Business Brief

Healthcare and Life Sciences
Graphics Solutions

Innovate in Healthcare and Life Sciences with Intel® Graphics Solutions

Unlock cutting-edge capabilities for healthcare and life sciences with open programming models based on Intel GPUs. Avoiding vendor lock-in from proprietary programming models helps reduce TCO in budget-tight environments while supporting innovation.

Key takeaways:

AI and compute-intensive workloads are increasingly prominent in the healthcare and life sciences sector. Intel enables these usages by complementing its CPU offerings with a unified set of GPU hardware at every scale:

- Intel® Iris® X^e and Intel® Arc™ integrated graphics
- Intel Arc graphics
- Intel Data Center GPU Flex Series
- Intel Data Center GPU Max Series

An open programming model based on **Intel's implementation of oneAPI** enables performance and productivity with cross-platform code that runs across GPUs, CPUs and other accelerator hardware, with no vendor lock-in.

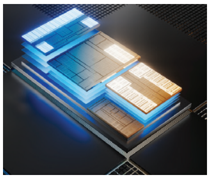
Even as data expansion transforms every industry, healthcare and life sciences face a unique scope of opportunities and challenges. Healthcare alone accounts for some 30% of the world's data production,[1] from digitalization across research, diagnostics and treatment. Data sources that range from genetic sequencing to medical imaging feed solutions for next-generation outcomes powered largely by visual computing and AI/machine learning and edge training/retraining. This evolution is helping drive increased deployment of GPUs in computer systems used across verticals, including healthcare and life sciences.

With U.S. healthcare spending at $4.5 trillion[2] — 17.3% of GDP — the AI healthcare market, valued at $11 billion in 2021, is projected to be worth $187 billion in 2030.[3]

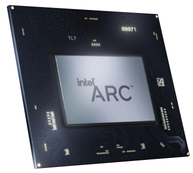
AI is being seamlessly added to every aspect of healthcare and life sciences, from medical imaging and patient monitoring to data analytics and drug discovery. At the same time, high-performance graphics and visual computing technologies are enabling visualization and display of rich datasets, helping make them more accessible and valuable to clinicians and researchers. These workloads are being processed at a growing range of locations across the distributed network, in the data center and cloud or at low-latency edge locations such as clinics, operating rooms and laboratories.

In the context of this rapid change, the openness and agility of healthcare and life sciences solutions is critical. Proprietary, licensed GPU coding models such as CUDA limit flexibility with vendor lock-in and lack of code portability across hardware that can drive up total cost of ownership (TCO). Flexible, heterogenous and open alternatives with supporting software toolkits for building solutions that embrace open standards and frameworks are paramount. Intel's offerings provide such an alternative, with a portfolio of integrated graphics in CPU and standalone discrete graphics cards for use from edge to the data center. This brief explores these offerings within the context of the healthcare and life sciences industries.

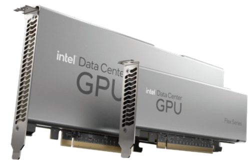
Intel integrated and discrete GPUs at every scale



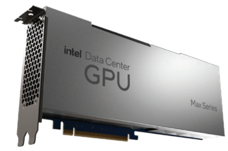
Intel® Iris® Xᵉ and Intel® Arc™ integrated graphics



Intel Arc GPUs



Intel Data Center GPU Flex Series



Intel Data Center GPU Max Series

embedded GPUs in Intel client CPUs

pair with Intel client and workstation CPUs

pair with Intel® Xeon® CPUs

Why are customers choosing Intel® Graphics solutions?

With Intel architecture, the emerging generation of AI-powered healthcare and life sciences solutions can deploy on any combination of existing infrastructure from client to edge to data center to cloud. Intel GPUs can share headroom with other heterogeneous solutions, enabling a flexible environment for high-speed innovation with low TCO.

- Avoids single vendor lock-in and adds another source for GPU hardware.
- Hardware neutral, performance-optimized programming with oneAPI.
- Flexible, cost-effective options at every scale.
- Scalability and portability from integrated GPUs all the way up to data center GPUs.
- Multiple virtualized instances possible in integrated and Data Center GPUs with memory management and software at no virtualization licensing cost.
- Interoperability for AI workloads across Intel platforms including CPUs, GPUs and Gaudi processors.
- Comprehensive customer support from world-class Intel engineers.

Hardware heterogeneity for healthcare and life sciences

For many years, solution developers have pursued code performance on GPUs with specialized programming methods such as CUDA. Such proprietary approaches to GPU programming can help meet demanding performance thresholds with specific hardware, but the code was not usable on other platforms. Resulting deployments were thus limited to the same closed development approach, without the ability to leverage available idle resources, including CPU headroom.

