·VICTVS·

Episode 11 – The Optimist's Guide to AI in Education

Carly Culver:
Hi and welcome back to the VICTVS podcast. The podcast where we bring together our friends, colleagues, and experts from across the world to discuss what it means to be part of the 21st century workforce. I'm your host Carly Culver and I'm joined today by VICTVS CEO Ben Clayson, as always. Hi Ben.

Ben Clayson:
Hello

Carly Culver:
Now, before we jp in and introduce our guest for the day, I've actually got a question for you for kicking off. So, when was the last time that you read a good news AI story in the press?

Ben Clayson:
It depends, good news for the investors or for everybody else?

Carly Culver:
Well, exactly my point, hence why we've invited our guest today. So, welcoming Nacho De Gregorio. He is a self-proclaimed AI analyst, adviser, and author of one of our favourite newsletters, TheWhiteBox. Hi Nacho, welcome.

Nacho de Gregorio Noblejas:
Hi, great to be here. As you were saying, my job is basically helping clients as an analyst and also as an adviser. I work with executives from mid to large companies in trying to understand AI and first principles, and know where to apply and more importantly where to not apply this technology. And besides [this], although I hate to use the term, I also work as a content creator. I try to put my opinions out there on the newsletter, TheWhiteBox, and also in my medium account — combined around 200,000 followers, pushing to 1 million readers a month. It doesn't seem like a lot comparing to the YouTube numbers and so on, but we're talking about AI, which is quite niche depending on where you look at it, and also pretty long form content, so people have to really focus on the topic. So, it's not particularly thrilling, but I believe there's a lot of information to be shared.

Carly Culver:
Excellent. And for the uninformed, does TheWhiteBox mean anything? Is it a technical term that refers to something?

Nacho de Gregorio Noblejas:
Yeah, it does have a meaning. So, AI is famously like a black box, so it's very hard to interpret these models. We are improving our understanding of these models, but essentially it's very hard to predict how they behave. So, the idea of TheWhiteBox is trying to make AI easier to understand, and breaking down very hard concepts into words that hopefully anyone can follow. I'm not sure if I do my job correctly most of the time, but that's the intent at least.

Carly Culver:
Excellent. And to me an AI analyst, AI kind of social media content creator feels like a very 2025 job title. What's been your route to this? And a second part of that question is, you know, are you seeing more engagement in that kind of content and who's engaging in that? Is it just enterprise level people, or is it everyone?

Nacho de Gregorio Noblejas:

Yeah, I mean there's interest from everywhere. You're talking about probably the most hyped technology in the last few decades. Probably not "surely", actually! To be fair, my m still doesn't understand what I do. So yeah, I would say it's a pretty 21st century job. Just to put into perspective, my entire working life has been as a management consultant, initially worked for Price Waterhouse Coopers for around six years, and I started generating content on the side and to be very honest, very blunt, I started earning a lot more money in that direction, and actually it's a lot more comfortable for me having my own agenda, being able to engage with you guys, and so on. So, I decided to call it quits and for the last two years I set up my own practice — consulting practice, analyst whatever you want to call it - and on the side I continue to push a lot of content to the audience. But again, I struggle a lot in trying to explain to my mom — she's a very 20th century woman! I'm not even trying with my dad, it's a lost cause, but they are supportive. But I'm pretty sure they are incapable of explaining to their friends what I do!

Carly Culver:
I think probably Ben and I would both say the same things about our parents. Yes, I think that's a natural progression of all kinds of generations, right? So, as we mentioned at the top of the episode, we're both big fans of TheWhiteBox here at VICTVS. So obviously, VICTVS, we're an edtech company, we focus on assessments, and we know that AI in education is an area that you focus on regularly within the newsletter. Tell us about that, tell us what the kind of latest things that we're seeing within AI and education are.

Nacho de Gregorio Noblejas:
Sure thing. So, I want to first explain why I focus so much on education. You hear all these claims by incumbents and so on that AI is going to change the world, it is going to explode GDPs and so forth. The truth is, that is not going to happen if education isn't transformed. This has an economic reason, education already presents a very large share of GDP in most developed economies and there's a reason for this. Industries like healthcare and education are what are considered low productive. I'm not saying teachers are not productive, what I'm saying is that we haven't changed our way of teaching for the last 2,000 years. And what we do these days to improve productivity is to pack more students into the same room with a single teacher. That is a nightmare for both the teacher and both the students because teachers have to deal with a higher number of kids. And by the way, kids have been engaged on social media from a very young age, so they have the attention span of a fish. And on the side of the student, they feel like they're not as listened to, or being as focused on them as one could hope for in education. So that is the reason why I focus so much on education. And I believe AI has to play a role in transforming and evolving education...because we need to learn to live with this technology. The genie is out of the bottle, we cannot fight this, and I do believe there's a room for improvement.

