POWER OF 2
SOLUTIONS POWERED BY PARTNERSHIP

Lenovo | NVIDIA

The Five Most Common IT Hardware Mistakes in AI

And Why Top AI Teams Avoid Them

AI is only as good as the IT infrastructure it runs on

With AI, you have the potential to deliver breakthrough innovations, solve intractable business problems, and strengthen the competitive advantage of your organisation.

But as smart as your algorithms may be, the results you can expect are only as good as the IT infrastructure they run on.

Despite this, you might be tempted to undertake AI projects without consulting your IT department in the belief it will save you time and help you get what you want.

But in our experience, this leads to mistakes and complexities that reduce the impact of your AI projects, increase their cost, and put organisational objectives at risk.

Here we share the five most common IT hardware mistakes you can make in AI and explain why you need specialist expertise to select the best IT hardware solution for your project.

Armed with this information, you'll get better results faster from your AI investment and contribute to the overall success of your organisation.

Happy reading!

From the Lenovo & NVIDIA team

Mistake #1
Processing power & accelerator

Mistake #2
Redundancy

Mistake #3
Insource or outsource

Mistake #4
Monitoring & management

Mistake #5
Core competence

Mistake #1

Processing power & accelerator

❌ **The ever-increasing size of AI models**, along with the amount of data needed to train them, creates intense resource demands that existing IT compute capabilities cannot meet. For this reason, your most fundamental barrier to the delivery of business value is the incorrect specification of processing power. Wrong specification can result in you being restricted to simpler models, longer training times and slower inference.

DOWNSIDE
RESTRICTED TO SIMPLER MODELS, LONGER TRAINING TIMES AND SLOWER INFERENCE

✅ **To meet the demands of AI**, specialised IT hardware units have been designed to accelerate these tasks. Depending on your application, you will likely need either a GPU, TPU, VPU, FPGA or ASIC. The correct specification can allow you to increase the number of parameters in your model, reduce training time from weeks to hours, and deliver inference in real-time.

UPSIDE
INCREASE THE NUMBER OF PARAMETERS IN YOUR MODEL AND REDUCE TRAINING TIME FROM WEEKS TO HOURS, AND DELIVER INFERENCE IN REAL-TIME

Mistake #2
Redundancy

❌ **A degree of redundancy is valuable** when planned into an overall environment, but it gets out of hand when each business function buys its own IT hardware dedicated to its own AI project. If your AI hardware is outside of IT's purview, it may lie idle and even get completely forgotten when employees are sick or leave their job, resulting in a costly waste of resources.

DOWNSIDE
WASTE OF RESOURCES

✅ **You will be far more resource-efficient** if you create an IT hardware cluster that all AI projects can access, taking advantage of virtualisation and containerisation to eliminate excess IT hardware. This will also reduce the energy required to run and cool your equipment, with the overall effect of helping your organisation reduce its environmental footprint. These resource savings also translate into financial savings, and you can avoid further expenses such as unnecessary software and OS licenses.

UPSIDE
LOWER SPEND ON IT HARDWARE, ENERGY AND SOFTWARE LICENSES, LOWER ENVIRONMENTAL IMPACT

Mistake #3

Insource or outsource

❌ **The promise of public cloud solutions** is convenience, but monthly costs quickly add up to significant sums. And if you develop your solution in a public cloud but deploy it on the edge, you should be aware of issues with connectivity, data privacy and latency, as well as complexity around monitoring and management.

DOWNSIDE
EXPENSIVE SOLUTIONS WITH POOR CONNECTIVITY, DATA PRIVACY & LATENCY

✅ **Buying your own IT hardware** will work out cheaper long-term. Or if you want to avoid capital expenditure, you can take advantage of pay-as-you-go or consumption-based options, such as Lenovo's TruScale. These enable you to scale your IT hardware according to your usage while retaining the peace of mind of on-premise security and control.

UPSIDE
A COST-EFFECTIVE SOLUTION WITH ON-PREMISE SECURITY & CONTROL

Mistake #4

Monitoring & management

❌ **Without an IT infrastructure specialist** on your AI team, your IT hardware may not be monitored and managed proactively and a reliable failover solution may not be in place. This can result in a higher risk of interruptions to AI jobs and disruption to your operation.

DOWNSIDE
HIGH RISK OF INTERRUPTIONS & BUSINESS DISRUPTION

✅ **With IT on your project,** you can be sure that your machines will be sufficiently monitored and managed, that routine maintenance will be performed during AI downtime, ensuring your AI application is available and reliable when you need it.