Today, almost half of developers target heterogeneous systems that incorporate more than one processor architecture, such as GPUs, CPUs and FPGAs.[4] Accordingly, medical device manufacturers, instrument equipment makers and others designing healthcare and life sciences solutions face growing hardware and software challenges. Inflexible requirements have produced untenable GPU supply limitations, with those basic economics driving up solution costs. Single-vendor proprietary lock-in and a closed software ecosystem make it more challenging still to satisfy evolving requirements. Contributing to the high expense is the rising licensing cost for GPU virtualization, creating TCO pressure on providers and implementors. Within the healthcare and life sciences industries, reducing cost and being able to explore alternatives is key in tight budget environments.

Intel takes an open ecosystem approach with its platforms including Intel GPUs that helps resolve those issues, freeing up platform choice to reduce switching costs, improve code portability, drive higher processor utilization and enable innovation. Hardware-neutral programming based on Intel oneAPI tools helps eliminate vendor lock-in and component supply issues. Flexible options offer balance between performance, cost and other factors, with scalability that ranges from integrated graphics to data center-class server GPUs. Multiple virtualized instances can be deployed on a single physical discrete GPU with no per-vGPU licensing costs (if used with Intel integrated GPUs or Data Center GPUs), and workloads are portable to CPUs and other hardware, taking advantage of otherwise-idle resources for outstanding flexibility and TCO.

The versatile value of Intel GPUs

Intel is helping drive greater GPU availability, diversity and flexibility with an extensive portfolio of integrated and discrete GPUs. Solutions can draw on the corresponding broad range of performance and scalability profiles for various use cases and deployment circumstances to optimize capabilities and TCO. All current Intel GPUs are based on the common Intel Xe architecture, which draws its name from the phrase "exascale everywhere." They consist of integrated GPUs based on Intel Iris Xe and Intel Arc graphics, as well as discrete GPUs that include Intel Arc GPUs for client and edge. They also include the Intel Data Center GPU Flex Series for deployment across the edge, cloud and data center.

Integrated graphics solutions

Intel Iris Xe and Intel Arc graphics

Intel Iris Xe and Intel Arc graphics built into Intel Core™ CPUs deliver cost-effective support for moderately demanding use cases such as medical PCs, portable ultrasound AI inference and many telehealth applications. These integrated graphics options can help reduce system cost and complexity by avoiding the need to purchase discrete GPU hardware, thus helping reduce initial equipment expense. Integrated graphics deliver power-consumption benefits compared to discrete GPUs as well, reducing long-term solution TCO.

In particular, Intel Core Ultra processors are engineered for high throughput and low latency on client and edge systems, available with integrated Intel Iris Xe or Intel Arc graphics for outstanding performance per watt and improved parallelism on workloads such as rendering medical images and AI inference. Learn more about Integrated Intel Iris Xe graphics and Intel Arc graphics.

*"The **Intel Core Ultra processor** has ushered in a new era of innovation in healthcare imaging. Our tests have revealed a remarkable **22% and 25% increase in AI performance throughput** for NerveTrack and Live ViewAssist/HeartAssist real-time ultrasound imaging applications respectively, compared to previous generations Intel Core processor paired with a competitive discrete GPU. This breakthrough, attributed in part to the **built-in Intel Arc GPU**, allows us to offer advanced AI features in next generation mid and entry level ultrasound devices **without the need for discrete GPUs**, resulting in more accessible and cost-effective cutting-edge imaging technology. The Intel Core Ultra processor is a game-changer, enhancing patient care, and opening new possibilities in the healthcare segment."*

– SungShik Baik, Principal Engineer, Samsung Medison

Example use cases for Intel® Integrated graphics

- Endoscopy
- Ultrasound imaging
- Portable X-ray imaging
- AI inference at the edge (device or PC)
- Telehealth, bedside monitoring with UI
- Medical tablets
- Interventional imaging
- Patient monitoring
- Scanner intelligence
- Robotics
- Real-time healthcare systems

Implementation studies

- Low-power, AI-based ultrasound leverages integrated Intel Iris Xe graphics for high performance per watt to improve battery life for portability and energy efficiency.
- AI-driven obstetric screening, optimized with Intel oneAPI tools, uses Iris Xe graphics for real-time recognition and tracking.