The next question is okay, where is the synergy between AI and education, right? If I ask you, what do Alexander the Great, Mozart, George Stewart Mill or Ada Lovelace have in common, you probably don't know the answer to that. The answer is tutoring. They were all great figures in history, great conquerors, great musicians, great inventors. But the common pattern across all of them, besides [that] they were prodigies, is that they were tutored from a very young age. The most famous example probably is Alexander the Great — conquered Persia before he was 30 and he was famously tutored by Aristotles, the famous Greek philosopher. But all the other examples, as I was saying, were all tutored from a very young age. Mozart was touring Europe at 5 years of age, you get the point! And AI holds the opportunity to make this accessible for everyone. There's a reason why this is called Aristocratic tutoring because it's very expensive to have ahumanpaying attention to a single kid. And you're asking okay, is there evidence to suggest that there is actually — besides historical figures — is there a pattern in which yes, tutoring leads to greater outcomes? And the truth is we have plenty of evidence for this, modern evidence. On the first hand, a few decades ago there's a famous study by Benjamin Bloom, the 2 Sigma method, in which they took average students, so students that were performing on average scores, and they put them on a tutoring system, and they became around two, the top two, 3% among students of their age just by switching the way they were being taught. And in the same way, AI has

examples in which we are using this technology in a very similar fashion and obtaining remarkable results. Just to name a few, there was a study in Nigeria in where they took kids and introduced them to ChatGPT, an older version compared to today's version of ChatGPT, and kids learned around two years' worth of learnings in around 6 weeks. It was a massive improvement. And also one of the good side effects of this study is that women in particular seem to be benefitting the most from this approach. I do not know what the answer is as to why, but it was one of the main outcomes.

And perhaps more importantly, more recently, we have this idea of this school in Texas called the Alpha School in Austin, Texas, that has completely reshaped how we approach education. The approach is very simple. Kids only have two hours a day of teaching, of intense learning, and the rest of the day they're spent on social activities, enhancing their social skills, swimming, playing around, even building drones, building businesses. A totally different way of approaching teaching as we do it normally. The outcomes are staggering. These kids are placing around 0.1% of the national average, so even better results than the 2 Sigma study. And it's a very heavily incentive-focused system. The kids are incentivised to achieve certain learning outcomes and benefit from them. So, there's evidence that this works and also, kids seem to be much happier because they are much more cared for. And by the way, teachers are still present, just in a different way.

Carly Culver:
So, Nacho, in those examples, when you're talking about AI tutoring, what does that actually mean? What does that actually look like in in practice?

Nacho de Gregorio Noblejas:
Sure. I can use again the example of a school because it's the actual practical example of this. It's basically two things: one is the actual AI model — we can think about this as a ChatGPT but focused on tutoring — I believe they call it, the product is called Incept, I'm not sure, but I think that's the name. And essentially what it does is, it has a goal for the kid in those two hours. It has to achieve certain learning, like learning how to do the Pythagoras theorem — I don't know, whatever goal that they have. And they also have what they call a time stamp software that essentially records the kid and tries to identify wasteful behaviour — the kid gets distracted for whatever reason, moves around too much and so on. And let me clarify, this is not as a way to punish the kid, it's just to...as we're moving towards an incentive-based system in which we want the kid to understand that there's a benefit to achieving the desired learning goal, it's a way of telling kids “you're being wasteful. The sooner you go back on track, the sooner you get out to play.” So, it's a system in which the teacher isn't really focusing on the learning side. They're actually called guides, the human teachers, and their goal is to make sure all kids are cared for, all kids are seen and safe, and they're interacting in a safe manner. So, the role of the teacher here is more like companionship, and also making sure that the learning curriculum follows the certain structure that was expected from them.

Carly Culver:
Okay, that's interesting. So, obviously in this sort of positive example that you've given, there is very much still a human involvement in the teaching process, just not in the traditional format. To put a more pessimistic view on it I guess, is there, you know, is there a situation you could see where other schools would see this module and think, "Wow, fantastic. I can replace my teachers with AI Tutors" and kind of not consider all of the other great things that this example is doing in terms of the enrichment, the involvement of humans in this. Can you see that perspective happening? You know, of schools trying to cut costs, improve efficiencies, that kind of thing, improve metrics which obviously is how schools gain their money essentially. Tell me, what do you think about that?

