep research in this space, and it's phenomenal having access to their wealth of knowledge. But I really like how I can connect the dots between different industries and how the

same problem might present in different ways. And I can pull from a solution I've applied in one place and have at least a starting point for a solution at another company. It's also been really helpful stepping into more of a leadership role that I've done many different types of technical jobs within the data stack. And so I can be a useful resource to a data engineer as well as a data scientist or an analyst.

01:03:35 And then, finally, especially recently, I think a big part of the head of data role is evaluating what you're going to build versus what you could buy. There are so many more options out there of white label B2B company selling machine learning solutions, and having that breadth of knowledge to say, "This is absolutely something that can be commoditized," and we can trust someone else to build it as long as it's not part of your secret sauce, you want to be able to make that decision with confidence. And I think it's really key as a leader right now to be able to do that.

Jon Krohn: 01:04:12 Nice. Yeah. I love that answer. And so last question here lending on your general broad experience, but in a previous interview, we pulled out from research, we pulled out during our research or specifically Serg, our researcher pulled out from an interview, you saying, "We're not all imposters. We're all just talented in different ways in different parts of this field." And that was a quote that I really liked. We're not all imposters. We're all just talented in different ways in different parts of this field.

01:04:42 And so I wanted to make sure that we got that quote into this episode. And it seems to me like that relates to the T-shaped people that you were talking about earlier on in the episode. But maybe beyond that, is there any advice you have for mitigating feelings of imposter syndrome that

so many of us have in the fast moving data science machine learning AI space?

- Emily Pastewka: 01:05:04 Naming it has been really helpful for me. So a specific example, at Rent the Runway, I joined to lead the machine learning group. It was pretty comfortable with ML data science analytics. I at Uber had had to build a lot of my own data pipelines in the early days. But by the time I left, we had a massive data engineering organization. The technologies had evolved. It was not my bread and butter to say the least. But as the most technical leader within the data org, I ended up leading data engineering when that director left during my tenure. And I was very upfront with the team. I have not done this job. I have immense respect for this job. I am very curious about it. Tell me when I'm taking up too much of your time going too deep, help me find the right level to zoom in. But don't be afraid to tell me when I need to learn something better.
- 01:05:55 And also being upfront with my leadership group that I needed a little extra space to learn some stuff, because it was not going to help anybody if I was just blindly decision making. So naming it has been really key to overcoming it. It's impossible to work yourself to death, or maybe not impossible. But I have not found that just working yourself to death in secret and hoping that you overcome it on your own is nearly as effective as, again, getting that support from the people who know the things you want to know.
- Jon Krohn: 01:06:28 Nice. I like the way that you framed that to the team acknowledging, yeah, this isn't where you cut your teeth coming up. And so getting that right balance of supporting them, but also learning. And that sounds really great. You sound like an amazing manager. I really want to work for you.



Emily Pastewka: 01:06:47 That's so kind.

Jon Krohn: 01:06:49 Oh, yeah. All right. And so one more technical question before I let you go is, are there any particular tools or technologies or approaches that you're particularly excited about today that you think our audience should hear about?

Emily Pastewka: 01:07:07 Having moved from larger to smaller and smaller companies, I have developed a very deep appreciation for some of the modern data stack components that help a smaller team scale themselves. And I think they're also helpful, even if you are in a larger company or working on a side project for you as an individual, just be able to move really quickly on a problem. So some of those people know really well, like on the D side, I think there have been massive improvements that I alluded to earlier. On the data science side, something that I've been really blown away by in the last few years is experimentation platforms coming out that do so much more than a traditional AB marketing test, being able to configure really complex hypothesis tests. I've personally used Statsig at a data team. And then, I've also spent a lot of time with the founders of Eppo, which is really beautiful platform that's also for experimentation.

01:08:10 I think as a data scientist, running experiments is really bread and butter, but it doesn't mean you want to maintain all of that code on your own. It's setting it up and explaining. Interpreting the results is the value add. And the faster you can do that, the more you can learn, which is what we all love in this role. Eppo, I actually learned about through an angel investor syndicate that I'm a part of, which is all heads of data at different sized companies in different industries. And that's been a really great professional resource for me to learn how other people are approaching this job, what tools they're learning.

Show Notes: <http://www.superdatascience.com/749>

01:08:49 And then as we collectively decide on startups that we should back, and these startups are coming to us not just for the tiny seed investment or early stage investment, but also because they're looking for your insights from heads of data. And so I'm getting the chance to advise these companies on the tools they're building. And a few more that are really early stage and that I'm really excited to see grow are some ones called Motif working on time series analysis automation. Time series data is so important for forecasting and for being able to modify consumer behavior real time. But it's not easy data to work with. Maintaining that chain of actions is typically a complex data set. And so they've got a really nice intuitive UI and set of tools.

Jon Krohn: 01:09:40 Is that a Sean Taylor's company?

Emily Pastewka: 01:09:41 It is. And his co-founder was at Uber for a long time.

Jon Krohn: 01:09:47 Oh, yeah, no kidding. Yeah. So he was episode number 617 of this very podcast, was focused on Sean Taylor and Motif.

Emily Pastewka: 01:09:54 I will have to go back and listen to that one. I'm not sure how I missed it.

Jon Krohn: 01:09:57 Yeah. 617. Nice.

Emily Pastewka: 01:10:00 You've probably had Chad Sanderson on too, if I had to guess.

Jon Krohn: 01:10:03 Chad Sanderson? No, I've never heard of Chad Sanderson. I'll have to compare notes on that afterward.

