2024 / WHITE PAPER

Healthcare and life sciences embrace AI, struggle with infrastructure

Artificial Intelligence brings significant advances to medicine and life sciences by accelerating processes like genome sequencing, drug discovery, drug optimization, and disease diagnosis. AI guides surgeons and radiologists through complex medical procedures, increasing overall productivity of skilled staff. However, wider adoption of AI requires a robust and scalable IT infrastructure.

AI is improving a wide range of medical specialties and surgical procedures. Intelligent diagnostics detect diseases at earlier stages, enhance patient care, and improve health outcomes. Many interesting developments stem from AI's unique ability to rapidly and accurately search through large data sets for patterns. AI is used to help discover gene mutations and their associated diseases and to develop treatments and drugs that are focused on patients who will best respond.

The use cases for AI in medicine are almost limitless, unlike the resources and time of human experts. The quality of results generated by AI depends on human's ability to train an adequate model and apply it efficiently, yet speed and accuracy of the tasks performed by an AI often compete for the same hardware resources.

In partnership with:

intel partner Titanium

PRODRIVE TECHNOLOGIES

Data analysts, developers, data scientists, and AI engineers should not need to choose between faster or more accurate outcomes. This paper explores emerging AI use cases in healthcare, addressing common challenges, and offers recommendations for overcoming them.

Training in the cloud, then local inferencing

AI training often requires massive processing power acting on huge data sets and is usually best performed in the cloud. This domain is dominated by large suppliers like AWS, Microsoft Azure, and Google Cloud. Users benefit from purpose-built services and gain access to extremely scalable infrastructure, which is crucial during phases involving extensive experimentation and optimization.

Inferencing, on the other hand, can efficiently utilize the trained model on-premises. It generally needs much less processing power and memory and requires access to sensitive patient information. For these reasons, AI processors are increasingly being built into on-premises medical equipment like hospital monitoring systems, endoscopes, X-ray machines, and MRI scanners.

Integrators often approach the shift from cloud to on-premises computing with reluctance. The lack of medical certification can hinder or even prevent integration of AI hardware in medical equipment. Prodrive Technologies provides medically certified hardware that conforms to standards such as IEC 60601-1 (basic safety and essential performance of medical equipment) and IEC 62304 (software life cycle processes of medical device software). Developers also often voice concerns that moving to on-premises means losing access to cloud software services. In close collaboration with partners, Prodrive provides software development kits and supports customers throughout an entire project to accelerate AI software deployment.

Prodrive Technologies offering includes IEC 60601-1 and IEC 62304 certified products

Deep learning helps sequence and analyze genomes

Some of the most exciting opportunities for AI lie in genome sequencing. The cost of processing patient genomes has dropped phenomenally over the last few years, with the price of sequencing a human genome dropping significantly below $1000 (Figure 1). Cloud-based AI training platforms working on increasingly large datasets will classify genetic variations and mutations with ever greater accuracy, forecasting disease development based on genetic and external factors. For instance, AI models will predict which patients are susceptible to specific environmental factors and will detect disease-linked genetic mutations directly from sequenced RNA datasets.