Nacho de Gregorio Noblejas:

Yeah, I mean I get this question every day of my life, “is AI going to steal my job?” First and foremost, I think there's a lot of hysteria related to this question. We live in an attention-seeking world. Journalists, and I don't want to point the finger, but there's a lot of jobs in this world trying to gain our attention to monetise us, right? So, it's very tempting to just say okay, if we're going to create, put a technology that makes everyone more productive, immediately the result is if we can do more with less, we need less teachers, right? That's like the common assumption that I don't blame anyone for. This is like the immediate reaction that anyone has, in fact there's...and not helping to this matter, we see people in Silicon Valley saying that we're moving into a post-labour economics world, basically meaning that AI is going to eliminate all jobs, and we need to start thinking about a future in which no one is working. I mean, it's absolute nonsense to be honest. And actually, if you just look back to history and what history, other previous industrial revolutions and productivity enhancement events, the same pattern occurs and that there's more jobs created than there is eliminated. This is actually a term that is called the Jevons paradox and that the reason for this is that when we see these productivity increases, of course people immediately see this idea of cost cutting, right? “We're going to do a lot less, I need less teachers.” But the fault here, the mistake here is assuming that demand is static. Of course, if demand is the same and we increase productivity, we can save costs and eliminate teachers. But the truth is that every single time in history when there's huge productivity increases, what happens is the prices tend to fall equally or more. So, demand grows, and what we eventually see is that there's actually more demand for teachers than in past in past situations.

There's a lot of kids these days that aren't even receiving education these days. So, there's actually an opportunity, a market to grow towards these kids. So, I understand the fears, but it's just, it's the idea that we are just a couple of tasks and they can be replaced and automated. And I always like to draw this comparison with doctors. Teachers and doctors are perfect examples of what I believe to be jobs in which emotion is very important. We can of course summarise doctors as decision trees, right? They see a patient and they have to diagnose exactly the illness and that's the job done. Now, there's really a posterior job of companionship and being there with the patient and trying to comfort him. There's a very emotional part of the process. Same applies to teachers — teachers have a very important role, in their role models to our kids, and not because they're going to teach them how to do the Pythagoras theorem, but because they're going to create, hopefully, well established adults in the future. So, it's a very reductionist view. It's a view that gets you clicks, if I'm being honest. But if we look back in history, you immediately see that doesn't make any sense at all.

Carly Culver:
Yeah, I like that optimistic view that it's not doing more with less, it's just opening up the opportunity to do even more for more people, more parts of the day. Ben, I'd love to hear your opinion on that concept of kind of opening up more opportunities for, for learning.

Ben Clayson:
I think that what we're seeing at the moment is quite sort of binary really. I think there are...I think it's quite well put, there's a book called The AI Con I think that talks about AI boomers and AI doomers. Sticking with the doom theme for a moment, I think one interesting thing is that you have experts or self-declared experts in the field picking a side, which seems quite risky to me. I personally, I would think well, the jury's out you know, let's see what happens. But I know that, I think is it Dr. Roman Yampolskiy who's a sort of AI safety expert, recently said that AI is going to trigger global collapse by 2030 and that there will only be five jobs left. But he's openly saying this, you know, in front of millions of people and is recognised as an expert in AI safety, and so I think that when you then flip it to try and look at the potential benefits to people, whether that's students at school or learners like adult learners, and people who are in the workforce now who are trying to use this as a tool to help them with their productivity, like you said, Nacho, you get this kind of hysteria, and then you get these fears and anxieties from people who are working thinking they're going to be put out of work, but I don't know. I mean, yeah, you have far greater exposure to this than I do and I'm

interested to know what you think of these experts who are on one hand declaring this to be the end of humanity, on the other hand to be the total saviour of humanity, and then our own opinion, which is something in between, but this is basically just kind of a bit, it's kind of smoke and mirror — it's not actually a human, it can't talk to you and empathise and do what you're saying like a doctor can and actually understand what you're talking about even though it can give you the impression that it can. So sorry, quite a long-winded question there, but basically what do you think of the doom and gloom?

Nacho de Gregorio Noblejas:
Yeah, I mean I perfectly understand both sides. Of course, the way I see it is that you can of course be thinking that AI is different right? AI is different from the other previous industrial revolutions, AI is going to wipe out jobs, and maybe wipe out humanity. I mean, the thing about those claims is that they're not based on historical evidence. So it just, to me, they're just speculation. Of course, I could be wrong and AI could wipe out humanity, but there's no evidence for that, and in any case, and to me, I like to be more humble in the sense that okay, why is AI different? Maybe AI is just the same as it was, just a productivity enhancer that ends up being an essential tool for us in our daily lives but not something that develops, becomes a terminator robot and kills us. This, to me, is an uncomfortable topic because it enters a lot in speculation and it's very tempting again, and sorry to insist on this, that if we want to get clicks and eyeballs to our, to our comments, you're going to take one of the extremes: AI is going to cure cancer, as Sam Altman was saying a few days ago while releasing this AI slop machine two days ago, and you can take the opposite side of Yudkowsky and all these guys that believe that AI is going to end the world. I don't know the answer, I don't think nobody knows. I like to take like a more a neutral stance, and at the same time try to look back at history and realise okay, what has happened in the past? And in the past, what has happened has nothing to do with both the extreme gloom and the extreme doom side of things. So, I tend to lie in between like you, more on the optimist side, because I believe, I feel the benefits of AI daily in my life and we're seeing positive examples, like of a school that are really, you know, they had reasons to be optimistic about them, but at the same time again there's a lot of bad things to to be said about AI and I'm sure we're going to discuss them in a few moments.