UPSIDE
HIGH AVAILABILITY & BUSINESS CONTINUITY

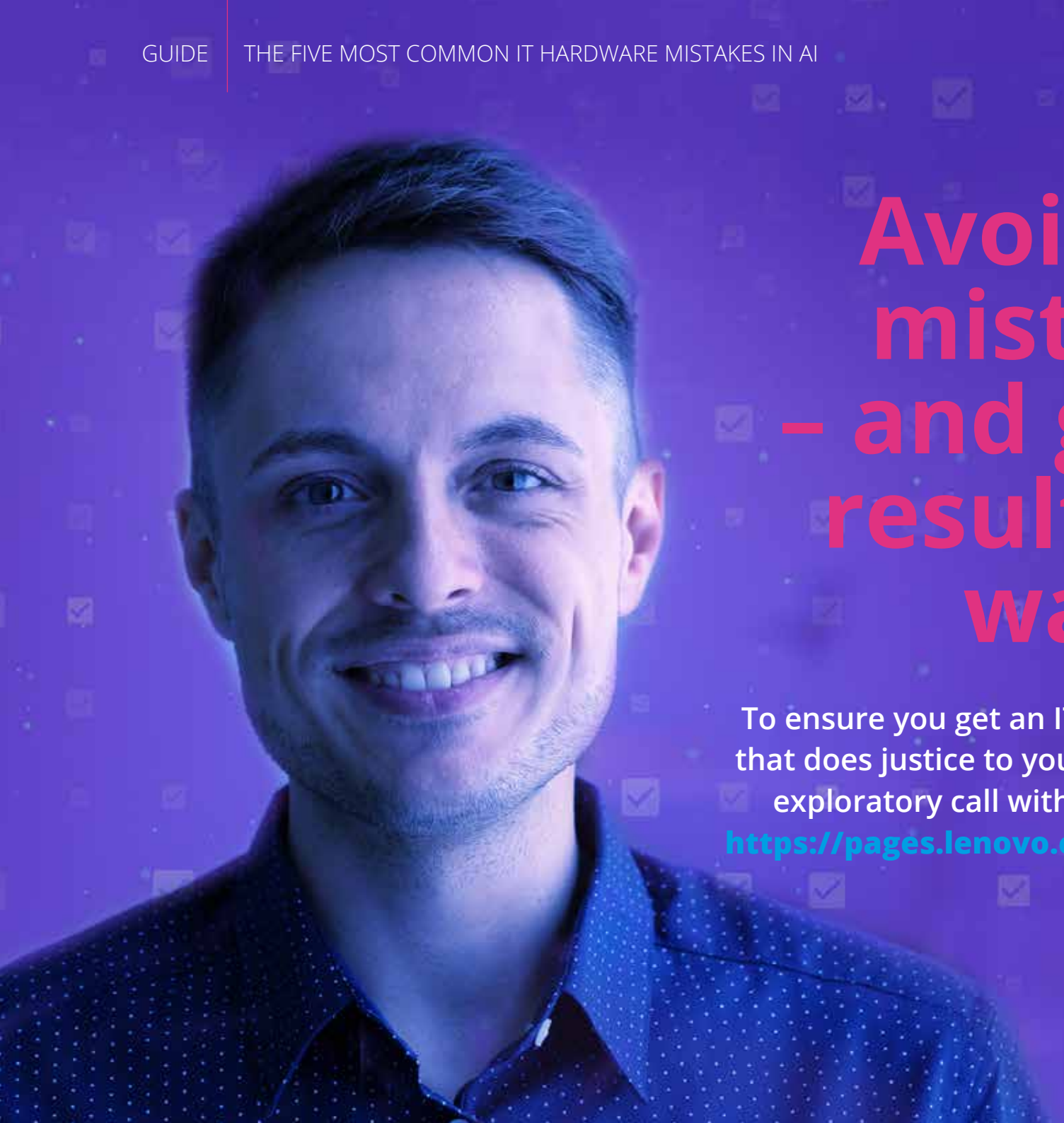
Mistake #5
Core competence

❌ **The reason you hire a data scientist** is to create AI. So when your data scientists spend time selecting, buying, installing, configuring, monitoring and maintaining IT hardware you are not getting the most out of them. A far better use of such a costly resource is to focus on solving the problem you've set them.

DOWNSIDE
WASTING TIME ON NON-CORE COMPETENCES

✅ **IT Infrastructure specialists in your IT department** can handle IT hardware related tasks to a high standard and cost-effectively. They can create an environment where all your data scientists need to do is create a virtual machine or run a container – saving you many hours of valuable time and wasted money.

UPSIDE
FOCUS YOUR AI RESOURCES WHERE THEY CAN DELIVER THE HIGHEST ROI



Avoid the mistakes – and get the results you want

To ensure you get an IT hardware environment that does justice to your AI investment, book an exploratory call with our specialist team at **https://pages.lenovo.com/mistakes-in-ai.html**

Why choose Lenovo & NVIDIA IT infrastructure for AI?

Top performance
Award-winning, world-record holding servers, accelerators, computing and storage – benchmarked by ML Perf

Comprehensive support
Workshops and partners to help you take your AI project from idea to implementation

Try before you buy
Test your algorithms and code in our AI Center of Excellence

Gain advanced AI skills
Join our Deep Learning Institute

Optimise your productivity
With Lenovo Intelligent Computing Orchestration

For information on products and solutions visit
www.powerof2.nvidia.lenovo.com/emea/

POWER OF 2
SOLUTIONS POWERED BY PARTNERSHIP

Lenovo | NVIDIA

About Lenovo

Focused on a bold vision to deliver smarter technology for all, Lenovo is developing world-changing technologies that create a more inclusive, trustworthy, and sustainable digital society. By designing,engineering and building the world's most complete portfolio of smart devices and IT infrastructure, we are also leading an Intelligent Transformation – to create better experiences and opportunities for millions of customers around the world. To find out more visit **www.lenovo.com**

About NVIDIA

NVIDIA's invention of the GPU in 1999 sparked the growth of the PC gaming market and has redefined modern computer graphics, high performance computing and artificial intelligence. The company's pioneering work in accelerated computing and AI is reshaping trillion-dollar industries, such as transportation, healthcare and manufacturing, and fueling the growth of many others. More information at **https://nvidianews.nvidia.com/**