Freddie Mac Single-Family Home Starts Here Podcast Episode Transcript:

Makes MEDA Sense

Announcer [00:00:01] Welcome to the Freddie Mac Single-Family "Home Starts Here Podcast," your connection to all the latest industry trends, insights and points of view on the mortgage market from Freddie Mac leaders and other industry experts. We'll be coming at you from the 2021 6th annual Freddie Mac Connect conference, where you can catch exclusive episodes throughout the event.

Michael Bradley [00:00:21] Hello, everyone, and welcome to "The Home Starts Here Podcast." It's my pleasure today to be hosting a discussion about artificial intelligence or A.I. with Frank Poiesz, Managing Director of Product Origination Technology at Black Knight, and Amy Gromowski Senior Leader Science and Analytics at CoreLogic. Frank and Amy are both working on developing and implementing A.I. Solutions for their clients. Welcome to you both.

Frank Poiesz [00:00:48] Hello, Michael.

Amy Gromowski [00:00:49] Thank you.

Michael Bradley [00:00:50] Today, we're going to discuss how lenders can leverage A.I. and what they need to do to mitigate the heightened risk posed by some of these solutions. Let me kick things off by giving you four reasons why you need to adopt A.I. First, A.I. models are just a whole lot more predictive than traditional statistical models, and they can make sense of a kind of data called "unstructured data." Think of text, voice, images, information from those forms. Second, other companies are already using A.I. to disrupt your business. And like them, you need to use A.I. to gain process efficiencies. Third, humans predictably make mistakes and A.I. can help us all work smarter. And finally, A.I. can help us identify new products and solutions for our clients. Now A.I. methods do elevate certain risks that need to be managed. One way of addressing these risks is by implementing strong governance. At Freddie Mac, for example, we have a corporate standard and ethical principles that must be followed when using A.I. And while there's a lot of buzz about A.I., it's essential that adoption be pointed squarely at real business problems. With that backdrop, let me now pose a couple of questions to our guests so that we can learn from their experiences. Amy and Frank, definitions, taxonomy and governance are all important aspects of developing solutions and mitigating A.I. risks. But what are the other considerations that clients need to consider? Amy, let's start with you.

Amy Gromowski [00:02:35] So clients need to consider when looking at different A.I. solutions, specifically around machine learning, quality of data, ground truth, the study design. These are all really important. So we hear about 'garbage in, garbage out' and what kind of data that we're using. In order to really evaluate the quality of a solution you need to look at what is available in order to be able to create the solution. So being able to know from the history what has happened or if you look at things like computer vision and extracting from imagery, that's a really hard problem to solve, not just because of the technology and the technique that's required from machine learning, but also just having the assignment of features and characteristics on the imagery to be able to create those algorithms. So that's what we would call "the ground truth" or knowing past behaviors to predict future behaviors. These are all really important considerations when you're looking at problems to solve. Do you have that data available? And then it's really about can I understand how the data is flowing? How would I be implementing this solution? And do I have the data that I need available in production? So, for example, market value, AVMs, this is really important in the lending and servicing space. We need to be able to understand that. We need to be able to understand the market value in a very timely fashion, let's say, daily. So to successfully be able to create these algorithms and then generate this market value daily, you need to be able to have your features in the model and have the data available in real time, pretty much and available to be able to calculate that

before and capture what's happening before the AVM is generated or calculated. So that's one of the considerations there. So that's all around the data science and the quality of that. The other really important factor is the business owner. Does the solution meet the business need and how do you evaluate that? Is it really solving the problem that's being asked of the model? And so working really closely with the business, with the customer to be able to get comfortable with how that model is designed and what problem it's solving for and is it reliable? These are all important considerations when you're developing these solutions.

Michael Bradley [00:05:25] Well thanks Amy, really great about the quality data and the solutions pointed towards a real business need. We'll talk about that a little bit later, but really good insights. Frank, how about you for the same question?

