

GenAI’s powerful growth is triggering a new breed of specialist data centres

July 2024

Johann Adjovi

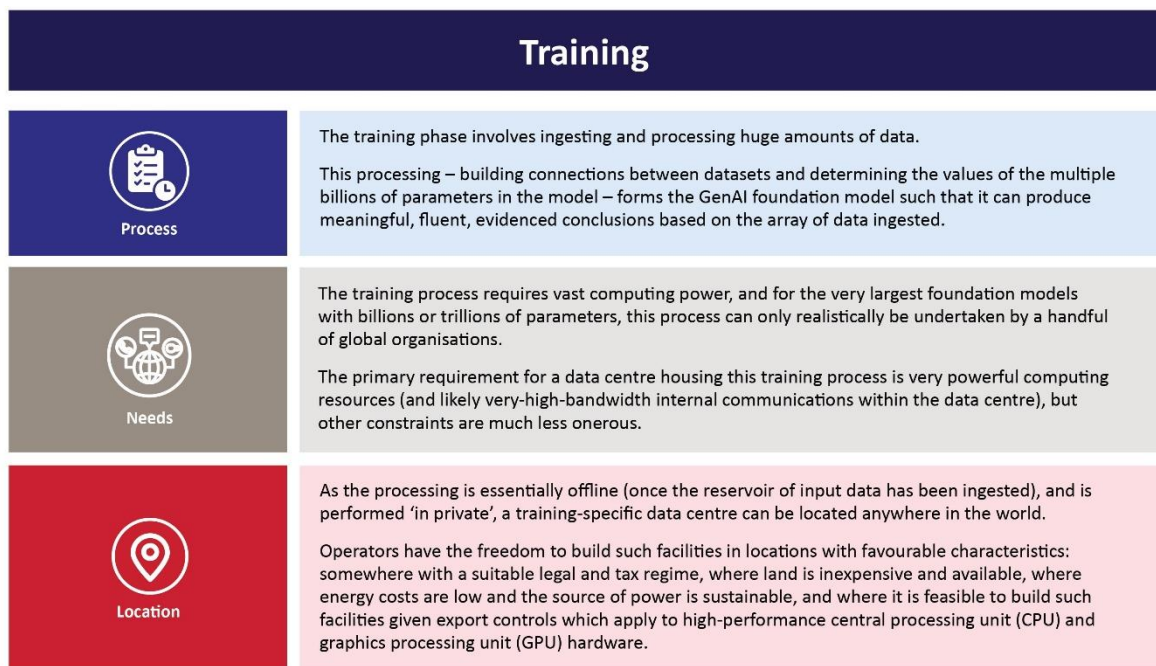
The booming AI ecosystem is driving change in every corner of the TMT landscape. While the rise in computational demand is fuelling growth in data-centre capacity in general, it opens up some very specific, immediate opportunities for differentiated data centres that are designed to meet the particular needs of different stages of the development of GenAI models.

AI training and AI inference have very different computational demands

There is already significant differentiation within the data-centre value chain. The evolving demand for AI offers operators (and their investors) an immediate opportunity to create and build on a new proposition: catering specifically for individual stages in the development of AI models.

