Drive your laboratory digital transformation with best-in-class cloud-based platform to support operational process from R&D, QA/QC through manufacturing.

- Accelerate your R&D for faster time to market
- Improve collaboration across your departments
- Enhance your decision-making capabilities
- Meet easily the increasing regulatory requirements

Powered by ORACLE® Cloud

www.agilab.com
A s we head into 2019, the Laboratory Informatics Guide reflects a shift in the way that laboratories are used and operated. The impact of new technologies and regulations are driving changes to the way that data is handled, stored, and analysed, leading to new challenges and innovation in laboratory software. On page 4 we have a contributed article from Darren Barrington-Light, from Thermo Fisher Scientific, which explores the importance of integrating LIMS systems into the pharmaceutical data chain. The article looks at some of the key features that LIMS can provide to pharmaceutical companies in strengthening data integrity.

On page 4 we have a feature which looks at the changing landscape of drug discovery. We speak to informatics experts about new regulation and the need to maintain a competitive edge through the application of new technology, such as predictive analytics. As drug discovery has shifted to a model that focuses on outsourcing informatics, software packages must also adapt to provide managers and decision makers with the ability to look across data created in-house but also through partners and contract research organisations.

Starting on page 12 you can find a feature which looks at changes in laboratory informatics software. We interview major informatics software providers to find out what they think will happen over the next 12 months. There is a particular focus on data and emerging technologies, such as AI and deep learning.

On page 16 we investigate the ‘lab of the future’. This began as an initiative from the Pistoia Alliance earlier this year, but has now been expanded to cover all software providers looking at the future of laboratory operations. The article focuses on new technologies such as the IoT, AI and deep learning, to find out how software providers are adapting to meet the demands of laboratories in the future.

In a contributed article on page 22 Dr Alexander Jarasch and Professor Martin Hrabě de Angelis explain that the increasing rates of data creation may call for innovative methods to help manage and analyse data. To overcome this challenge the researchers are investigating the use of graph technology.

On page 24 we have an article from Mark Newton, from Heatland QA, on the dangers of research misconduct. Mark reports that this could be a huge concern for regulatory authorities, eventually forcing major changes in the way laboratory data is summarised, reviewed and analysed, leading to new challenges and innovation in laboratory software.
Maintaining data integrity

Darren Barrington-Light, senior manager product marketing for Thermo Fisher Scientific, explains the importance of integrating a LIMS software package into the pharmaceutical data chain.

In today’s world, where drug development integrates science and technology, consumer safety is paramount in the pharmaceutical industry. Any drug released must adhere to strict regulatory standards and have robust data to prove the results are an accurate reflection of the development process.

The maintenance and assurance of data accuracy and reliability across processes, including compound synthesis, production optimisation, quality control and product release, are critical aspects of drug development. Without managing data integrity across all of the phases of discovery and development, process errors can go unchecked, procedure and activity tracking becomes unattainable, and multitudes of data can be lost. This can result in sub-par products that may be a hazard to human health.

Data integrity is defined as the completeness, accuracy and consistency of data, indicated by the absence of inconsistencies or alterations in prescribed methods. Companies strive for data integrity and are required to maintain it through compliance with regulations set by government agencies and safety organisations. As the pharmaceutical industry develops medicines, they are held to the highest standards of data management compared to other industries.

Ensuring data integrity is not a simple task. With such a complex chain of events and multiple groups working together, preserving data integrity across an entire process is an enormous challenge. Consider the need to track every lab employee’s activity, every instrument calibration, use and reading, every compound analysis step and every quality control point. The data collection alone can quickly become a significant undertaking. Taking an informatics view of data acquisition and management builds in and automates data integrity.

Guiding principles

The pharmaceutical industry generates a huge amount of information through the research, discovery and development of drug products. From gaining a new understanding of metabolic pathways that influence drug efficacy, to synthesising new compounds discovered to be key in a new treatment, labs constantly generate data. Due to this regular influx of information, labs can adhere to principles that help maintain quality in processes and output.

Regulatory agencies, including the US Food and Drug Administration (FDA) and the European Medicines Agency (EMA) recommend five key principles to assist labs in good industry practices. The set of principles that goes by the acronym ALCOA governs the way that data can be generated, collected and managed. According to ALCOA, all data associated with a pharmaceutical product should be attributable, legitimate, contemporaneous, original and accurate, in addition to complete, consistent and enduring.
Following these principles provides an opportunity for pharmaceutical labs to ensure a high quality standard that can stand the test of time and regulatory reviews.

On top of this desire to keep data consistent with ALCOA, labs are required to comply with regulations set in place to create an assurance of quality among regulators and consumers. Regulation compliance can only be confirmed with regular audits and reviews of laboratory processes and data management systems. In order to successfully demonstrate that a lab is in compliance, evidence of a monitored production chain across the complex phases of drug development must be readily available.

The data collection web

Obtaining accurate and integrated data across all networks within the pharmaceutical development, manufacturing and supply chains can mean the difference between success and failure. Data comes from everywhere within a pharmaceutical network and creates a web of information that can be organised into one location for analysis. Following ALCOA principles and complying with regulations on data management allows the industry to maintain public trust and brand value.

The need to have good data collection and management systems is central to any lab’s priorities. If a lab analyst records out-of-specification data from a quality control check in an uncontrolled worksheet, then validates and passes the product on official reporting documents, the data chain has been broken. Even if a second quality control check validated the product, the data was not managed appropriately, leading to disputable results and a product that cannot move forward. Aggregating instrument use and all activities, monitoring automated SOPs, and using electronic laboratory notebooks (ELNs) can significantly ease not only time spent in data management, but also in accountability and accuracy.

To err is to be human

Human error is inevitable in any circumstance, whether in a lab or in the office. Companies can introduce automated systems and regulated processes to minimise the effects of error on a firm’s bottom line. But, in the pharmaceutical industry, human error and tendency for variability can be detrimental. Mistakes that create downstream inconsistencies can ultimately result in an ineffective or unsafe product.

