# Foresight: The CPA Podcast

## Season 6 Episode 4: Can AI make sustainability reporting, sustainable?

**Neil Morrison:** Welcome to Foresight, the CPA podcast. I'm Neil Morrison. Back on our first episode this season, Zohaib Akhtar said something that really caught our attention. The conversation was about how AI will change the role of the CFO. And one example Zohaib gave was around sustainability reporting.

**Zohaib Akhtar:** So, when you look at environmental sustainability, essentially what they would do is they would use generative AI to understand the regulations, make it more, I would say, reachable across the organization, bring about more understanding.

**Neil Morrison:** Zohaib went on to talk about how AI could maybe then suggest solutions. So, for example, ways to cut down on transport costs. Now, this was just a prediction. He was giving an example of maybe how AI could be used, but we wanted to know if AI is, in fact, being used this way. And the answer is yes, sort of. Manifest Climate offers an AI-powered solution that helps companies compare their performance on sustainability measures against the various regulatory requirements. It also helps compare the company's performance on these measures to other similar companies. Ryan Myers is the Chief Product Officer at Manifest Climate. He's been working in the climate change space for over a decade, and the transformation he's seen in a short period of time is remarkable. These days, he's bringing the latest AI to these challenges, but it wasn't so long ago that all he was bringing to these companies was basic awareness of the issue.

**Ryan Myers:** At the time, I was coming out of school, and I was like, okay, sustainability is a big buzzword. Everyone's using that term, so it must be relatively sophisticated in the way it's used. And I remember getting on a webinar, this is where I was pretty surprised, getting on a webinar, and it was for sustainability professionals, by sustainability professionals. And at one point, the guy who was leading it said, "Now, I know it might feel like decision-makers don't appreciate the importance that you're trying to put on these ideas, on these different CSR or sustainability-related initiatives, but it may not be next week or next month, but at some point, they'll understand the value of this and come around and start executing on some of these decisions." And that was the moment where I was like, wow, that is the advice to a hundred sustainability professionals across the continent. "They may not get it, but just keep pushing because someday they'll get it."

**Neil Morrison:** Yeah, that's where it was 12 years ago. How much more developed is the space now?

**Ryan Myers:** As you can imagine, it's changed a ton. I think actually COVID also helped with that a little bit because it made people realize that these risks for things are really real, and we are very connected as a planet. So, what happens in China affects us in Canada and affects the states. And everyone kind of knew that, but it just made it more tangible. And so, the risks are much more real than they were before. The narrative has shifted from it's coming to it's here. And so, I think that it does present material risks for investors. And so, that's why you're seeing more regulation coming out that's mandating disclosure of climate risk management.

**Neil Morrison:** What are some of the challenges that small and medium-sized businesses face when they're trying to figure out whether there's alignment with ISSB requirements?

**Ryan Meyers:** It just comes down to resources. So, a lot of small and medium-sized businesses don't have people that are dedicated to focusing on sustainability or focusing on climate risk management. And so, when a new regulation like IFRS 2, ISSB, or even the ones that the SEC is talking about, or there's regulation that's rolled out in the UK and New Zealand, when these things come out, I think a lot of companies start with, "We have no idea even how to approach this. We don't know where we're at. We don't know where to start." It just takes time, money, resources that a lot of small and medium-sized businesses, and to be honest, even larger businesses also, a lot of them don't have. So, I think that's the main problem.

**Neil Morrison:** Yeah, which is where Manifest Climate comes in. You have an AI solution that analyzes gaps between the company's performance and all of these various requirements. I was wondering, if I'm a CPA using the solution, what does it look like?

**Ryan Meyers:** So, we have an AI that underpins our software that looks at 50 different data points aligned to different regulatory standards. So, each data point maps to one or more of these regulatory standards. And then we have the analysis completely automated with the use of AI. So, that means that as a user, you can basically give us a few of your documents and within hours, or technically minutes, but we say hours, you can have all your files analyzed, and we can benchmark you against all of these different disclosure standards across these jurisdictions. So, in terms of output, what that means is that when you log into the software, we map your disclosure, or lack thereof, against whatever regulatory standard you care about. So, if it's IFRSS 2, then you click that, and we've basically got all the different sections of IFRSS 2, and then whether we found information from your disclosures or not, and if we found it, where we found it.