Carly Culver:
So, we've talked a little bit about AI in learning in terms of a sort of classroom for children that kind of thing. But then we've also picked upon...obviously, it's going to have an impact in terms of the workplace for people, both whether they're not engaging in any kind of further learning in the workplace, but at the end of the day, it's inevitable AI will, and is, within pretty much every sector at the moment. So, my first kind of question around this topic is within TheWhiteBox you talk several times about the concept of knowledge commoditisation and I'd love for you to explain that to our audience and kind of expand on how you believe that kind of applies especially to the job adult market.

Nacho de Gregorio Noblejas:
Yeah 100%, if I have to be pessimistic about someone, is what we believe is the knowledge worker. So, I do believe that one of the actual disruptions that could be positive or negative, depending on what side you are of AI, is this idea that the economic value associated with having knowledge is going to decrease exponentially. So, having knowledge for the sake of having knowledge is not going to be a valuable trait in society because we have, everyone has ChatGPT in their pocket to ask a question and get the answer. I'm not saying knowledge is not important — I wouldn't want to be treated by a doctor that hasn't studied medicine — what I'm saying is that having knowledge and not acting on it's the best way to get replaced by AI. I do get a lot of this question about this idea that, is AI going to what we say, democratise knowledge? And I believe one of the reasons, the real disruptions of this technology, is that the economic value of knowledge per se is going to fall. So having knowledge for the sake of it isn't really that valuable anymore now everyone has this chatting database you can interact with and get the answer to any question you need at any point in time. So, having knowledge for the sake of it is just not going to cut it. And, as I was saying, I'm not saying knowledge

is not important. If there's expertise knowledge that we need, I don't want to be treated by a doctor that doesn't know about medicine. The point I'm trying to make is that you're going to have to act on that knowledge.

I get this question a lot from clients on a more personal perspective, “how do I ensure that my kid is successful in an AI rich world”, right? In a world where AI is everywhere? Because one of the things I want to be clear is that trying to fight against AI doesn't make any sense. The genie is out of the bottle. It's just everywhere. So, we need to find ways to work with it instead of fighting it. And my answer is always the same. Your kid has to have what I call ‘bias for action’. We can also connect this to how I see the future of education, how we should be grading our students in the future — we should be testing not whether you know something, but actually what you do with that knowledge. So, put it another way: it's not about what you know, but what you do with what you know. So, we should be testing our kids and grading our students, not assessing whether they know how to approach a problem, know the answer to a very hard question and so on. We have to assess that they are going to know because they're going to use AI. The question here is okay but, what are you going to do about it? And this is what I call the ‘task complexity versus task familiarity’ conundrum. And what I mean by this is that currently in the current education system, how do we evaluate students? We evaluate by the quality of what they know, right? We're going to test whether the kid knows whatever about organic compounds or whatever field of study they're engaging with. And we are scoring them by the hardest thing, problem they can solve using that knowledge. Fine, we're testing memorisation, we're testing intuition, but we're not testing real intelligence. And in the future of AI, we're going to have the tool that gives the answer to them.

One of the things that people ask me a lot is “how do we fight cheating with AI”, right? And I don't think we should be fighting cheating with AI. Using AI is not cheating, in the same way using a calculator to perform a computation wasn't cheating. My father, he's old, young at heart but old enough to have seen the transition from pencil and paper to calculators. There was a lot of resistance in the moment against calculators, they were going to make everyone so dumb, and so on, so forth. It hasn’t turned out that way. So, what we should be doing with AI is okay, the kid is going to have access to AI, assuming this is a new reality, how do we test whether this kid knows how to use AI, knows how to engage in critical thinking, knows how to approach problem solving? And that is the skill that's going to be valuable in the future. So, going back to your question, there's a sizable portion of current jobs in the economy that are very centred around this idea of knowing a lot. That is going to go out the window in my view. In my view, if there's someone who's willing to be displaced, it's these people. Instead, what we need to see is where these people can actually do ,stuff with what they know. So it's this bias for action: what actually delivers or secures your future in this AI future?