Discrete graphics solutions

Intel Arc GPUs

Demanding client and edge use cases such as AI inference with larger models, media processing, image visualization and rendering and AI-based surgical multi-screen display can be supported by Intel Arc GPUs in PCIe cards or MXM module form factors. The platform features not only hardware for processing of vector graphics but also Intel X^e Matrix Extensions (XMX) engine to accelerate AI inference, and Intel X^e Media Engine to improve throughput for transcoding and other media-processing workloads.

Intel Arc GPUs also include hardware acceleration for ray tracing, a requirement of volumetric rendering used for medical visualization. The support next-generation interactivity based on robust user interfaces and displays with graphics acceleration and multiple ray-tracing units. Learn more about Intel Arc GPUs.

Intel Data Center GPU Flex Series

Intel's general-purpose data center GPU Flex Series is optimized for media stream density and quality as well as AI analytics performance. Infused with server capabilities, Intel Flex Series GPUs enable high levels of reliability, availability and scalability. Available in two versions — Flex 140 for maximum density and Flex 170 for high neural network compute performance — these GPUs are the ideal choice for use on high-performance medical equipment to support AI visual inference, media processing, remote display delivery, real-time image visualization and rendering, image reconstruction, image pre/post processing and low-latency compression/decompression. Learn more about the Intel Data Center GPU Flex Series.

Example use cases for Intel® Arc™ GPUs

Healthcare

- Medical imaging: lighter visual computing and AI:
  - AI-enhanced endoscopy and ultrasound
  - X-ray
  - Low-end CT
  - Low-to-medium tier mammography
  - Interventional imaging
  - Medical image rendering workstations
  - Multi-screen, AI-based surgical display
  - AI-based patient monitoring
  - Virtual nursing
  - Virtual ICU
  - LLMs at the edge
  - AI on Voice for healthcare
  - Portable MRI/CT/X-ray

Lab & Life Sciences

- Lab diagnostics (graphics display functions and lighter data analysis workloads)
- Cell analytics (lighter workloads)
- Digital pathology (lighter workloads)
- Visual inspection in pharma (lighter workloads)

Implementation study: Intel® Arc™ GPUs

Medical imaging AI inference solution uses Intel Arc GPUs to reduce TCO and increase flexibility across diagnosis and treatment scenarios.

Example use cases for Intel® Data Center GPU Flex Series

Healthcare

- Medical imaging: Heavier workloads for image visualization, rendering, pre/post processing, denoising, reconstruction, and AI:
  - Mid - high end CT
  - MRI
  - Medium-to-high tier mammography
  - PET
  - Molecular imaging / nuclear imaging
  - Image-guided therapy
  - Radiation therapy
- Data-center class health IT: Telehealth, remote radiology, edge platforms, etc.
- Generative AI-based medical applications
- Digital twin applications

Lab & Life Sciences

- DNA sequencer (entry-level)
- Lab diagnostics (heavier data analysis workloads)
- Cell analytics (heavier workloads)
- Digital pathology (heavier workloads)
- Visual inspection in pharma (heavier workloads)

Intel Data Center GPU Max Series

Intel's highest performing, highest density Max Series GPUs are designed to take on the most challenging high-performance computing (HPC) and AI workloads. Equipped with high bandwidth memory, Max GPUs excel at complex, data-intensive use cases such as drug discovery, data processing for cryogenic electron microscopy, high-end DNA sequencing and photon counting CT. Learn more about the Intel Data Center GPU Max Series.

Example use cases for Intel® Data Center GPU Max Series

Healthcare

- Next-generation photon counting CT
- Larger-parameter generative AI-based healthcare applications
- Digital twin applications

Lab & Life Sciences

- DNA sequencer (high-end)
- Drug discovery and biophysical data analysis
- Cryogenic electron microscopy

Open software for sustained business advantage

To obtain optimum performance, the GPUs are integrated in tandem with the associated open programming model. The strategy of a vendor-neutral platform and ecosystem is powered by Intel oneAPI development tools. oneAPI provides freedom of hardware choice with a flexible alternative to proprietary model lock-in. Solution developers can shift among hardware including GPUs, CPUs, FPGAs and other accelerators without significant recoding. Tools to help drive performance and productivity with optimized TCO include the following:

- **Migrate CUDA code to SYCL** with the Intel DPC++ Compatibility Tool to create a single source code for easy portability to multiple vendor architectures, including Intel GPUs. This tool automates most of the process and saves significant time on ongoing code maintenance. Codeplay offers additional support to optimize code for heterogeneous, cross-vendor architectures.
- **Realize the value of open hardware** with Intel oneAPI components including analyzers, debuggers, AI tools, frameworks and libraries. Leverage Intel's cutting-edge, built-in hardware features and broad portfolio of software tools to optimize and maximize performance on Intel GPUs.
- **Optimize performance and resource utilization** with Intel® VTune™ Profiler (part of the oneAPI Base Toolkit), a code analysis tool for developers. VTune is used to detect performance bottlenecks, assess parallel code efficiency, understand memory access patterns and refine GPU operations.