**Neil Morrison:** And how ou're doing relative to that?

**Ryan Meyers:** Yeah, yeah, exactly. Yeah.

**Neil Morrison:** And does this give you a sense of whether you're pretty close to being compliant, you're really far from being compliant? Does it give you a green, yellow, or red light? How does it let you know how you're doing?

**Ryan Meyers:** Yeah, so right now, this is a question that we have sometimes as well, of how much is enough, but we don't say that we're a compliance tool, and we want to be very specific with our language. So, we don't use the word "compliance." We talk about getting you an understanding of your alignment against these different standards. Partly because with 50 data points, we give you a pretty decent idea of how you're doing A, compared to a standard, especially if you're starting out, and B, compared to over a thousand other companies. So, it's easy to see for companies like you or your average in your sector, where you are falling behind and where you are ahead. But this kind of brings us to how deep can AI go and what is the role of AI relative to people? And at least with our software, we want to give you A, an understanding of how you compare against the standards, how you compare against your peers, and then what you should do next, and what are examples of what scenario analysis looks like, and what your peers are doing related to that. But we don't go down to the specific line item in the standards. We still need people to basically take the output and then go into those sections and go, "Okay, for this one, we're doing this or not doing it," and take it from there.

**Neil Morrison:** But it sounds like it's giving you a sense of maybe things that you could be doing. It would be the starting point for conversations about, "Okay, we've got some serious gaps here. How do we approach this?" It can really accelerate that conversation.

**Ryan Meyers:** Totally. And I think a big part of it as well is companies come to us and they say, or our users actually say, "Okay, we use it to conduct the gap assessment, but then we need to demonstrate internally why we're behind or why us being behind is important and how we can fill those gaps." And what we get with the software is specific examples of what our peers are, not just whether they're doing scenario analysis, but the details around what they're doing related to scenario analysis. So, our base AI extracts 50 data points, but we actually have a secondary level of AI that uses LLMs, which I'm sure we're going to talk about in a moment, where we extract a whole bunch more data points. So, we essentially cover 50, you can think of it as 50 mutually exclusive areas across these regulatory standards. And then we use LLMs to add some depth to each of those 50 areas, so we can extract more data points. And you can understand, "Okay, not just is there coverage here, but what are people doing exactly, and what should we be doing at our company?"

**Neil Morrison:** Okay, let's get to the large language model part of this. So, this is where you're getting some depth to it. You've got all the numbers, and now it's giving you some feedback, some more in-depth feedback. In some ways, it's now doing more of the consulting side of things, I'm guessing, is what that is.

**Ryan Meyers:** Yes and no. It's important to understand that LLMs are good at some things and they're bad at some things. And I'll give you an example. They're really good at summarization. They can be okay at extracting very unambiguous data points, very clear data points, but they can be unreliable at best at making judgment calls. And so the problem with getting them to go a step further in terms of looking at regulatory compliance at the line item level is that sometimes there is a judgment call. Sometimes there's a topic that is a bit squishy that requires expertise, and LLMs are not very good at that. So that's why, just, they can't entirely do the role of what a consultant necessarily could do, or I should say what a human could do. If you have somebody like a CPA who's going down to the line item level, because it doesn't have great reliable judgment.

**Neil Morrison:** So, it can summarize; it's very good at summarizing all of the massive data that it's analyzing, and it can point out glaring, perfectly obvious things, but it does it faster than you having to go through it yourself. But when it comes to judgment calls and things like that, at the end of the day, that requires a human, a CPA ideally, to be looking at it and coming up with a plan.

**Ryan Meyers:** Yeah, and that's why I was implying earlier that we've got two levels to our AI, the level of depth where we're extracting very specific data points. We use an LLM; we also use an LM to do summarization, but the base level of our AI that we use to extract those initial 50 data points for coverage against the regulatory standard is not an LLM. We trained a more traditional, what's called natural language processing or NLP algorithm, that was based on hundreds of thousands of data points from our team. So, we use a base AI that is expert-driven, which is also, I think, where humans still come into the picture for some things.

**Neil Morrison:** So, I think you've actually answered the question. One of the problems with LLMs and ChatGPT is it tends to hallucinate; it comes up with ideas that are not there. I know this, even when I get it to summarize an interview that I've done, one of the things I often get, I find this fascinating, I ask it to give me a really great quote from the interview that I've done, and it just makes up quotes, and I'll ask it, "Is that made up?" And it says, "Yeah, sorry, I made that up." So, is that a problem? Is that an issue...