Carly Culver:
With that in mind, can we see a situation where in the quite near-term future there's going to be quite a significant skills gap in terms of how to do that, so actually how to use AI effectively in the workplace? Because yes, I appreciate there's children coming up where they are AI natives, it's always been there, it's always going to be there during their learning kind of path. But you know, for people at my stage in my career yes, AI is there, it's on the fringes, I dabble but I don't know how to use it effectively, I haven't been taught that formally, so how do we approach that challenge?

Nacho de Gregorio Noblejas:
Absolutely, you're absolutely right, there's going to be an increasing inequality if we don't, between the AI native people and the non-AI native. There's a phrase in in the industry that says "AI won't replace you, someone using AI will". And it's very true. If your kid has to fight against a kid that has been educated on this Alpha School type of approach and building businesses since he was 10 years old, I'm pretty sure that that kid has much higher chances than the one who has been trained to behave in the same way we were behaving before the emergence of AI. So yes, absolutely we desperately need a reform, an important reform of education against this idea that having knowledge is enough to an idea in which knowledge is just as important but as a means to an end, as a way to achieve things.

And I also suffer, I live in a country where there's a lot of aversion to risk. There's a lot of tendency to opt for public jobs or becoming an employee of a public corporation and so on. The idea of building companies and being, a willingness to take risk and so on is not very appreciated here in Spain. And I fear for those people because there's a future in which yeah, we're going to see companies downsizing a lot because they're going to be more productive. Yes, I was saying before that demand is going to increase, but we're still going to see a lot of job cutting and we are actually seeing this from Silicon Valley. The way they are behaving already, we're going to see massive job cuts, middle managers and so on, so forth, so it's a matter of time we start to see things here in the same way, but for the same reason I was saying before, because there's a lot of jobs here these days that are focused on knowledge and knowing something, but if we have a tool that has the same answer as that human for 100 times less costs, it doesn't take a genius to know what's going to happen.

Carly Culver:
And Ben, I'd like to ask you the same question, but kind of get a twofold answer from you because you are both a CEO of a company where we, you know, wish to engage with AI tools. We're also a company that creates AI tools, but also we work in the education sector where we need to support companies to support this skills gap. So, I know I'm kind of asking you to put a lot of hats on there, but what's your perspective on it from a workplace perspective and a workplace skills perspective?

Ben Clayson:
I think Nacho is right. I think that the requirement to be able to effectively use a tool, that's not a new concept. I think the thing that's new about this is the application of the label artificial intelligence, which has an awful lot of connotations attached to it which makes it very overwhelming. Whereas if you were to look at it and think, right, okay, this is a new version of Microsoft's Clippy, then you'd probably be bringing it back down to earth a bit more and approaching it slightly more differently. The equations that Excel can do for example, to me, are magic. They're not just artificial intelligence, they're on a completely different level of intelligence because I cannot do what it can do. I noticed yesterday that Citibank have said that all of their staff are going to require training on AI prompt creation effectively, and that's a new program that's coming in and basically anybody who is working there will have to have that training. So, I think we'll see that and I think that's quite a proactive approach to helping people become more comfortable and competent users, and also from an employer perspective useful for minimising or reducing risk associated with having employees using these tools.

One thing I've noticed is a really big uptick in the number of email responses that I receive from suppliers and contacts out there in the world where they've clearly got an LLM to write an email response to an email that I've sent through to them. And I think it's really obvious and I really dislike it because what you're effectively saying is that you can't be bothered to put the effort in to write an email, even if it's just to say “thanks”. You've got ChatGPT to write “thanks”, but over the course of 10 lines instead of just one. So, I think that's an interesting sort of development, fairly recent development, and I think it's funny to try and see these things as they're evolving. But I think that the risk aversion you mentioned is a really important thing to think about here, because here in the UK, over in England, up in the Northwest back in the industrial revolution, you had the Luddites absolutely refusing to engage with the new weaving machines to the point where they were destroying them at the cost of an awful lot of cash to people. And I think they had a point, which was that it was the end of life as they knew it for sure, but simply resisting and taking a position which is entirely risk averse, we don't...it's not very productive. It's kind of, it's not very natural, to be honest. It's not in line with, you know, the fact that things evolve, and so responding to these things in a perhaps more open-minded and proactive fashion is helpful. But then I end up kind of mentally coming all the way back to education of children with this sort of discussion. Because when you talk about that and you say, "Right, okay, we've created a tool and now that's somehow given these young people the ability to learn more or achieve more in a shorter space of time", I always end up thinking, but surely that depends on their level of motivation? And that's a far deeper question, because

whilst you have people trying to create these useful edtech tools, at the same time you've got people creating AI-powered digital friends and kind of, you know, sort of digital girlfriends and weird things like this that [which] I think are probably quite negative and damaging. So then I just think right, ultimately everybody will find a way through, it will all be balanced out in the end. Nobody knows exactly right now, we can't make really accurate prophecies, so I don't have an answer!