Frank Poiesz [00:05:40] So for lenders who are looking to adopt this kind of table stakes, that lenders need to consider in just looking at the auction from square one. And by that, I mean, what exactly should I apply A.I. machine learning to? Michael, you mentioned a couple of problems that that you look at in terms of doing predictive models and Amy talked a lot about operational solutions. In all of those cases, there's a clear distinction between A.I. and machine learning and traditional software that lenders need to understand. So traditional software is always relies on conditional logic - if a thing is true, then do this. A.I. machine learning is all about probabilities. Basically, given a bunch of data, the A.I. has to make a prediction about some desired outcome. That requires a different mindset for a lender when they think about how to use A.I. You can't think of things in terms of absolutes, you have to look at things in terms of the gray. What is the possibility that I'm going to get a good outcome and how do I measure the veracity of the data that I'm getting from an A.I.? So at the same time, A.I. is incredibly useful on problems that include a lot of data, as Amy mentioned, a lot of data and data that's highly variable. So let's see, a lot of data and highly variable, that sounds like the mortgage industry generally. Everybody deals with that every day. So no, we don't think that A.I. and machine learning are a solution for, say, underwriting loans, which has tons and tons of rules associated with it today - though, in the long run, who knows what our friends at Freddie Mac will invent for us. But in the short run, looking at things like that document problem is important. So to expand on what Amy said, if you think about a document, say consider a bank statement, if a bank statement comes to your shop, some human has to look through it, review transactions, do some keyboarding, hope they get the keyboarding right and then use the data from that document in a way that's productive for the underwriting. The way computer vision models work is they look at that document and say, 'OK, here's a bunch of data. I know I'm looking for a table. What characteristics do tables have and how do I try to identify them?' And then it finds a table and says, 'yeah, I think' and yes, I'm generalizing but the model basically says what the probability is that it found what you're looking for and then you use those probabilities to understand what data you should accept and what data you should not. So what lenders should be thinking about is what is that a problem they're looking at yields only to this kind of solution. That's to say no, 'if then' logic really won't solve it, because up to the minute I need to look at my pipeline data to try to optimize where I send my pipeline. And I just can't do that with 'if then else' logic because too many variables are involved, there's too much data. So they're the kind of solutions that the industry is focusing on, things that yield well to probability-based solutions.

Michael Bradley [00:09:00] Neat, thanks Frank. In our terminology, we would say that you have some sort of objective function that you're trying to solve or you think you're trying to get to the top of the mountain. But there's different roads to get to the top of the mountain and you have to be in a position to decide among those different roads, which might be the best one to meet the business need. So it's a really nice summary Frank, thank you. Now let's transition to the so-called black box problem that plagues A.I. solutions. Clients need to understand and trust the models if they're going to implement them. Our experience has been if the business clients don't understand it, they're not going to implement. So, Frank, can you speak to the importance of explainability in building A.I. solutions?

Frank Poiesz [00:09:48] Sure. So if you think about those problems that Amy and I both discussed, you really can't devise a test that gives you one hundred percent solution, because with traditional software, you can say 'write a test case that says if this happens, this is the outcome.' In the case of A.I., the test is going to be 'I've got a hundred things coming in, how many of them were right? Or a similar probabilistic answer to the question. So talking with the vendors about explainability is a critical part of understanding the solutions that you should be applying to your business. What's the difference between testing and explainability? Well, it's just that. Testing gives you a totally Boolean yes or no, you passed the test. Explainability is more looking at and making a judgment about whether the information provided by the vendor, about how the A.I. is constructed, where it gets its data, how closely the data is guarded and curated, and then how the

A.I. comes up with its answers, needs to be coupled with looking at output and say, 'OK, the output is it looks like it's correct most of the time.' So if, say, for example, you were going to evaluate of a voice to text solution, you would certainly expect some transcription errors from that voice to tech solution. So make sure you have a perspective on how much proof you need for that software to be explainable as the solution at the level of quality that you require.

