

“We experience good days and bad days with AI”

A CONVERSATION WITH NIC BUCHANAN*

What are the capabilities of AI?

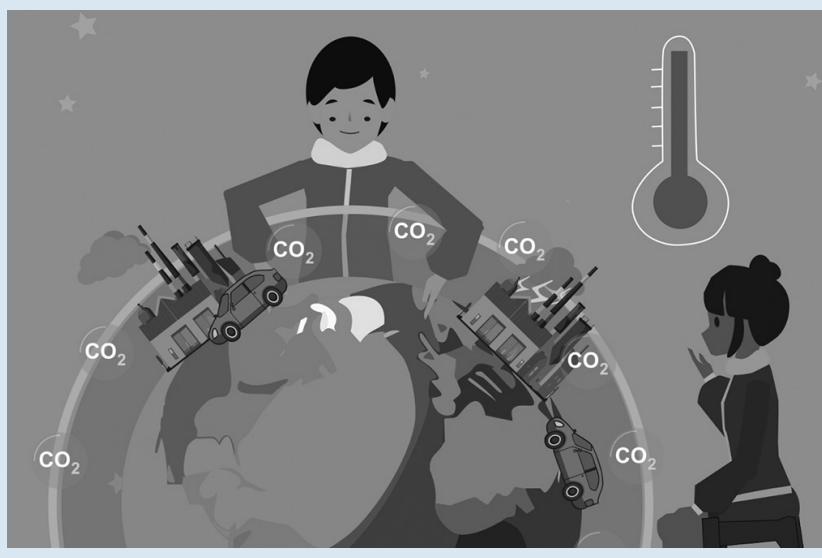
Buchanan: AI seems capable to do almost anything that humans can do. While there are limitations, I believe the problems will be resolved and the technology will work. AI is still learning and encountering new inventions and new software every week. So, my answer is: anything.

You are working on The Climate Dictionary as part of the project “For Us, No Planet B!”¹, where the voice over will be translated in as many languages as possible. Can you share your experience selecting the right software and how AI plays a role, especially in translation?

Buchanan: The journey began with excitement and enthusiasm, and that excitement still exists. However, it took considerable time to find the right software and not just to follow the hype and advertising of the latest, newest, most expensive software. Many new tools promise a lot but often don't deliver. We would buy licences for software that promised to be able to do everything, but it didn't work as expected because the programmers still hadn't fixed all the issues. The most memorable part of the beginning of the journey was figuring out which software to use and which software would be able to deliver results usable for TV or social media. For our project, the process became a lot easier when we were told that the content was produced for online use and that it didn't need to have broadcast quality. This meant we didn't need the top refinement on all the audio settings.

The Climate Dictionary

Simple and appealing animations (max. 1.30 min.) explain basic concepts of global warming. The animations were created by teams from different countries (e.g.: carbon footprint and fast fashion (Brazil) or fossil fuels and CO₂ (South Africa, see illustration)) and are made available online in numerous languages.



When that restriction was removed, it opened up a lot of doors with simpler software that just got the job done instead of trying to be everything to everyone.

In visual AI, particularly animation, it's the same principle. It is about defining your end use and then you choose your software accordingly. It's been 2 years now, and the journey has been fascinating, I've been loving it.

Can you give us a step-by-step guide on the key milestones of your journey with AI and the software you used?

Buchanan: When we started, the text-to-video tool *HeyGen* was the one programme that everyone was

discussing. They had even signed up new Hollywood actors to lend their voices to the software. We tried to create simple explainer-video voiceovers out of them, but it wasn't working.

After several iterations, we ended up with *ElevenLabs* which at the time felt like the holy grail of software. However, we've since had a couple of changes in the last months and were then using *Resemble*, a tool I hadn't even heard of a year ago. We kept exploring and adapting. While *Resemble* was our choice for a while, we found ourselves back with *HeyGen* and *ElevenLabs* again. *Resemble* is not being used currently. Luckily, programming isn't our job. There are other people out there trying

INTERVIEW

to understand and address our needs. We have conducted a lot of workshops with different software companies to convey what we need. This communication is valuable for both parties – when developers understand our goals, it becomes easier to find solutions. For example, we aim to convert a German voice into Portuguese or Spanish and achieving that requires very specific tools.

