TechScape: How cheap, outsourced labour in Africa is shaping AI English
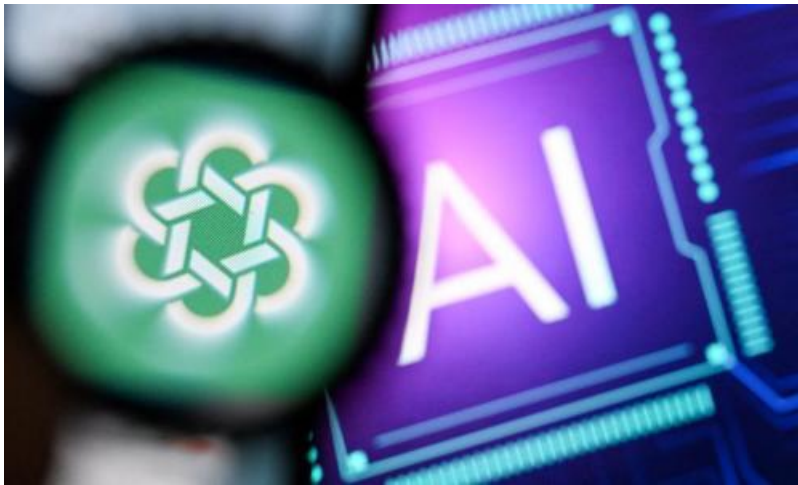
Publication Date: 2024-04-16

Author: Alex Hern

Section: Technology

Tags: Technology, TechScape newsletter, ChatGPT, Gadgets, Artificial intelligence (AI), Language, Nigeria, Africa, newsletters

Article URL: https://www.theguardian.com/technology/2024/apr/16/techscape-ai-gadgest-humane-ai-pin-chatgpt

We're witnessing the birth of AI-ese, and it's not what anyone could have guessed. Let's delve deeper. If you've spent enough time using AI assistants, you'll have noticed a certain quality to the responses generated. Without a concerted effort to break the systems out of their default register, the text they spit out is, while grammatically and semantically sound, ineffably generated. Some of the tells are obvious. The fawning obsequiousness of a wild language model hammered into line through reinforcement learning with human feedback marks chatbots out. Which is the right outcome: eagerness to please and general optimism are good traits to have in anyone (or anything) working as an assistant. Similarly, the domains where the systems fear to tread mark them out. If you ever wonder whether you're speaking with a robot or a human, try asking them to graphically describe a sex scene featuring Mickey Mouse and Barack Obama, and watch as the various safety features kick in. Other tells are less noticeable in isolation. Sometimes, the system is too good for its own good: A tendency to offer both sides of an argument in a single response, an aversion to single-sentence replies, even the generally flawless spelling and grammar are all what we'll shortly come to think of as "robotic writing". And sometimes, the tells are idiosyncratic. In late March, AI influencer Jeremy Nguyen, at the Swinburne University of Technology in Melbourne, highlighted one: ChatGPT's tendency to use the word "delve" in responses. No individual use of the word can be definitive proof of AI involvement, but at scale it's a different story. When half a percent of all articles on research site PubMed contain the word "delve" – 10 to 100 times more than did a few years ago – it's hard to conclude anything other than an awful lot of medical researchers using the technology to, at best, augment their writing. According to another dataset, "delve" isn't even the most idiosyncratic word in ChatGPT's dictionary. "Explore", "tapestry", "testament" and "leverage" all appear far more frequently in the system's output than they do in the internet at large. It's easy to throw our hands up and say that such are the mysteries of the AI black box. But the overuse of "delve" isn't a random roll of the dice. Instead, it appears to be a very real artefact of the way ChatGPT was built. A brief explanation of how things work: GPT-4 is a large language model. It is a truly mammoth work of statistics, taking a dataset that seems to close to "every piece of written English on the internet" and using it to create a gigantic glob of data that spits out the next word in a sentence. But an LLM is raw. It is tricky to wrangle into a useful form, hard to prevent going off the rails and requires genuine skill to use well. Turning it into a chatbot requires an extra step, the aforementioned reinforcement learning with human feedback: RLHF. An army of human testers are given access to the raw LLM, and instructed to put it through its paces: asking questions, giving instructions and providing feedback. Sometimes, that feedback is as simple as a thumbs up or thumbs down, but sometimes it's more advanced, even