Michael Bradley [00:11:28] Great. Yeah. So, for example, in our space, if you're building a default kind of model and you need the businesspersons expecting is LTV goes up, the probability of default would go up. As FICO score goes up, the probability of default would go down. And if they don't see those kinds of relationships mapped out from the data to the output, then they're not comfortable at all with the model and they're not going to buy into it. So you have to figure out ways to make it clear to those business partners to lower their blood pressure so that they'll be willing to go move forward.

Amy Gromowski [00:12:04] Yeah, I completely agree with all of that. I think it's important for the end users, lenders, servicers, myself, where we're more in the vendor space. You have real things around legal and regulation and customer inquiries and so questions that need to be answered and you need to be comfortable, you know, with what's in the model and how does it perform in certain areas, in certain ways; geographical representations might be one. You know a statistician might argue in its purest form that explainability, as long as it's predictable, really isn't important, right? If you can prove that, you know, over time out of sample that the model holds true and it performs well, that you can't pick apart something that has hundreds of variables going in and say, I can explain every single piece of it. But in the data science world, it is important. It is important to be able to say, in general, these sorts of variables behave this way and you can be comfortable with that. Let's say, you know, that there isn't any disparate impact happening. That is really important in in the real world, right? We're not just an exercise theoretically. So being able to break out those data layers and explain here's what's going into the model while still protecting IP is very important. And I found, you know, being able to do that and just give some level of visibility into here's what's going in, here are some of the general behaviors in the model and that they make sense, such as a habitual refinancer, is more likely to refinance in the future. These are just common things that you can see and understand right from just real-world experience. And then being able to represent that geographically, visualizations really help as well. So completely agree, explainability is important and, you know, if you are looking and working with a partner in the data science world, it's important that you're comfortable and that your data science provider and solution provider is open to those discussions and sharing some level of information.

Michael Bradley [00:14:22] Well, that's great Amy, thanks. And, certainly, at Freddie we're steeped in fair lending analytics and many lenders don't understand that fair lending actually applies from the time you start to solicit for your loans all the way to the property gets disposed of. So fair lending is definitely an important topic and that's one of the things, in addition to our business partners, we need to be able to explain the results to our legal compliance people that oversee us from a fair lending aspect as well. It's a great point. That's amazing

Frank Poiesz [00:15:00] One more perspective. So I often think about how my parents taught me to cross the street. They always said, "look both ways." So look upstream when you talk with a provider at how they gather the data and how they manage the data to be sure it's consistent and refreshed or how they manage the whole pipeline for data to come through and turn into the A.I. solution. And then look downstream by making sure that there is a set of quality control checks. So this is a lot like lending quality control. You should have tests that say, “did the outcome", and Amy mentioned this and so did Michael, "did the outcome look like what I expected and how often did it not look like what I expected?" And when that frequency of false positives or bad answers gets too high you want to talk to the vendor, but don't expect a hundred percent.

Michael Bradley [00:15:59] Right, ok great. Good advice. Now there is the idea of going from concept to production can be pretty challenging with A.I. solutions. So, Amy, if lenders want to advance and scale A.I., what besides governance and expertise do they need to think about?

Amy Gromowski [00:16:17] Yeah, I would say, you know, A.I. can be big? And it means a lot of different things actually to different people. My advice is start small and start small, but plan for the future. So understanding all of the different use cases and applications of A.I. Looking from a business perspective what's going to be, you know, the best value, if you

