# What have we learnt about generative AI and journalism

For six years I have been investigating the value of artificial intelligence for news publishers with our global JournalismAI project at the LSE, supported by the Google News Initiative. It's been an exciting journey that has been shared by hundreds of journalists around the world who are intrigued by the benefits of machine learning, automation, and data discovery. But in the last 18 months the technology has made a leap and suddenly everyone is talking about and playing with 'generative' AI such as ChatGPT and DALL-E. So what happens next and what should you do about it?

This article will take you through a strategy of approaching generative AI. It will seek to avoid both the dystopian hype about robots taking over and the marketing froth. My professional opinion is pretty simple. Yes, this is a significant 'step-change' in the technology, but it's not a miracle. Yes, it opens up incredible creative potential and an opportunity for efficiency gains. But it is also unreliable and rapidly evolving. It contains profound flaws and risks to the user, to publishers and to society in general. So approach with caution, but also with informed enthusiasm.
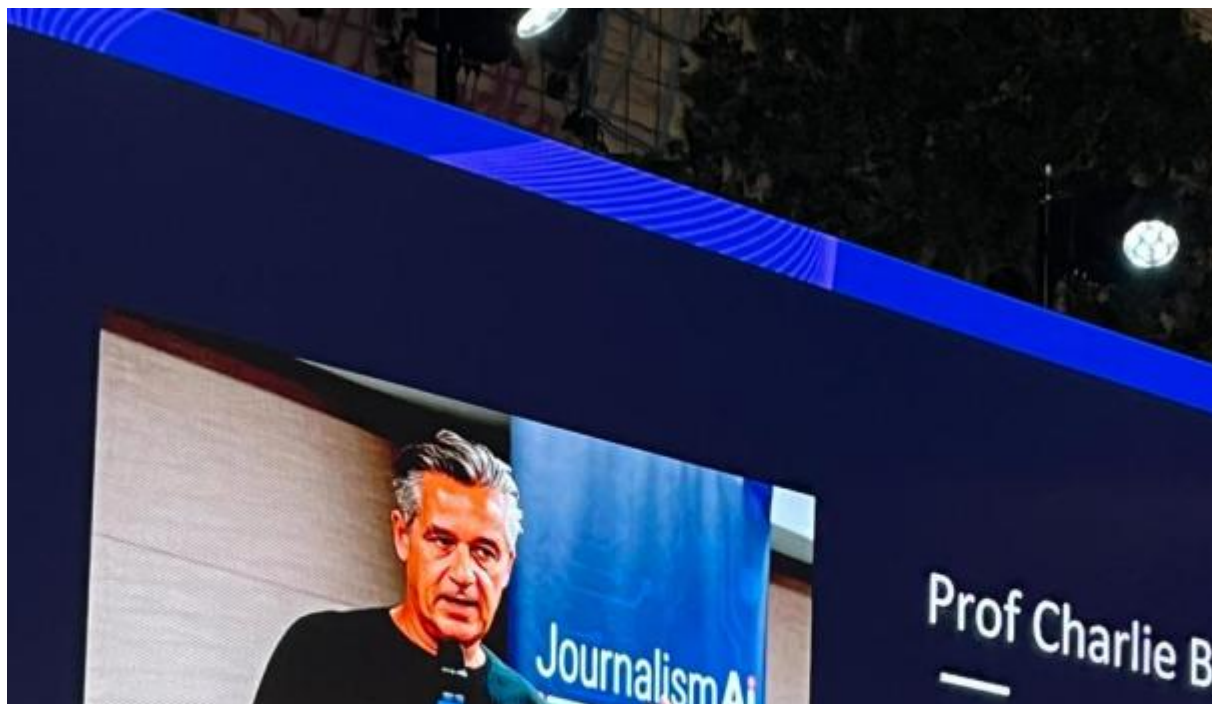
JournalismAI on tour

The Difference

So what is the difference between 'generative AI' and that bundle of technologies we called Artificial Intelligence before? This is not a technical article, but in brief, old AI was either supervised learning where an algorithm learns from labelled data – it is trained by

the programmers. It then makes predictions about new data that it is given. So you search for an image of a cat. It gives you an image of a cat because it's been taught to do that. It is a language prediction machine, not a truth machine. It doesn't 'know' anything.

GPT stands for "Generative pre-trained transformer". Generative AI is focused on generating new data. It uses statistical models to learn the underlying patterns of a given dataset and then it generates new data points. It does this by using vast 'Large Language Models' that enable it to respond coherently to the questions, or 'prompts' that you give it. It seeks to predict the answer you want. It does not 'know' or 'understand' the answers it gives. It is not sentient. But it is brilliant at giving the appearance that it does.

Programmes such as ChatGPT or image generators such as Midjourney are successful partly because they have a brilliant UX. You simply give it a prompt and something happens. If you have tried it, you can't fail to be impressed, excited, perhaps even a little spooked by its ability to create responses to everything from requests for jokes or poems to answers to the meaning of life.

As you will have seen or read, it is possible to reach its limits pretty quickly. A programme like ChatGPT or Google's Gemini is designed to tell you what it can't do. It will seek to remind you that it is only a piece of software, not a human being. If you ask it a simple question it will give you a pretty sensible, orthodox answer. If you engage in a sustained dialogue with complex ideas it will struggle. It might be AI but it is not truly 'intelligent'. It is prone to 'hallucinations'. These are not deliberate 'lies' because generative AI has no concept of the 'truth'. It is only trying to predict what the answer should be. And so it might make up false facts or fake sources because it predicts they should exist. More on those flaws later.

