My, Myself, and AI Podcast

Fashioning the Perfect Fit With AI: Stitch Fix's Jeff Cooper

Sam Ransbotham: How can humans and generative AI work together to ensure we're dressing for success? Find out on today's episode.

Jeff Cooper: I'm Jeff Cooper from Stitch Fix, and you are listening to *Me, Myself, and AI*.

Sam Ransbotham: Welcome to *Me, Myself, and AI*, a podcast on artificial intelligence in business. Each episode, we introduce you to someone innovating with AI. I'm Sam Ransbotham, professor of analytics at Boston College. I'm also the AI and business strategy guest editor at *MIT Sloan Management Review*.

Shervin Khodabandeh: And I'm Shervin Khodabandeh, senior partner with BCG and one of the leaders of our AI business. Together, *MIT SMR* and BCG have been researching and publishing on AI since 2017, interviewing hundreds of practitioners and surveying thousands of companies on what it takes to build and to deploy and scale AI capabilities, and really transform the way organizations operate.

Sam Ransbotham: Hi, everyone. Today, Shervin and I are talking with Jeff Cooper, senior data science director at Stitch Fix. Jeff, thanks for taking the time to talk with us.

Jeff Cooper: Thanks so much for inviting me. I'm really excited to be here.

Sam Ransbotham: Let's get to some fashion basics. What is Stitch Fix?

Jeff Cooper: Stitch Fix is an online personal styling service. We serve more than 3 million clients — that's what we call them — in women's, men's, and kids' sizes. What we're trying to do is help people get dressed: to offer the most convenient way to find clothes that you love, clothes that you might've found yourself. And we can find a great way to put them together with other things, find clothes that you might not have chosen for yourself, and help you push the boundaries of your style.

We have a unique art and science approach. When you sign up, you're matched with one of our thousands of style experts. Our stylists work together with our data science team and tools that we provide them to help find clothes for you. We send [clothes] to you: You keep what you like, you send back anything that you don't want for free. We've been doing this for over a decade now. We just had our 13th birthday yesterday, and we just passed 100 million Fixes. At this point, we have spent a lot of time thinking about how to serve our clients the best, how to blend great algorithms and data science with our style experts' intuition and understanding, and so forth. We really believe this model is a way to help clients find what they need.

Shervin Khodabandeh: What a great example of humans and machines working together. Tell us more about how that actually happens. What does the machine do? What does the human do? How do they work together?

Jeff Cooper: We're really passionate about this model. We've been at it for a long time, really since the beginning. As you can imagine, for any retailer, the idea of sending people things that they haven't specifically chosen for themselves, that they can return them for free, feels a little risky. In order to do it well — as we think we do — you really have to know your customers incredibly well.

So, we have certain things that we think hard about doing. We have to have great learnings about our customers. We ask them many questions. Our customers are interested in talking to us about their style. They're here to help be styled. We think really hard about all the ways that we gather feedback. When you try on an item, whether you keep it or return it, we ask for a lot of feedback. We get a lot back: 85% of clients leave feedback on items. We ask for a lot of questions up front, and we ask some questions. And typically, clients [make] many requests as they go, shipment by shipment.

Then you have the question of, what do you do with it? A big piece of this is the human approach. Our stylists get to know our clients. We have tools where our stylists can see the history of all the people that they've worked with, all the feedback that they've given, all of the ratings. Within those tools, our stylists also get recommendations from our own internal systems about what our systems think might be great for that client. That's really where our machine learning and AI comes in.

We think of our tools on the machine learning and data science side as a great way to help our human stylists get in the ballpark. For any given customer, there might be thousands or tens of thousands of items that might, in principle, be appropriate for them. For the person working with a client and trying to serve them in a timely fashion, it's hard to go through every single thing in the inventory and think about what might be the perfect fit. So a lot of what we do is help our stylists narrow down, with data and algorithms, to a set of items that we think are pretty good, that respect the client's requests for both in [terms of] their style, and also for that particular Fix. If they're shopping for a particular occasion, our algorithms can interpret that from the request, and our stylists can see that.