Because of the extreme pressures put on labs dealing with drug discovery, development and production, automation becomes even more important and integral in each process. Adopting electronic SOPs and ELNs, and integrating these with a scientific data management system (SDMS), is the first step in increasing accountability and minimising errors. Electronic processes such as these create automatic check points within each phase of a protocol or longer term procedure, so any inconsistency observed by the system at any point in a process can be caught and corrected before the next step.

This is extremely valuable for manual procedures, including working with instrumentation, running through a QC protocol, or measuring values. For example, users not adhering to SOPs with analytical techniques such as GC-MS and LC-MS used in quality control can affect the data integrity chain. Manual integrations of peaks in chromatograms can be inconsistent if not done automatically. Implementing suitable controls contained in modern chromatography data systems (CDS) platforms delivers reproducible results using sophisticated identification and integration algorithms. Integrated laboratory information systems (LIMS) can include automated CDS in this case to enter, store, track and trace data for

Without managing data integrity across all of the phases of discovery and development, process errors can go unchecked, procedure and activity tracking becomes unattainable, and multitudes of data can be lost.

The success of a connected lab

By adopting integrated data systems, from LIMS and CDS to ELNs and SDMS, pharmaceutical companies immediately solve consistency issues, can clearly define protocols and accountability, and centralise data collection and maintenance. These systems can enhance data quality throughout the development process, ensuring good laboratory practice (GLP) and good manufacturing practice (GMP).

Integrated informatics is useful in synchronising and centralising activities for members of a lab network. These platforms can increase operational efficiency, effectiveness and flexibility, improve brand integrity and reduce costs and issues. By easing data retrieval for further analysis and review, adverse events can be raised before they affect the development pipeline. Corrective actions and quick decisions can be made before an error can amplify or even occur. These benefits result in higher accuracy, faster processing and better visibility.

A LIMS can securely archive instrument readings, method parameters and test data associated with a particular sample, in addition to any user interactions made with the software. As a result, every activity can be electronically documented, along with the identity of the individual who performed it, ensuring complete transparency throughout an entire workflow.

Providing comprehensive audit trails and fully searchable workflows can help limit the need to manually record data or inaccurately enter information. A search for activities involving too few steps, or those that have been interrupted or aborted, can quickly reveal actions which can subsequently be used by supervisors to investigate nonconforming procedures.

Analysis of integrated LIMS data can assist in forecasting demand and planning. Examining usable data allows managers to audit protocols or update regulations to comply with changes.

Managing control

Implementing and integrating various process workflows keeps a lab running at its highest

efficiency, and can ease reviews, maintain audit trails and allow better control of experiments. Process traceability makes data retrieval simple. In reviewing a process for compliance and quality, data can be searched and the relevant information easily recalled. Not only can this safeguard robust protocols and ensure adherence to regulatory standards, but can immediately call out any inconsistencies or user error. By integrating systems into one centralised LIMS, reviews become a simple process that only assists in proving quality in processes, instead of alerting regulators to inaccurate results.

In order to ensure superior data quality and that ALCOA principles are adhered to, integrated data systems are the key to success. An integrated LIMS works within the pharmaceutical data chain to streamline and automate procedures, and improve governance of manual input. Using connected systems provides traceability across product workflow that instils end-user trust in product safety. While regulatory authorities continue to raise expectations for data integrity, pharmaceutical manufacturers are increasingly looking to exceed basic data management by using integrated informatics to track entire processes from discovery to market.

www.scientific-computing.com/lig2019
Drug development is facing change – both from technological pressures, such as the use of AI and machine learning, plus new regulations which are driving sweeping changes to the way electronic records are created and stored for clinical trials.

Informatics software providers are now beginning to provide software which can help to facilitate this change, but also manage and derive insights from the vast quantities of data that is being recorded.

Semantic search, metadata, indexing and more sophisticated software packages for predictive analytics, and even AI and machine learning approaches, are being used to fill the gap between traditional methods and the data-driven paradigms that are being used in today’s laboratories and research centres, as well as pharma biotech and contract research companies.

The need to control and derive insights from data is becoming crucial in order to maintain a competitive advantage. This is driving a focus on data exchange, digital transformation and analytical capabilities that can help to provide more value from data and, at the same time, help organisations to work within more stringent regulations.

Andrew Anderson, vice president, innovation and informatics strategy at ACD/Labs, said: ‘My training just before taking this role was in looking at industrial insights, and translating the innovation plans that come from those insights. What I see in drug development is that, if you go back 20 years when a new drug was matriculating through a clinical process, there would be an asset investment based on the probability or likelihood of what the rate of approval might be.’

This approach required pharma companies to create major assets, such as manufacturing plants that companies could use to take raw material and create a finished product. However,
THE POWER OF AUTOMATED COMPLIANCE-READY CDS AND DATA MANAGEMENT

Effectively manage the life-cycle of your Empower™ data with the enhanced data management capabilities of NuGenesis™ LMS.

CAPTURE
data from instruments and scientists

CATALOG
according to protocols/projects

MANAGE
entire data lifecycle

FIND/VIEW
critical data in seconds

Improve and manage the availability of your chromatography results

►►Learn more at www.waters.com/poweroftwo
There has been an increasing interest in outsourcing your supply chain and that has been realised fairly effectively across a variety of industries, not just pharma, but other industries as well.

Over time and with costs and attrition rates rising, pharma companies began to move towards a more risk-conscious model. This relies on outsourcing different aspects of the drug development process, in order to reduce the risk and the level of investment required.

The cost of externalisation

'Over the last 20 years another trend that we are all aware of is attrition, or what I would call unanticipated attrition. In a lot of ways those specialised assets – to make a particular compound from manufacturing, formulation, distribution – can be completely unusable if they are so specialised that they are built for a specific purpose. There has been an increasing interest in outsourcing a supply chain, and that has been realised fairly effectively across a variety of industries, not just pharma but other industries as well,' said Anderson.

