SCIENTIFIC COMPUTING WORLD

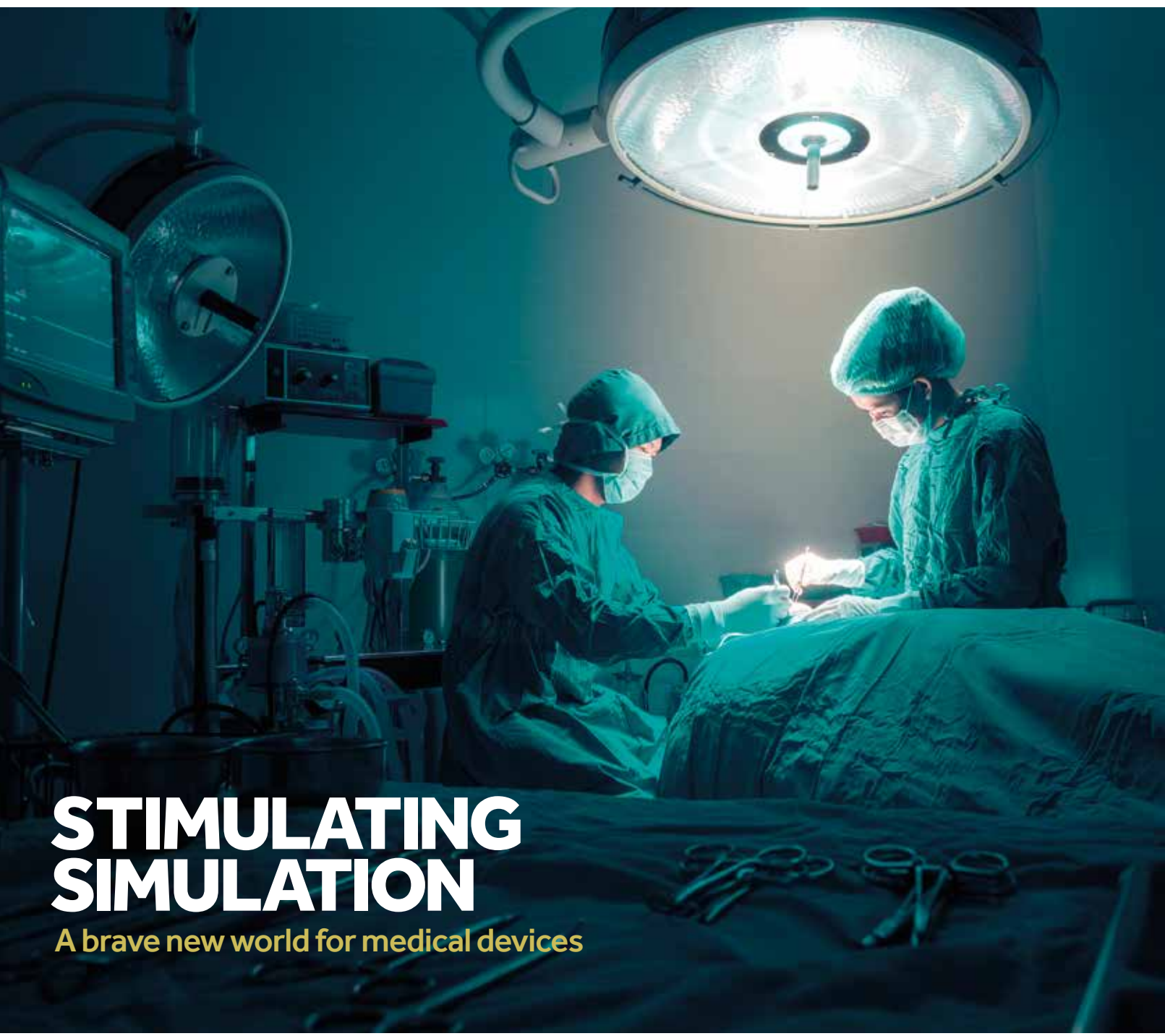
Computing solutions for scientists and engineers

December 2018/January 2019
Issue #163

High performance computing
A cloudy future

Laboratory informatics
Artificial healthcare

Modelling and simulation
Engineering hits the catwalk



STIMULATING SIMULATION

A brave new world for medical devices

www.scientific-computing.com

Pharma&Biotech

Lonza

The MODA™ Platform

The Missing Piece in Your Lab Systems Portfolio

MODA-EM™ Software for QC Micro – Implement, Validate, Integrate Seamlessly.

Lonza's MODA-EM™ Software is purpose-built for QC Microbiology's unique challenges. MODA™ Software delivers better business alignment with QC, at a lower total cost of ownership versus customizing your existing LIMS or other lab systems.

- Reduce investigation and reporting time with rapid analysis and visualization of Micro data
- Manage, schedule and track large EM and utility sample volumes
- Reduce errors while improving compliance and data integrity
- Deliver measurable return on investment from time, cost and compliance savings

Find the missing piece in your lab systems portfolio.
Visit **www.lonza.com/moda** or contact us at **moda@lonza.com**

www.lonza.com/moda

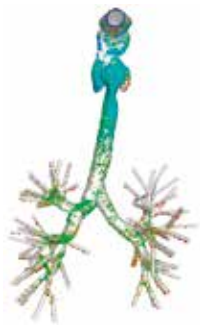
Contents

December 2018/January 2019 ● Issue 163

High performance computing

Laboratory informatics

Modelling and simulation



Resources

Editorial and administrative team
Editor: Robert Roe
robert.roe@europascience.com
Managing editor: Tim Gillett
editor.scw@europascience.com
Specialist reporters: Sophia Ktori, Clare Sansom,
Design: David Houghton, Zöe Andrews

Advertising team
Advertising manager: Mike Nelson
mike.nelson@europascience.com +44 (0) 1223 221039
Production manager: David Houghton
david.houghton@europascience.com +44 (0) 1223 221034

Corporate team
Managing director: Warren Clark

Scientific Computing World is published by Europa Science Ltd, 4 Signet Court, Cambridge, CB5 8LA ● ISSN 1744-8026
Tel: +44 (0) 1223 211170 ● Fax: +44 (0) 1223 213385
Web: www.researchinformation.info

Subscriptions: Free registration is available to qualifying individuals. Register online at www.scientific-computing.com Subscriptions £180 a year for six issues to readers outside registration requirements. Single issue £20. Orders to ESL, SCW Circulation, 4 Signet Court, Swann Road, Cambridge, CB5 8LA, UK. Tel: +44 (0)1223 221 030. Fax: +44 (0)1223 213 385.

US copies: Scientific Computing World (ISSN 1356-7853/USPS No 018-753) is published bi-monthly for £180 per year by Europa Science Ltd, and distributed in the USA by DSW, 75 Aberdeen Rd, Emigsville PA 17318-0437. Periodicals postage paid at Emigsville PA. Postmaster: Send address corrections to Scientific Computing World PO Box 437, Emigsville, PA 17318-0437.

Cover: Xmee/Shutterstock.com

Subscribe for free at www.scientific-computing.com/subscribe

Robert Roe
Editor

Artificial gets real

In final issue of 2018, AI continues to be a strong theme across all three core sections of the magazine as the technology weaves its way into every facet of scientific research.

On page 4 we have a feature looking at the falling costs and increasing use cases for HPC in the cloud, which is driving adoption of this technology. On page 7 we take a look at Google's DeepMind, which has been creating some interest as it recently announced a new AI technique to predict protein structure.

On page 8 we report from the SC18 conference in Dallas, Texas. This year the conference broke records for attendance and celebrated its 30th anniversary. As in previous years the exhibition provided a host of new products and services for the HPC industry, and there was also an increased AI presence in 2018.

Our laboratory informatics coverage begins on page 10 with the conclusion of Sophia Ktori's two-part feature on the use of AI in healthcare. In this issue Sophia looks at some of the fundamental healthcare applications that can be assisted by implementation of AI technology.

On page 14 we have a report from the Global Altair Technology Conference, held in Paris, France. Altair announced the acquisition of SimSolid at the event and also called for greater integration of multiphsyics across their software portfolio.

On page 16 we have an article from Toshiba's Andrew Shields on the importance of quantum cryptography in ensuring data security in the future. On page 18 Gemma Church focusses on the use of modelling and simulation technology in the design and validation of medical devices.

@scwmagazine

SCIENTIFIC COMPUTING WORLD

Cloud coverage increases

ROBERT ROE LOOKS AT THE CHANGING WAYS THAT THE HPC INDUSTRY USES CLOUD COMPUTING TECHNOLOGY

Cloud usage models are changing as the price of cloud is lowered and specialist services are developed for HPC users. This is taking the form of cloud-based disaster recovery services, testbed services and bare metal HPC and AI infrastructure, which is now being developed to support HPC users.

Mahesh Pancholi, business development manager at OCF, commented that the cost of cloud is falling and this is making users more aware of the potential for cloud in HPC. 'There is definitely a change in attitude towards cloud. I think there is also increasing variety of use cases because public cloud providers have gone beyond just selling their spare cycles.

'They have realised just how much of a business opportunity there is and so they are starting to offer more niche and specialised services and these are the kind of things that cover the needs of HPC or research computing users.'

Launch of cloud-based HPC services from companies such as vScaler, Oracle, IBM, Xtreme-D and many others combined with increasingly HPC-friendly hardware installed in main cloud providers is creating a more favourable environment for HPC users.

While some workloads have already been deployed on cloud resources, many HPC applications require a specific level of performance that requires HPC specific technologies such as high bandwidth networking, fast storage and large memory nodes. It has taken time for these services to become more widely available and this has helped to bring costs down, which opens up new possibilities for HPC services.

The signs point to increasing use of cloud by HPC users over the coming years. As the number of possible use cases grows, and the cost of using cloud falls, it is likely that HPC users will increasingly make use of cloud technologies. Naoki Shibata, founder and CEO at Xtreme-D, a HPC and AI cloud computing provider based in Tokyo, Japan, noted that Hyperion's recent research into the cloud computing points to considerable growth in the market.

'Judging by the trending of HPC in the cloud, we see continued fast growth of the cloud HPC market. Hyperion finds that whereas in 2011, 13 per cent of HPC sites used cloud, in 2018, this figure is at 64 per cent. This still allows for accelerated growth because so far just seven to eight per cent of the work is being done using cloud. With greater availability, flexibility, and ease of access, we conclude that a much greater portion of HPC work will be migrated to the cloud in the coming months and years,' stated Shibata.

New ways to use the cloud

Disaster recovery technology in HPC can be incredibly expensive as it requires that hardware is kept in reserve in case of emergency. OCF initially set out to create a cloud based service that could be used for disaster recovery which could reduce the cost of reserving additional hardware. The system keeps the infrastructure to procure nodes running so that users can quickly setup a cluster if disaster strikes.

'Generally with a cluster you have some management nodes that make sure everything is running; you have a scheduler available and it can be used to deploy new hardware as it comes in. We are taking that approach and putting that into the cloud. You will have your management software running and then as you need you can spin up additional nodes' said Pancholi.

This provides users with the safety of having cloud resources available to them quickly without having to pay for the majority of the resource unless it is needed. Pancholi stressed that this

"Hyperion finds that whereas in 2011, 13 per cent of HPC sites used cloud, in 2018, this figure is at 64 per cent"

Valery Brozhinsky/Shutterstock.com

is not just for HPC users, however, as it can be used to recreate your internal infrastructure so researchers can look at novel technologies.