Intel's open software approach and the associated single, portable codebase across multiple architectures reduce development costs, lower code maintenance and speed time-to-market. Using Intel tools as part of the broader open ecosystem enables developers to improve performance on Intel hardware while maintaining the hardware-vendor neutrality of the codebase.

Enabling deep learning development at GE HealthCare

"Using Intel® oneAPI Base Toolkit, we have successfully implemented GEHC's proprietary TrueFidelity DL, a deep learning image reconstruction algorithm available across much of the company's CT portfolio. TrueFidelity DL images achieve exceptional clarity at low dose. The open source SYCL compiler provides near entitlement AI/DL inferencing performance for several NVIDIA GPU devices including popular Tesla T4, A4000, and A5000 GPUs across all data types FP32, TF32, FP16. Based on GEHC experience with OpenCL, code portability is crucial to protect our SW development investment and re-use the SW across different platforms and vendors. We also found that cuDNN to oneDNN porting is very straight forward. oneAPI is an open source project with strong commercial support from the 3rd parties. It ensures continuity of support and bug fixes, and we look forward to continuing our collaboration with Intel."

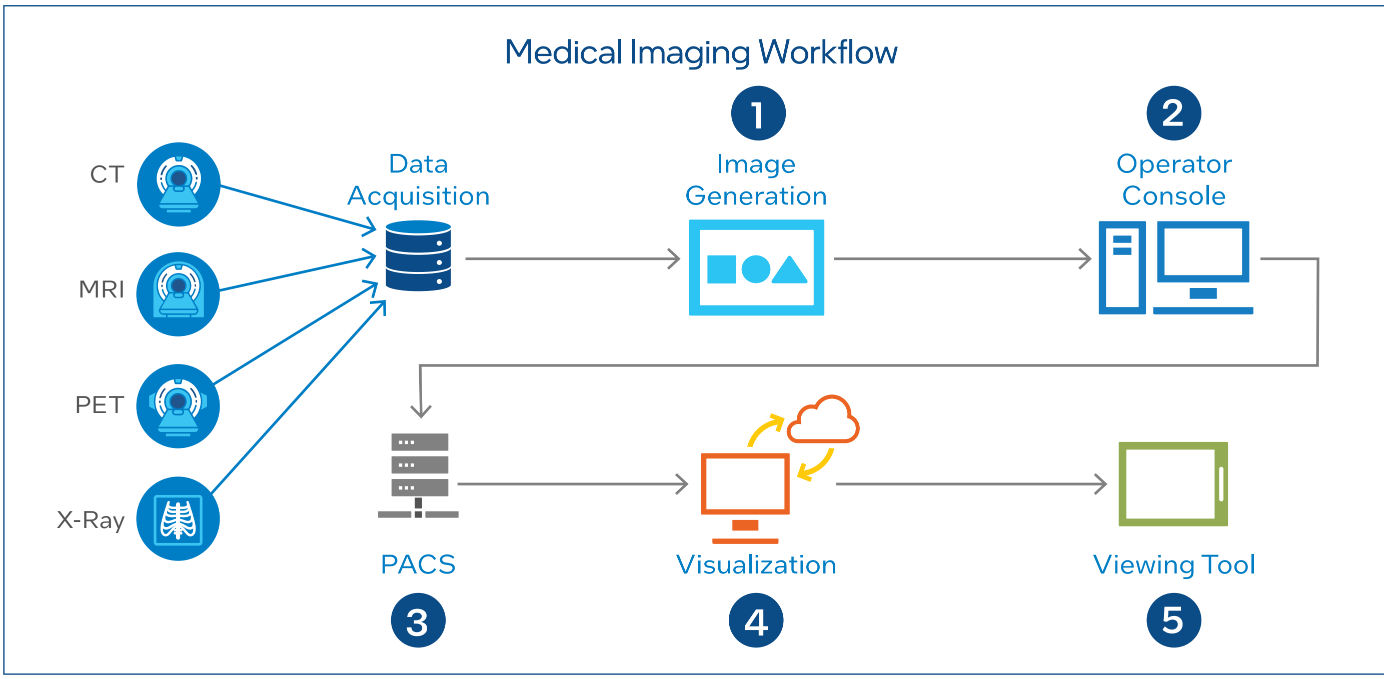
Implementation study: Intel® oneAPI and AI tools

Accelerated healthcare diagnostics with oneAPI and AI tools enables AI-powered solutions for usages as diverse as medical imaging, diagnosing mental health disorders, cancer and lung disease.