'But with that asset flexibility, being able to hire CMOs for certain unit operations in a manufacturing process, or a variety of CMOs to provide different functions, such as manufacturing or quality assessment, I have even recently heard of outsourcing compliance, regulatory compliance, in particular,' he added.

One thing that Anderson says could be very useful in overcoming the challenges in maintaining control of data in the face of an increasingly outsourced business model, is data exchange. 'Data access points can be limited, just based on the nature of the data that is being acquired, the effort to summarise that data, the ability to take those summaries and interpret the data to make effective decisions.'

David Wang, general manager of informatics at Perkin Elmer, provided some details on the regulations that are affecting drug development processes. 'On the clinical side there [have been a] whole set of regulations that I think started in the European Union.

'The regulations, known as ICH E6(R2), are a set of guidelines that aim to significantly increase the probability that pharmaceutical and biological manufacturers would be capable of delivering strong analytics, along with their clinical trials. The regulations look at the processes of recording and managing data through electronic means when carrying out clinical trials.

'The reason I bring that up is even though that was published in 2016, Europe enacted it in full force during 2017, and the US in 2018. One effect is that the major manufacturers we are working with want to get a much more sophisticated analytical understanding of their clinical data than previously,' said Wang.

He noted that externalisation and outsourcing meant that data was not always easily accessible by the company that is creating the new drug. In the past, some critical data may be held by CROs. 'As the number of expensive biologics or more narrow indication drugs hit the market, the regulatory agencies are rightfully advocating for patients and asking for greater transparency,' said Wang.

Whereas in the past it was good enough to say that a treatment helped to solve a problem in a percentage of a given population, now regulators want pharma and biotech companies to be able to provide more contextual information, such as the subset of people who were affected, and what characteristics meant that the treatment proved effective.

Wang noted that in the past the organisation 'may not immediately have that information at their fingertips. They might have to check with the CRO that performed that research, and the regulatory agency may have to wait a week or two. Today, that is not seen as a sufficiently rapid response for consumers. 'I think it is a very positive development that the regulatory authorities are now asking pharmaceutical and biotech companies to have a direct hand in understanding the details and context of clinical trials data for patients,' said Wang.

The next challenge is the fact that usually you will have a portfolio of clinical projects, central suppliers, manufactured products or approved products that correspond to potential options to investors in your virtualised supply chain.

In addition to data exchange, Anderson also noted that digital transformation would be a key step to better controlling and deriving value from data created in drug development. Abstracting data generated at a manufacturers site can make it very difficult to retain context and robust data management practices. Creating fully-digital pipelines to feed data back to the organisation helps to meet regulation and competition challenges, while also maintaining the low-risk outsourcing-based business model.

'Being able to take data as it is generated and stream that data to decision makers directly. It needs to be stored, managed and collected in ways that allow decision makers to look across their portfolio of projects, both clinical and manufacturing projects, as well as

Data-driven drug development
ABBOTT INFORMATICS IS YOUR TOTAL SOLUTION PARTNER

For more than 30 years, STARLIMS solutions have served customers around the world, helping improve the reliability of laboratory sampling processes, manage complex testing workflows and analytical methods, support compliance with global regulatory requirements and industry standards, and provide comprehensive reporting, monitoring and analysis capabilities.

Visit www.informatics.abbott and discover how Abbott Informatics’ STARLIMS Integrated Solution can support your organization maintain a competitive edge.
potential suppliers, in this virtualised network of manufacturers;’ said Anderson.

While the terms may change, the sentiment that usable data must be at the heart of modern scientific processes is a sentiment shared not only by Anderson but also Jabe Wilson, consulting director, text and data analytics for Elsevier.

Wilson noted that it is not just in drug development that these changes are being felt, as Elsevier shifts to meet the demands of modern science. ‘We have been involved in not just publishing books and journals, but in creating databases and indexing content for the better part of 30 years. It is a long heritage at Elsevier, in terms of curating information and making it usable for scientists. That naturally flows into science software,’ said Wilson.

‘A lot of work that gets done is taking the fundamental research in science and adding semantics and other contextual data. Elsevier has done a lot of work around the creation of indexing software, and creating the right dictionaries and ontologies to extract information,’ added Wilson.

Wilson also commented on drug development. ‘On the other side of the fence what has been going on in drug development is a lot of the low-hanging fruit has been taken, and the diseased areas that are being looked at actually have much more complex mechanisms. Many times we have to address not one protein or gene but multiple targets.

‘Many diseases or conditions that would have been familiar in the past are actually historical terms, because disease has macro level symptoms that we would describe as disease but actually they are made up of multiple different conditions, which have different mechanisms and often you need to understand individual biological makeup – the phenotype of individuals – in order to understand how to address their disease conditions,’ said Wilson.

‘What we are doing at Elsevier – and many people are struggling with this – is making sure that the semantic data and semantic indexing can throw its hands around all that complex data and put it in a place where people can access it. The other thing is, as well as semantic information, we are now looking at things like machine learning. Semantic data allows you to create features which you can then feed into neural networks to look for patterns in these huge amounts of data.’

The rise of AI in drug development

Wilson notes that while Elsevier may have traditionally been seen as a journal publisher, the company’s focus is shifting towards information and analytical services for science. ‘We are focusing on the business of doing the research. We are using those analytical tools to answer R&D questions.’

Wilson explained that AI and machine learning are powerful tools, but that does not mean that traditional methods should be taken for granted. ‘In drug development there are a lot of predictive analytics tools which have been used for the last couple of decades. So when people start talking about machine learning, there is a lot you can do with traditional statistical models or traditional data science.

‘However, what we are starting to see is machine learning models being used more routinely and becoming part of the arsenal of tools used by researchers. One of the challenges we do see, though, is what you might call socialising that within the work environment,’ Wilson continued.

‘Getting people comfortable with the tools and able to interpret, or have a sense of, what kind of confidence they should have in the output. Those aspects are perhaps something that we don’t talk about quite so much, but the social aspects are probably more important, in some senses, than the actual computing power,’ added Wilson.