Pancholi gave an example of a recent customer who had become responsible for not just the HPC resources but a whole range of research computing services. 'They have got people looking into novel areas like IoT and AI and the size of your team to run those services means that you have to be quite clear about where you put those resources. For cutting-edge research you often need to put a disproportionate amount of resources to set it up and keep it running for a small number of people,' said Pancholi.

'If you leverage a public cloud resource for that it makes it easy to investigate and see if it is worthwhile to invest that time and money in the underlying infrastructure without having to move people away from the main services that everyone else is relying on,' Pancholi added.

However OCF's cloud replication service does more than just provide users with cloud availability as it has been designed to provide users with a replica of their own internal cluster. In the past this has been prohibitively expensive. Without doubling up on servers, it is not possible to recreate a cluster ready for when there is a major issue.

Pancholi states: 'We have seen a trend towards HPC becoming more and more accepted as a critical service for research institutions. From a business perspective we have also seen changes in how these institutions are run.

They are moving away from a home-grown IT directive to more commercially aware CIO types. The first thing they want to do is make sure they have disaster recovery coverage across the organisation for their key services.'

OCF aimed to come up with a way to provide the reassurance of having a cluster for disaster recovery available without the huge costs. This meant keeping the software and management infrastructure available 24 hours a day, so that users can quickly provision the nodes to replicate their own internal HPC cluster. 'The work that we have been doing with public cloud providers led us to believe that we could start to do something that is a proper replica of your cluster in the cloud,' said Pancholi.

'What we have aimed to do with our Cloud Replication Service is to actually build – from the ground up – a replica of your nodes and your software stack and provide a way for your data live data to be available to you in the private cloud'

'There have been attempts to try and fill this void before and predominantly people have tried to host software in the cloud and say well you do mainly Ansys stuff on your HPC cluster so you can come and do Ansys in the cloud. Now my personal experience of running a HPC cluster for a university helps me understand that there is a big difference between running a piece of software on two different infrastructures. There are so many different parameters that can change things but actually what you need is something that is installed as close as possible to the original system.'

'What we have aimed to do with our Cloud Replication Service is to actually build – from the ground up – a replica of your nodes and your software stack and provide a way for your data live data to be available to you in the private cloud,' said Pancholi.

'We think this is a very novel approach and the cost implication of having that on standby instead of buying a second cluster is something that we have found to be very attractive proposition when we have spoken to people.'

Bare-metal future

While some companies develop specialist services for disaster recovery and testing new software, others are looking to make cloud an avenue for traditional HPC workloads. There are several companies now offering so called 'bare metal' cloud servers or platforms which are designed to deliver performance that is comparable to an in-house HPC cluster. →

"You create a service and put it out there and then find that people will find novel ways to use it that you hadn't considered because it fits a niche that they have"

Home Art/Shutterstock.com

→ Xtreme-D has launched its Xtreme-Stargate, a gateway appliance that delivers a new way to access and manage high-performance cloud resources at SC18, the largest US HPC conference, held in November each year.

The Xtreme-Stargate system is a new HPC and AI cloud platform from Xtreme-D. It provides high-performance computing and graphics processing and the company aims to provide a cost effective platform for both simulation and data analysis. The company has already announced a Proof of Concept (PoC) programme, which has been tested with several early adopters.

Xtreme-D has announced that public services will start at the end of 2018 in Japan, collaborating with Sakura Internet as the major Japanese-focused HPC and datacentre provider. Services in Europe and the US are expected to commence in 2019.

Xtreme-D's founder and CEO Shibata commented that the main benefit of using the companies Xtreme-D DNA product is in establishing a quick optimal HPC virtual cluster on public cloud. 'It saves hours of paid use on the cloud that it would have taken to create the configuration without the use of Xtreme-D's HPC templates as a means for gaining access.'

'The tool also provides ongoing status of the job being processed, thus allowing better budget control of spending on the cloud and alerts for action in case resources are not sufficient so that work done will not be lost. Xtreme-D Stargate adds a higher level of data security through hardware-based data transfer. The ease of access and the resulting cost savings help accelerate the use of cloud for HPC workloads,' comments Dr David Barkai, technical advisor, Xtreme-D.

Shibata also notes that the bare metal infrastructure provides a unified environment that helps to reduce the burden of managing multiple types of computing resources.

'It is difficult to manage a hybrid cloud approach for an HPC solution. Customers incur a double cost and need to put in twice the effort in order to manage both on-premise and cloud implementations,' comments Shibata. 'There is also the significant cost of uploading and downloading data to and from the cloud datacentre. Therefore, the HPC native cloud customer requires a separate workload setup for each environment (cloud and on-premise).'

'The bare metal approach allows customers to manage a single environment. For example, AI and Deep Learning are best suited for cloud computing, but it is very difficult to size the infrastructure for Deep Learning. Xtreme-Stargate provides elastic AI and HPC-focused infrastructure, and is thus the best solution for both HPC and AI. It addresses both data I/O and general purpose computing,' adds Barkai.

Serendipitous adoption of cloud services

When OCF set out to create the cloud replication service the idea was fairly simple, to create service that could be used for disaster recovery that could mimic an in-house cluster. The service provides an opportunity to massively reduce downtime and give users the opportunity to get their applications up and running as quickly as possible.

However, as Pancholi notes, by designing a robust service they had created something that could be used for a number of different use cases. 'Initially the thought was "what if my cluster goes bang?" and that is how we came up with the idea but actually we started to realise that you can utilise it for testing or you can utilise it for providing a workspace for a group of users that are new or hasn't necessarily got access rights to your cluster,' stated Pancholi.

'In the testing scenario let's say there is a critical update to a critical piece of software for your cluster. Other than deploying that on your private cluster, you could test that out on the cloud. You can see, as close as is possible, the consequences of that update and if you are happy with it you can deploy it across the main cluster.

'It's not going to be 100 per cent all of the time – nothing can be but it certainly provides a level of assurance that you didn't have previously.'

Pancholi explained other use cases such as to create a workspace for a group of new users or as a testbed for new software or testing a new architecture. 'You create a service and put it out there, and then find that people will find novel ways to use it that you hadn't considered because it fits a niche that they have,' adds Pancholi.

OCF's Dean commented that these kinds of cloud services can be particularly suited for those kind of testbed activities. Even in the short time that this service has been developed OCF has already found users testing AI and IoT research, which can be carried out in this way.

'You can imagine that with these technologies they are cool and interesting technologies – a bit of a buzzword at the minute – when you are talking to an organisation a lot of people will say they want to do it but you don't know what the uptake is going to be like until you build something,' said Dean.

'With computing resources being per-built for you in the cloud you can grant access to that and then you can see how many of your users actually start using these services. Then you can start doing some analysis to find out if it will be viable to bring this service in-house,' added Dean.

DeepMind predicts protein structure

DeepMind has announced a new tool in AI research

DeepMind's new AI research system, AlphaFold, builds on years of genomics research by using data to predict protein structure.

AlphaFold has been developed over the last two years but is built on many years of prior research, using vast genomic data.

This technology could have significant implications for healthcare and medicine as it will enable scientists to gain insight into the way that diseases develop and possible preventions. The ability to predict a protein's shape is useful to scientists because it is fundamental to understanding its role within the body, as well as diagnosing and treating diseases believed to be caused by misfolded proteins, such as Alzheimer's, Parkinson's, Huntington's and cystic fibrosis.

A protein's properties are determined by its 3D structure. For example, antibody proteins that make up our immune systems are 'Y-shaped', and are akin to unique hooks. By latching on to viruses and bacteria, antibody proteins are able to detect and tag disease-causing microorganisms for extermination.

Similarly, collagen proteins are shaped like cords, which transmit tension between cartilage, ligaments, bones, and skin. Other types of proteins include CRISPR and Cas9, which act like scissors and cut and paste DNA; antifreeze proteins, whose 3D structure allows them to bind to ice crystals and prevent organisms from freezing; and ribosomes that act like a programmed assembly line, which help build proteins themselves.

But figuring out the 3D shape of a protein purely from its genetic sequence is a complex task that scientists have found challenging for decades. The challenge is that DNA only contains information about the sequence of a protein's building blocks called amino acid residues, which form long chains. Predicting how those chains will fold into the intricate 3D structure of a protein is what's known as the 'protein folding problem'.

An understanding of protein folding will also assist in protein design, which could unlock a number of benefits. For example, advances in biodegradable enzymes could help manage pollutants like plastic and oil, helping us break down waste in ways that are more friendly to our environment. In fact, researchers have already begun engineering bacteria to secrete proteins that will make waste biodegradable, and easier to process.

To catalyse research and measure progress on the newest methods for improving the accuracy of predictions, a biennial global competition called the Community Wide Experiment on the Critical Assessment of Techniques for Protein Structure Prediction (CASP) was established in 1994, and has become the gold standard for assessing techniques.

How can AI make a difference?

Over the past five decades, scientists have been able to determine shapes of proteins in labs using experimental techniques like cryo-electron microscopy, nuclear magnetic resonance or X-ray crystallography, but each method depends on a lot of trial and error, which can take years and cost tens of thousands of dollars per structure. This is why biologists are turning to AI methods as an alternative to this long and laborious process for difficult proteins.

Fortunately, the field of genomics is quite rich in data thanks to the rapid reduction in the cost of genetic sequencing. As a result, deep learning approaches to the prediction problem that rely on genomic data have become increasingly popular in the last few years.

> Using these scoring functions we were able to search the protein landscape

The team focused specifically on the hard problem of modelling target shapes from scratch, without using previously solved proteins as templates. We achieved a high degree of accuracy when predicting the physical properties of a protein structure, and then used two distinct methods to construct predictions of full protein structures.

Both of these methods relied on deep neural networks that are trained to predict properties of the protein from its genetic sequence. The properties our networks predict are: (a) the distances between pairs of amino acids and (b) the angles between chemical bonds that connect those amino acids. The first development is an advance on commonly used techniques that estimate whether pairs of amino acids are near each other.

The team of researchers trained a neural network to predict a separate distribution of distances between every pair of residues in a protein. These probabilities were then combined into a score that estimates how accurate a proposed protein structure is. We also trained a separate neural network that uses all distances in aggregate to estimate how close the proposed structure is to the right answer

Using these scoring functions we were able to search the protein landscape to find structures that matched our predictions. Our first method built on techniques commonly used in structural biology, and repeatedly replaced pieces of a protein structure with new protein fragments. The researcher trained a generative neural network to invent new fragments, which were used to continually improve the score of the proposed protein structure.

The second method optimised scores through gradient descent – a mathematical technique often used in machine learning for making small improvements – which resulted in highly accurate structures. This technique was applied to entire protein chains rather than to pieces that must be folded separately before being assembled, reducing the complexity of the prediction process.

Oiling the wheels of HPC

Robert Roe reports on new technology and 30 years of the US supercomputing conference at SC18 in Dallas

Dibrova/Shutterstock.com

As the HPC industry descended on Dallas, Texas for this year's instalment of the US supercomputing conference SC18, there was significant enthusiasm for the future of the industry – which is making gains through the rise in the use of AI.

'From our volunteers to our exhibitors to our students and attendees – SC18 was inspirational,' said SC18 general chair Ralph McEldowney. 'Whether it was in technical sessions or on the exhibit floor, SC18 inspired people with the best in research, technology, and information sharing.'

