White Paper

Researchers Using AI: Don't Let Data Access Derail Clinical Breakthroughs

Sponsored by: Pure Storage

Cynthia Burghard
February 2018

IDC OPINION

The power of artificial intelligence (AI) is enabling clinical breakthroughs that identify biomarkers without invasive procedures, diagnose skin cancer with a photograph, predict adverse clinical events, and recommend treatments based on current literature. Getting these innovations to market requires access to large, complex data sets to train the AI models. One of the gating factors in getting clinical insights to the bedside is the data access challenge researchers face. The large data sets required to train the AI algorithms must be read for each AI model training run. In traditional legacy hard disk drive (HDD) arrays, this results in delays for researchers and tedious, time-consuming work for IT staff. All-flash arrays (AFAs) allow AI researchers to use their data to increase their models' accuracy without latency, saving both time and money.

The power of artificial intelligence (AI) is enabling clinical breakthroughs that identify biomarkers without invasive procedures, diagnose skin cancer with a photograph, predict adverse clinical events, and recommend treatments based on current literature.

THE ENVIRONMENT AND MARKET DYNAMICS

Economic uncertainty amid narrow margins in healthcare is keeping hospital CEOs up at night. In a recent Deloitte survey (**www2.deloitte.com/us/en/pages/life-sciences-and-health-care/articles/health-system-ceos.html**), 85% of CEOs cite the uncertainty of Medicaid funding as top of mind. The resulting increase in uncompensated care could certainly cause a financial crisis for many hospitals. CEOs at academic medical centers, children's hospitals, and others in the survey raised concerns about cuts to the National Institutes of Health (NIH), Health Resource and Service Administration (HRSA), or the Food and Drug Administration (FDA) as they rely heavily on these funding sources.

In highly competitive environments, more favorable terms in risk contracts are going to healthcare providers that can demonstrate value and clinical excellence. The ability to win grants and deliver research results in a timely fashion may well impact reimbursement albeit indirectly. With the likelihood of dwindling government-sponsored grant funds, the competition to attract and retain talented researchers will only grow. Offering researchers an environment that facilitates more timely delivery of research results, thus improving their personal reputations, can only help in recruitment and retention.

At the heart of all clinical research is the improvement of the health of patients, and there is an ever-growing body of evidence of the positive impact of AI in healthcare. For example, AI is being applied in cardiology where deep learning is being used with MRIs to measure the volume of blood transported with each contraction of the heart. Cardiologists typically need 30–60 minutes to calculate the volume of a ventricle or an auricle; AI finds the answer within seconds. The results of this research conducted at Massachusetts General Hospital resulted in the FDA's approval for Arterys MICA in 2017 as reported in Forbes Magazine, January 2016. The FDA also recently approved AliveCor's KardiaBand EKG reader, the first medical device accessory for the Apple Watch.

The pressure to bring clinical innovation and operational improvements to health systems more quickly and efficiently will only increase. Time to discovery must be shortened as pressures to improve quality and manage healthcare costs continue to escalate. A competitive advantage exists for those that can leverage the power of AI to demonstrate clinical excellence and operational efficiencies.