As AI and machine learning tools become more ubiquitous they are finding their way into many facets of our daily lives – the same is becoming true of science and research. From search engines to smart speakers found in many people’s homes or even Netflix or Spotify accounts, we can now see the effect that AI is having on our daily lives.

It may take a little bit longer for these technologies to permeate into laboratory practices – particularly for highly regulated industries such as drug development, which can be more reluctant to embrace technological change.

Wilson noted we are beginning to see predictive analytics ‘where people want to get some support, in terms of how to fabricate a new chemical or understand its potential properties. Alternatively, they might want to identify what genes and proteins are associated with a specific disease condition,’ said Wilson.

‘Generally, people have much higher expectations, in terms of the specificity of the information that is going to be delivered to them. I think that is really one of the key elements to all of this, in terms of user needs, because, in a sense, AI and semantic technology should sit in the background, it should be a tool that supports the workflow teams and delivers them insights and data, but they don’t necessarily need to be able to programme neural networks themselves. They just want to get the right information at the right time,’ concluded Wilson.
Discover more with CDD Vault

CDD Vault is a complete platform for drug discovery informatics, hosted through an intuitive web interface.

Helps your project team manage, analyze, and present chemical structures and biological assay data.

Smart Software Saves Time®

**Electronic Lab Notebook**
Now available as an ELN
Capture, Search, Protect & Work Together
CDD Vault is the ultimate electronic lab notebook, integrated directly with the chemical and biological assay data repositories and our full analysis and visualization environment. Designed for project teams, you can archive and search all of your experiments with ease and collaborate securely.

**Inventory**
Keep track of compounds
Debit, Track & Locate
CDD Vault is an integrated inventory system for your compounds and reagents. Stay informed of the current status of your inventory whether in-house or shared externally.

**Activity and Registration**
Store & organize your data
Store, Search, Analyze & Report
CDD Vault is a modern web application for your chemical registration, assay data management and SAR analysis. It lets you organize drug discovery data and collaborate across project teams. It is simple to use and extremely secure.

**Visualization**
Plot data sets & mine them
Interact, Calculate, Compare & Publish
CDD Vault is a dynamic analysis tool for drug discovery data. It lets you plot and analyze large data sets to find interesting patterns, activity hotspots and outliers. With publication quality graphics, it is easy to illustrate and share results across project teams.

cddvault.com
What are the biggest factors driving change in today’s laboratories?

Jacqueline Barberena, global marketing and product director at Abbott Informatics:

‘The biggest factors driving change in today’s laboratories can be summarised by: expected year-on-year increases in performance and productivity; more stringent data integrity and compliance requirements; end-to-end integration of workflow processes; and the ability to track sample QA/QC requirements. Additionally, newer technologies, such as mobile capabilities, apps, and cloud deployments are becoming more widespread as lab users demand the ability to access secure data any time, anywhere.

‘More than ever, as data grows exponentially daily, organisations need solutions to manage these growing amounts of data, but that also go beyond data management. This can be accomplished through advanced analytics, which will enable them to make business critical decisions in a timely fashion, despite the enormous amounts of data and information flowing throughout the organisation. Laboratory informatics solutions are changing to support customers facing the above-mentioned challenges.’
Andrew Anderson, vice president for innovation and informatics strategy at ACD Labs:

‘The driving factors for change in laboratory workflows and software are the transition from document-driven knowledge management to a data-driven paradigm. The ability for any data source (from design to planning, to execution, to analysis, to decision/conclusions). Another key change is the introduction of ‘virtual collaboration’ the ability for scientists to effectively communicate/collaborate with colleagues in other parts of the world.

‘Historically, a human would ‘associate’ data to the appropriate data repository. In order for a data-driven decision making paradigm to be effective, data sources, repositories, and decision support interfaces must be integrated without human intervention. Furthermore, virtual collaboration also required digitisation of data in a way that it can be manipulated and reviewed, ideally, as though it were collected “on premises”. While we are still seeing these transitions in effect, informatic workloads are currently increasing to support such integrations and data sharing technologies – with the envisioned payoff being less human work over time.

‘We believe that any application software must afford data integration directly (or as much as possible). While there are many applications that integrate some data sources, many data types are relegated to abstracted representations. This is especially pertinent for analytical characterisation data, which is at the heart of critical decisions made in research and development on a daily basis. LIMS and ELN, in particular, often refer to analytical experiments, but the rich data that is acquired to support decisions is often abstracted to either pictorial representations or alphanumeric summaries (e.g., area percent = X, identity was confirmed by NMR, etc.). Future systems, including LIMS and ELNs, must afford more functional integration components to provide direct access to the analytical data that can be reviewed and re-examined easily – either conforming to relevant data standards when possible, or able to easily support custom integration.’

Daniela Jansen, director of product marketing at Dassault Systèmes:

‘Software and hardware vendors are moving towards a platform offering, but in most cases it is a closed, proprietary platform. Platforms can provide data continuity and traceability, and support decision making across the product lifecycle. A platform also allows to move away from point to point integrations that are cumbersome and expensive to maintain.

‘Platform-based systems allow for a substantially different architectural set-up: small dedicated agile applications can be plugged into the platform in a modular fashion, services like reporting or instrument interfacing can be used by the different applications. This approach avoids overlap and redundancy of capabilities, improving user experience and work efficiency. It also lowers the cost of ownership.

‘The increased adoption of cloud technology is based on both, the natural progression from legacy systems as well as the need to transform laboratory operations. Many legacy systems are about to run out of support, so organisations are looking for replacements. Due to the technological advancements, firms have the opportunity to not only move to a newer system but to also take a new approach to their laboratory informatics landscape.

‘They can move away from rigid monolithic systems that often come along with duplication in capabilities, to a platform-based approach with modular applications and common services allowing them to transform the way labs are working today. And they can leverage the cloud technology many of the new solutions are offering, allowing them to adapt the scale, cost and capabilities of their deployment to their current needs. In times of ongoing merger, acquisitions and divestments, this is a compelling value proposition.’