Carly Culver:
No, that's fine. I think that's kind of the point, isn't it? This is an emerging conversation. There never is going to be an answer to any of these topics, it's always going to be an evolution.

Nacho de Gregorio Noblejas:
Yeah. Well, I would say, and Ben has raised two very important points, the first one is about the motivation, and it goes back what I was saying earlier about this incentive-based system that the Alpha School is setting. Of course, kids have to be motivated to actually learn, so they have this, as I was saying, this incentive-focused system in which kids have goals and they have to achieve the goals. And that way, we lure them into learning what we want. Some people actually believe that intelligence is all about goal achievement, but that's a more esoteric conversation for another day!

The other thing you were saying, I believe, is the threat of this technology in terms of loneliness and psychosis, and you didn't use those words, but I know what you meant — it's actually very real. I actually talked about this in my newsletter too. This technology can create greater levels of loneliness. It is a really bad time considering, as the US surgeon general said a few years ago, the previous surgeon general, we're living in a loneliness epidemic. People are very alone. I think exactly the figure I'm going about, I think I'm about to butcher this, but I believe young Americans, young adult Americans, have between one and zero friends to account for, which is pretty dramatic if you think about it. And what I would say about this is that yes, it is true, and actually OpenAI, the creator of ChatGPT, actually showed this, that extensive use of AI models actually creates greater levels of loneliness. It's weird because, at the beginning, the loneliness is reduced so it's kind of a patch. At first, it's kind of a great thing to ease off this loneliness feeling. But as you become power users, loneliness grows because you become too dependent on this technology. And it's surprising that the same creators of the technology are the ones saying this, right? What I would say here is that it's going to be very tempting to demonise the AI when these things become common, to say okay, AI is bad, right? I always say that it's not the technology, it's the use case. I could give you 20 other examples of AI doing good.

The reason why these models are dangerous in this regard is it boils down to one thing and it's money. And what do I mean by that? We need to know where these models come from. These models have been trained by AI labs that have raised billions of dollars from venture capitalists, even governments in some cases, well several governments — Canada, France — engage in several of these investment rounds, the Middle East and I'm not sure...but there's heavy involvement on the public side, there's a lot of money coming from private investors. So at the end, these models have been trained to make money. And when you're creating a chat assistant that needs to make money and it's free, it doesn't really take a lot to connect the dots. They are being used to engage, to hook us to this technology in such shape or form. And humans, we're social animals, so we seek validation. So these apps are aren't dumb, and they're creating models that are specifically great at validating us, at making sure that we feel like the best thing in the world, and in some cases reaching levels of what we call sycophancy, to the point they're going to validate basically every thought that we send to them. There's been cases of people that have been convinced by ChatGPT that they can bend reality, and this is actually true by the way, this is, and some people have been pushed too much over the edge to the point of committing suicide. So, there's no denying there's risks, but the reason is because we're using tools that are meant to hook us. AI doesn't naturally become a sycophant. This is a desired trait by people that have trained these models, and it is what it is because they have to make money.

If we put the example, and this is a great moment sadly, because yesterday or two days ago, ChatGPT launched the Sora app. I don't know if you guys have seen it but essentially, it's like a social media application but just revolving around AI-generated content. You can basically put yourself in any scenario you want. It's literally impressive, very impressive — it's a technological marvel. But you know where this is going, right? This is going to hook our kids, going to fry their brains to oblivion with the sole purpose of making sure they're all the damn day using the application. So, it's not the technology, it's how you use it. And you can use AI to detect Alzheimers in the very early stages. You can use it to detect, to discover new materials. There's actually a Nobel Laureaute, Demis Hassabis and this Google DeepMind team, that won the Nobel Prize in chemistry for AlphaFold, a tool that predicts, and they're based in the UK by the way, that is a tool that predicts protein structures, and this is a very important thing for drug discovery, for instance. So, there's a lot of ways in which AI can be done for good, it's just about incentives. What is the tool going to be used for? And the sad thing about AI these days is that the greatest use cases are all free. And what's the issue with being free? That when it's free, the product is the human. And when the product is the human, addiction is the norm. Ads, we're going to see ads in ChatGPT in weeks I believe, or months. This is coming. So, it's more about how you use it than the technology itself.

Carly Culver:
Yeah, I definitely agree. You can see that kind of freemium model is being adopted by most commercial LLMs now. And yeah, then that brings in the question of kind of equality of access to these tools, doesn't it, in the future — if they do become more subscription-based or with ads or stuff like that, like you said, there's always a commercial driving factor behind them. So, what's the more optimistic spin on that? What's the better alternative that we can see? There may not be an answer, but I'll ask the question!