will, for your money. So it may not be the biggest impact, but if you think about it's small, it allows us to learn and iterate quickly, so we can get better, you might want to start there and while still delivering some business value. So at CoreLogic, if I think about when we started computer vision, we have all this great imagery that's available to us. We understood there was a lot of value in the market and being able to extract content from this imagery. But we also wanted to make sure that we started small and we started with just some proof of concept. What would it take? Do we have the ground truth needed to do this? And if we don't, how do we go after that in a cost-efficient way? These are all things we want to look at and then you want to prove how can I actually do this and work with your clients and your customers to say what is the value of this to you? And then as you start to move into more of a, you know, deployment, full scale operationalizing these models, then you can look at a full text stack? What would future state look like, let's say, in four years? Because A.I. is expensive, the technology is expensive, and it's also always changing and there's always new options available. So you want to have some sort of plan in place around what the future will look like, but have that plan be flexible as technology evolves and changes. And. Also, that you can thin slice, I call it like 'thin slice out' just portions of it that you can tackle in lower costs while consistently delivering value to your business and to your customers ultimately and staying competitive and being able to really fully see the advantage of A.I. without getting into a really big, costly project that can weigh you down and ultimately hold you back, ironically, which is the opposite of what you're trying to do. So I would say start with that. There's additional things just around, you know, what kind of performance you see from your models and how you continue to iterate. But the biggest thing I would say is start small, plan for the future, and incrementally iterate through.

Michael Bradley [00:18:58] Great. Frank, did you have anything to add to that?

Frank Poiesz [00:19:01] Yeah. And in fact, I'll reinforce what Amy said in her opening remarks. There's a step that lenders can take today and every day going forward to be prepared for a future that is more A.I., machine learning driven and has more ability to scale. And that is, be really, really good at data. So making sure that, you know, what data that you have is good and you know where it is and you know how to make sure you safeguard the quality of your data for the foreseeable future. More data is always good to data scientists since you don't know where the next application of A.I. will be. Chances are good you've already got data that will contribute to the success of a future AI project. So be really strong data stewards and you're going to be more successful in the business today because God knows, Freddie Mac wants clean data from you, but also remember that it's going to serve your purposes when these kinds of applications become more prevalent.

Michael Bradley [00:20:05] Amy and Frank, do you have any closing thoughts?

Frank Poiesz [00:20:08] I guess my closing thoughts are that we have to continue to be vigilant in our regulated environment to meet the standards that Michael talked about earlier. We have to be careful about how we apply any new technology and this technology has the ability to do things that have never been done before, whether good or not. So, yes, we have to be vigilant, but this industry has a tremendous opportunity in A.I. So we shouldn't be daunted by the need to be good at compliance. If we were daunted by compliance, we wouldn't be mortgage bankers. So yes, be very focused on it, on not only compliance, but making sure you protect your own processes so that they are positively not negatively impacted by A.I., but don't run for the hills, Be engaged with the market for A.I. solutions. Give some of them a try and I think that in the long run, everybody will be better for it.

Michael Bradley [00:21:07] Great, thanks. Amy?

Amy Gromowski [00:21:09] Yeah, that's great. I completely agree with Frank. I would also just expand on that. I think there isn't one size fits all when it comes to, you know, A.I. solutions - different providers or even just internally you'll be able to do different things. It kind of gets back to, you know, the data and the quality of the data and the depth and the breadth of the data and the different problems that that data can solve. You know, I think there's a lot of different expertise levels and types of solutions that you can leverage to bring into your organization or build internally. And those things will evolve and change over time. So continuing to understand what that market landscape looks like, the value that it brings to your business and to your customers, and being in a position to pivot, as needed, to expand and to grow the capability

internally as well is really important. And then putting the right framework around all of that, how to assess and continually assess those providers or what you're doing internally when it makes sense to pivot and all of the governance around the quality of the models and meeting the expectations of regulators and all of that, just that framework then is important and you'll be able to grow as the market and technology grows and changes.

Michael Bradley [00:22:33] Well, thanks again, Amy and Frank, for a great discussion, sharing your experience around A.I. with the audience. I'm sure they really enjoyed it. Thank you.

Amy Gromowski [00:22:42] Thanks, Michael.

Frank Poiesz [00:22:43] Thanks, Michael.

Announcer [00:22:44] Thank you for listening to this special edition, 2021 Freddie Mac Connect Conference episode of "The Home Starts Here Podcast." After the conference stay tuned and subscribe to catch exclusive interviews with key industry leaders and experts. Home Starts Here is available wherever you listen to your podcasts. We appreciate you rating, reviewing and sharing with your network.