How will the development of new technology, such as AI, deep learning, or the IOT, impact the development of lab software?

Barberena, Abbott Informatics:

‘There is no doubt that technology is and will impact the way labs operate. There are a few technologies that are a game changer, as they can shape labs in the near future.

‘Machine-learning solutions – over the next few years every app, application and service will highly likely incorporate AI at some level. In the LIMS sector, this has already been applied to add business value to organisations, in the form of advanced analytics and advanced user experiences. Big data analytics can help organisations innovate and run their businesses in a much more optimised fashion. Being able to see and understand the lab’s data is crucial for organisations to make better, more informed decisions.

‘Speech recognition – one of the strategic technology trends, as identified by 2018 Gartner report, are the so-called conversational platforms that “will drive a paradigm shift in which the burden of translating intent shifts from user to computer”.

‘Augmented reality – this is a technology that can amplify human performance and human experiences. In laboratories, AR coupled with holographic technology can add an entirely new dimension, supporting to solve real problems and increase efficiency by displaying, for example, relevant holographic information on top of Lab physical layout.

‘A few other examples on usage could be: imagine you wear protective glasses when executing a test method; if those were smart glasses, they could display the method SOP as a hologram, to help follow step-by-step instructions. If you wear protective gloves, which make it awkward to use a laptop or touch-screen device, how about having the option to use gestures and voice commands to interact with the LIMS software, via AR holographic screens? AR is here to stay. Applications are virtually limitless.

‘Blockchain – this will certainly continue to impact the development of lab software. When managing data from concept to consumer, as many labs do, multiple parties need access to the same sources of data at the same time. Having access to review, confirm, and edit this data simultaneously, via Blockchain, will continue to enhance lab’s capabilities.

‘Cloud – there is strong trend towards externalisation of IT and R&D work. Naturally, cloud deployment supports this trend and, according to Gartner, by 2020, 80 per cent of the laboratory software solution will operate on the cloud. This is a great point and time to influence an organisation’s cloud strategy.

‘Mobile – in three to four years, operating in the lab without mobile devices will be considered ‘unheard of’; just like running a business without email is inconceivable in today’s world. More than ever, users spend less time in the office and expect to be able to work anytime anywhere, and have access to all their
work resources. A LIMS with mobile-capable features can provide users access to the lab from virtually anywhere.

‘UX – the expectation from software nowadays is very high. This is impacted due to Consumerisation. Organisations should stick with products that invest in UX as it will help reduce support and training.

‘LEAN – this is a growing trend that can bring laboratory optimisation and performance to a whole new level. It helps identify waste (those additional activities we do everyday and do not contribute to the final outcome).

‘In today’s high-capacity laboratories, the number of tasks ongoing at any one time and the amount of data being generated on a daily basis has swelled exponentially. Additionally, as businesses expand globally, now, more than ever, organisations need laboratory software that is capable of managing large amounts of data in real time, and that is able to integrate with numerous and disparate systems and instruments.’

Anderson, ACD Labs:
‘From our perspective, the promise of these technologies is well established – from both a productivity and innovation perspective. We believe that two sustained efforts must be undertaken to fully realise the benefits of such emerging technology:

‘Increased efforts in connecting all data sources to, broadly-speaking, prediction interfaces. Currently, AI/ML/DL can be limited based on the training data types available. Infrastructure development – connecting high-volume data sources to prediction interfaces will require significant investment.

‘We envision a future where data, application functionality, and prediction capability are available on demand, in whatever interface a scientist chooses. Moreover, if this future is realised, there are two benefits.

‘We believe that the amount of effort that scientists undertake to prepare summaries of their effort will be dramatically reduced. The time spent shuttling between applications will also be dramatically reduced.

‘The effort to design, plan, execute, analyse, and then summarise experiments is currently supported by a large number of monolithic applications.

‘In the case of synthesis laboratories, a scientist uses separate synthesis design, inventory, instrument control, and decision support software interfaces, all in one day. The future of laboratory software, as described above, reduces the number of interfaces substantially.’

Jansen, Dassault Systèmes:
‘In conversations with our customers, we have identified time-to-market being the ultimate driver for change. Personalised health and the desire for more precision therapies is changing the way that they are developed. Knowledge capitalisation is basic for leveraging of new and existing knowledge in the lab and next-gen manufacturing with moving from large batches towards continuous manufacturing has a deep impact on the analytical instruments and methods used, as well as on the related data analytics. And total quality efforts are attempting to make compliance and quality an asset, instead of a cost.

‘Laboratory informatics need to allow users to work not only in a more efficient and cost-effective way while remaining compliant, but they also need to provide the flexibility to adapt to completely new ways of working. They need to be able to deliver contextualised data in real-time for faster decision making.

‘Machine learning technology itself doesn’t need to be adapted. It just needs to be used in the right way. It is more about identifying the right data and providing the data in the right format to be leveraged. Data needs to be standardised and contextualised for meaningful outcomes. Dassault Systèmes does provide the tools today, and in order to help our customers to leverage their data, we are actively engaged in data standardisation projects of consortia like Allotrope or the Pistoia Alliance.’

How might the tools or workloads a typical user might deal with change over time?

Barberena, Abbott Informatics:
‘In today’s competitive landscape, organisations need to take a holistic approach to their lab processes and identify areas of waste.

‘The future of laboratory software will evolve into more fully integrated solutions, from concept to consumer, as well as customer portals in which the systems, lab techs and clients are more tightly integrated, with the end goal of entering data once and enabling it to flow through all of the business processes, through to either the end-user or client. ’Most organisations choose ‘the best in breed’ solution for a given function, for example: best stability study module, best environmental monitoring module, best inventory manager, equipment manager, spec manager, experiments notebook. What these organisations don’t realise is that they need to develop interfaces to these solutions, which is costly not only to build but also to maintain. It is also challenging to work with multiple vendors, because when an issue arises, vendor one will point to vendor two and vice versa,

In order for a data-driven decision-making paradigm to be effective, data sources, repositories, and decision support interfaces must be integrated without human intervention

Jansen, Dassault Systèmes:
‘Systems consolidation and convergence will provide laboratory users with a new experience and transformation in efficiency. Traditional systems will start being replaced by more agile, user-friendly cloud-based lab informatics applications that are based on a holistic open-platform approach.