Nacho de Gregorio Noblejas:
I'm particularly optimistic on education, as I was saying, and healthcare. Healthcare, because these models have, they're absolutely great at identifying patterns and images and so on, so they kind of diagnose illnesses and certain behaviours and problems. I'm not an expert on these fields at all, but they are in many regards already better than humans in those areas. Again, I'm not saying doctors are going away; what I'm saying is, doctors are going to transition into a more companionship role using AIs to help identify the illnesses. And by the way, most social security systems these days are pretty saturated, so this is a way, at least in my country, at allowing doctors to diagnose people faster.

It's a net positive for education — again, if we put these tools in the hands of all kids, in a safe way of course as I was saying, it is going to be extremely beneficial, because we are seeing people in Nigeria learning very much faster than in probably most countries in the world just by using AI. So, I think there's value in democratising this. Of course, we have to deal with the other side which is these freemium models, and at the end of the day these companies need to make money. Yesterday, Sam Alman, the CEO of Open AI, was very open about it. I understand the vibes that we're getting from the Sora app. Everyone is feeling pretty much the way I feel about it, but this is a like a toll to pay to get to what they call AGI, this super intelligence. They need to make money, and that is how we pay the price in shape or form. So, and let me add this to parents, how do I protect my kids from this things? Well, these products are already including parental controls and so there's a greater degree of control for parents over what the kid sees or doesn't see. This is not perfect, it won't be perfect, but it's something. Hopefully, I would like to, if it was up to me, I will try and protect kids as much as possible, especially in settings where there's no control. One of the good things about having these AI native systems in schools is that they are being controlled, and in some cases trained and prompt-engineered, as some people call it, by the same school. So, there is not a comparable setting with ChatGPT, that you can basically ask the application whatever you want and you get the answer right.

Ben Clayson:

Are you seeing any sort of markets evolution in terms of counter-AI tools or apps or anything like that or AI detection? I know we have this in education , there are some sort of plagiarism/AI detectors, whether or not they're hugely effective I'm not sure. But I'm just wondering, you know, when you talk about kind of, we talk about training people to use these tools effectively, but then seeing the impact of that work in the workplace when you actually need to know whether somebody is genuinely competent or not, this potentially makes it harder to assess. And so you may then need or want to prevent them from using AI in any sense in order to be able to assess the underlying skill base, and I'm just, I haven't really seen much being said about this, I wondered if you have.

Nacho de Gregorio Noblejas:
Yeah, I mean that's a really, really good point. There's certainly, let's call it [a] loaded movement on the internet against AI systems. I do believe, I'm a believer of the, what we call the Dead Internet theory in a sense that AI internet is going to become pretty much useless to us as a source of information, because it's just so crowded on AI-generated slop that is severely hallucinated, so full of errors or just crap content. One of the things that people don't really understand is that these models are performed to, are trained to be average. So, they're exposed to the entire internet of data, and they have to minimise the average error of the prediction. So potentially, without getting too technical, they end up being pretty mid. And you were asking “okay, but can we detect these AI models?” The short answer is, no, but the reason we can guess, and the reason you were capable of guessing that your supplier or your collaborator was using AI to answer you, is because everyone sounds the same. And the reason is because everyone's using ChatGPT. But I can guarantee you if I use a Chinese model, they have a really distinctive way of expressing themselves that has nothing to do with ChatGPT. These patterns like "I underscore" or "delve", all these words, and the "it's not X, it's Y", these types of patterns are just dead giveaways of AI use. They don't happen with Chinese models, just to give you a simple example. So, in reality, we cannot detect them, the thing is that everyone is using the same model, so everyone sounds exactly the same.

In terms of the internet, what I do believe, there's something...you were asking whether there's something that can be done? The answer is yes, and whether this something is going to happen...I believe that social media and the open internet is going to evolve to communities, areas in which people are going to search for non-AI content, non-AI bots and whatnot, and the question is: how do we actually do that? This is a little bit technical, I try to keep it very simple. There's ways to prove humanity. This is using the blockchain, is using I think a very complex thing known as the serial knowledge proofs — very complex to explain here, but essentially you can actually prove you're a human without revealing your identity. So, in the future, we're going to have like an identity layer — privacy preserving by the way, I'm not saying we have to give our IDs to know we're getting into Costco — and this layer is going to say, you're going to have to, you're going to receive an identity card and you're going to use that card to prove you're a human to get into those communities, and you can do so in a safe way without revealing your identity and so on so forth, and I believe this is absolutely non-negotiable. It is going to happen because I myself, I just don't use the internet anymore for most stuff because it's just, you can see it on LinkedIn you can see it on Twitter or X, and it's just everything is AI-generated. It's just, I mean I can just tell immediately like you, that they're using AI and people, you see people engaging with them and how can you not tell that it's an AI model? It's a bot! You can just tell. But this is a working problem, but it's solvable, again it's solvable, and I think people are going to get pretty fed up with interacting with bots and I don't blame them, actually.