SC18 highlighted progress in HPC, the rise of AI technology and several new HPC technologies on display during the exhibition. AMD's 2018 resurgence continued at SC18 announcing new CPU and GPU products alongside a number of contract wins for large-scale HPC systems. The US strengthened its position in the Top500 after regaining the top spot in June this year. And finally the SC cluster competition saw the Chinese Tsinghua University take the top spot in the overall category of the competition in Dallas.

Chinese students take top honours

The Student Cluster Competition (SCC) was first held in 2007 to provide high-performance computing experience to undergraduate and high school students. Today the student teams are comprised on six students and at least one advisor. Hardware is donated by vendor partners and the student teams design and build clusters, which need to be configured and optimised to run specific codes across the 48-hour completion.

This year the Tsinghua University from China took the top spot overall. The Linpack crown was taken by Nanyang Technological University in Singapore with a score of 56.51 Teraflops. The HPCG was also won by Tsinghua University with a score of 1,985.97 Gigaflops.

As the competition limits teams to 3,000 watts of power for their HPC cluster, most teams opt to use GPU technology. For example, Tsinghua University used eight Nvidia V100 GPUs in their competition entry.

In a blog post from Nvidia, Bu-sung Lee, team leader and faculty adviser at Nanyang University stated: 'If you don't have GPUs, best of luck. It's essential.'

However, it was not all new technology at SC this year. There was also reflection on the past as this year marks the 30th anniversary of the annual international conference of high performance computing, networking, storage and analysis. It celebrates the contributions of researchers, scientists and HPC users who have helped to advance this industry over the last 30 years.

New technology on the show floor

AMD has had a particularly strong 12 months, from plucky outsider to a growing force within the HPC industry. The launch of the Epyc processors gave the company a good start but along with new products at SC18 the company also announced several new contracts for AMD-based HPC systems.

These systems provide the groundwork for stronger adoption of AMD in the HPC ecosystem as these new systems provide real benchmarking opportunities to give potential users an idea of how these systems work at scale in real production environments.

AMD announced that The US Department of Energy's NERSC, Cray and HAAS F1 Racing and The High-Performance Computing Center of the University of Stuttgart (HLRS) will all be using AMD Epyc processors for their upcoming systems. AMD also pushed its technologies into the cloud with announcements around the deployment of Epyc processors in MicrosoftAzure.

'It's been a fantastic year in the supercomputing space as we further expanded the ecosystem for AMD Epyc processors while securing multiple wins that leverage the benefits AMD Epyc

> "It's been a fantastic year in the supercomputing space as we further expanded the ecosystem for AMD Epyc processors while securing multiple wins that leverage the benefits AMD Epyc processors have on HPC workloads"


Dallas Skyline

processors have on HPC workloads,' said Mark Papermaster, senior vice president and chief technology officer, AMD. 'As the HPC industry approaches exascale systems, we're at the beginning of a new era of heterogeneous computer that requires a combination of CPU, GPU and software that only AMD can deliver. We're excited to have fantastic customers leading the charge with our Radeon Instinct accelerators, AMD EPYC processors and the ROCm open software platform.

Cray announced a new computing platform called Shasta which is set to replace its XC50 systems. Shasta is an entirely new design aimed at exascale performance, data-centric workloads, and increased competition of processor architectures. The National Energy Research Scientific Computing Center (NERSC) announced just ahead of SC18 that it has chosen a Cray 'Shasta' supercomputer for its NERSC-9 system, named 'Perlmutter,' in 2020. The program contract is valued at $146 million, one of the largest in Cray's history, and will feature a 32-cabinet Shasta system.

Cray seems to acknowledge the shifting demands of HPC applications as the new system is positioned to take advantage of the growing trend for a single system to handle converged modelling, simulation, AI, and analytics workloads. The system allows users to mix and match processor architectures in the same system (X86, ARM, GPUs), as well as choose system interconnects from Cray (Slingshot), Intel (Omni-Path) or Mellanox (Infiniband).

'Our scientists gather massive amounts of data from scientific instruments like telescopes and detectors that our supercomputers analyse every day,' said Dr Sudip Dosanjh, director of the NERSC Center at Lawrence Berkeley National Laboratory. 'The Shasta system's ease of use and adaptability to modern workflows and applications will allow us to broaden access to supercomputing and enable a whole new pool of users. The ability to bring this data into the supercomputer will allow us to quickly and efficiently scale and reduce overall time to discovery.'

'Cray is widely seen as one of only a few HPC vendors worldwide that is capable of aggressive technology innovation at the system architecture level,' said Steve Conway, Hyperion Research senior vice president of research. 'Cray's Shasta architecture closely matches the wish list that leading HPC users have for the exascale era, but didn't expect to be available this soon. This is truly a breakthrough achievement.'

Top500 shows US gathering pace

The newest release of the Top500 list of the fastest supercomputers also took place at SC. While the DOS system Summit system from Oak Ridge National Laboratory (ORNL) was still first place there was significant movement in the top ten places. The US DOE system, Sierra, at Lawrence Livermore National Laboratory (LLNL) took second place however the overall list showed that China is still edging out the US in the total number of systems that made the list.

Summit widened its lead as the number one system, improving its High Performance Linpack (HPL) performance from 122.3 to 143.5 Pflops since its debut in June 2018. Sierra also added to its HPL results from the previous list. The system increased its score from 71.6 to 94.6 Pflops, enough to take the number two position. Both are IBM-built supercomputers, powered by Power9 CPUs and Nvidia V100 GPUs.

Sierra's ascendance pushed China's Sunway TaihuLight supercomputer, installed at the National Supercomputing Center in Wuxi, into third place. Prior to last June, it had held the top position on the Top500 list for two years with its HPL performance of 93.0 petaflops. TaihuLight was developed by China's National Research Center of Parallel Computer Engineering & Technology (NRCPC).

The share of Top500 installations in China continues to rise, with the country now claiming 227 systems (45 per cent of the total). The number of supercomputers that call the US home continues to decline, reaching an all-time low of 109 (22 per cent of the total). However, systems in the US are, on average, more powerful, resulting in an aggregate system performance of 38 per cent, compared to 31 per cent for China.

> "Whether it was in technical sessions or on the exhibit floor, SC18 inspired people with the best in research, technology, and information sharing"

Applying AI

Elnur/Shutterstock.com

SOPHIA KTORI COMPLETES HER TWO-PART SERIES ON THE USE OF ARTIFICIAL INTELLIGENCE IN HEALTHCARE RESEARCH

Artificial intelligence isn't all about smarter, more accurate diagnostics, predictive or personalised medicine, points out Lee Harland, founder and CSO at Cambridge, UK-based SciBite. There are far more fundamental applications. 'There's a phenomenal amount of data coming out of laboratory machines, and this is causing data management problems that are undermining the achievable value of that data.'

The availability of vast amounts of increasingly detailed data has, in parallel, spawned something of a philosophical change in how scientists perceive that data, Harland suggests. 'Lab work and clinical practices have traditionally been very hands on, and data was a bit of an inconvenience.' But now that so much R&D and clinical testing is automated, data has become completely central to results and decision making. 'Speak to the pharmaceutical industry and it's evident that data is now at the very heart of what they do, and much of their remit is to generate data that drives product development.'

Bringing in AI

Using AI to clean up and make sense of messy text-based information means that AI algorithms used downstream to analyse that information can then make full use of what is put in front of them, Harland continues. 'In a drug discovery environment being able to access and use information on the biology of a drug target, its potential side effects, comparisons with other molecules in the same class, and structural data, depends on the clarity and interoperability of data available in databases, and generated by instrumentation.'

From the user's perspective, and starting from first principles, the first stage is knowing where to find your data. Sounds simple, but Harland cites one company that acknowledged holding 148 different databases. 'And even once you know where your data is, you must know what to search for,' he says. The painkiller acetaminophen, for example, is known as paracetamol in the UK, but in the US its everyday name is Tylenol. Search for the wrong term and your software may not recognise that the two mean the same thing and two different people may obtain two sets of very different results.'

SciBite is addressing these sorts of first layer issues using AI as the foundation for tackling the problem of synonymy – paracetamol vs Tylenol – but going far beyond the capabilities of traditional search engines that treat words as just strings of letters, with no intrinsic meaning. Harland said: 'Many drugs have multiple trade names, for example, and every gene has many different names. It's one thing generating software that can recognise multiple terms, but with AI we train machines to understand that Tylenol and paracetamol, may be different words, but they are, in effect, the same drug, and once you can train a machine to recognise what a word "is", then those words start to have meaning.'

Understanding scientific meaning

SciBite's AI algorithms can be thought of as the plumbing that underpins text-based systems, such as a laboratory notebook or assay registration system. They allow software to understand scientific meaning, Harland suggests.

'Text itself is pretty much useless to a computer,' he said, 'but turn that text into data, and suddenly it becomes usable content. Think about a published scientific paper, which may contain a huge amount of scientific information, but is completely unusable from an analytical perspective. It's just a collection of words in a specific order.

'Take that same paper and run it through our software, however, and it outputs data that is interoperable with other datasets, and can be turned into more structured, machine-readable data that will work with downstream analytical algorithms.'

At the heart of SciBite's semantic analytics software suite is TERMite (term

identification, tagging and extraction), a named entity recognition (NER) and extraction API, which scans document text in real time, at about two million words a second.

'It can find a word, let's say Viagra, and it knows that Viagra is a drug. By looking at word usage and proximity of words in that document, the system can then figure out what Viagra is used for, and then apply that knowledge to find and extract information on Viagra from millions of other documents.'

TERMite has been developed as a system which pharma companies can 'plug-into' their existing analytical software, Harland explains.

'Just as you find a spellchecker inside a word processing software, our TERMite application can sit inside scientific data applications, making them instantly more intelligent.'

SciBite has also generated more than 100 ontologies containing many millions of synonyms across topics including genes, drugs, diseases, adverse events,

> "Once you can train a machine to recognise what a word 'is', then those words start to have meaning"

all of which are delivered through TERMite. The firm's TExpress software, which also works with TERMite-processed data, goes a step further and is able to find and extract semantic patterns of biomedical notation within sentences, such as text that describes how a specific gene defect leads to a certain disease.

'Over the next few years, as we work to improve the quality of our data even further, AI will be able to ask more sophisticated questions, such as "why" someone is looking at changes to features in cells. When we get to this point, AI will be able to add even further depth of insight, because it understands the "why" of that question. In the cell recognition example, this might be because we are looking for compounds that can treat cells by generating the changes we are looking for. And then we can start to use AI to look for similarities in the biology of how different compounds work.'

Costs and serendipity in drug discovery

The drug discovery process represents a huge financial and resource drain on the industry, and has historically endured a high candidate attrition rate, comments Andrew Hopkins, CEO of UK-based Exscientia. 'Traditional drug discovery operations account for about 35 per cent of the total cost of bringing a drug to market, and you may have to run 20 drug discovery projects, each one costing $15-20 million, even in the early, preclinical stage, just to get one molecule that will ultimately stand a chance of FDA approval.'

