BioSpectrum
the business of Bi
24 years of leveraging the business of Bio
www.biospectrumindia.co

Suffering From GIGO? Turn to GRDs Instead, for better healthcare innovations

05 January 2024 | Views | By Jassi Chadha, President and CEO, Axtria

In the pharmaceutical space, generative AI-ready datasets or GRDs have the potential to revolutionise research and advance personalised medicine



Here's the Problem…

Let me get straight to the point. Many pharma and life sciences companies that dove into the Generative AI space are now dealing with a significant problem: GIGO. For the uninitiated, that means "garbage in, garbage out." It's an old computing term that has taken on new life amid healthcare's burgeoning AI enhancements. Very simply, it means training your artificial intelligence models on bad data means you'll get bad answers in the end.

Seems obvious, doesn't it? So why do so many pharma companies fall into that trap? Was it the excitement to get on the GenAI train? Was it budgeting reasons? No matter why they did it, healthcare companies are now finding out that their GenAI setups are suffering from hallucinations, and their output isn't quite good enough for guiding next-best-actions for field reps or chatbot responses to potential clients and patients.

It's very simply a matter of incomplete data at the onset. The idiom is "garbage in, garbage out," but the reality is that the data you put in is massively valuable. It is in no way garbage; it's just incomplete. You may be using structured data like those found in Excel sheets, customer databases, clinical trial results, and a myriad of other sources. But you're missing out on the gold mine of results and insights: unstructured data. This includes doctors' notes, recordings from patient exams, image scans, wearable data streams – any data that simply doesn't fit into a neat spreadsheet.

You need to find a way to get that treasure trove of information into your GenAI setup. And you do that by turning it into "generative AI-ready datasets" or GRDs.

Generative AI-Ready Datasets

By converting data into the kind that's ready for AI analysis, you can have GRDs that become far superior for applications like natural language generation and creative work. GRDs allow generative models like ChatGPT to learn and craft content that resembles the input data.

In the pharmaceutical space, GRDs have the potential to revolutionise research and advance personalised medicine. From an operational standpoint, GRDs can help build the template for future custom training models. They can help refine models already in use. And they can bolster bespoke medicine by using individual patient preferences and health information. Of course, this is all done with the utmost adherence to privacy and ethical guidelines.

Sounds Familiar, But It's Not

Introducing natural language interactions provides real potential, from faster data retrieval to a laser-like focus on a topic to time saved. The upside is limited only by your creativity. But simply adding a fancy chatbot to your framework won't cut it.

None of these benefits will be possible if your information assets aren't prepared. You need to turn them into generative AI-ready data.

We know that GenAI and the large language models (LLMs) that power it are prone to inheriting biases from the training data. Those biases are two-fold. First, it can only report on the information it was fed. If your statistics were flawed, if you had the wrong dates, if you underestimated the number of prescriptions, whatever your mistakes were at the start – those mistakes will carry over to the results.

But the second and more capacious bias is the style of answer those models will give. If you're not feeding it the unstructured data, you can't expect it to provide similar results. In other words, to paraphrase the idiom at the start of this piece: "human in, human out." If you train your GenAI model with doctor's notes, you can give results the way a doctor would say it. Feeding it the "human touch" will make the responses far easier for us in the physical world to understand.

Conclusion

It's not a suggestion at this point. To get the most out of your expensive AI setup, you must feed it the best information possible. That means no longer ignoring the hard-to-analyze unstructured information. You must convert it to generative AI-ready datasets. All data is valuable, so why leave half of it on the table? Use it, and you'll uncover the hidden gems of insights you already possess. Forget "GIGO." With GRD, you'll have "insights in, better patient results out." And that gives us all a brighter future for healthcare innovation.

Jassi Chadha, President and CEO, Axtria