What have been the biggest challenges?

Buchanan: One significant challenge has been the concept of “hallucination”, a phenomenon across all AI, where the programme doesn’t understand itself. An English term we often use is “brain fart”, a momentary lapse which can happen to anyone when they’re tired. I think AI gets tired as well occasionally. These hallucinations meant we might receive 30 seconds of the voice we ordered and paid for, but then it would suddenly go back to the original voice.

It’s challenging not to understand how and why these glitches occur and even software engineers sometimes can’t explain them. We suspected that internet connections and speed in Africa played a role, but the engineers denied that this affects performance and claimed it only slows things down. So, we experience good days and bad days with AI.

The 2 main challenges are maintaining quality and addressing these hallucinations. They have decreased over time, but are still there.

What translation difficulties do you face?

Buchanan: Hallucinations also impact translation. *ElevenLabs* was quite effective for internal translations, with only a few mistakes when converting English audio into German. That was one of its stronger points. However, *Resemble* uses *Google Translate* and sometimes adds extra words or sentences that weren’t in the script.

How do you define quality in AI-generated language translation?

Buchanan: It is actually a paradox: we’re using AI to achieve a human-like voice. We want the human feel in each AI-generated voice and yet we don’t want to use human voices due to time and cost constraints. And yet we go all the way through the computer and the software to try and get back to a human voice. I think it’s a strange journey. Concerning the translations, being someone who only speaks 2 languages, I don’t even try to judge or gauge the quality of the translation. I focus on the sound of the voice and whether there’s some emotion in their performance. I believe the next generation of AI voice software will allow us to incorporate warmth and various emotions like anger, similar to what we could do with human voice recordings in the past. But it will all come down to the sound editing and how much we can tweak these audio files, whether they’re human or computer-generated. It’s a strange journey of moving from humans to computers and back to human-like results again.

You’re also involved with animation and AI, aren’t you?

Buchanan: Yes. There are numerous programmes that allow us to input text prompts to generate visuals, which can then be converted into short motion clips. There are 2 sides to it: one involves taking an existing animation, giving the characters to AI, and asking it to create different scenarios like climbing a tree or running. The other involves asking AI to generate entirely new characters, scenes and totally new concepts for animation.

However, I’ve found that the technology for animation is lagging behind voiceover AI although we see impressive demos of Hollywood-quality live-action movie-shoot productions, sometimes created in someone’s garage. Again, it’s growing, it’s evolving, it’s improving, but it’s not yet the magic cure for all our problems.

Will there be an AI-based chapter in next year’s The Dictionary of Climate Crisis?

Buchanan: I believe there will be. Our team has decided to use human-generated concept art and then put it into AI to fill in the gaps. Animation, like any film, consists of 25 pictures per second. Remember the old Disney days of manually drawing and folding through and seeing all the movements. AI can help a lot here because you don’t have to draw 25 frames for the second anymore. You can draw one frame and instruct the AI to animate the next 5 seconds, producing results in less than a minute. It may not always come out right as AI hallucinates, but it can achieve what would have taken a human animator a few days – and probably there would have been mistakes as well because we’re human. Then it would have taken another few days to fix it.

It’s not a creative process, but it’s a tool that we must use if we want to express our creativity and get something on TV or onto a movie screen without the massive Hollywood budgets that we don’t have access to. AI therefore offers the opportunity to produce great, high-quality programmes for children even with smaller budgets, and to make them accessible to a global target group through AI-generated language translation. ■

NOTE

¹ <https://prixjeunesse.de/share-pool/for-us-no-planet-b/>; https://izi.br.de/english/publication/television/37_2024_E/What_we_can_do.pdf [20.2.25]

*Nic Buchanan is CEO of the 2D animation studio Pixcomm, South Africa.