Then the last mile is handled by our stylists, who know about fashion trends in a way that our algorithms still don't know. They know about the human, emotional connection that they've made with their client, the specifics of the occasion they might be asking for. They can really help them figure out: "Ah, this would be the best thing for you to try on at this time."

We have a lot of tools built on the data science side that we arm our stylists with to help them find the best assortment of things that they can send our clients to.

Sam Ransbotham: 85% of customers leaving feedback seems huge. That doesn't seem normal for feedback, but I guess that makes sense because you've got this situation where it's in the customer's interest to let you know as much as they can.

Jeff Cooper: That's exactly right. We are a retailer with a unique model, and that puts some constraints on us but also offers us a lot of power and a different kind of relationship that we can have with our clients. We think that direct relationship is the most important feature of our model, so we do a lot within our product and in our communications to clients to keep that feedback loop going.

We, as you said, have very, very, very high feedback rates. And again, these are even for things that people aren't keeping. Typically, if you're returning something to a big-box retailer, you're not necessarily going to leave detailed feedback if you're sending it back. But for us, our clients know, "Hey, that helps my stylist out. I'm working with a person here, and I'm working with a set of tools. If I tell them more about what worked or didn't work, then they learn about me faster." And that also helps our stylists and our tools evolve with our clients.

Something that you loved a couple years ago, or even a couple seasons ago, might not work for you anymore. Or you might feel like, "Hey, the trends in the place I'm working at have moved on," or, "I've started a new job. I want to try something new." Getting that feedback is a really great way for our clients to communicate with us and help us keep our understanding of their style really fresh.

Shervin Khodabandeh: Give us a sense of the scale here. You mentioned 3 million customers. How many items?

Jeff Cooper: We're a full-size, full-spectrum clothing and apparel retailer, for apparel, shoes, and accessories. We ship a couple hundred thousand Fixes a week, and we have a couple thousand stylists employed.

There are many moving parts to help this business scale from where it started with our founder originally just putting these Fixes together in her apartment. That's a big part of what our tools and automation are about — taking a model that is bespoke and human, and enabling this kind of connection. We're empowering the stylists so that they can scale that connection to many, many clients. It enables our business to scale across the millions of clients that we're helping to get dressed.

Shervin Khodabandeh: I have to imagine that generative AI must be pretty high on your radar, too, with its more cognitively advanced capabilities. Can you comment on that?

Jeff Cooper: Very much so. We're really excited about all of the new advancements over the last several years. One of the great things about having a great relationship with our customers, and [having] a lot of data about our clients, is that we can make that data even more valuable with technological advancements. The data we've collected one, two, three, four years ago becomes more and more valuable to us as new models and new kinds of machine learning and AI are developed. We can apply those tools to the data that we already have and help fine-tune those models with that data, [while] thinking about how to train new products on the data that we already have.

We have been excited about generative AI for some time. We started working on our outfit completion model for what we think of as a generative AI process: Our stylists were teaching a model on which items go together in order to help it build outfits in a fully automated way. If you go to our site and you've shopped with us, you'll see a popular feature called Complete Your Looks, which helps pick out items that you've kept, that we know you own and like, and pair them with other things that might be interesting to you. Our clients can shop for those themselves right on the site using a feature called Freestyle, or they can save them for their stylists to notice and talk to their stylists about. "Yeah, I loved the way that this looked," and so forth.

Creating new content using deep learning models based on top of our existing personalization engines was some of our early forays. We also, very early on, got excited about large language models. We've spent time with them for smaller-scale projects, for things like crafting dynamic ad copy or helping to improve our product description pages across our site.

More recently, we've made better use out of the new models. We have a really exciting new feature with our stylists using generative AI. Our stylists write personalized notes for each client every time they ship a Fix. And we've rolled out a new feature with OpenAI's GPT-4 that enables stylists to choose from a template. This is an optional tool where they can get some of the introductory, common-to-many-Fixes language out of the way [that also includes the data fed into the model] about what that customer likes and the items that are in that Fix, and so forth. [This is] a time-saver to enable our stylists to write those notes faster, with candidate language [they can use].