amounting to writing a model response for the next step of training to learn from. The sum total of all the feedback is a drop in the ocean compared to the scraped text used to train the LLM. But it's expensive. Hundreds of thousands of hours of work goes into providing enough feedback to turn an LLM into a useful chatbot, and that means the large AI companies outsource the work to parts of the global south, where anglophonic knowledge workers are cheap to hire. From last year: The images pop up in Mophat Okinyi's mind when he's alone, or when he's about to sleep. Okinyi, a former content moderator for OpenAI's ChatGPT in Nairobi, Kenya, is one of four people in that role who have filed a petition to the Kenyan government calling for an investigation into what they describe as exploitative conditions for contractors reviewing the content that powers artificial intelligence programs. I said "delve" was overused by ChatGPT compared to the internet at large. But there's one part of the internet where "delve" is a much more common word: the African web. In Nigeria, "delve" is much more frequently used in business English than it is in England or the US. So the workers training their systems provided examples of input and output that used the same language, eventually ending up with an AI system that writes slightly like an African. And that's the final indignity. If AI-ese sounds like African English, then African English sounds like AI-ese. Calling people a "bot" is already a schoolyard insult (ask your kids; it's a Fortnite thing); how much worse will it get when a significant chunk of humanity sounds like the AI systems they were paid to train? AI hardware is here The world of atoms moves more slowly than the world of bits. The November 2022 launch of ChatGPT led to a flurry of activity. But where digital competitors launched in a matter of weeks, we're only now starting to see the physical ramifications of the AI revolution. On Monday, AI-search-engine-for-your-mind startup Limitless revealed its first physical product, a $99 pendant that you wear on your shirt to record, well, everything. From the Verge: The $99 device is meant to be with you all the time … and uses beam-forming tech to more clearly record the person speaking to you and not the rest of the coffee shop or auditorium. Limitless can do a lot to help you keep track of conversations. What was that new app someone mentioned in the board meeting? What restaurant did Shannon say we should go to next time? Where did I leave off with Jake when we met two weeks ago? In theory, Limitless can get that data and use AI models to get it back to you any time you ask. It's a genuinely exciting space to cover because no one actually knows what AI hardware should be. Limitless has one answer; Rabbit has a very different one, with its R1: R1 is built as an intuitive companion device that saves users time. While phones have evolved into all-encompassing personal entertainment devices in recent years, r1 is positioned as a standalone hardware portal to cut through distractions and help users handle their everyday digital tasks smarter, more efficiently, and more delightfully. Looking like a small, square smartphone, the R1 is a push-button partner to an AI agent which, the company says, can be trained to carry out tasks on your behalf. The physical object, designed by renowned consultancy Teenage Engineering, looks delectable, but the whole thing rides on whether the AI agent at its heart can actually be trusted. At its best, it could bring powerful AI assistants into our daily lives; at its worst, it would just make you nostalgic for Siri. And the worst is not impossible. Humane is the first major company to get AI hardware to market, with its AI Pin – and it's not gone well. From the Verge's review: As the overall state of AI improves, the AI Pin will probably get better, and I'm bullish on AI's long-term ability to do a lot of fiddly things on our behalf. But there are too many basic things it can't do, too many things it doesn't do well enough, and too many things it does well but only sometimes that I'm hard-pressed to name a single thing it's genuinely good at. None of this – not the hardware, not the software, not even GPT-4 – is ready yet. The AI pin isn't going to be the last piece of AI hardware we see, then. But it might be Humane's last. If you want to read the complete version of the newsletter please subscribe to receive TechScape in your inbox every Tuesday.