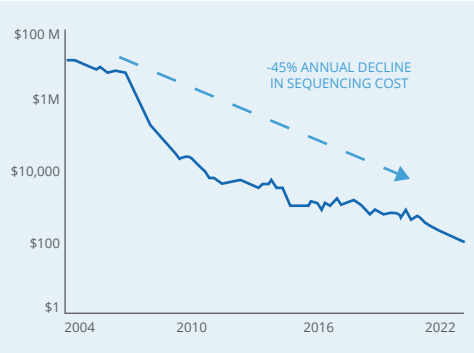
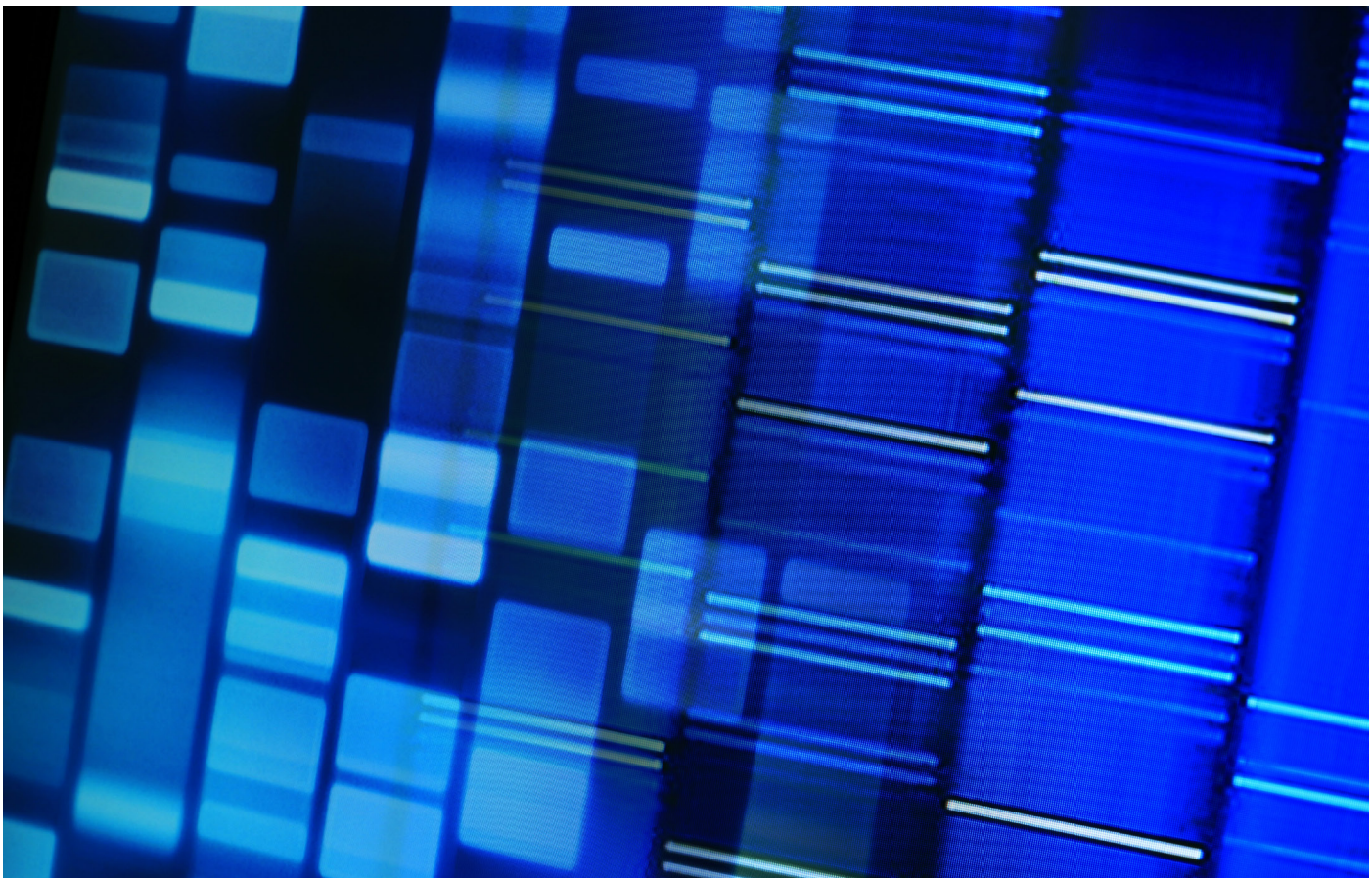


Figure 1: Cost of sequencing human genome (log scale)

Source: National Human Genome Research Institute, Illumina, Complete Genomics. Data as of 12.31.2022

Sequencing a whole human genome generates 100GB of raw data, but the output can be shrunk to less than 1% of the original size with processing. Repetitive and predictable sequencing is well suited for onsite processing, which eliminates the need for high-bandwidth internet and cloud access. The ability to perform in-situ analysis shortens diagnostic time, especially in remote locations with less developed IT infrastructure.

Cloud-based AI training platforms will classify genetic variations with greater accuracy



Genome sequencing for patients visiting their doctors will increasingly be done locally for early diagnosis. The 5th Generation Intel® Xeon® scalable processors aid in this by offering a 30% increase in genome mapping speed compared to the previous generation and a massive 250% increase compared to 5-year-old hardware. Although patients usually wait several days or even weeks for sequencing results, the time can be cut to less than a day with massive processing power of multiple sequencers working in parallel.