There has always been a large element of serendipity in the early stages of drug discovery to find promising 'hits', he notes, but the experience and insight of the scientist driving each project shouldn't be →

There's lots of data (large and small) in the life sciences but much of it is inaccessible, unstructured and of low quality. With the shift towards ML/AI approaches the quantity and quality of data must improve to ensure that trust in algorithmic output is high

→ underestimated. Exscientia has built an AI-driven drug design platform that automates the design and in silico assessment and optimisation of potentially millions of compounds against specified targets, to select the most promising candidates for further development. Steered by what the firm's CEO Andrew Hopkins terms seasoned [human] drug hunters, the platform's algorithms learn from the existing wealth of experimental, structural and 'omics' data that is already available on targets, diseases, and compound activity, and new experimentally derived data to bolster the learning dataset even further. Through this process the platform can design and then optimise candidate structures against designated targets, through design-make-test cycles.

It's a project-focused process that Hopkins maintains is faster than traditional high throughput screening-based approaches, and is significantly more likely to generate candidates that will ultimately succeed in the clinic. 'Exscientia's starting point was the premise that algorithmic methods can improve design efficiency through evolutionary approaches. What we asked was: how can we increase the efficiency and success rate of searching chemical space to design and optimise better drug candidates?'

Marrying human intuition with AI

Traditional drug design is founded on human interpretation of available data, the formulation of a hypothesis, and the chemical structures that may have the predicted properties against the desired target, Hopkins continues. 'This is a largely intuitive process, where you may make up to a couple of thousand molecules to solve individual problems.'

The firm's AI-driven platforms can effectively design and pre-evaluate millions of compounds to predict efficacy, selectivity and ADME – absorption, distribution, metabolism and excretion – against any selected targets. It's an active learning approach, rather than a deep learning approach, Hopkins says. 'Active learning methods are about asking which experiment will provide us with the most information to answer a question.' By asking the right questions, it can learn faster and generate a better design process.

> "Over the next few years, as we work to improve the quality of our data even further, AI will be able to ask more sophisticated questions, such as 'why' someone is looking at changes to features in cells"

A 'full-stack' drug discovery capabilities

To expand its in-house laboratory capabilities, Exscientia recently acquired UK biophysics specialist Kinetic Discovery, which has added protein engineering, biophysical screening and structural biology expertise to Exscientia's own drug design, pharmacology and computational platforms. Exscientia had been working with Kinetic Discovery through an ongoing drug discovery partnership with Evotec, and says the company is a perfect fit with its existing in-house capabilities.

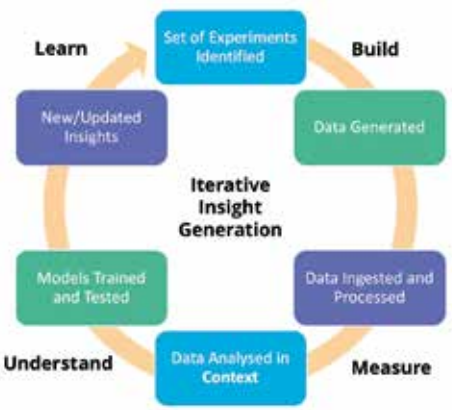
In combination with a recently constructed laboratory at expanded premises on the Oxford Science Park, the acquisition of Kinetic Discovery has effectively transformed Exscientia into a 'full stack' AI-driven drug discovery firm that can go from gene to clinical candidate for any druggable target selected, Hopkins claims. 'With the Kinetic Discovery acquisition, we now have in-house capacity to develop any assays, solve our own crystal structures and be in a strong position to build our own internal portfolio of drug candidates,' Hopkins notes.

'We spent the first five years focused on technical and market validation of the approach in real-world drug discovery projects with the industry, and we are now in a position to scale the platform up.' And with four AI-designed preclinical candidates now being developed by partners and in-house, Hopkins anticipates that the first of these will enter the clinic during 2019, adding a further layer of validation to the platform.

The success of the pharma and biopharma industries – and the drugs and diagnostics that they develop – thus ultimately relies on the experimental data that they generate, the analysis and interpretation of that data, and subsequent decisions made.

Tobias Kloepper, CEO at Aigenpulse, suggests it's a workflow that should incorporate all relevant experimental and relational data generated enterprise-wide at all points in discovery and development.

Aigenpulse has developed a modular, machine learning-driven platform that puts all of that experimental data and metadata in context. It applies analytical algorithms to underpin key questions with data, and enable efficient human interpretation and decision making. Historically, this level of information exploitation has not been practical, because key data often falls by the wayside. 'Scientists haven't had adequate computational tools,' Kloepper comments. 'They may store experimental



Iterative insight generation requires data, relevant context and modelling

output in Excel spreadsheets or flat-files. Merging multiple files in different formats becomes challenging and it is difficult, at that stage, to ask contextual questions of such data.'

More than just digitising data

Addressing the deficit is much more than just digitising experimental results, Kloepper continues. Digitising data is a fundamental goal for every industry, from retail to manufacturing. 'But in the life sciences you also have to keep up with a constantly evolving research environment that is developing new assays and new ways of using existing assays. Any intelligent digital platform needs to be able to evolve alongside that research process.'

The Aigenpulse platform employs machine learning to engage that analytical process so that scientists can have more confidence in their interpretation of those analyses. It's a concept for which about 80 per cent of the work involves getting the data into the right structured, contextual environment and the other 20 per cent is the machine learning to derive insight from that data, Kloepper notes. The two must go hand in hand, and there's little point in asking your algorithm to answer a question if you don't have reliable high-quality data available and structured appropriately.

Provide your algorithms with a robust dataset and the job becomes more seamless and accurate. Importantly, the Aigenpulse platform can work with data in several formats across a company, from laboratory information management systems (LIMS), electronic laboratory notebooks (ELNs) and other data repositories, to results of sample analyses captured using proprietary software or output in proprietary data formats. 'As well as contextualising that data, the Aigenpulse platform is designed to remove noise from such data, which allows the software to more accurately model patterns.'

"AI isn't going to make scientists redundant, but it will help scientists find the best answers to the questions they ask. In the next 10-20 years we will see every scientist being able to use machine learning algorithms as routinely as they carry out common laboratory assays today"

Advanced analytics can empower decision making when the algorithmic output is packaged in tools and software which are easy to use and interpret

Often such processes are about 95 per cent automated, but interfaces built into the software allow scientists to validate data to support the algorithm, Kloepper states. 'Our control vocabularies harmonise data, spanning gene expression datasets, common assays such a ELISA and FACS, or mass spectrometry analyses.' The software then looks at the structured information from every perspective, including disease, targets and compounds, so that the algorithm can learn and derive answers to specific questions set by scientists and biostatisticians.

Persistence of analytics

Try and do that with other platforms, or when all you have is PDFs and Excel files, and there will always be issues with data mapping and data matching, says Satnam Surae, chief product officer at Aigenpulse. 'We enable that persistence of analytics, which extends to running a machine learning pipeline on all of the data available to the whole company if necessary, or scaling things down to the level of individual experiments. Importantly, by retaining the data in the state it was output and at the time it was derived, you can compare models at different time points.'

The Aigenpulse platform provides scientists with a web-interface for their data and analytical results, which reduces complexity and optimises usability. 'Scientists can pull up the bits of data that they want, select the model or method that they want to run, put in the parameters, and click to set the analysis running,' Surae adds. 'The back end does all the work, and the output is displayed at the front end, in the most appropriate form, and in a matter of a few clicks.'

The Aigenpulse platform can be integrated into existing IT infrastructure on clients' premises or in the cloud, and it can be precisely configured to match the requirements of each client. 'Our aim is to help scientists derive greater insights into their research, through their data generated, and support them to ultimately develop better drug candidates faster and with less attrition,' Kloepper states.

Concerns that intelligent software will put jobs at risk are unfounded, he believes. 'AI isn't going to make scientists redundant, but what it will do is help scientists find the best answers to the questions that they ask. In the next 10 to 20 years we will see every scientist being able to use machine learning algorithms as routinely as they carry out common laboratory assays today.

'Scientists are very open to new technologies, and AI-driven tools will enable them to be more data driven in their decisions. And, ultimately, this will help industry develop more effective, safer drugs, faster.'

Simulation driven design in Paris

Robert Roe reports on developments in multiphysics simulation at the Global Altair Technology Conference

As the modelling and simulation industry arrived in Paris for the Global Altair Technology Conference at the end of October, attendees were treated to three days of conference content focusing on innovation and new technologies based on simulation-driven design.

This year's event focused on sharing applications of simulation-driven innovation from technology leaders and industry executives from all over the world, including keynotes from Ferrari, Jaguar Land Rover and Zaha Hadid Architects.

The keynotes were complemented by a host of technical sessions and presentations on the latest technology trends such as IoT and the digital twin, e-mobility, electric vehicle design – and the impact of IoT, AI and machine learning on influencing the future of design.

'Simulation, optimisation, cloud-based HPC, and the Internet of Things are coming together to transform how innovative products are brought to market,' commented James Scapa, Altair's founder, chairman, and CEO. 'As we move forward into an era of machine learning and artificial intelligence, opportunities will grow for the convergence of technology and emotion into great designs.'

This year's ATC in Paris had more than 800 participants and offered more than 140 presentations. The keynotes from Harvard Business School, Columbia University, JLR, Ferrari, Volvo Cars, LG, ClassNK, and Team Tao featured topics such as simulation-driven innovation in marine, electronics, artificial intelligence, e-mobility, autonomous drones (for deep sea applications), NVH and vehicle architecture (C123). One of the highlights, Jim Scapa announced in his keynote presentation, was the acquisition of Simsolid.

Altair acquires Simsolid

Simsolid works on full-fidelity CAD assemblies to provide fast, accurate, and robust structural simulation without requiring geometry simplification, cleanup, or meshing. Its underlying technology is based largely on the work of Dr Victor Apanovitch, a former professor at Belarus Polytechnic University and the cofounder of Simsolid Corporation.

'We believe Simsolid is a revolutionary technological breakthrough which will have a profound impact for product design,' said Scapa. 'It's incredibly fast, accurate, and robust and we believe a game changer for our industry.'

The Simsolid computational engine is a commercial implementation of novel and unpublished mathematics based on extensions to the theory of external approximations. The software controls solution accuracy using multi-pass adaptive analysis. Large and complex assemblies can be solved rapidly even on laptop computers.

> "As we move forward into an era of machine learning and artificial intelligence, opportunities will grow for the convergence of technology and emotion into great designs"

Dr Uwe Schramm, Altair's chief technical officer, noted that: 'We are very serious about solution accuracy. Others have tried to accelerate the interface between CAD and simulation by degrading the mathematical robustness. It is our feeling that by rapidly moving forward with the methods in Simsolid and expanding them across applications we can have a real effect on how design gets done while maintaining our high standards for computational excellence.'

Simulation-driven design

The first day of the three-day conference featured in-depth industry seminars with customer presentations from BMW, OHB, Alstom Transport, Protiq, Renishaw, Valeo, Scania, EOS, VibroTech Engineering, and alongside these presentations Altair experts provided insights into topics such as design for manufacturing, lead time reduction, fatigue simulation, design for AM, e-mobility, system simulation, digital twins and IoT.

Dave Harcourt/Shutterstock.com

> "We really see that we need to focus on mixed multiphysics simulations where we add more and more physics to our simulation portfolio"

On all three days the technology showcase of the event featured solutions of more than 20 Altair Partner Alliance members, software and hardware allies as well as two stages for product demos: one for the latest versions of the Altair HyperWorks and Altair Inspire product suites – including Simsolid – and one for Altair Partner Alliance member demos.