Don't Let Outdated IT Infrastructure Impede Clinical Excellence

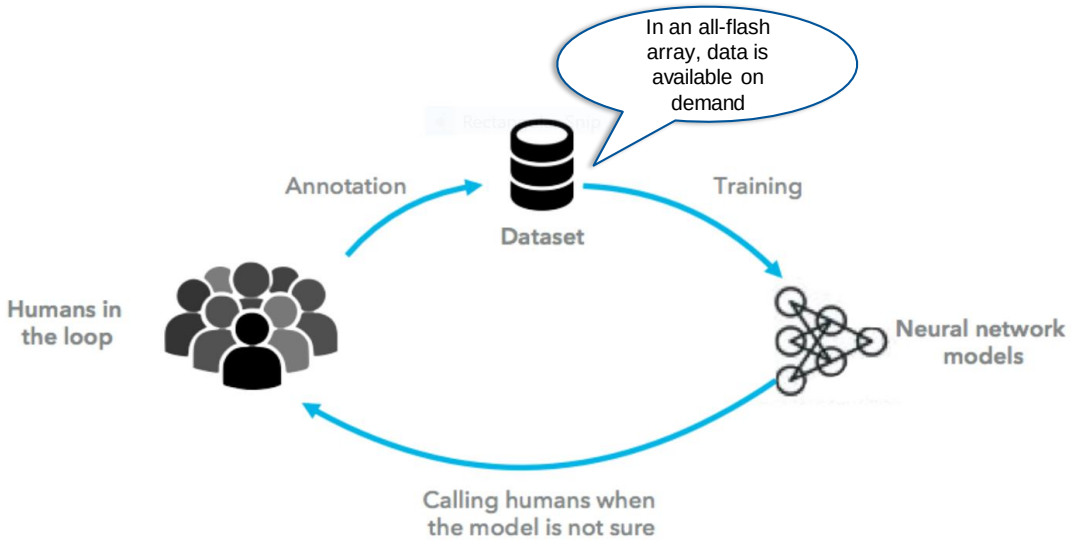
Organizations that successfully make the shift to value-based healthcare will recognize data as a strategic asset that must be harvested, aggregated, and analyzed to optimize clinical, financial, and operational results. Organizations also understand the importance of creating a 360-degree view of the patient that includes data generated between office visits and inpatient admissions. This data includes biometric readings from consumer-grade medical devices or wearable activity trackers, social determinants of health, and other environmental data such as weather and pollen counts.

IDC estimates that the lifetime data volume for a typical patient is approximately 1,100TB. Progress in high-definition and 3D medical imaging and videos; continuous advances in bioinformatics, as in the case of genomic sequencing; emerging digital pathology systems; and the widespread adoption and continuing growth of technologies such as picture archiving and communication systems (PACS) and radiology information systems are among the key causes of this volume explosion. In turn, healthcare organizations are not only investing in big data and analytics but also looking to AI and machine learning to glean insights from all this data and support clinical decision making at the point of care.

Not only does the process of developing and validating AI models require vast amounts of data but, by the very nature of medicine, the "learning process" is time consuming and fraught with delays when researchers cannot get access to their data sets on an "as needed" basis. With the predicted loss of government grant funding, it is critical that researchers manage funds effectively. Paying IT staff to load data is not an optimal use of funds; hiring additional researchers would seem to be a higher value proposition (see Figure 1).

FIGURE 1

AI Model Training Process



Note: Figure is from *Building Game Changing Data Sets,* a presentation by Augustin Marty, CEO, Deepomatic.

Source: Deepomatic, October 2016

New all-flash array designs promise to address storage performance, maintenance, and availability issues across all industries and have taken over from legacy designs that were originally built around spinning hard disk drive and hybrid disk flash approaches. Today's AFAs are entirely solid state, leveraging NAND flash, 3D NAND, and more modern technologies such as 3D XPoint as they become available. There are no mechanical spinning disks present, dramatically reducing read/write latency and preventing many types of failures that typically plague storage teams. Data processing is available at wire speed with high availability, sometimes up to 99.9999%. Reducing read/write times by an order of magnitude or more while increasing uptime greatly improves users' efficiency and productivity.

Highly flash-optimized storage platforms offer compelling advantages in performance, efficiency, economics, and availability relative to disk and hybrid storage platforms and overcome many of their shortfalls, namely complexity and input/output processing. While HDD capacity has increased over the years, the mechanical nature of HDDs imposes an upper limit on performance, and the technical workarounds needed to achieve large capacity requirements add complexity. Unlike disk technology, flash technology continues to advance at a rapid pace, resulting in even denser flash media, more cost-effective options, and greater technical performance in datacenters. The use of all-flash array resolves one of the key barriers to bringing clinical innovation to market – on-demand access to data increases productivity in training AI algorithms.