**Ryan Meyers:** Absolutely,

**Neil Morrison:** For you are humans at the end of the day needing to verify that data that they're getting from you?

**Ryan Meyers:** Yeah. So that is definitely a problem. And yeah, I was listening to actually a past couple of episodes leading up to this and listening to your interview with Cathy who said, "it'll skew towards completion over accuracy." And I was just like, I think I said out loud when I heard that, I was like, "Yes," I wrote it down. But it's totally true. And I'll give you an example. I was working on a prototype of a new feature where we can, this isn't in our software yet, but I sometimes prototype with different things. One of them was what we're calling document builder where we can take all the data in our system and generate, we can auto-generate a climate-related disclosure for you. And while I was prototyping this, part of my prompt was only use information that is specific to this client or information that's given to you to help stop hallucination. But then after that, I said, use as many names and details as possible in the output. And it was only a couple of days later that I'm going through this output and I had two different outputs from two different prompts, and I'm reading them and I went, "Wait a second. The name of the head of ESG in this one is different than the name of the head of ESG in this one." And I was like, "Did I feed it different data? Is it a different company? Did I make a mistake?" And I went in and I was like, "Oh." And I googled it and I was like, "Oh, neither of these people exist." It's like making up these names because I told it to use as many names as possible. So, I was like, "Okay, I've taken that line out." So it is, and that's part of the reason by the way, that we're not auto-generating disclosures quite yet because we've actually gotten close to solving at least the major part of that problem. But it is a risk. Hallucination is definitely a risk. And so that's why we're trying to use it for elements instead of generating content, summarizing and distilling ideas is a really big one. For example, one of the ways that we're using that is sure, we've got dozens of pages of content that might be tagged automatically as related to a certain topic, one of those 50 topics that I mentioned that companies are doing, and we've got over a thousand companies in our database, but how do we make sense of that for our end users? And that's actually always been a problem for us, and we can use LLMs for that now. And so that's what we're doing. And so we're basically saying, take this 12 pages of information or 20 pages of information, and instead of our users having to go through all these different pages and try and take out the relevant information, we can just say, distill this down to three sentences with all the most important information related to this topic. And boom, now we got that. Now we can surface that in a searchable database, which is why people who maybe don't have climate expertise can come in and pretty quickly get an understanding of what's the norm, what's best practice, what could they be doing, all that kind of stuff. So LLMs are really good for that kind of stuff.

**Neil Morrison:** Right, just don't ask it for advice, basically.

**Ryan Meyers:** Don't ask it to give you a whole bunch of content and then not expect it to hallucinate because it's at least a risk.

**Neil Morrison:** So, I know what's going to scare CPAs about this is you're uploading; there's data that's going in. A lot of it is public, but some of it may not be public. It's going out; it's going into the system. So, there are two concerns there. One is the security of that data, how secure is it? And secondly, you're doing this comparison, and I may be misunderstanding this, but I get a comparison to other companies. So, is my data from my company suddenly available to other companies, the stuff that I have disclosed and maybe I'm not ready to disclose it publicly?

**Ryan Meyers:** Yeah, great questions. So first, in terms of security, yeah, we have SOC 2 Type 1, I guess, certification right now. By the time this actually gets released, we might have SOC 2 Type 2 because our lead developer has been working on that for some time, and we wanted to get it for February. So yeah, we have a pretty strong level of security in terms of data security. In terms of privacy, keep in mind that the majority of our analysis, actually, our entire analysis that's done automatically by us is on documents from your website or documents from Edgar or Cedar. So it's already publicly available information. That is the analysis that we then share across clients. So if you're populated in our product across other clients, then it's all your publicly available information that's been parsed. If you're uploading draft documents that are internal, if you're uploading private documents, that can be labeled as such. If you're putting in your climate profile information about what you're doing, but you're not disclosing, that's in your own partitioned-off area, and there's no sharing of that information.

**Neil Morrison:** There's no way that someone else is seeing that. There's no other companies when they're looking at a comparison when they're doing their own comparison, like, and Company X is planning on doing this.

**Ryan Meyers:** No, exactly. And

**Neil Morrison:** There's no way that can be leaked through. There's no way for the LLM to leak that through or anything.