Altair CTOs Dr Uwe Schramm and James Dagg delivered a keynote presentation entitled: 'The World is Built on Engineering', which illustrated the power of Altair's simulation technology through several practical examples applied to some of the most common product development challenges.

Starting with simulation in the early phases of concept design, designers can take advantage of CFD simulation to explore and validate how different designs impact the aerodynamic performance of a vehicle. Conference attendees were shown how complex, detailed nonlinear airframe simulations can be modelled and performed quickly and accurately. The presenters will also explain how multiphysics problems like e-powertrain development can be solved with a model-based development approach, which considers electrical and electro-magnetic simulation of systems and sub-systems.

During the presentation Schramm noted that Altair has been developing solver solutions since around 1992 when Optistuct development was started. Then in 2007 the company acquired Radioss technology which was 'the real kick off for solver development'.

'In the last five years with the acquisition of electromagnetics technologies Feko and Flux we really can now simulate a complex world of systems with system simulation tools, manufacturing simulation and so on,' stated Schramm.

'If you look at the capabilities that we are developing for the design of electric motors and electrical machines such as self-driving cars, you can see that we provide an entire solutions portfolio throughout development through the engineering V. Going down in the development phase and going up in the verification and validation phase,' Schramm added.

'We really see that we need to focus on mixed multiphysics simulations where we add more and more physics to our simulation portfolio. Then of course computational performance is key, so you really want to optimise your structures and products so we need additional performance. And then of course you need pervasive optimisation, adding optimisation to any of these tools can help you to make quick design decisions,' said Schramm.

Schramm noted that engineering is becoming increasingly pervasive. 'Engineering is in almost everything we touch, a part of every human experience. It's in the planes and trains we travel in, the cars we drive and the appliances we use every day.'

However, engineering advanced products requires a meticulous, multiphysics development process, in which products are simulated, optimised, and verified many times to ensure the highest quality. Enabled by a broad simulation portfolio, and the effective utilisation of high-performance computing, Altair's platform aims to address all simulation needs throughout the product lifecycle.

'We have assembled an entire multiphysics systems simulation portfolio starting with structural, crash and safety impact simulation, thermal analysis, fluid flow, manufacturing and electromagnetic simulation,' said Schramm. 'We are not creating the mother of all solvers in this technology, we are connecting the multiphysics into the design problem and link this through optimisation technology.'

Quantum Cryptography: a new age in data security

Dr Andrew Shields highlights the importance of quantum cryptography in ensuring the future of data security

Yurchanka Siarhei/Shutterstock.com

90 per cent of all data created has been done so in the past two years, with proliferation now reaching the point that 2.5 quintillion bytes are generated every single day. Fuelled by greater mobility and emerging technologies, such as 5G and the Internet of Things (IoT), this data explosion will only continue and accelerate. With much of this data containing sensitive information, the challenge for organisations is in ensuring its secure transmission.

IT decision-makers evidently recognise this challenge, with the latest research from Toshiba revealing that, for over half (52 per cent) of businesses in Europe, data security is a top three investment priority for the year ahead. Yet despite this, businesses often fall behind in keeping themselves up-to-date and ahead of cyber-security trends and developments.

Enter quantum cryptography, which, by harnessing the principles of quantum physics, has the ability to usher in a new age of secure online communication. But what exactly makes quantum cryptography so secure, how can it fill the gaps in online defences, and when will businesses be able to take advantage?

Quantum cryptography provides a secure method for generating and distributing secret keys between two parties via an optical network. The inherent unpredictability in the state of particles, such as electrons or photons, enables quantum cryptography to be used to generate the random numbers needed for cryptographic applications. By harnessing this, it is possible to share a secret digital key that can be used for encrypting or authenticating information via streams of encoded single photons, which are sent through an optical communication network.

The future of quantum cryptography

The emergence of quantum computers will ultimately render much of today's encryption unsafe. Currently, it is widely considered that public key encryption is an essential part of data security, but that's being challenged by new attack strategies. Today's security challenges and tomorrow's security fears are driving the adoption of reliable quantum cryptography solutions and services to enable better data security. As a result, the global quantum cryptography market is forecast to grow from $285.7 million in 2017 to $943.7 million by 2022, a CAGR of 27 per cent according to Research and Markets.

Yet quantum cryptography's arrival is not as close to fruition nor as widely acknowledged within relevant fields. This doesn't detract from its essential and unparalleled value, as we move into the quantum age, so how far away from quantum cryptography are we, and what needs to be progressed to achieve this?

> "The emergence of quantum computers will ultimately render much of today's encryption unsafe"

Quantum cryptography in practice

Despite not yet being commercially available, scientists are beginning to deploy the technology to demonstrate its benefits. Toshiba's Cambridge Research Laboratory recently published a paper explaining a breakthrough made using a protocol known as Twin-Field QKD, extending the range of QKD to more than 500 kilometres of standard telecom fibre. This opens up the potential for secure communication between cities such as London, Paris, Dublin, Manchester and Amsterdam. Further still, large collaborative ventures, such as the Innovate UK EQUIP project and the EU Commission's Horizon 2020 programme, are also working to make QKD a valuable and accessible tool for the enterprise.

Planning for the future

Quantum cryptography has great potential to become the key technology for protecting communication infrastructure from cyber-attacks, and putting businesses on the front foot when protecting operation-critical information. It won't be long before we reach the point where corporations and government agencies will be able to utilise the technology to greatly advance their capabilities and day-to-day operations. However, if quantum-resistant public-key cryptography is not ready by this point, we simply won't be able to maintain current levels of information security.

Quantum-resistant public-key cryptography is clearly vital then, in ensuring we're able to continue life as normal in the age of quantum computing, and to live and work as conveniently as we do now. Unlike other existing security solutions, quantum cryptography is secure from all future advances in mathematics and computing, even from the number crunching abilities of a quantum computer.

Dr Andrew Shields is assistant managing director at The Cambridge Research Laboratory of Toshiba Research Europe

Subscribe for free*

SCIENTIFIC COMPUTING WORLD

Do you compute?

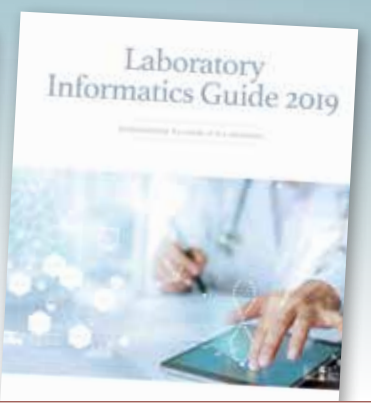
The only global publication for scientists and engineers using computing and software in their daily work

Do you subscribe? Register for free now!

scientific-computing.com/subscribe

*Registration required

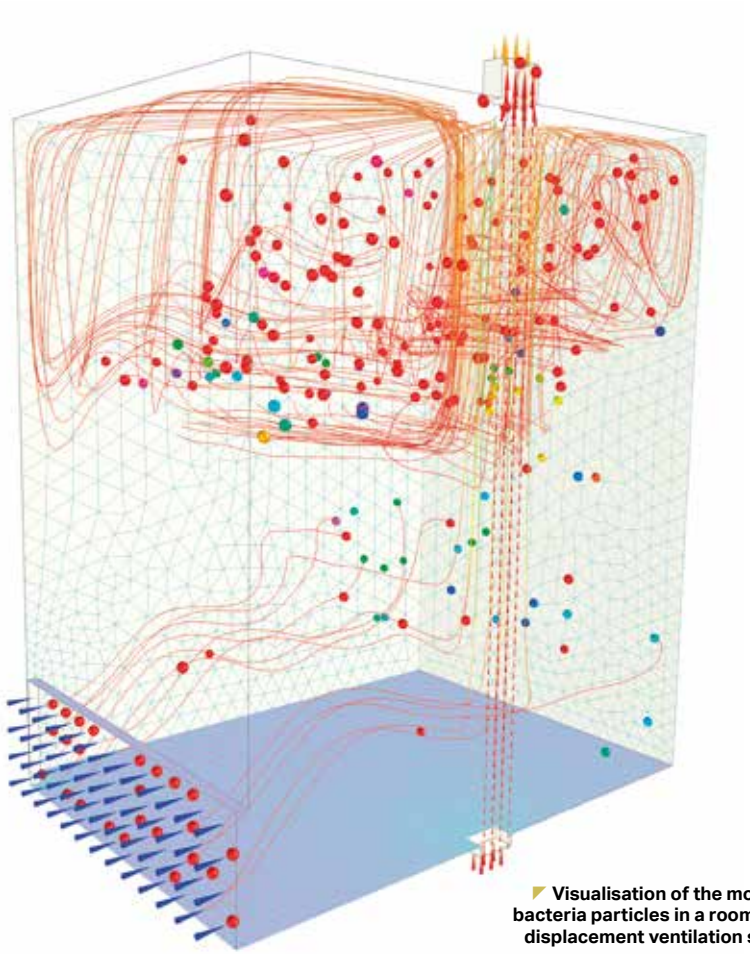
Also published by Europa Science



Model Medical Devices

GEMMA CHURCH REVEALS HOW SIMULATION AND MODELLING ARE AIDING THE DESIGN AND DEVELOPMENT OF A RANGE OF MEDICAL DEVICES

Simulation example made using Comsol Multiphysics software and provided courtesy of Comsol



Visualisation of the motion of bacteria particles in a room with a displacement ventilation system

The healthcare industry has been slow to adopt engineering simulation compared to other sectors.

However, an intrinsic need for modelling and simulation has become increasingly apparent as medical technologies continue to advance. Valerio Marra, marketing director at Comsol, explained: 'In an industry where safety is of paramount importance, the capability to investigate different scenarios by specifying boundary conditions, material properties and physiological mechanisms allows for early and harmless correction of design mistakes.'

The medical devices sector is now taking steps to integrate modelling and simulation into its design and development processes. Paul Goossens, vice president of engineering solutions at Maplesoft, said: 'While system-level modelling is something that companies in other industries, such as aerospace and automotive, have been using for decades, it is becoming a more popular tool in the medical device market.'

Goossens added: 'One of the driving factors is the current increase in safety and functionality issues within the medical device industry, and the growing concern surrounding product recalls. Products are removed from the market for several reasons and the complexity of many of these devices requires an early diagnosis in the design process, which is where early insights and diagnosis in system behaviour can allow for safer products without the time and costs associated with the standard cycles of medical device testing.'

Simulation and modelling reduce the lengthy timescales and high costs often associated with medical device developments. Medtronic, which makes such instruments, recently claimed the use of computer models helped reduce its time to market by two years and cut the cost of clinical trials by $10 million for a specific treatment.

Many practical issues associated with real medical tests can be overcome. Marra added: 'Especially for medical devices, the fact that simulation results can be accessed in locations where it would be impractical (if not impossible) to place sensors on a physical prototype, or in the human body, is greatly appreciated.'