Extraordinary Potential

The potential for publishers is quite extraordinary. Here is an [article with some great use cases.](#) We have already seen how publishers can use previous AI technologies in creative ways. The Washington Post's Heliograf automatically writes articles on simple stories such as sports scores or election results. Reuters' News Tracer allows it to spot, track and help verify news stories breaking on social media. The New York Times' Project Feels automatically generates personalised summaries of stories that seek to

understand the emotional tone of an article.

Generative AI offers the opportunity to expand that exponentially. But at the moment, I would never use current generative AI tools to publish directly without human oversight. It is just too risky. Instead, it is best to think about generative AI as a set of tools that might supplement your workflow and augment your newsgathering, content creation and distribution. Some of the gains will be invisible to consumers. Coders say that generative AI is giving them massive gains in programming efficiency. Video and audio editors say that it is giving them clever shortcuts.

Journalists tell me that ChatGPT and its equivalents are very useful for helping them create. You can use it to brainstorm ideas, to try out different styles or formats. You can give it a relatively large text – perhaps an academic article – and ask it to summarise it in plain English or bullet points. But for anything at all complex you have to check the outcome. You can't rely on it. Perhaps it is not so different to using content from social media, websites or even copy from news agencies or other media. It is not infallible.

The same applies to image generation. Generative AI pictures are completely 'made up'. They do not use existing imagery directly. They are predicting what your prompt has suggested. This might make them useful for graphics, marketing, or article illustrations. They might well be more interesting than stock photos. But the whole point of news imagery that reports on what is really happening in the world is that it must be authentic. In that sense, this is a pretty straightforward distinction that publishers have been making for a long time.

So, enormous potential benefits, but also serious risks. I see three kinds. Firstly, the universal dangers of generative AI. The data sets can be biased or incomplete. The algorithm can 'hallucinate' giving false answers or sources. Then secondly there are the general risks to publishing and the news media. Who is responsible if it makes a mistake? Is it invading privacy or appropriating other people's content? (including publishers') Is it potentially an unfair competitor that will also supercharge the spread of misinformation and propaganda? And then thirdly, the specific dangers for journalists or publishers. I've already raised a few of those in this article. And generally, those can be dealt with by observing the usual best journalist practice. Edit. Check everything, have multiple sources, and then review the results.

So what should be your strategy to benefit from generative AI? My advice is not so different to approaching more basic forms of AI.

Everyone In Your Organisation Needs Some Generative AI knowledge

Make sure everyone in your organisation has some basic knowledge about Generative AI. There are loads of introductory articles and courses. Your people need to know about AI and generative AI, because they will definitely be using it, in the future they will use it more, and the world they report on will be using it too.

Assign Responsibility for Generative AI

Identify people in your organisation who can take responsibility for getting across this technology and it's fast-moving development. You need to understand how it might relate to your workflows, mission and business model. How should you change your hiring policy and skills-sets? It is vital that they work across the organisation, gathering insights and views from ALL staff. Address concerns and spread best practice.

Explore

Explore existing use cases. This about what problems it might solve or ameliorate. What processes might it enhance? What things need to be kept 'human'? Play around the tools to build your confidence and AI literacy. Create a culture of confidence, not fear, around generative AI.

Iterate and Review

Start small. Try and then try again. Do it but then review. Is it saving time? Is it accurate? Does it enhance the quality of the content and the user experience? Could you develop new products or systems?

Guidelines

Mitigate ethical and editorial risks. Think through the potential dangers to your reputation. Put in place guidelines – there are good examples out there already that are sensible and helpful. But think it through for yourselves and involve all staff in that process. Every publisher is different. Guidelines can give confidence to your staff but also send a positive message to advertisers/funders and the audience.

This is yet another challenge isn't it? You are already busting a gut just to keep afloat and thriving. And here we go again. First you went online all those years ago. Then you had to deal with social media about a decade ago. Perhaps you are now truly 'digital first'. Then along came artificial intelligence and perhaps you used it to improve your subscription processes or to personalise newsletters? Now once again, there's a new technological kid in town demanding attention and threatening to disrupt everything.

A top digital news executive said to me recently, we have to ask if we have the desire, time and skills to handle generative AI. Well, if you got this far through all the years of change, then perhaps you have. The cliché is that we exaggerate the short-term effect of these technology leaps, and underestimate their long-term impact. That might well be true in this case. You do have time to see how this plays out. It is going to evolve quickly, though regulators might well step in to constrain its development. Some news organisations will do deals with the LLMs. Others may find their fundamental business model is under threat. But regardless of what happens, it is vital to pay attention to generative AI and to start the process right now of thinking through how it might change your working life and your business.

Charlie Beckett is a professor in the Department of Media and Communications at the LSE. He is director of the LSE journalism think-tank Polis and leads the [Polis/LSE Journalism and AI project](#)

JournalismAI at LSE is an open, global project with an array of free training courses, case studies, articles, research reports and tools to help you explore AI and publishing. We have a network of 10,000 people around the world facing the same challenges as you. [Please join us](#)!

A version of this article first appeared in [inpublishing magazine](#)