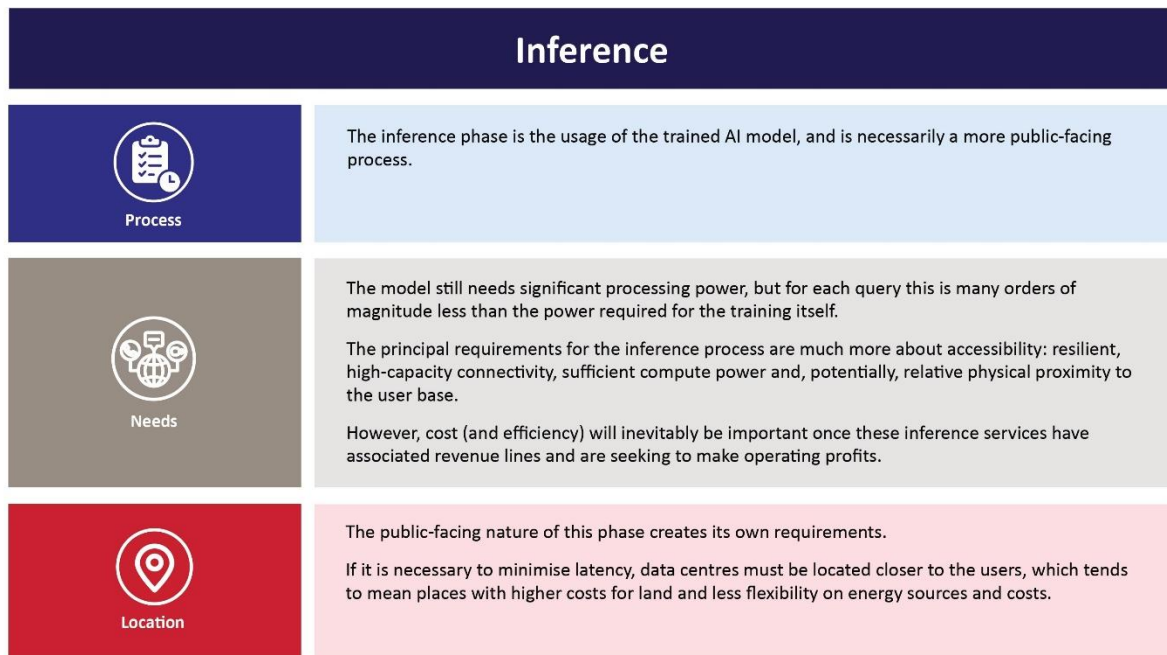
The process of *training* a generative artificial intelligence (GenAI) foundation model has particular requirements in terms of processing power and energy consumption; the needs of the *inference* process are very different.

Figure 1: Overview of the AI training process



Source: Analysys Mason

Figure 2: Overview of the AI inference process



Source: Analysys Mason

Two distinct types of AI-based opportunities therefore emerge:

- Hyperscale data centres with very high computing intensity designed for efficiency, which are especially well suited to the training phase of AI models. Low-cost, high-power facilities take care of the brute-force training process, which generates little or no direct revenue in itself.
- Smaller-scale data centres may offer a better profile for the processes of fine-tuning and inference. These facilities can be smaller but may need to be more widely distributed, highly resilient and scaled according to local demand. Ideally, they would have high-bandwidth, low-latency connectivity.

Differentiation across the data-centre value chain creates immediate opportunities for data-centre operators (and their investors) to build facilities that can meet each specific (and growing) need. It also produces a geographical opportunity for governments seeking to position their country as a centre of excellence in GenAI.

Nordic countries would be positive candidates for the deployment of hyperscale data centres for AI training, given their access to plentiful renewable energy and fresh water for cooling, as well as sparsely populated regions. In the Middle East, Oman (Muscat, Salalah) or Bahrain (Manama) could be attractive hosts for AI training data centres owing to the abundance of land, sufficient data-centre know-how and low-cost power (although investment in low-carbon energy generation will be a prerequisite for many potential investors).

Major cloud markets such as UAE (Dubai, Abu Dhabi), Saudi Arabia (Riyadh, Dammam) or Qatar (Doha) could be good locations for AI inference data centres. In Europe, the major financial centres of Frankfurt, London, Amsterdam and Paris might also be natural candidates for such data centres, though availability of power and land could pose constraints.

If governments and policy makers are to position their countries as GenAI leaders, they must understand and adapt their strengths and weaknesses as AI hubs, as well as cultivating positive relationships with leading international players.

In the longer term, ecosystem requirements are likely to change and players should monitor developments

The speed of change in AI creates uncertainty, and stakeholders must keep refining their thinking as the AI ecosystem develops. In addition to evolving demand for data-centre capability, there are changes relating to regulation and the handling of data, developments in the chipsets and cooling technology, as well as changing sustainability criteria. For the providers and investors that stay abreast of these developments, there are opportunities to serve and create a new market of specialist data-centre services.

About us

Analysys Mason is using its deep domain knowledge in the TMT industry to support stakeholders across the sector to take advantage of the opportunities brought about by GenAI, and to navigate the uncertainties of the current volatile period. For more information, please contact Johann Adjovi (Partner).