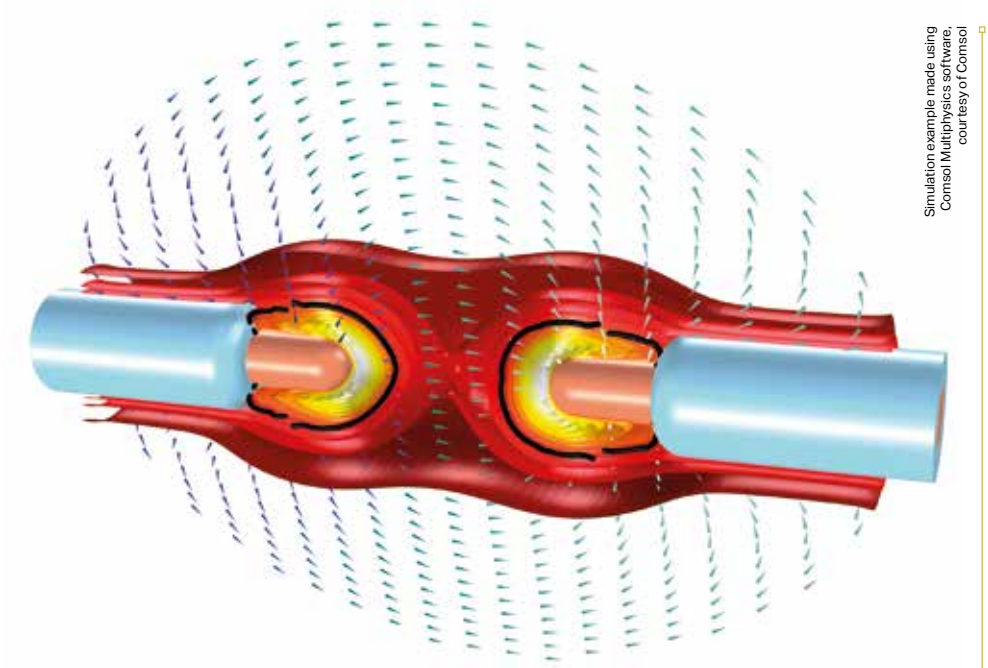
Multiphysics simulation can boost medical device design in many ways, according to Marra. It reduces the need for physical prototyping, makes available the measurements of any modelled variable at any point in a medical device and its surroundings, and provides high-fidelity medical device modelling.

However, a range of challenges still needs to be addressed and managing the complexities of both the human body and the medical devices is paramount.

Regarding the complexity of the medical devices, Goossens explained: 'A typical approach for the development of multi-domain systems often carries the risk of high costs and time-consuming re-engineering, due to the lack of interoperability between different domains.

'From powering systems to the mechanical, electrical and fluid components involved in medical devices, these multi-domain systems present many challenges, not only because of the complexity in modelling the related sub-systems, but also when it comes to interfacing each of these sub-models into one single integrated model,' Goossens added.

When it comes to the complexities of the human body, Thierry Marchal, global industry director for sports and healthcare at Ansys, said: 'This never-ending challenge requires more research and investment to improve the models in their accuracy, long-term predictability and



Simulation example made using Comsol Multiphysics software, courtesy of Comsol

Multiphysics simulation of tumour ablation. The localised heating of malignant tissue is achieved through the insertion of a four-armed electric probe. This model couples the bioheat and electric field equations, and models the temperature field in the tissue

ensure we fully understand the limitations of the current models.'

The human body is also a very different environment compared to the systems regularly simulated in established industries, such as aerospace or automotive. Milad Mafi, product marketing manager at SimScale, said: 'The biggest challenge from a simulation setup perspective is definitely geometry modelling. Simulation has its roots in classical mechanical engineering and all concepts and processes developed in recent years are designed for technical geometries and materials.'

A good example is the material characterisation of steel compared to bone. Mafi explained: 'While metallic materials can be described by isotropic material models, this is often not possible with biological materials. The bone gains high stiffness from the complex microstructures, which move in geometric scales of a few micrometres and cannot be resolved with current methods. It is, therefore, necessary to supply replacement models that are both technically and medically suitable.'

Restrictive regulations

These models also need to be validated to demonstrate that the model does accurately predict what will happen in the human body for various patients. Ansys recently collaborated with medical device companies, under the guidance of the FDA and ASME regulatory bodies, to address this point and develop new standards for the verification and validation of medical devices.

> "While system-level modelling is something that companies in other industries such as aerospace and automotive have been using for decades, it is becoming a more popular tool in the medical device market"

Such close collaboration between simulation companies and regulatory authorities could help accelerate adoption in the medical devices market. Ansys has also collaborated with numerous hospitals around the world and regulatory authorities in Europe, the USA and Asia. Marchal said: 'This [collaboration with regulators] is helping us understand what these local agencies and European Notified Bodies require in terms of model validation and results format to accept simulation results as evidence for their respective regulatory approval.'

It's not just collaboration with regulators that will help drive further adoption of simulation and modelling in the medical devices arena. Simulation experts and device designers are also collaborating at a growing rate.

Such cross-discipline work is facilitated through simulation applications, according to Marra, who added: 'Such apps allow the users to input the parameters needed to get the information they need, the stress in a bone structure or the temperature rise during an ablation process, without having to deal with the complexity of a multiphysics model, something their fellow simulation specialists took care of for them.'

SimScale has also developed a production-ready SaaS application for engineering simulation. It provides instant access to computational fluid dynamics and finite element analysis via a user-friendly web application. Mafi explained: 'In general, our users use simulation with SimScale to test cardiovascular stents, hip joint or arm prostheses, disposable pumps, in vivo blood flow or laboratory equipment.'

Abundant applications

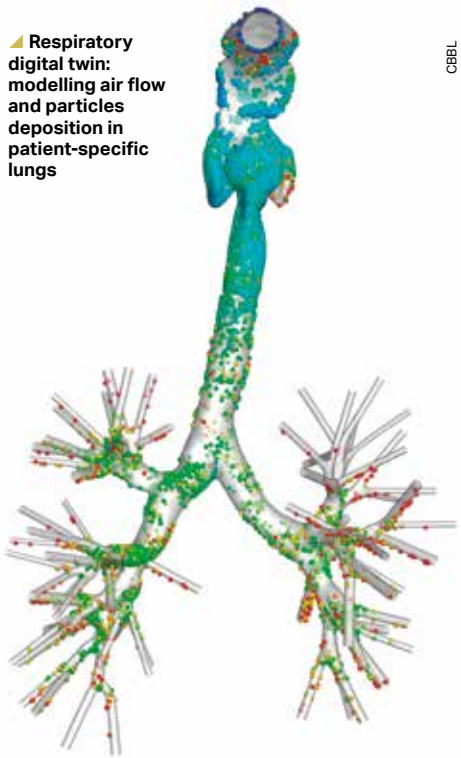
A wealth of application areas are now emerging in the medical devices sector. Mafi added: 'In my opinion, the leading star is the availability of patient-specific implants. Especially in the treatment of diseases of the cardiovascular system, such as arteriosclerosis but also in the treatment with joint replacement, simulation offers an unbelievable potential. FEA and CFD analysis play a decisive role in this context and will increasingly act as a catalyst in the coming years.'

Simulation and modelling are also helping the healthcare industry develop artificial biocompatible organs, such as an artificial heart, kidney and pancreas. Marchal explained: 'These devices are expected to be implanted into patients for an extended time. Thanks to simulation, their size, weight and necessary energy to support them are now compatible with patients' and medical professionals' expectation. This solution may soon solve the problem of lacking organ donors.'

Another emerging area is the adoption of simulation in surgical planning. Marchal explained: 'Creating a computer model of part of a patient gives the surgeon the luxury to test different surgical approaches, identify potential problems and select the best method before entering the operating room.

'This new approach is greatly relieving the stress from the surgeons, who can now quietly investigate various solutions without the stress of the patient waiting on the table.

'This also improves the outcome of the surgery, helping reduce the recovery →



Respiratory digital twin: modelling air flow and particles deposition in patient-specific lungs

CBBL

→ period and sometimes helping avoid dramatic situations,' Marchal added.

Rise of the robots

The use of robotics in medical devices is another evolving area, according to Goossens. In the biomechanics field, a team of researchers led by Dr Andrew James Smith, at York University, have used MapleSim to model autonomous battery operations in humanoid robots and electrical assistive devices.

Goossens said: 'The research group undertook the task of determining at what point in the transitions between sitting and standing energy can be regenerated in an orthosis or prosthesis, much like how a hybrid vehicle regenerates energy during braking by drawing it from the motor for re-use in the vehicle's operation.'

Biomechanical data from human trials was used to provide the desired trajectories for the simulations in a multi-domain model in MapleSim, allowing the robotic model to mimic human movements when transitioning between sitting and standing positions. Goossens said: 'The group's findings have a meaningful application for prostheses and orthoses design, and determining the most efficient battery autonomy means the operation time of these devices can be extended, and smaller, lighter batteries can be used, reducing their bulk. Ultimately, a more efficient device can reduce joint loads during standing-to-sitting for users – critical for people suffering from joint diseases.'

Robotic-assisted surgery is another developing field, which allows precision robotic tools to act like a surgeon's arms, hands and fingers, allowing surgeons to reach areas the human hand cannot reach without making large incisions.

Surgeons using haptic technology can control the movement of a robotic arm, but when they strike an obstacle during a surgical procedure, they feel the force of the obstacle against their hand and can take steps to avoid it.

Engineers at Quanser used Maple to develop the controllers both for the robot motion and for the haptic feedback to a surgeon's hands. The engineers first modelled the behaviour of the mechanism and, using Maple, developed the systems of differential equations that modelled the kinematics and dynamics of the system.

Goossens said: 'Once this was done, the team could very quickly test their model by solving the equations of motion, and develop the control strategies within the Maple environment. The work in this area has far-reaching and exciting implications for the future of robotically-assisted surgery, such as brain microsurgery, nanosurgery and telesurgery. The time for remote surgery with a sensory element is definitely here.'

Digital twins

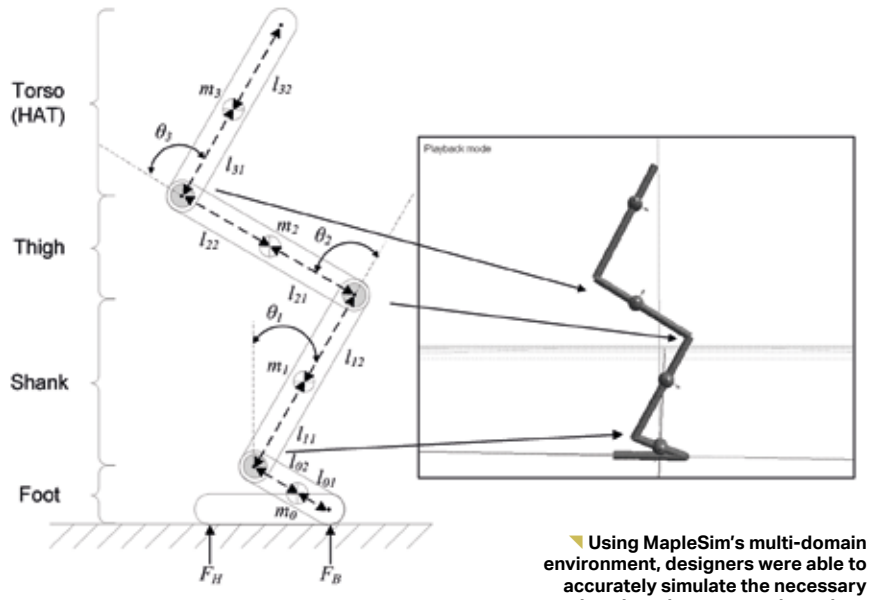
Digital twins are a popular emerging application area in the medical devices arena. Here, a computer model of the patient can test new devices before the real patient receives the device.