Prodrive Technologies' Zeus Xeon 5th Gen scalable series rack server is a great choice for an efficient backbone of advanced AI healthcare systems, such as gene sequencers and analyzers. Combining two 5th Gen Intel® Xeon® scalable processors with up to six double-width GPUs or up to 18 single-width PCIe cards, it is one of the densest air-cooled servers available. The server works efficiently whether used standalone or in clusters and can support both small sequencers and large whole-genome sequencing. Such applications further benefit from super-fast, 63GB/s 5th generation PCIe interface and 8-channel DDR5 memory controller operating at speeds up to 5600 MT/s.

By offering high reliability, low maintenance, and a long life cycle, the Zeus series servers provide a solid foundation for customizations. The options are not limited to selecting the main computing components but can extend to implementing custom BIOS, deploying specific software revisions, and incorporating hardware adjustments to the electronics for optimized cost and performance.

> "The advances in computing hardware give us wonderful opportunities to increase diagnosis speed, and smart AI algorithms will become weapons to combat invasive diseases."

- Bartosz Straszak
Systems Architect at Prodrive Technologies

Zeus Xeon 5th Gen Scalable series rack servers form the backbone of AI infrastructure

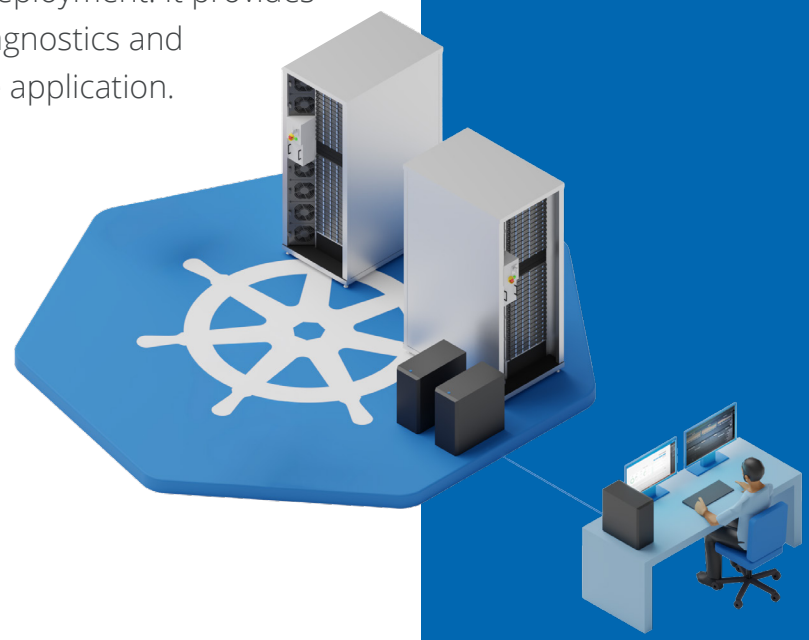
Accelerating drug discovery

New drug research is extremely expensive, partially because so many potential candidates fail. Genetic analysis can now help by revealing mutations that cause diseases or fuel the growth of cancers. Targeting drugs based on the genes linked to the diseases allows for the development of drug combinations and treatments and helps predict which patients will respond best. Pharmaceutical manufacturers have been trialing AI-developed drugs for diseases such as pulmonary fibrosis, dermatitis, ulcerative colitis, ovarian cancer, and various inflammatory and autoimmune conditions.

Accelerated drug discovery requires vast processing resources, and thus is usually done in the cloud. Cloud-based services can be massively scaled, utilizing hundreds of servers located in public data centers. However, cloud utilization brings its own issues, including potential privacy and security challenges.

Setting up on-premises data centers creates the burden of managing individual machines that are shared between multiple applications and organizing computing tasks to make the best use of available resources.

Orchestration platforms such as Prodrive Kubernetes Services (PKS) aid infrastructure teams by providing abstraction layers for the hardware, allowing them to focus on their algorithms. PKS offers not just containerization but an entire orchestration suite aimed at supporting software deployment. It provides comprehensive management and diagnostics and scales the system to the needs of the application.



Prodrive Kubernetes Services supports software deployment and simplifies system management and diagnostics

Monitoring life support and documenting patient data

Continuously checking patients' temperatures, oxygen levels, heart rates, and other vital signs on hospital bedside monitors is an important application of local AI systems.