‘Organisations will start leveraging the Internet of Things (IoT) in the laboratory. This will provide them with more data – delivering more data and more insight, as well as more reliable data of higher quality and integrity.

‘IoT will allow users to work more efficiently in the lab, as it will remove many time consuming non-value adding steps from the workflows, as the ‘things’ are not only limited to lab equipment but can also include wearable devices, google glasses, biometric bracelets, motion sensors, location beacons etc. At the same time, it will improve data integrity and quality, as data transfer from and to the devices is automated. And through the introduction of more sensors, labs will be able to generate more data faster, that can improve and accelerate decision making.’
CDD Vault helps lab users overcome data challenges in modern drug discovery workflows

We had spreadsheets all over the place, and data from different projects that were just separated in different folders. It got to the point where we didn’t know where to put our data, or where to later find it. This is one of the common issues that has led global biotechs of all size to Collaborative Drug Discovery, a software provider for research and development data management.

Drug discovery is data driven, and that data underpins every scientific and commercial decision, which could ultimately spell the difference between the success and failure of a research and developmental program for new drugs. Yet in today’s labs the handling and management of data doesn’t necessarily maximise its value or usability.

Scientists commonly store and manage their data in insecure, often difficult to find disparate documents and spreadsheets. While this method might be okay for a lone scientist working in a vacuum, it is not likely to represent a smart approach for collaborative scientists working in drug discovery or in other chemical or biological fields that rely on the ability to store, recall, process and share large amounts of data quickly.

CDD Vault acts as a central smart warehouse for all drug discovery data, explains Kellan Gregory, the informatics firm’s head of product excellence. ‘Our platform offers a comprehensive set of core utilities to allow lab members to access all of their results data, in its contextual format.’ This means the ability to handle any type of data. ‘As well as being able to capture numbers and text, we can also capture the native file in line with the data.’ A formula builder, tools for activity and physical chemistry property calculations, the ELN and dynamic visualisation tools then provide extra layers of intuitive analyses.

Julio Martin is director and head of the Kinetoplastid Discovery Performance Unit (DPU) at GlaxoSmithKline (GSK) R&D’s Tres Cantos Open Lab Foundation, a ground-breaking PPP initiative set up at GSK’s dedicated diseases of the developing world (DDW) research facility at Tres Cantos in Madrid. The Open Lab Foundation supports collaboration by giving external partners access to GSK compounds, infrastructure and drug discovery expertise, with a view to accelerating research in multiple areas from target discovery and validation, to compound screening lead identification, and optimisation. The organisation chose CDD Vault as its data management platform for all Tres Cantos members to access all of their results data, in line with the data. ‘A formula builder, tools for activity and physical chemistry property calculations, the ELN and dynamic visualisation tools then provide extra layers of intuitive analyses.

for any additional configurations required at any point. The ability to become productive in a short amount of time is really big plus point for our customers.’

And with UK government figures released in April 2018 indicating that more than four in 10 of all UK businesses – and 72 per cent of large businesses – suffered a cyber breach or attack in the last 12 months, CDD Vault is highly secure. Built into industry standard SSAE 16 Type II certified cloud storage, and with 2-factor authentication and IP tracking on the ground, CDD Vault is designed to minimise the chance of a hack either from the outside, or from inside the client’s organisation.

James Moe is president, CEO and co-founder of Oligomerix, a small biotech company interested in understanding the role of tau protein in neurodegenerative diseases, and the discovery and development of treatments for Alzheimer’s disease and related tauopathies. Moe comments: ‘We’re using CDD Vault as a way of storing our chemical structures, as a way of searching them, as a way of storing all of our assay data. We’re also using it for doing calculations, for analysing our data, for creating reports and communicating the data to others, and then also, very importantly, for working securely with collaborators. So it’s been instrumental for all of those purposes. Another primary concern is having a database where we have better security over our molecules.’


Using CDD Vault means we can put more of our internal resources into scientific research, rather than have increased costs associated with setting up, maintaining and upgrading complex platforms

Overcoming data challenges
Scientists are now beginning to use new technologies such as the internet of things (IoT), artificial intelligence (AI) and machine learning (ML) in their daily workflows. What was once a niche option to fulfil a particular challenge is now becoming more ubiquitous in many lab environments. In response, informatics software providers are designing software tools that can help integrate these technologies into the workflows, and software platforms used by scientists to assist their laboratory operations.

A trend that applies to almost all lab-based industries is the increasing amounts of data – either generated in the laboratory or available through previous work by the organisation and its partners, or through public data repositories. This increase in data helps to drive innovations, such as AI. Deep learning (DL), for example, requires huge datasets to properly train a neural network, so that it can predict information with the required accuracy and precision needed for scientific research.

IoT is a source of increased data, as the sensors or other IoT devices provide constant streams of data which need to be analysed and interpreted. If they are not used directly for scientific data, but maintenance or some other use, then the data still must at least be stored and managed, so it can provide contextual or maintenance information if not directly used to provide raw data for laboratory experiments.

Data is also increasing from other sources. Previous shifts in laboratory operations, designed to meet regulation or to provide more insight from available data sources, mean that many laboratories have begun to store more data. This data comes from a variety of sources such as; experimental data, including ‘failed’ experiments, organisation partners and collaborations; meta data and contextual data from laboratory instruments; AI or DL.
There’s so much great knowledge buried in your data – problem is, much of it remains isolated and undiscovered in Research, in Clinical, in Manufacturing, in scientists’ lab notebooks themselves. Our Informatics solutions connect all your disparate islands of information through a unique, comprehensive combination of instrument control, industry-leading informatics applications, and enhanced services – so you can not only manage your data, you can help drive the discovery process. This is how unconnected data becomes smart data. PerkinElmer Informatics solutions: Because when your data connects, you can go where your science takes you.

www.perkinelmer.com/informatics
we want to explore the range of things, as well as doing some practical activities in the short to medium term,' added Lynch.