**Ryan Meyers:** Yeah, great question. So then for the LLM, we only send that information to an LLM when it's used in some form of analysis to produce data for you specifically. So we'll never send it, generate data points, and then show that to other companies. It's only, the private information stays private. In terms of the LLM itself, we use what's called, it's  too technical, but we use what's called a stateless API endpoint. So it essentially means that we go through our own endpoint through Microsoft Azure, we hit the model in a private server, and then the model itself is stateless. And that means it doesn't retain any information. It just gives the result and sends it back. So a lot of companies have banned the use of ChatGPT because you submit a prompt there with information, and then OpenAI has access to all of that information. That's not what we use. We use a much more secure and private approach to get results from LLMs.

**Neil Morrison:** Okay. Interesting. Last question here. In the future, is it possible that AI will continuously monitor an organization's performance on all of these various metrics without input from humans? You just put it in, it's there, working in the background, always watching it, and then just spitting out information on an ongoing basis. We've talked a lot on this program about the continuous audit. I'm wondering if that's something that can happen around these sorts of disclosure requirements or around sustainability requirements.

**Ryan Meyers:** There are AIs out there that are already producing numbers automatically in real time. The question is, how reliable are the numbers? And so that's why we're putting quite a large amount of effort into not just producing automated analysis, but into making sure that the quality is as close to perfect as it can be and is at least as performant as people. The models are getting better, though. And this is, GPT-4 is not the end; it's just the beginning. Right. And same, there's Google's Gemini, and there's a new version of Gemini coming out soon. GPT-4 and a half will be coming out. So it is only going to get better for sure. And I think if I were a CPA, the way I would be thinking about this is, this is a great tool to be used, it's only going to get better. It's not going anywhere. So I think it's really important to gain some sort of exposure to it. I would start with just using ChatGPT, but if your company's not already working on it, they should be building an internal version of GPT that is secure for your data. And I think everybody should be learning how to use it. Because I used to interview people and part of the interview that I used to do would basically be getting people to say that they'd Google something and then I would ask them what they would Google. And I think that understanding how to use the best tools that are available to get to the best outcome is really important. And GPT is the next thing, but there's still experts that are required to be in the mix. That's why we're not using it for our base AI. It's all expert-driven. And I do think that even using GPT or any LLM, actually, for that matter, you still need experts to be in the mix to review and understand their strengths and weaknesses and make sure that people are picking up the slack on the weaknesses. So yeah, I think there's a lot there. But as a CPA, I would definitely be trying to get your hands on, try to get some exposure to it as quickly as possible.

**Neil Morrison:** It's been fascinating. I really appreciate you taking all this time to talk to me.

**Ryan Meyers:** Yeah, my pleasure. This has been fun.

**Neil Morrison:** Ryan Myers is the Chief Product Officer at Manifest Climate.  On our next episode, we have another case study. This time, it's AI and tax advisory. One company out of Toronto has developed a platform that uses a Chat GPT style interface to offer tax advice. The software is able to dig into proprietary sources of the latest developments in tax regulations. Benjamin Alarie is the CEO of Blue J. He's also a tax law professor at the University of Toronto. I asked him if tax professionals should feel a little uneasy about this virtual advisor.

**Benjamin Alarie:** If somebody feels that nervousness, it's right to register that nervousness and notice that nervousness and be mindful of that nervousness. But, what I would also say is this technology is really a boon for professionals who provide tax advice. It puts a floor under them; it gives them new abilities to see things, to notice things, to understand things at a deeper level. And the underlying complexity of tax is going to continue to increase to keep pace with our technological ability to understand and interpret and apply it in new circumstances. And so, I'm

not particularly worried as somebody who's very much in the thick of it. I can totally understand, Neil, why you're posing the question, because it seems like it's so easy for our minds to do some pattern matching and think, "Oh, this technology is improving at an accelerating rate and soon, maybe there won't be a role for me here in providing tax advisory services." But I think that's the wrong way to connect the dots, at least for the foreseeable future. Perhaps in some far away time that will be the case but I think for now everything that I'm seeing points to tax advisors being supported, and helped, and enabled with this technology.

**Neil Morrison:** That's Ben Alarie, the CEO of Blue J speaking on our next episode. And that's it for this episode of Foresight, the CPA podcast. If you like what you heard, please give us a five-star rating or review wherever you get your podcasts and share it through your networks. It really helps others to find us. Foresight is produced for CPA Canada by PodCraft Productions, and please note the views expressed by our guests are theirs alone and do not necessarily reflect the views of CPA Canada. Thanks so much for listening. I'm Neil Morrison.