What All-Flash Storage Means to Clinical Researchers and Their IT Colleagues

There is growing consensus that legacy HDD storage hamstrings researchers and creates unnecessary and tedious tasks for IT staff. The size of data sets and the complexity of the data that must be moved for each iteration of AI model training are examples of the operational inefficiencies that can be resolved through AFA. Data loading is a redundant, time-consuming, low-value activity.

Clinical Researchers

The field of clinical research is highly competitive. Not only are health systems competing to recruit and retain talent, but researchers across the globe are also competing for government funds and private sector grants to move machine learning research forward. To be successful in this market, researchers must demonstrate clinical excellence through the volume, criticality, and frequency of published results. Health systems that continue to struggle to find adequate financial resources rely heavily on grant funds, as do individual researchers. Not only are funds at stake, but the reputation of both the institution and the individual is at risk if the discovery process is hampered by data access. Careers as well as future funding depend on bringing clinical breakthroughs to market ahead of the competition.

Clinical researchers and data scientists depend on IT staff to move their research forward. Researchers not only depend on IT but also may have to wait for resources behind higher-priority activities. IT infrastructure must not be a barrier to bringing improvements in patient care to the market. Sought-after clinical researchers and data scientists will be easier to recruit and retain in an efficient environment that allows them to advance themselves and their organization.

IT Staff

As an IT professional, one of the least desirable tasks is to load and unload data for analytics. Doing it once is sufficient training; it is not a skill that provides career advancement opportunities — it is time consuming and tedious. In an era of digital transformation with so many advanced technologies in use, IT professionals are looking for opportunities to use new technology to advance their careers. Healthcare organizations have a difficult time recruiting and retaining talented IT professionals, particularly if the IT staff are saddled with managing legacy infrastructure and being viewed as the bottleneck to clinical innovation.

Clinical researchers and data scientists are under pressure to deliver research results for reasons identified previously and to deliver faster than others. Having to wait in an IT queue to get access to data to refine and train AI models creates unnecessary friction between IT and researchers. Unfortunately, IT is often labeled as a bottleneck by business. Freeing up IT staff from low-value tasks results in improved responsiveness. IT can be the hero to data scientists and researchers by demonstrating that IT infrastructure and data access will not be a barrier to advancing AI in healthcare and will facilitate advances in clinical innovation.

Before the days of self-service business intelligence, the time delays between the business' request for reports and IT's delivery of reports were a tremendous source of friction, which was eliminated when end users could develop and run their own reports. Self-service eliminated that friction; AFA can relieve the friction for clinical researchers.

The quest to achieve operational efficiency in healthcare is a top priority. Replacing HDD storage with AFA has been documented to have a lower TCO than legacy storage and improves operational efficiency, introduces more flexible buying models (storage as a service), and brings data to physicians at the point of care.

PARTING THOUGHTS

The economic impact of grant-funded clinical research on a hospital is significant, and speed from bench to bedside critical for clinical innovation. IT infrastructure must not be a barrier to clinical excellence. Access to data in real time is required to advance clinical excellence and can be accomplished with the introduction of AFA.

When evaluating the switch from HDD or hybrid storage to AFA, it is important to understand the underlying IT infrastructure improvements and the savings associated with AFA. For the clinical researcher, moving to AFA means data is available in real time and the friction between IT and research due to the latency of data availability is removed. Ideas and hypotheses for clinical breakthroughs will be validated and get to market faster. The ability to recruit and retain talented researchers will improve competitive positioning and access to grant funding, which can all be realized through the adoption of AFA.

Health systems can change the perception of IT from supplier to partner in the delivery of clinical excellence to the market.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

Global Headquarters

5 Speen Street
Framingham, MA 01701
USA
508.872.8200
Twitter: @IDC
idc-community.com
www.idc.com