## Interviewee 5

**To what extent is the construction industry prepared to embrace the transformative technologies of Industry 4.0?**

I don't know how much prepared we are but let's say that we are very much aware of many problems that we have such as interoperability, such as the way stakeholders are not very well connected around a project, either in the design process of a project or the management of the building phases. So, we are very much aware that compared to different industries we are behind in terms of streamlining of processes and using smart tools, and many of us are not educated in working with data or in thinking of IT solutions or new technologies. Most of the designers, the architects, and the engineers are just educated within the expertise of that specific field. Most of the students of engineering and architecture know the theory but they know very little practice. They really have to wait for their first work experience to use specific tools. So, in that sense we are not very much prepared in terms of training to work with tools and software.

Some of us for reasons of personal interest try to change this, and those will be the people you see on internet using Rhino and Grasshopper, for example, and trying to hack their workflow because they are simply tired of manually repeating boring tasks, or perhaps they see that with modern tools we can do things that other tools don't do. But those are a minority of people and also very few offices demand this kind of skill set. I mostly see a change coming from the bottom up, it's not the companies that try to change, it's people from the bottom.

**As a key technology in Industry 4.0, what major benefits can AI bring to the industry?**

Most traditional tools are very heavy in terms of the knowledge that you need to get insights and to get validated results — as we say: garbage in, garbage out. And most of them, depending on the type of analysis, can take a lot of time to compute and to give a result. The way we see AI tools really helping and making a change is that we can get much faster results with better accuracy. For example, with the platform that we created, we developed tools to predict the daylight performance inside buildings, tools to predict the wind comfort around the building, and tools to predict the energy use of a building, and all of those problems require expertise in what input to feed and what output to expect, and what kind of research questions to ask. But with machine learning what we are doing now is to train machine learning models using various data sets that can within seconds predict the performance of daylight energy or wind comfort, for example, so it can really bring earlier in the design processes a lot of insights that usually come later when the engineers are engaged. So, very early on when we have the most flexibility, we have a better understanding of different variables, and we can shape early on the design with more security. So, it's about reducing risk basically and having a faster and a more iterative process allowing us to try and fail many more times.

**What are the disadvantages or risks, if any, of adopting AI in the construction industry?**

We see a lot of tools on the internet, if you follow LinkedIn or YouTube you see there's always something new coming. But I'm kind of trying to be a critic of what we are trying to answer with those tools, because it's great to make a lot of fancy tools but, what are you trying to solve? Did you make something just because you could? Are you really trying to solve a specific problem? And then, with AI everything relies on training data so it’s important to be critical about the quality of the training data that you have so you can trust the results, and to have enough validate data. A machine learning model will be able to predict well only what it has seen before so you need a lot of diversity in the training data. If you ask something very different, the model will always give you an answer, but you need to be aware that the model will not know how wrong it is.

I also want to add that because those tools are basically tools to simplify things, to make things faster, or to give a broader audience access to knowledge, we see a lot of non-experts claiming that they are experts because they have the right tools. Just because you give an architect a tool to predict something related to engineering it does not mean that he can claim to know the right answer for sure. Trained engineers are usually taught to only use simulation tools to validate an assumption. You should always have a very good insight and a very good intuition of the problem you're trying to solve and only use simulation tools to validate your assumption. But now with all this digitalization, and all this data, and all these new tools, I see this is completely forgotten and so many people have the approach of “oh, if I don't know something I'm going to simulate it and I trust the results.” But how can you trust the results if you don't have an intuition in the first place? I think this is the most dangerous aspect.

**What do you think are the special characteristics and challenges of data produced by the AEC industry?**

How we see things is that everyone creates data in some form, and it's usually stored somewhere on servers. But to make use of this data you need to sort it, and to filter it, and I think probably 90% or more of everything we produce is just produced to serve its first purpose. We produce very little data with the intention of using it for machine learning application. So, there is a big job if we are interested in data. The most time-consuming task with machine learning is to filter, cleanse, and prepare the data because the training of the data is basically automatic, it's a script, or it's a function, or it's a software that will train the machine learning model. So, 90% of the job is to produce quality data. I think that almost none of the data that we have is ready to be trained. So, we need a lot of data scientists and people who know how to handle data to make it ready for machine learning application.

