

Legal services

Latham & Watkins buys Nvidia servers to set up in-house AI systems

Law firm is customising open-weight models as alternative to OpenAI and Anthropic

KAYE WIGGINS — NEW YORK

Law firm Latham & Watkins has bought its own Nvidia GPU servers and started customising AI models, an expensive and technically demanding set-up that shows how much the technology is changing Big Law's business model.

Latham, the US's second-largest law firm with \$8.3bn in revenue last year, has in recent years purchased several servers, each of which holds multiple GPUs, the powerful computer chips that run AI models.

It is the first public example of a big law firm buying its own AI hardware and fine-tuning models. It makes Latham responsible for cyber security and operating the servers but gives it the ability to keep clients' sensitive data entirely within its own systems.

"Sometimes we may have information that is so sensitive, client information that we really want to protect, we don't want to put it to any cloud vendor," said Rene Mendoza, the firm's chief information officer.

He added that "with a lot of consumption costs coming" for how AI usage is priced, the firm wanted to have "flexibility".

Legal work, much of which is characterised by large volumes of information and formulaic or repetitive tasks, is among the sectors that AI is disrupting the most rapidly. Wall Street banks are pushing large law firms to cut fees because technology speeds up tasks.

But law firms have a disincen-

tive to make big investments in AI because upfront capital expenditure cuts into the profits distributed annually to equity partners. Firms have tended to subscribe to AI software services and cloud infrastructure instead.

Latham's decision to buy its own hardware marks a radical change. Its engineers are tailoring existing open-weight AI models to its needs, something more commonly associated with tech companies than law.

"I've been in a lot of conversations about AI, but I haven't heard of this," said Joy Heath Rush, chief executive of the International Legal Technology Association, a membership group for legal technology specialists. "It took me by surprise, and I don't get taken by surprise a lot."

The move stands in contrast to other law firms' approaches. Latham's larger rival Kirkland & Ellis is working with Palantir to develop AI tools, while A&O Shearman has worked with legal tech start-up Harvey, and Freshfields this year announced a deal to work with Anthropic.

Latham is still using AI products and services from large tech companies alongside building its own technology. It declined to say how much it has invested or is planning to invest in the new approach, which it has been working on for the past few years.

The cost of buying and running GPU servers on that



Latham is the first public example of a big law firm buying its own AI hardware and fine-tuning models © Latham & Watkins

scale and employing the specialists needed to run, maintain and protect the system can run into the tens of millions of dollars per year, potentially rising to hundreds of millions as technology advances.

Latham leases space in a data centre facility that is locked and only accessible by its own staff. The firm's team of machine learning engineers and software engineers are fine-tuning Nvidia Nemotron 3 open-weight models.

Open models, unlike those from OpenAI and Anthropic, make their design public and allow users to download and tailor them and run them using their own hardware.

Nvidia boss Jensen Huang has said he hopes open models will help bring in a wider range of business customers for AI and related hardware beyond tech groups.

Latham said it employed more than 900 technology specialists. "Over the past few years, the increase in the number of professional technology staff, innovation lawyers, machine learning engineers, AI engineers and coding lawyers has been significant," said Michael Rubin, a partner at Latham and chair of its AI strategy committee.

The approach allows Latham lawyers to use a wider range of models than ChatGPT, Claude and Gemini. They could decide whether to use the firm's own AI systems or commercially available tools based on the nature of the task, Mendoza said.

It also gives Latham alternatives if providers such as OpenAI and Anthropic raise their prices or terms. "We are not hitching our wagon to one particular company," he said.