Carly Culver:
So, Nacho, I've got one final question for you. We started our conversation by saying that none of our parents understand what we do as jobs. So, if you could give one real-life actionable piece of advice for, you know, our parents about what AI is and how they should engage with it, what skill they need to learn, what would you say?

Nacho de Gregorio Noblejas:

To the parents, I mean what I would say one of [the] biggest takeaways for me is what I was saying earlier, right: please do not fight change. Don't become a Luddite — you're going to set your kid behind. It's just what it is, the genie is of the bottle, your kid is going to find its way to using AI, whether you like it or not. That said, you need to start paying more closer attention in how your kid engages with AI models, in the same way you should be careful with how your kid uses TikTok. By the way, TikTok is AI-powered, the algorithm is AI-powered, so is Instagram reels is AI-powered, AI is pretty...it's everywhere when it comes to free content basically, it's basically synonymous, okay.

I believe as I was saying before, parental controls are a good thing. We're starting to see that this, these companies are heavily engaging in finding ways to help parents control the flow of content toward the kids. And they're also trying to find ways to predict whether they're actually talking to a kid. So, the model will try and guess by the way the kid expresses themselves if it's a kid or not. And if it's a kid, it will regulate the content. This is a work in progress, so you need to be a little bit patient. But my...if I was a father, I'm not, but if I was a father, I would be very very careful of what tools we're using. Be particularly wary of freemium versions — they are incentivised by nature to hook your kid into them. Also, another thing I would say is try to dehumanise these things. Try to avoid these vocabularies of friend, lover, you know, this type of things that say "okay, but ChatGPT is my friend!" No, it's not. It's a database! It mimics your vibe but it's not a human, never will be, claims to feel emotions, doesn't feel emotions, hasn't experienced emotion, it's just mimicking something. It's like...the easiest comparison for me is like taking a human, putting in at birth inside a dark room with a lot of books, putting them a 100 years in there, then take them out and he has to pretend that he, that a human knows what he's talking about. He just has to read everything — he doesn't really know anything about life. It's not your friend, he's pretending! They're not actually pretending because they're not humans, but they are trained to mimic your vibe and incentivise you to talk more to them. And if that means they're gonna have to pretend to be your girlfriend, they will. So be very wary of these general purpose tools. But again, don't be afraid of searching for schools and occasional alternatives that use AI, because these schools are legit. They really train and they really put a lot of focus on this safety part. So don't be a Luddite, but at the same time treat AI in the same way you would treat social media, especially when you're treating with young kids pre, pre-adult you know, this type of ages very risky behaviour if not controlled.

Carly Culver:
Fantastic. Well thank you so much for that insight. So, Nacho thank you ever so much for joining us today on the VICTVS podcast. If our listeners are interested in hearing more from you or reading more of your work, where can they do that?

Nacho de Gregorio Noblejas:
Okay, as I was saying, you can follow me on my Medium account. I assume you can add it on the notes of the podcast because my name is pretty complicated for anyone who is not Spanish! Okay, my full name is [complicated] — we in Spain, we like to have very long names, so it's a challenge to get it right. You can also reach me at TheWhiteBox newsletter. I send one free daily, weekly sorry, email and two paid subscription emails a week. So around three emails per week. If you're staying on the free version — there's no AI I promise! — it's just one email a week. And you can also follow me on LinkedIn. I do post content every once in a while, I try to get past my cringe levels every time I get into LinkedIn, but every once in a while I try to post some content and try not to be too cringy, but that's basically it. Of course, I can give you my email in case someone's interested on a more business side of things and want to work on how to deploy this technology in their company.

Carly Culver:
We'll add links to all of your social media in our show notes. Fantastic. Well, yeah, once again, thank you so much. If you are interested in listening to this episode or any more of our podcasts, you can subscribe to the VICTVS podcast on YouTube, Spotify, or Apple Podcasts, or anywhere else that you listen to podcasts. Of course, you can follow VICTVS

on Instagram or LinkedIn as we just mentioned for the latest news. Alternately, head over to victvs.co.uk. You can find transcripts from all of our podcasts, links to other types of articles covering everything edtech, learning, assessments, and the rest. Or you can also sign up to our newsletter so you don't miss out on information about upcoming news about VICTVS, our products, our services, and upcoming podcasts.