Emily Pastewka: 01:10:08 Well, I'll just mention, another one I'm really excited about is this company called Gable, which is really early stage, but they're working on the interface between engineering and data engineering, this concept of a data

contract where you can protect the quality of the data you're reading into your warehouse. I think we all resonate with the value that would add if done extremely well.

01:10:32 And then on the ML side, a company that I've been talking to for a long time and I've personally been playing with the software just because it's so enjoyable to work with is called Baseten. They started as a pretty general purpose ML infrastructure company. And now, they're focused more on helping people get LLMs and other generative AI models into production in minutes instead of days or weeks. And really, I couldn't believe how approachable it was as somebody who has not gone super deep on LLMs. That's been a really nice way to feel like I could get more comfortable with the space. It's not something I want to fall behind on. So having ways to stay on top of the research via good tooling is really great for me.

Jon Krohn: 01:11:28 Nice. Well, some great recommendations there, thanks to your involvement in the InvestInData executive angel investors group. I guess that's something that some of our listeners might want to check out and maybe apply to. How do you get involved in that?

Emily Pastewka: 01:11:41 Yeah. I would say you can reach out to me directly. And we're certainly looking for more voices from diverse backgrounds. Like I said, it's having that cross cut of the whole industry and the people in it that hopefully makes us better than a coin flip.

Jon Krohn: 01:12:00 Nice. Yeah. Hopefully. And yeah, some great companies named there will be sure to include all of them in the show notes. Before I let guests go, I always ask for a book recommendation as you knew because you've been listening to episodes. And so I know you're well prepared for that question. But something that I wanted to digress

on speaking about tools and technologies that people could be excited about, prior to hitting the record button, you were telling me about an app called Libby that I guess people at least in the US could be using to be getting their books, I guess, basically free of charge.

- Emily Pastewka: 01:12:33 Completely free of charge. I love Libby. It is an app that connects you to the digital contents of your local library. So you need a library card, and then you enter your library card details, and you can download audio books and ebooks. And I do both. It's very, very handy. It keeps me listening to books way more than I would otherwise.
- 01:12:59 And a book I listened to very recently is what I'm going to recommend here. It's the biography of Leonardo Da Vinci by Walter Isaacson. He has had a number of phenomenal biographies making bestseller lists recently of tech founders. But I really loved the Da Vinci biography because Da Vinci is the ultimate generalist. It's the definition of the Renaissance man. And I think seeing what an impact that type of person can make on so many disciplines was very inspiring for me, again, as somebody who has struggled with imposter syndrome and wondering, "Will I ever know something deeply enough?" I'll never be Da Vinci, but it's nice to know that we remember him as an artist. In fact, his contributions had he published them to anatomy of the human body, to vision, to so many fields was just incredible. And I think it's really an inspiring read about just curiosity and pursuing the things that interest you.
- Jon Krohn: 01:14:00 Yeah, like designs of a helicopter without having the motor or the infrastructure, but those kinds of concepts, wild the breadth of ideas that Da Vinci had, for sure. It sounds like a fascinating read. And so for people who want to reach out to you to get involved in InvestInData or to talk about opportunities at Palmetto or clean energy in general, or maybe other questions that they have, or



they just want to follow your thoughts, what's the best way to do that?

- Emily Pastewka: 01:14:27 My LinkedIn is probably the best way to do that right now. Haven't been very active on any of the social media. And I do maintain a personal website. It's a GitHub pages. So emilypastewka.github.io. I have some posts in the work. So it's a little long content, but there is a page with more opportunities or ways to get in touch on there.
- Jon Krohn: 01:14:53 Nice. All right. We'll be sure to include those in the show notes. Emily, thank you for reaching out after all these years. It was great to have you on the show to be immersed in your wisdom and to hear you speak so clearly about every topic that you covered. It was really a joy. Thank you so much, Emily.
- Emily Pastewka: 01:15:10 Thanks so much for having me.
- Jon Krohn: 01:15:17 Hope you enjoyed today's episode. In it, Emily filled us in on the myriad ways data science and AI can help us shift to clean energy, such as through helping people make better clean energy purchase decisions, providing clean energy financing, detecting hardware maintenance issues, and automatically sending energy where it's needed most as part of a smart grid. She also talked about the breadth of roles needed to bring cleantech data and AI products to life, including systems engineers, software engineers, data engineers, data scientists, ML researchers, and data analysts.
- 01:15:45 As always, you can get all the show notes, including the transcript for this episode, the video recording, any materials mentioned on the show, the URLs for Emily's social media profiles, as well as my own at superdatascience.com/749. Thanks to Ivana, Mario, Natalie, Serg, Sylvia, Zara, and Kirill on the Super Data Science team for producing another marvelous episode for

Show Notes: <http://www.superdatascience.com/749>



us today. For enabling that super team to create this free podcast for you, we are deeply grateful to our sponsors. You can support the show by checking out our sponsors' links, which are in the show notes.

01:16:14 Otherwise, feel free to share the podcast and particular episodes you like with people you think might like them. Please do review on your favorite podcasting app or on YouTube, and subscribe if you haven't already and all those other kinds of great ways of engaging with us on the show. But most importantly, we hope you just keep on tuning in. I'm so grateful to have you listening, and I hope I can continue to make episodes you love for years and years to come. Until next time, keep on rocking it out there, and I'm looking forward to enjoying another round of the Super Data Science Podcast with you very soon.