Goossens explained: '[A digital twin] creates the perfect companion for diagnostics, maintenance and product innovations. Modern design tools can create model-driven digital twins to assist in all stages of product design, and because it doesn't need test data to predict behaviour, it can be used for conceptual design before any physical prototype of a device has been constructed.'

> "This never-ending challenge requires more research and investment to improve the models in their accuracy, long-term predictability and ensure we fully understand the limitations of the current models"

This has positive implications for the patient, as Marchal explained: 'This allows medical professionals to implant medical devices and assess the impact on your body, or they can give your digital twin some medicines and assess how your body would react to specific doses.

'This is opening the door to predictive medicine which will be able to predict what could happen to you, and also preventative medicine able to suggest the treatment to prevent health issues before they impact your life. When digital twins become a widespread reality, we expect that the cost of healthcare will dramatically decrease, while each of us will enjoy a longer and healthier life,' Marchal added.



Using MapleSim's multi-domain environment, designers were able to accurately simulate the necessary motions in order to properly analyse battery autonomy in assistive devices

*Registration required

Webcasts
now available online

Creating a connected ecosystem to gain insights across R&D

Sponsored by ThermoFisher SCIENTIFIC

As data continues to increase exponentially, many organisations are turning to the latest advancements in digital technology to respond to the evolving needs of R&D scientists. For many organisations, a key challenge is managing the complexity and the number of systems in place that generate multiple types of complex data.

Ensuring Data Integrity & Facilitating Compliance with ISO 17025

Sponsored by ThermoFisher SCIENTIFIC

Today's quality landscape is incredibly complex with multiple data analysis, informatics, and enterprise systems. To ensure quality results, regulators across industries are focusing on data integrity, data security and consistent laboratory practices.

Laboratory 4.0: Moving Beyond Digitalization in the Lab

Sponsored by BIOVIA

Today's laboratories are becoming increasingly complex, with ever more data being generated and captured. At the same time, regulatory oversight is stronger than ever and places new compliance burdens on everyday operations.

Leveraging Machine Learning for Decision Making in the Materials Sciences

Sponsored by BIOVIA

Machine learning and Big Data analytics offer asignificant opportunities to improve R&D in the materials sciences, providing scientists with a new set of tools to analyze their data.

www.scientific-computing.com/webcasts

SCIENTIFIC COMPUTING WORLD

HIGH-PERFORMANCE COMPUTING

University of Birmingham to deploy AI cluster

Elnur/Shutterstock.com

Researchers at the University of Birmingham will soon be able to carry out research on the largest IBM POWER9 Artificial Intelligence (AI) cluster in the UK, as the university has now announced the deployment alongside HPC integrator OCF. OCF and the university will integrate a total of 11 IBM POWER9-based IBM Power Systems servers into its existing high-performance computing (HPC) infrastructure, the Birmingham Environment for Academic Research (BEAR).

Birmingham initially deployed two IBM Power Systems AC922 servers, powered by POWER9 CPUs with the industry's only CPU-to-GPU Nvidia NVLink interconnect, in September. However, the advanced research computing (ARC) team soon realised it needed more computational power tailored to increasing AI workloads generated by the university's researchers, delivering computational vision analysis and to solve life science challenges, such as cancer diagnosis.

'It's very important to us, as a research-led institution, that we are at the forefront of data research, which means we are always looking at ways to make AI quicker and more accessible for our researchers,' said Simon Thompson, research computing infrastructure architect at the University of Birmingham.

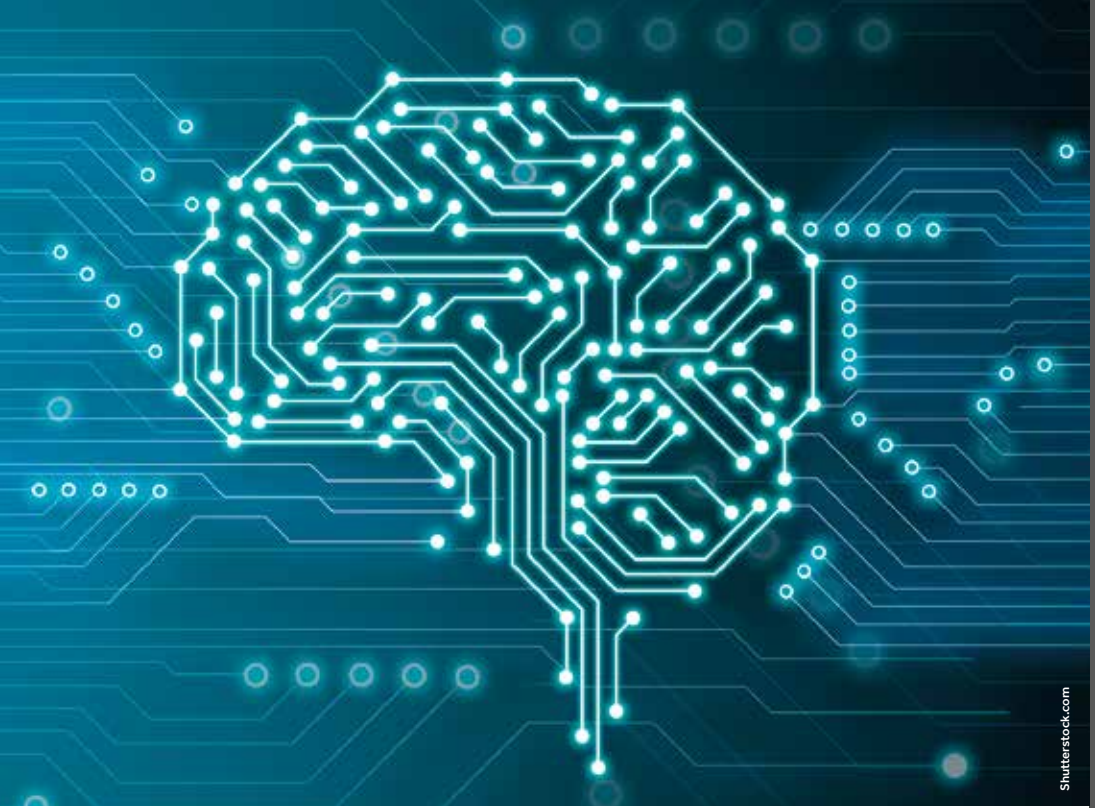
'With the sheer amount of data, the common questions from researchers are how can we analyse it fast enough and how can we make the process even quicker? With our early deployment of the two IBM POWER9 servers, we have seen what is possible. By scaling up, we can keep pace with escalating demand, and offer computational capacity and capability to attract leading researchers.'

The university will now add an additional nine IBM Power Systems AC922 warm water-cooled nodes, each equipped with four Nvidia Tesla V100 16GB Tensor Core GPUs, 1TB of system memory, dual 18 core POWER9 CPUs and Mellanox 100Gb EDR InfiniBand. The system uses IBM PowerAI Enterprise software, unlocking potential for accelerated computing, capitalising on the largest IBM POWER9 cluster in the UK. IBM will also support the new systems by providing comprehensive training and support to Birmingham's researchers in partnership with ARC.

This significant enhancement to BEAR means a more powerful and versatile computing environment for researchers. For example, fellows from The Alan Turing Institute looking at early diagnosis of – and new therapies for – heart disease and cancer, will use AI to run faster diagnostics.

In contrast, researchers in the physical sciences are similarly using machine learning and data science approaches to quantify the 4D (3D plus time) microstructures of advanced materials collected at national large synchrotron facilities, such as the Diamond Light Source. This research expects to use the large model support provided by IBM PowerAI software to analyse TBs of data generated daily; currently an almost impossible task.

'We are thrilled the university has decided to invest in building the UK's largest POWER9 AI cluster', said Simon Robertson, director, IBM Servers, UK & Ireland. Julian Fielden, managing director of OCF, added: 'The University is leading the way with this impressive project and will continue to attract world-class researchers with this type of innovation.'



HIGH-PERFORMANCE COMPUTING

Inspur and Intel collaborate to accelerate AI innovation

Intel and Inspur have announced a new collaboration, X-Plan, which aims to accelerate AI innovation.

The collaboration covers several items, one of which is the joint innovation and software stack optimisation of AI servers around Intel's next-generation chips, including the Purley Platform's new micro-architecture Cascade Lake-SP and neural network processors NNP-L and NNP-I.

Based on Intel VNNI's technology and Inspur's experience in AI computing markets, both parties will provide powerful and flexible AI accelerated computing and customisable computing architecture systems, so as to help global AI users cope with computing challenges and deliver faster innovations.

'Currently, the trends of AI industrialisation are becoming more and more obvious. AI computing determines the speed of AI innovation. X-Plan will, via AI computing product innovation and talent training, inject the vitality of continuous innovation into global AI R&D institutions and commercial companies to promote the improvement and development of industrial AI,' said Liu Jun, GM of AI and HPC at Inspur.

The Inspur and Intel joint AI laboratory will bring together engineering and R&D teams from both companies and focus on innovation in AI computing, algorithms and applications, to provide strong R&D support for the X-Plan. In addition, Inspur and Intel plan to collaborate extensively with universities and research institutions to promote the development and improvement of AI ecosystems through talent training.

'AI has become an important form of computing for the present and the future. Its development depends on a good ecosystem. The X-Plan will become an important part of Intel's global AI computing ecosystem, with the goal of providing the global customers with efficient AI computing products and building a good talent development environment,' said Nash Palaniswamy, senior director of global technical computing at Intel.

Inspur has been investing heavily in the development of AI infrastructures in four layers, including a comprehensive computing platform, a complete management and performance suite, optimised deep learning frameworks, and end-to-end, agile, cost-efficient AI solutions for its industry customers. Inspur has become the most important AI server supplier for Baidu, Alibaba and Tencent, and helps AI customers increase compute performance in voice, image, IVA, search, networking etc.

In IDC's 2017 China AI Infrastructure Market Survey Report, Inspur had a 57 per cent market share in AI servers.

MODELLING AND SIMULATION

Altair to acquire Datawatch

Altair has announced a definitive merger agreement under which Altair has agreed to acquire Datawatch. Altair will pay $13.10 per share in cash, representing a fully diluted equity value of approximately $176 million. The transaction was unanimously approved by the Boards of Directors of both companies.

James Scapa, Altair's founder, chairman, and chief executive officer, commented: 'Bringing Datawatch into Altair should result in a powerful offering consistent with our vision to transform product design and decision making by applying simulation, data science and optimisation throughout product lifecycles. We see a convergence of simulation with the application of machine learning technology to live and historical sensor data as essential to creating better products, marketing them efficiently, and optimising their in-service performance. Datawatch is a great team of people with best-in-class products, and we look forward to them joining us.'

Altair believes the acquisition of Datawatch will be useful in a number of different markets that want to leverage data analytics and data science technologies. Datawatch's solutions, which include data prep, data prediction, and real-time high-volume data visualisation technologies, are highly relevant and applicable to almost any company and vertical market. Altair also reports that there is opportunity to cross-sell Datawatch products into Altair's primarily manufacturing customer base.

Michael Morrison, chief executive officer of Datawatch, added: 'The Datawatch team is excited to join Altair and benefit from its long track record of success with developing and bringing to market highly differentiated software technology across diverse industry verticals. We feel great about the cultural alignment and look forward to driving continued innovation in our market-leading solutions as an integral part of Altair's vision.'