Use case example: Medical imaging

GPUs can be leveraged in healthcare applications to speed up computations necessary for diagnostics and screening. However, as described in this brief, competing GPUs may be associated with vendor lock-in. Based on open programming model, Intel's discrete GPUs have the potential to improve performance, energy efficiency and TCO across the breadth of healthcare and life sciences implementations.

Medical imaging provides an instructive example, where use cases are generating exponentially more data than in years past, as denser, higher-quality images are used for more sophisticated analysis, including through AI. This segment is growing at 6.2% through 2030, from $38.16 billion in 2022 to a projected $61.51 billion in 2030.[5] Accordingly, the need for faster data processing, 3D image rendering, AI inferencing and enhanced display of high-resolution medical images is exploding.



Common medical imaging modalities such as CT, MRI, PET and X-ray use a common workflow from data acquisition and image generation through assessment by clinical professionals, as shown here. Intel GPUs may be implemented throughout that workflow:

1. **Image generation (including image preprocessing and image reconstruction)**: In this step, depending on the system configuration, either Intel Data Center GPU Flex or Intel Arc GPUs can be used for image-reconstruction algorithms, including deep learning-based reconstruction algorithms.
2. **Operator console:** This conventionally non-intensive step, which consists primarily of augmenting data with metadata, can be improved with AI methods, often handled by an integrated GPU based on Intel Iris Xᵉ or Intel Arc.
3. **Picture Archiving and Communication Systems (PACS):** Intel Data Center GPU Flex may support remote viewing and editing of large and complex radiology images, in conjunction with general-purpose virtual desktop infrastructure (VDI) to improve productivity by radiologists and clinical staff.
4. **Image visualization/rendering:** 2D and 3D image rendering is a compute-intensive, critical step in medical-imaging workflows. Intel Arc discrete GPUs handle rendering tasks from the simple to the photorealistic with low latency.
5. **Viewing tool:** Depending on the platform that the radiologist or other medical professional uses to view the images, integrated graphics based on Intel Iris Xᵉ or Intel Arc graphics or discrete GPUs based on Intel Arc graphics may be appropriate to optimize image quality and viewing experience.

Conclusion

The market potential for healthcare and life sciences solutions will be satisfied largely by GPUs, especially for AI and visual computing. Industry requirements are no longer compatible with closed, proprietary programming models. Interoperability is the way of the future. Intel GPUs — both integrated and discrete — are the foundation for an open, ecosystem-driven alternative. With scale-up from the client and edge to the data center and cloud, Intel GPUs support outstanding performance and productivity with Intel oneAPI tools. The combination of Intel hardware and software technologies enables a single codebase to be deployed across heterogeneous hardware, for TCO that powers innovation.

Intel Health and Life Sciences

Intel Graphics Solutions

More information

- Where to buy: Intel Data Center GPUs and Intel Arc GPUs
- Profiles in success: How customers are using Intel discrete graphics
- Learn more: Migrating to oneAPI
- Digitize clinical and research innovation: AI in healthcare and life sciences
- Get results with Intel Data Center GPUs: Deliver More, Faster, Now
- Reach out to Intel's Chief Architect for Intel Medical Imaging and Digital Health: Beenish Zia

Solution provided by:

intel®

[1] Fortune, March 19, 2018. "Tech's Next Big Wave: Big Data Meets Biology." https://fortune.com/2018/03/19/big-data-digital-health-tech/.

[2] Centers for Medicare & Medicaid Services, National Health Expenditure Data, Historical. https://www.cms.gov/data-research/statistics-trends-and-reports/national-health-expenditure-data/historical.

[3] Statista, September 28, 2023. "Artificial intelligence (AI) in healthcare market size worldwide from 2021 to 2030." https://www.statista.com/statistics/1334826/ai-in-healthcare-market-size-worldwide/.

[4] Evans Data Global Development Survey Report 22.1, June 2022.

[5] Medical Imaging Market Size. Fortune Business Insights. August 2023. https://www.fortunebusinessinsights.com/industry-reports/medical-imaging-equipment-market-100382.

0224/MK/MESH/352243-001US