AI can also improve patient engagement and help monitor post-operative patient- and site-level data, enhancing adherence and compliance to treatments. This, in turn, improves outcomes and reduces hospital readmission rates.

Automatic documenting can reduce administrative work, allowing doctors to focus on patient care. While embedded platforms and single-board computers handle basic monitoring, more complex AI operations are delegated to a separate technical room equipped with workstations and 19-inch rack-based equipment. The 5th Generation Intel® Xeon® Scalable processors provide significant advantages in this environment, delivering a 22% speed improvement[1] in language-based tasks.

NOTE 1 22% faster BioGPT fine-tuning compared to 4th generation of CPUs

Image enhancement and augmented reality

Another field where local AI is bringing advances is image enhancement. AI sharpens outlines and enhances colors and contrast to reveal faint structures that might otherwise be missed. It can even highlight tools or objects of interest. Augmented Reality (AR) takes this further by combining computer-generated graphics with real images. Especially

Augmented reality aids medical personnel by providing reference data

useful in surgeries, AR provides navigation, reference data, and highlights areas of concern. Three-dimensional, labeled, and dimensioned models are projected onto actual organs to outline any abnormalities. Procedures are less invasive, and the process is also ideal for clinical training in specialties like orthopedics and neurosurgery.

Live video editing is often performed on uncompressed 4K streams to avoid compromising quality and introducing visual artifacts. Processing is done onsite to prevent delays in displaying images and disruptions related to networking and distributed computing. Equipment is typically placed in a technical room adjacent to the operating theatre or examination room. The AVIDIS platform from Prodrive Technologies connects video sources and endpoints using an IP-based protocol, handling real-time video distribution and processing. Multiple uncompressed video streams with up to 4K resolution are transferred and routed over an Ethernet-based network. With features such as cropping, scaling, interlacing, and overlay available out of the box, the system easily integrates with IT infrastructure for further AI enhancements.

System builders do not need to immediately resort to dedicated, high-end hardware to handle additional tasks. Many processors integrate smaller GPUs perfectly capable of tasks like video encoding. Prodrive's Poseidon 13th Gen Series industrial PCs address the needs of a cost-efficient, robust AI platform. The product line includes 1U (single-height) and 3U rack versions with increased expansion options (extra drive bays, RAID and PCIe slots). More compact form factors are available on request.

Intel Iris® Xe graphics integrated in 13th gen. processors include up to 96 Execution Units (EUs) that can be used for tasks such as user interface generation, encoding, and simple AI. Dedicated GPUs or FPGA add-in cards can be installed if more computing power is needed. Where space and cost-efficiency are priorities, Prodrive assists with custom, single-board designs that combine general-purpose CPUs with application-specific FPGAs sized according to the computing needs.

AVIDIS (top)

Poseidon 13th Gen Series industrial PC (bottom)

Expanding AI use will bring advanced and afforable care

AI is expanding the range and effectiveness of treatments that medical professionals can offer. As many countries face a shortage of medical personnel, AI is becoming a valuable tool to boost productivity. Medical and surgical equipment is increasingly embedding the hardware needed to run AI efficiently, with a projected market size of a hundred billion dollars within four years (Figure 2).

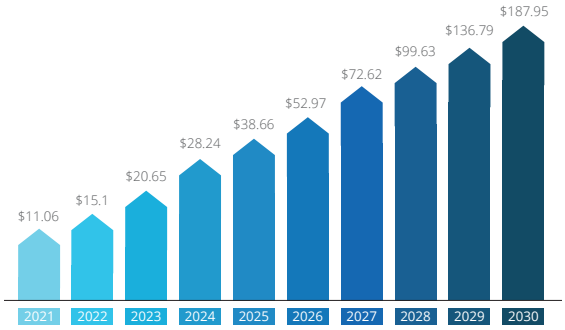


Figure 2: Predicted growth of AI in the healthcare market

Source: Precedence Research 2021 - 2030

> "AI is pushing every computing boundary. Increasing the performance of hardware combined with smart AI algorithms allows us to analyse data on the fly that, a decade ago, we could at best only store. This use of processing power will make healthcare and quality of life better and more affordable for an ever-increasing number of people."

- Bartosz Straszak
Systems Architect at Prodrive Technologies

Please contact us to explore how we can tailor AI solutions to meet your specific needs: contact@prodrive-technologies.com