Creating the lab of the future

Several firms are already embracing technologies that could shape the future of laboratories. DNA Nexus is a Pistoia Alliance member that focuses on trying to streamline and accelerate genomic science with its cloud-based informatics and data management platform. It aims to help companies leverage the engineering frameworks to train and understand models without needing to be engineers themselves,' added Daly.

Transcriptic, another Pistoia Alliance member has designed the Transcriptic Common Lab Environment (TCLE), a cloud-based platform which acts a laboratory operating system. This allows protocols to be shared and also connect instruments and robotic arms, which can be managed remotely.

Yvonne Linney, CEO of Transcriptic, explained that the company was created around the frustration of having to repeat experiments as part of routine lab work that was carried out by the founder during an internship in a professor's lab. 'He was doing bio-engineering and part of the internship involved taking readings from various analytical instruments at whatever time of day or night,' said Linney.

'He thought “there has got to be a better way of doing this,” and so he started to think about some type of remote access labs and instruments, so the initial idea was about connecting devices and robotic arms to a cloud-based infrastructure,' said Linney.

Certara is also developing technology for changing workloads in the laboratory. While the company is known for drug development consultancy, it also produces software which aims to accelerate the entire drug discovery process.

The company is now focusing on the development of modelling and simulation solutions which can help further enhance its users' capabilities to take new drugs to market. Thomas Kerbusch, PhD, president of Certara strategic consulting services, said: "The development cycle is being expedited and positively impacted by new technology. At Certara, we are focused on the use of new technologies, especially modelling and simulation (M&S). The use of M&S has been encouraged by global regulators, to the point it is now essential to modern drug development. ‘M&S, to some extent, is used in almost every drug approval today today. It is used across the entire development cycle – from assessing drug safety and efficacy, determining dose selection, and addressing the needs of special populations – to making go/no-go R&D and commercial decisions – and assessing alternative formulations and opportunities for successful
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

For Research Use Only. Not for use in diagnostic procedures. © 2018 Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise specified. AD80142-EN1218
new drug indications. While this increased use is well underway, we see the opportunity for M&S to grow by three- or four-fold during the next decade,’ added Kerbusch.

‘Software today needs to be more modular and interconnectable than ever. That need for a holistic ecosystem underpins our development programs, allowing users to incorporate multiple types of software in their work, while providing an infrastructure that enables them to best organise, share, communicate and leverage their work. For example, we recently purchased a product called Pirana, a flexible, extendible pharmacometrics workbench that provides modellers with structure, tools and a graphical user interface to facilitate the iterative processes used to create models and perform simulations. That product leverages tools that exist today and those that will emerge in the future,’ said Kerbusch.

‘The diversity of our users’ needs drives our development strategy, which ensures that we provide them with robust products. Developing a fully-supported, maintained and validated software product requires a sustained commitment, investment and supporting infrastructure, including resources dedicated to software training, education around coding infrastructure, including resources dedicated to software training, education around coding language and use-case proficiency, customer software and license support, ongoing maintenance, and validation for compliance with 21 CFR Part 11 requirements,’ said Kerbusch.

Rise of the machines

At Transcriptic, the vision for the future of the laboratory is based on automating menial processes to free scientists to focus on more important tasks. However, that is not to say all experiments should be carried out by robotic arms; the system uses cloud-based informatics software to pass protocols to humans and robotic systems. The added benefit of this is processes and protocols can be meticulously maintained, so an organisation can ensure scientific procedure is carried out correctly over multiple users, labs or even different sites across the world.

‘The vision that we have got is really the development of a programmatic laboratory. You could call this a robotic lab or a cloud lab but basically converting all those important instructions into a format, i.e. code, that can be easily transferred – remove the human interpretation and either the human or the machine is doing things in exactly the same way,’ said Linney. ‘You don't necessarily need a full robotic arm to use the Transcriptic system; in fact we have a combination of robotic interfaces and human interfaces.

Transcriptic uses its own software system called the 'TCLE' to control the Robotic Cloud Lab. ‘It is the overlying process that is involved in running that lab. We really think about it as an operating system, in the same way that you would have a similar system for your computer. All instructions and data are fed into TCLE, which is then fed into the cloud and passed down to either the instruments or the scientist.

‘We have developed these work cells which are a collection of devices and, like other automated systems, some might be connected to a robotic arm. All that is run through TCLE, which is the programmatic interface over the top,’ added Linney. ‘Everything is connected through that lab, whether it is connected to a robotic arm or through a human interface. All the instruments are connected through very small computer interfaces, then those instruments are tied into the operating system.’

While the idea of automation has been around for some time, Linney suggests that previously this was used for more specialist uses, as a system would only be suitable for a single instrument or experiment. ‘What is different, and I used to run the automation division at Agilent, what we were finding even there was that people didn’t have the resources and the funding anymore – this was large pharma, as well as smaller companies. These organisations did not want to go out and buy a set of equipment to complete a single process over and over again, because those processes were likely to change. This led to redundant instrumentation.

‘The way things are going is to think more about how we can produce a more generic automated lab system that can be used for a lot of different experiments, rather than a single type of protocol over and over again,’ added Linney.

‘The whole area of our idea of the LotF, where there is a combination of humans and instruments connected together, is an area that is of great interest to everybody. We are thinking about how to ensure that everybody is working from the same interface, from the sharing of information and the sharing of data, so it really is consistent with where it has been produced,’ concluded Linney.

AI in the lab

AI and related disciplines, such as machine learning and deep learning, are making an impact on how scientific research is conducted.

‘We foresee that AI-enabled data extraction and structuring will start to enhance this process [drug development]. But not until the onset of actionable deep-learning methodologies can we expect to reduce the human interpretation element of creating databases. Until that time, AI will facilitate repetitive tasks, like extraction, to speed up the creation and update of databases.’