That's a great example of the kind of approach that we love in AI. We're taking something that is our human connection, and we're making it faster and easier and more scalable for our stylists. This has saved close to 20% of note-writing time for our stylists, which is a big savings at our scale, and our stylists have been really thrilled with how this feature has rolled out.

Shervin Khodabandeh: It allows them to focus on their strengths, which might not necessarily be note-writing but is much more in design and picking the right assortment and all that.

Jeff Cooper: Exactly.

Sam Ransbotham: This seems fundamentally different, Shervin, than many of the guests we've talked to. If I'm shopping for a battery, I know what battery I want; I just need to find it. So I need to communicate to the company what I want. But in this case, I don't know what I want. From Polanyi's Paradox, we know how to do things we can't explain. The example I always hear is pool. You can shoot pool without knowing trigonometry. Well, in this case, how can we tell these models how to behave if we don't ourselves know what we want or like?

You're in an interesting scenario here where you're in, like you said, a discovery relationship. People like me don't know what style we want, but I know what I hate when I see it. That seems really a fundamentally different way of working.

Shervin Khodabandeh: It's more open-ended.

Sam Ransbotham: Yes, it's much more open-ended versus destination-oriented.

Shervin Khodabandeh: I would say that, in my view, isn't this a design problem where you don't know what you're designing, exactly what shape it should be? It could be automotive, it could be art, but there are a lot of parameters and boundary conditions. You have choices. It's not a risk problem like, "This is a fraudulent transaction. Don't authorize it," or, "This is the right offer for this customer at this moment. Send him this out of these three promotions." This is different because it's so open-ended. Maybe there is not just a global optimum [choice]; maybe there are many.

You were talking about an ongoing relationship. If I'm in a relationship with Sam, I'm not trying to optimize every single interaction. I'm just trying to have a good relationship.

Jeff Cooper: What you said, Shervin, resonates so much about this being something where we don't know what the *right* — the quote/unquote "right" — end goal should be, and it's quite difficult to design. On the data science team, we think a lot about needing an objective function on these models. What does it need to be? We have a lot of debates on the team about exactly how to model our client happiness and satisfaction in a way that the models can steer in the right direction.

One of the reasons we're so passionate about this combination of humans and ML is that, first of all, it enables us to solve some of those thorny problems by saying, "Well, the humans will do some piece of the company's objective function, and the models will do some piece of the company's objective function. And both of them will contribute the things that they're best at so that we can help make our overall client outcomes the best."

A really interesting thing, when thinking about design space and how machine learning models can help with these fundamentally creative problems, is we see both patterns within our usage where the models are helping our stylists get in the ballpark. Then our stylists narrow it down and find the last mile. But we also see patterns of the other kind, where our stylists are fundamentally describing some core constraints, and then our models are nailing down exactly where they want to land.

Our outfit model is a good example. We spent a lot of time with our stylists helping them train the model. A lot of the training is about building guardrails into that model that say, "You're never going to have this kind of pants go with this kind of jacket. These are pajama pants. They cannot go with a nice blouse" — these are fundamental guardrails, in hard business logic but also in repeated training and helping the model understand the core concepts.

In many parts of this creative process, there are places for both the machine to provide the core search space that people work within and the humans to set out the core search space that the models are then working in. Which one you use depends a lot on the specifics of the product feature that you're trying to design and the scale. For something like our outfit model, we're trying to create tens of millions of outfits a day for our clients. We cannot have human beings put all of those together every time. For our Fixes, we must have our stylists be really involved in that process because that's one of our core promises.

Depending on the kind of feature, the kind of scale you're working with, there is a spectrum of possible interactions between the human and the AI model that can help the company produce the best outcome.

Shervin Khodabandeh: That makes a lot of sense. When you make the comment of, "This is a pajama top that doesn't go with this" — it might not now but could at some point. Right?

Jeff Cooper: Exactly.

Shervin Khodabandeh: It seems like it's an ongoing dialectic, maybe a trialectic, between the stylists, the machine, and the customer.