LABORATORY INFORMATICS

ACD/Labs' announces software update

Shutterstock.com

ACD/Labs has announced significant updates across its ACD/Spectrus and Percepta platforms. Version 2018.1 delivers key tools for analytical and process chemists in pharmaceutical development, NMR spectroscopists, mass spectrometrists, plus LC and GC method development specialists.

'Scientific organisations expect a wealth of innovative capabilities in the applications they use to help them solve their most complex scientific challenges. But, they also expect those applications to be continually updated and improved,' said Andrew Anderson, vice president of innovation and informatics strategy at ACD/Labs. 'This software release offers a wealth of new and valuable functionality – we believe it represents a commitment to our customers to help them improve their productivity, provide deep scientific insight, and facilitate decision-making.'

ACD/Labs' v2018.1 software release includes the following updates:

Analytical Knowledge Management in Pharmaceutical Development – taking into account specific feedback from customers, v2018.1 of LuminataMetaSense include multiple requested features for both products. Luminata now includes automated calculation of impurity carryover across stages, and improved handling of process schema imported from third-party drawing packages – including reagents and reactants. MetaSense users can now identify a greater number of metabolites through the addition of user-defined reactions, expand the scope of projects by analysing data from multiple different species (human, mouse, rat, dog) together, and accelerate reporting by utilising customisable reporting templates.

NMR and MS dereplication – two new features have been introduced that streamline identification of 'known unknowns' in both NMR and MS analysis. The Known Structure Identification NMR workflow, available as an add-on to both ACD/NMR Workbook Suite and ACD/Spectrus Processor, quickly determines whether experimental 13C spectra match predicted signals of known compounds. Similarly, improvements to the Intelligent Component Recognition workflow in ACD/MS Workbook Suite enables efficient identification of LC and GC known unknowns through batch spectral searching. Additional key NMR and MS updates include improved peak picking and multiplet assignment accuracy for 13C NMR spectra of fluorinated compounds, and upgraded MS Tree control to better visualise multiple MS spectra together.

Usability improvements to method development solutions – in response to customers communicating their chromatography software needs, ACD/Labs has simplified the v2018.1 product line by consolidating functionality from across various chromatography applications into two comprehensive method development and optimisation solutions: the improved ACD/AutoChrom, and the new ACD/Method Selection Suite.

In particular, Method Selection Suite combines physicochemical property predictions with method optimisation tools to define better starting conditions, estimate retention times, and optimise key separation parameters, all according to quality by design (QbD) principles.

Physicochemical and ADME/Tox Property Prediction – v2018.1 of the Percepta platform, which provides in silico predictions of physchem, ADME and toxicity properties, includes significant improvements to the logP structure database (>1700 compounds added) to deliver more reliable and accurate logP estimates for novel compounds.

Plus, this expansion also improves prediction tools that require logP – logD, oral bioavailability, blood-brain barrier penetration, and more.

LABORATORY INFORMATICS

Autoscribe announces advanced AQC and multi-sample visualisation

Autoscribe Informatics has announced the release of v5.4.6 of Matrix Gemini LIMS. New features include improved visualisation of data and information, especially hierarchical data structures and storage locations including freezers, well plates and other storage types, and advanced analytical quality control (AQC) functions.

The new Container View allows container types to be represented in the LIMS, including well plates, storage racks, cryogenic storage boxes and chemical inventory storage facilities. Each point, or location, within the container grid can represent a sample (or item) and may be configured to show any information about the item that is desired.

The location of individual items, or groups of items, within the storage system may be changed using drag and drop features. The system will also allow entire containers (and their contents) to be moved. All changes can be recorded in an audit trail for later recall, providing a complete audit history of the sample location during storage. The Container view, in conjunction with the Matrix Tree view functionality, allows laboratories to define and manage any hierarchically defined information within the laboratory.

The Container View is not limited to displaying items in a storage facility. Inherent flexibility means it can be used to visualise many types of information, for example, data heat maps or the mapping of contamination hot spots in a facility.

It can also be used for resource mapping and allocation, including the creation of Kanban boards. Samples, instruments or people (or any resources) can be defined and allocated using the built-in drag and drop methodology. Like all Matrix Gemini LIMS features, the container view is configured using Matrix Gemini's powerful built-in graphical configuration tools, allowing maximum flexibility, without the need for software coding skills.

Advanced AQC provides highly flexible run sheet functionality, allowing users to create run sheet templates that meet the needs of today's analytical laboratory. AQC samples including blanks, duplicates, replicates, spikes, controls, spike duplicates and control duplicates and their position within a run, can be defined based on user defined patterns.

Advanced AQC allows complex rules to be used when placing control samples among standard samples within a run sheet.

These rules include set positioning of individual QC samples, leading and trailing sets of QC samples, as well as repeating patterns and randomised positioning within sets of unknown samples.

Advanced AQC allows the laboratory to maximise sample throughput, while meeting regulatory and quality standards, and without compromising the quality or accuracy of data.

It can also drive improved efficiency and maximise the use of resources across the laboratory.

Matrix Gemini v5.4.6 is now available for immediate delivery to all Autoscribe Informatics customers.

LABORATORY INFORMATICS

IDBS continues expansion with key appointments

IDBS, a provider of enterprise scientific informatics platforms, has announced the appointment of Graeme Dennis as pre-clinical pharma commercial director, Abhay Kini as director of product management and April Pisek as solutions consultant.

These appointments strengthen IDBS' ability to drive the creation of labs of the future, the company says, as well as its preclinical footprint in the scientific informatics market – furthering its position as a provider of data management solutions and services to preclinical and bioanalytical organisations.

Christian Marcazzo, IDBS general manager said: 'Attracting talent like Graeme, Abhay and April underlines the importance of bringing scientific talent to our product and go-to-market strategy. These appointments bring us deep domain expertise and a fundamental understanding of how technology helps our customers discover and develop new therapeutics.'

Graeme Dennis will be responsible for driving IDBS' expansion in scientific informatics strategy, implementation and integration. Graeme has a 15-year track record in pre-clinical pharma and discovery informatics, holding senior roles at Accenture's LabAnswer, Dotmatics, Vanderbilt University Medical Center and Harvard University. He has a Bachelor of Science degree in chemistry, from Vanderbilt University, and is based in Boston, MA.

Abhay Kini will help deliver IDBS' vision of a digital lab, driving product marketing, management and strategy. Abhay's career spans more than 15 years in R&D and consulting services in life sciences, holding positions at Waters Corporation, Medidata Solutions and Oracle. He has a PhD in chemistry from Purdue University in West Lafayette, Indiana, and pursued post-doctoral research in the molecular bases of olfaction at Columbia University.

April Pisek brings a decade of expertise in quantitative bioanalysis and drug development support for small biotech and large pharma companies. Prior to joining IDBS, April worked at AIT Bioscience where she held a succession of positions including ELN administrator, building templates from sample prep to anti-drug antibody. She has a Bachelor of Science degree in chemistry from Ball Start University and is based in Indianapolis, Indiana.

These appointments follow the earlier announcement this year of Marcazzo as general manager.

Dennis said: 'I'm excited to join IDBS – its application of digital technologies that transform scientific and laboratory processes has set the standard in scientific informatics solutions.'

Kini said: 'I am very pleased to have the opportunity to join IDBS. I look forward to bringing my experience and skills to a team that is committed to delivering the highest levels of customer satisfaction across all aspects of the business.'

While Pisek added: 'When it comes to supporting drug development, IDBS' technology and its commitment to delivering a high-quality service defined from the customer's point of view is a big reason why I joined the company. I am delighted to be part of the team.'

Suppliers' Directory

Find the suppliers you need quickly and easily

www.scientific-computing.com/suppliers

HIGH-PERFORMANCE COMPUTING

Altair Engineering, Inc
(248) 614-2400
info@altair.com
www.altair.com
■ Software

Boston Limited
+44 (0)1727 876 100
sales@boston.co.uk
www.boston.co.uk
■ Systems Integrator

Calyos SA
0032 4 75 73 54 65
newbusiness@calyos-tm.com
www.calyos-tm.com
■ Cooling

Cluster Vision
+31 20 407 7550
marketing@clustervision.com
www.clustervision.com
■ Software

Eurotech Spa
+39 0433 485 411
welcome@eurotech.com
www.eurotech.com
■ Systems Integrator

Numerical Algorithms Group (NAG)
+44 (0)1865 511 245
katie.ohare@nag.co.uk
www.nag.co.uk
■ Software

OCF
+44 (0) 114 257 2200
scw@ocf.co.uk
www.ocf.co.uk
■ Systems Integrator

LABORATORY INFORMATICS

Abbott Informatics
+44 (0)161 711 0340
AISalesUK@abbott.com
www.abbottinformatics.com
■ LIMS
■ Scientific Document Management Systems

Accelerated Technology Laboratories
+1 910 673 8165 (outside of US)
800 565 LIMS (5467) US and Canada
info@atlab.com
www.atlab.com
■ Bioinformatics
■ LIMS

Amphora Research Systems
+44 (0)845 230 0160
info@amphora-research.com
www.amphora-research.com
■ Electronic Laboratory Notebooks

Autoscribe
+44 (0) 118 984 0610
info@autoscribeinformatics.com
www.autoscribeinformatics.com
■ Data Management
■ Electronic Laboratory Notebooks
■ LIMS
■ Scientific Document Management Systems

Biovia
+ 1 858 799 5000
www.3ds.com/products-services/biovia
■ LIMS
■ Electronic Laboratory Notebooks

Eusoft
+39 080 5426799
info@eusoft.it
www.eusoft.it
■ LIMS

IDBS
info@idbs.com
www.idbs.com
■ Electronic Laboratory Notebooks

InfoChem
+49 (0)8958 3002
info@infochem.de
www.infochem.de
■ Cheminformatics

LabWare
infoEU@labware.com
www.labware.com
■ LIMS

Osthus GmbH
+49 241 943 140
office@osthus.de
www.osthus.de
■ Electronic Laboratory Notebooks
■ Cheminformatics
■ LIMS

Siemens
+1 322 536 2139
industrial-it.swe@siemens.com
www.siemens.com/simaticit-rdsuite
www.siemens.com/industrial-it/lims
■ Electronic Laboratory Notebooks
■ LIMS

Thermo Fisher Scientific
+44 (0)161 942 3000
marketing.informatics@thermofisher.com
www.thermoscientific.com/informatics
■ LIMS
■ Scientific Document Management Systems
■ Chromatography Data Systems
■ Spectroscopy Software

MODELLING AND SIMULATION

Altair Engineering, Inc
(248) 614-2400
info@altair.com
www.altair.com
■ Software

Integrated Engineering Software
+1 204 632 5636
info@integratedsoft.com
www.integratedsoft.com
■ Mathematics, Simulation and Modelling

LABWARE 7

LIMS and ELN together in a single integrated software platform.
A laboratory automation solution for the entire enterprise.

LABWARE
Results Count

Offices worldwide supporting customers
in more than 100 countries

www.labware.com

Leverage the power of digital science for smoother, faster research results.

Every innovation starts with an idea. And the right digital science platform, can help that idea become a reality. Thermo Fisher™ Platform for Science™ software connects your people, instruments, equipment and scientific apps, so you can work at the bench with the confidence that your data is always ready to analyze, scale and share.

Find out more at **thermofisher.com/digitalscience**

ThermoFisher SCIENTIFIC