In my experience for example, I worked in different architecture offices and we used to save BIM or IFC files of every project that we have but it's just stored there for legal reasons because we are obliged to keep the data of our clients for three years, let's say, and if you don't have the intention of using that data later for a specific occasion you will never care about how you save that data in the first place. Whenever you need specific information, you will always have to dig for it and to format it the way you wish to use it, so that's one problem.

Another major problem is that we have too many commercial software that basically do the same thing, but in different manners and they save their data in different formats, for example, we have many IFC viewer software. And we see very little open file formats and even if formats are open, we see very little that are standardized. IFC for example, is an open format but there are many ways to save an IFC file. So, if you can represent the same thing in different manners even if it's open, you still need a different translator or a different reader to access it in each case.

**Which aspects of data quality — accuracy, completeness, timeliness etc.— do you think that the industry needs to focus on?**

I think accuracy, because what I'm trying to say is that even if you have a lot of data sets, you need to go under the hood and to know how this data was produced. It doesn't only matter to know if this building has been performing a certain way, you need to be sure that you have a lot of data confirming this, and you need to know who made this validation, what were the parameters that went into the analysis to create this output. Machine learning often takes a lot of output as a new input for training, so the question becomes can you trust those outputs? Who made them? What software was used? What was the premises? Did it follow standards?

For example, in our application of machine learning for daylight prediction we are using thousands and thousands of cases where daylighting was simulated to train the model, and we need to be sure that the simulation was done properly according to what engineers would say is good quality. So, unless you produce the data yourself, I don't think you can trust it completely.

**To what extent do you think that AI adoption is tied to BIM adoption?**

I think that their purposes are different. For BIM to be well implemented you need every single stakeholder to be using it, whereas AI, it's not about an ecosystem, it's not about a community, it can be just a feature somewhere. I think BIM has a much more difficult implementation curve because everyone has to work on it together, whereas with AI you can start on your own and make use of it right away.

**What data skills do think are necessary for adopting AI technologies within the AEC industry?**

I think some background in data science and data management and working with data sets, and I believe the skills to generate yourself the data that you wish for machine learning. I don't really believe in buying data or finding data and using it. I don't think that if you don't create it yourself that you can be sure of it. I don't think we can easily cross-use data because it's still very siloed very field specific.

**Do you think that it is sufficient that only a few professionals or technicians develop a special proficiency in data skills?**

I think that every sub-industry or every niche should have people skilled to produce and generate data. I don't think that Autodesk, for example, can produce any relevant machine learning application for engineers. I think only engineers can do it for themselves because only they can trust the data they are using. We don't need everyone in every niche to do it, but in every single niche we need a group of people that take actions on it because I don't believe that I can interfere in someone else's expertise. So, to produce those machine learning models I think you need specific knowledge and knowhow. But anyone could use them, that's the magic. We just need about 10% to create valid models and the remaining 90% can use them.

**How would you describe the proficiency level of data literate professionals? What must they be able to accomplish?**

Like I said, basic proficiency in computer science, in data science, being able to know and to program at least in Python, being able to work with databases, being able to generate the data for the problem that you're trying to solve yourself. So, if you are looking to automate a floor plan for a building you need to have knowledge about floor plans in the first place, to be trained as an architect, and you need to work with big databases and with some programming.

**What about regular practitioners, designers, and engineers? Don’t you think that they need to acquire some data skills?**

Like I said, we are creating tools so that anyone can create better buildings using our machine learning applications, but I'm still very particular all the time about teaching people to ask the right questions in the first place because I see many people using tools and they trust whatever result they get from the machine. So, I think even if you have a good tool, you still should know how the tool is made what it's trying to solve so you can be a little bit critical about it. So, I think that just because technology is becoming more efficient and intelligent, people should not become less intelligent.

**Do you have any further recommendations or final thoughts on the subject?**

I don't expect end users to do things differently. I don't expect architects or engineers to be to be doing anything very differently, just to be a bit more critical and curious. We need to be getting together and developing things together more often, because even in my field I see many small groups kind of doing the same thing. We don't necessarily have to compete always, but we can also work together toward the same goal. So, I think being transparent and sharing findings is important, to not create more silos because our problem is that there are too many silos. Maybe people are concerned about protecting their intellectual property and earning money off it by creating products, but that can definitely be done without creating silos.