Jeff Cooper: If you want to make it a little more complex — as we love to do in data — it's really a four-point problem where the fourth is wider fashion trends, exactly to your point. We have this ongoing evolution of what our customers are seeing out in the market, what they're seeing out in fashion, out in the world, and what our stylists are seeing as up-and-coming trends that our customers might not be aware of — or might be aware of but don't think they are right for. Our stylists can see something within our clients and say, "Actually, I think you would look great in this." Our models can help pick up some of those trends in the data among other, similar customers. So it does end up being this really interesting set of conversations between all of those points.

Sam Ransbotham: When you were talking about "pajamas don't fit with this," even I know not to wear dark socks with sandals, but it seems it must have been frustrating for your stylists to have to teach a model all those things that we take for granted. But you started with an individual human.

Shervin Khodabandeh: Do you wear light socks, just so we know?

Sam Ransbotham: Wait a second, you're telling me there's no right sock choice for a sandal? My fashion world is thrown asunder here.

Jeff Cooper: It's all about confidence. If you know what you're trying to go for, Sam, you can wear it.

Sam Ransbotham: You should follow me to set the new trends. You started off with a very human world, and now you're in a very augmented world. Over the course of these 13 years, it seems like there must have been some frustrating toddler years where your stylists have to be saying, "I cannot believe this stupid model put this together like this." How did you work through that?

Jeff Cooper: I wasn't here at the very beginning of Stitch Fix, but I've heard about plenty of stories in developing the models. You get started where you can. The process of getting stylists comfortable with the scores that our models are producing is an ongoing one that we're always talking about. In the early [days], you had basic recommendation models, and even 10, 12 years ago, people had a good sense of a simple scoring system: It says these other clients who have bought similar things might also be interested in this kind of thing. That is going to help you narrow down to a set of items that might be useful. That's something that any stylist, really anybody working in retail can understand. We've just layered on improvements and complexity since then, working really closely with our stylists where they request a lot of features, both changes to the model or additional information that might be helpful to them.

One of the spaces we've been working hard on and considering where it might be useful involves customers who have been with us for dozens and dozens of Fixes. To have a stylist come in and look at all of the feedback they've given over years potentially can be really complicated. With our new generative tools, we have the possibility of creating summaries of those things and compressing some of that information a bit further. In this case, it's almost like you have a stylist working alongside a partner that can help do some of the extra work.

How we think about talking to our stylists about this score is a really complicated problem for any human in a loop kind of system. It's not one that we've solved. We do a lot of training with our stylists. A really big advance for us in the last couple of years was moving to a single unified recommendation model. One of the toddler steps we took [moved us from] "Here's a machine learning model for women only for Fixes. Here's a machine learning model for women only for the Freestyle portion of the site, the kind of clients shopping on their own. Here's another different model only for men," to having even several different models that might be used at different points in the Fix journey.

A big advance was to help bring all of those models together into a single centralized place, where we can gather all of the information about all of our clients, and now, day-to-day, client to client, stylists can feel like, "OK, this model always knows all of the information about the client," as opposed to, "Oh, when they're shopping over here on this part of a site, it doesn't know things that I, as the stylist, know that this customer has bought."

Sam Ransbotham: That seems really easy to say, but really hard to do.

Jeff Cooper: A lot of it comes down to this explainability question: This interaction that we have between stylists, customers, and models — to take the social portion out of it for a minute — any machine learning system has to face the question of explainability of the people that are using it or getting outputs from it often need to understand something about why these things were generated. That's a hard enough problem to solve just when you're talking directly to a customer. If I look at my recommendations on another retail site, I might be like, "What? Why is this being recommended to me? I don't quite understand." Many different people have tried to solve this problem in different ways.

We have the additional complexity of our stylists needing to understand where these recommendations come from, and our stylists needing to explain those recommendations to our clients. So, we need to find ways for our stylists to have a sense of the model's thought process, in some sense, and then for them to also be able to explain why these things might have, we think as a kind of human plus model combo, been a particularly good choice for our clients?

