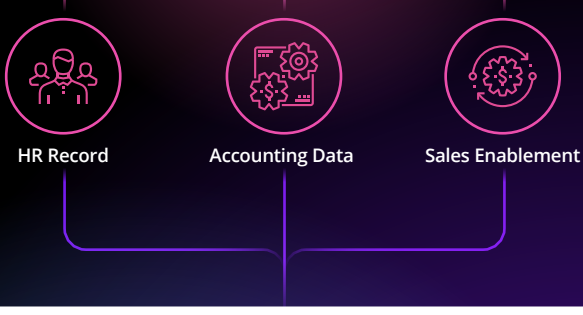
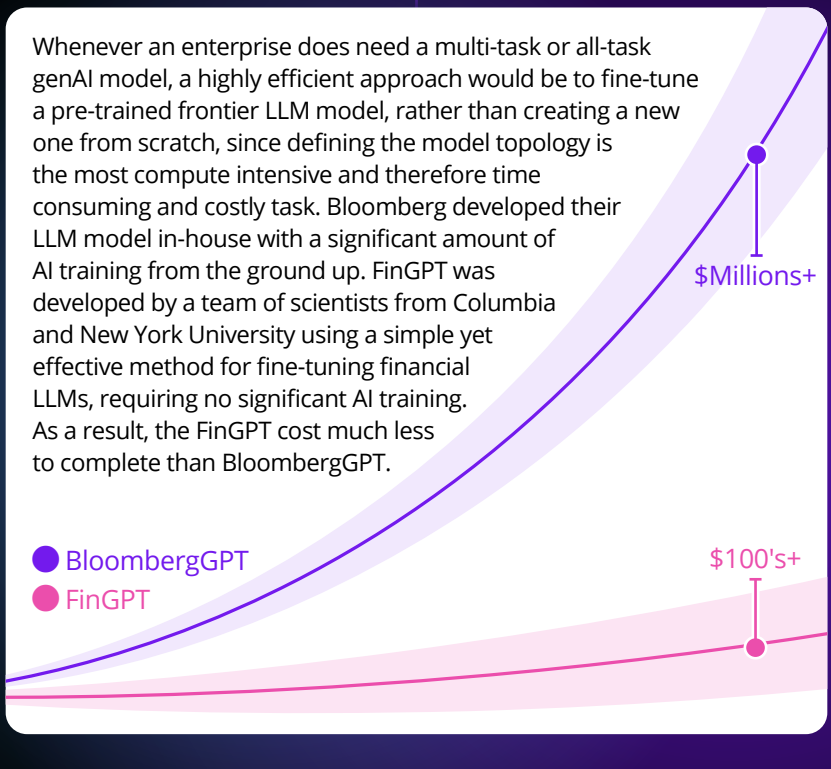
Brought to you by Informa Tech and commissioned by:

Generative AI At Our Fingertips

Most enterprises do not need a multi-task or all-task generative AI (genAI) application but rather a task- or domain-specific high-accuracy model. These are more likely to have a lower computing requirement and can be deployed on existing infrastructure, thus saving costs. At the same time smaller models are easier to manage, can be designed with daily updates from a real-time database, thus enabling better and more accurate results.



Whenever an enterprise does need a multi-task or all-task genAI model, a highly efficient approach would be to fine-tune a pre-trained frontier LLM model, rather than creating a new one from scratch, since defining the model topology is the most compute intensive and therefore time consuming and costly task. Bloomberg developed their LLM model in-house with a significant amount of AI training from the ground up. FinGPT was developed by a team of scientists from Columbia and New York University using a simple yet effective method for fine-tuning financial LLMs, requiring no significant AI training. As a result, the FinGPT cost much less to complete than BloombergGPT.



GenAI does not require a million-dollar supercomputer

GenAI models can even be trained on CPUs

BOLT2.5B

Generative AI model trained on **10 4th Gen Intel Xeon CPUs in 10 days**

Sun	Mon	Tue	Wed	Thu	Fri	Sat
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31				

10X

REDUCTION IN INVESTMENT

To find out more, read the full eBook here

www.intel.com/content/www/us/en/content-details/793579/generative-ai-at-our-fingertips.html

Omdia.com
Omdia OmdiaHQ

Intel.com
Intel Intel