That's something that, it was a great example, we think our stylists are still really, really, really expert at. It's quite difficult to beat, even with advanced language models, the power of a person who knows their domain well and can talk through why for you, as an individual, this piece might be the best.

Shervin Khodabandeh: We talked about the wide gamut of items and customers and data, and you have, you said, several thousand stylists. How is AI helping them learn from each other? When you were talking about, "our stylists," I'm thinking it's not a homogeneous group of people.

Jeff Cooper: That's right.

Shervin Khodabandeh: They have different tastes and they could learn from each other, or they could challenge each other. How are you doing that?

Jeff Cooper: We trained them on the latest and greatest for our machine learning models and our tools; what are the things to be aware of this season or as new merchandise and apparel rolls in for the current new month? Much of that training is done at the human level to help them understand, "Here's the things to look out for, here are the things that are going to seem new." A great example here was the rollout of this generative AI note-writing template, where the training varied a good bit depending on: Are you someone who has been writing those for many, many years on your own? Are you someone who has seen other attempts that we've made to do note-writing tools, or are you coming into this fresh?

Our research suggests that our clients are looking for more interaction with our stylists as humans. We think that's the really exciting next frontier, to help our clients understand, from the very beginning, that we can talk to them about why this stylist has been paired with them. Any style expert is going to have a sense of the customer that they really resonate with, the fashion trends that they really resonate with. We have all of that information, and that's a really exciting area that we're thinking about finding ways to surface better to our clients.

Shervin Khodabandeh: Tell us how you ended up where you are. What was the journey like?

Sam Ransbotham: I don't see a lot of cognitive neuroscience in this so far. Is it there?

Jeff Cooper: We have a wonderful team of people, many of whom came from scientific backgrounds. I'm sure, as many of your guests have talked to you about, a background in academic science ends up being a wonderful set of experiences to learn about how to interact with real data.

We have a bit of a running gag at Stitch Fix. Our people with social science backgrounds — like myself coming from psychology, other people who have come from perhaps economics or other social science backgrounds, our partners who have physical science backgrounds — you get somewhat different exposure working with data. If you come from astronomy or geology or chemistry, you might have a sense of how you expect data to behave. Stars, they are a little bit different from each other, but they —

Sam Ransbotham: They follow some rules.

Jeff Cooper: Yes, they follow rules.

Sam Ransbotham: Humans don't.

Jeff Cooper: If you spent your academic background cutting your teeth on working with college undergrads or little kids or even grownups, you understand variability in data at a more visceral level than you might otherwise.

I got into data science in part because I'm interested in people, I'm interested in human behavior. I just think people are the most interesting, complicated things in the world. That's why I got into psychology in the first place. And data science ends up being the field where there is the most data about what people do. A lot of what I think about day-to-day still really resembles thinking about our core theories about decision-making that I was doing back in grad school.

You were saying earlier, Sam, that it can be hard to figure out what the objective function is for fashion or for an outfit to put together. If you've spent time in psychology, that's all you used to think about, those kinds of problems.

Everything you would do is trying to take this messy, amorphous, human concept and turn it into some kind of mathematical model just to be able to measure it and quantify it.

You have to get really comfortable in data science when you're working with real customers, especially in businesses where you are working directly with customers who are not going to do exactly what you think they're going to do and are not looking for exactly what you think they're looking for. You have to get comfortable with the idea that you need to take something very squishy, something difficult to render into numbers, and find a way to turn it into numbers so that you can measure it.

Data science is a wonderful field for the art of using math, using statistical modeling, using high-precision computing tools to actually say something interesting about these really complicated, hard-to-predict things about how we feel and decide and how we feel. They seem very difficult to put into numbers, but when you put them into numbers, you can often learn a lot.

Shervin Khodabandeh: That's a great answer. It opens our minds that data science doesn't have exactness necessarily. In fact, the sort of amorphousness that you're talking about and the spectrum of possible, very good solutions to something, and it's a great tool for open-ended problems, which are actually, all of human problems are open-ended that way.

Jeff Cooper: Statistics and machine learning are both fields fundamentally about dealing with variability in data. They are about dealing with problems where you do the same thing and something different happens, and nothing is better for variability in data than fashion. Two people look at the same thing and one of them thinks, "That is amazing," and one of them thinks, "Eh, not for me." Being able to take that very human variability and turn it into something that you can approach with numbers — try and make some predictions about, try and summarize at scale — is an incredibly fascinating problem.

Sam Ransbotham: Shervin, are you five-questioning?

Shervin Khodabandeh: Jeff, do you know about the five questions?

Jeff Cooper: I don't think I do.

Shervin Khodabandeh: That's wonderful. That's supposed to be the way.

Sam Ransbotham: Shervin likes when people get blindsided.

Shervin Khodabandeh: We have a short segment where we ask you five random questions.

Jeff Cooper: I like that you've described this as a verb: "Are you five-questioning?"

Shervin Khodabandeh: Tell us the first thing that comes to your mind. What do you see as the biggest opportunities for AI right now?

Jeff Cooper: How to get it to work with humans.

Shervin Khodabandeh: Wonderful. What is the biggest misconception about AI? What do people get wrong?

Jeff Cooper: That it's smarter than humans.

Shervin Khodabandeh: What was the first career you wanted? What did you want to be when you grew up?

Jeff Cooper: An astronaut.

Shervin Khodabandeh: When is there too much AI?

Jeff Cooper: When it doesn't leave space for people.

Shervin Khodabandeh: What is the one thing you wish AI could do right now that it can't?

Jeff Cooper: Operate in the physical world more. We're very excited about — when I say "we," I mean those of us in the data science and our AI community — all we can do with information and language. It's fantastic, super important. There are so many more opportunities to help people when we think about not just robots but automation and physical automation starting to be more linked to these more cognitively powerful models that we can interact with in a more human way.

We think a lot of what has been so interesting to everybody about the large language model moment and the big advances is that it offers the opportunity to interact with these automated systems much more like you would interact with a person. That's what people want to do. And so unlocking the ability for us to interact with automated systems that can operate in the physical world more in a more human way, I think is an area I'm really excited about.

Sam Ransbotham: It makes a lot of sense. Jeff, we appreciate the time you spent with us.

Jeff Cooper: It's been so fun.

Sam Ransbotham: It's really fascinating how you're using AI to help your stylists and your customers learn more about themselves. In this case, everything you've mentioned has been learning — bidirectional learning, even. I hadn't appreciated that symbiotic relationship before today. Your models are exploring a space and your stylists are helping the models explore the space. That feedback and loop seems really important.

I also hadn't appreciated the complexity. These things always sound so simple: "Oh, yeah. Use some AI to solve this problem," but the devil's in the details. In this case, the devil doesn't wear Prada; the devil wears silicon.

Jeff Cooper: That was very good.

Sam Ransbotham: I like the phrase about getting started where you can. I think that's a good phrase — Shervin, it's one we can pick up on. You get started where you can. Thanks so much for talking with us today, Jeff.

Jeff Cooper: It's been so great. Thanks for inviting me.

Shervin Khodabandeh: The devil wears silicon. That was really good. Did you just come up with that right now, or did you work on that?

Sam Ransbotham: No, I wish.

Shervin Khodabandeh: That was very good.

Sam Ransbotham: Thanks for joining us today. Next time, Shervin and I talk with an AI startup founder who starred in a recent Spielberg film. Get ready, Player One, grab your popcorn, and tune in in two weeks.

Allison Ryder: Thanks for listening to *Me, Myself, and AI*. We believe, like you, that the conversation about AI implementation doesn't start and stop with this podcast. That's why we've created a group on LinkedIn specifically for listeners like you. It's called AI for Leaders, and if you join us, you can chat with show creators and hosts, ask your own questions, share your insights, and gain access to valuable resources about AI implementation from *MIT SMR* and BCG. You can access it by visiting mitsmr.com/AIforLeaders. We'll put that link in the show notes, and we hope to see you there.

BCG + MIT Sloan Management Review