Certara has already begun preparing scientific data that the business is curating and organising publicly-available trial data, so that it can be used in quantitative model-based assessments for drug development, market access/formulary, and patient care decisions. Certara is currently working on databases across more than 40 therapeutic areas.

‘Additionally, we have an AI-technology called ClinGenuity that is used for creating clinical study reports and building narratives, along with the redaction of patient–protected information in clinical documents in advance of publishing,’ added Kerbusch.

Not every laboratory needs to be fully automated, based in the cloud, with extensive AI workflows. But the maturation of these technologies for laboratory users is crucial to creating the labs of the future. By adopting technology carefully, based on the needs of lab users and their workflows, laboratory managers can streamline, and accelerate, the way that scientists carry out research.
The Edge

Software for DMPK, ADME, Pharmacology, In vitro, In vivo...

REQUEST
PLAN
TEST
ANALYSE
REPORT

BioRails
- Track and optimize R&D
- Plan and schedule experimentation
- Generate plates and sample files
- Drive robots and instruments
- Apply quality control standards
- Store and search results
- ELN & workflow support
- Manage and order inventory
- Progress candidates faster

Morphit
- Import and analyse raw data
- Visualise and interact with results
- Generate publication quality reports

www.edge-ka.com • info@edge-ka.com • +44 (0) 2380 411 098

• Single solution with LIMS, ELN and SDMS
• Powerful workflow management
• Flexible and fast deployment
• Future proof technology (100% web, mobile, cloud)

MORE THAN LIMS WITH LES

iLES iVention Lab Execution System

MAIL OR CALL US NOW FOR A FREE PILOT AND PRODUCT DEMONSTRATION

info@ivention.nl / +31 38 452 83 75

iVention Lab Automation

www.ivention.nl
How connecting data may lead to discoveries in medical research

Dr Alexander Jarasch and Professor Martin Hrabe de Angelis explain that novel research methods produce tremendous amounts of data that cannot be analysed with classic analysis tools – so scientists need to look for new approaches, such as graph technology.

With its ‘Grand Challenge’, the UK has set a target of using data, AI and ‘innovation’ to transform the prevention, early diagnosis and treatment of diseases like cancer, diabetes, heart disease and dementia to prevent a potential 25,000 deaths a year.

Similar ambitions exist in many other countries – this is also a major question for all researchers worldwide at the moment, including in Germany. If we are serious about dealing with the challenges they represent for patients and society and healthcare systems as a whole, we need to study these diseases in much more depth, in order to provide novel methods for prevention and treatment of diabetes.

I believe these new technologies will be crucial in gaining new insights into the workings and causes of these chronic conditions and diseases. The problem faced by everyone trying to do this is that the analysis methods we have been relying on may have reached their limits due to the vast amount of data produced by novel research methods (e.g. omics). The really promising avenue is to use big data levels of data, so as to combine and better connect data.

**Integrate and link together more and more data points**

That’s complicated by the fact that nowadays, research – especially in Life Sciences – is not limited to one technology or one discipline.

The German Centre For Diabetes Research, where we work, is a multi-centre organisation that combines all the different data that originates from different studies, reports, surveys and research projects from different locations in Germany. So we have masses of data from clinical trials and patient information, and our data covers various disciplines, from studies on molecular level...
to pathway analyses and animal models.

To answer the interesting and suggestive biomedical questions about diabetes, we have to connect this data and look for new insights, patterns and correlations. That's because we realise it is no longer enough to answer a biological or medical question from one direction, we need to integrate and link more and more data.

This is the next step, not just in biomedicine, but also in the healthcare sector, which is increasingly turning from general blockbuster drugs to individualised treatment or precision medicine. For this to progress, it is necessary to network significantly more and, above all, look at as many aspects of the problem as we can. This is why the DZD and other researchers think graph databases – the technology that powered the Paradise Papers investigation – could help in the prevention, discovery of new subtypes, early diagnosis and treatment of major illnesses.

It's important to know that we aren't just using Excel or standard business relational (table) databases any more – we add a whole new layer here with graph databases. The standard technology we use in each of our research locations in Germany is a relational database, as well as spreadsheets and documents files. But once we realised more and more of that data is connected, we started looking for a solution to bring our data in relation to each other, and create an overall context for our research.

Relational databases have their merit. However, we needed something to bring these data silos together and uncover connections – to be able to jump from one data point to another is crucial for us. That's why we turned to graph technology.

Diabetes is a metabolic disease, but it's not sufficient for researchers to only look through metabolic data. They also have to take into account data of other disciplines, such as genomics or proteomics. In the human body, everything is connected in metabolic pathways; a gene encodes a protein that is active in a metabolic pathway and metabolises a metabolite, which in turn is able to regulate another gene. In a way, our metabolism is a network of thousands of components that are connected with each other, which is a graph data model.

**Link diabetes research with Alzheimer's**

That's why it's so important to be able to uncover these connections and to create a new layer of analysis on top of this data, so we use technology from the graph database world called Neo4j. The great thing about Neo4j is that it has a visual interface we can use for queries and experimentation. We are using it to deepen our 'map' of diabetes – to uncover hidden relationships and pursue the resulting new questions.

For example, we do a lot of important research on animal models to study processes and then compare them to humans, so there is a lot of animal data from mice and pigs. This can generate a hypothesis we want to pursue – for example, 'In the pig model is the prediabetes type X due to causes A and B?' Is this regulated similarly? Are there similar processes?

We think we can link the molecular human data from the basic research with the highly standardised animal model data. In a graph representation, abnormalities, patterns or connections can then be recognised, which will then lead to further research questions. In the long term, it would also be interesting if data from diabetes research could also be used in other areas, such as cancer or Alzheimer's research, to uncover possible connections.

Graph isn't the only advanced technology we see as being useful. For example, we will definitely use machine learning techniques with graph software to identify unknown patterns, for example to try to identify (new) subtypes of diabetes we find discussed in the literature. Another example is Natural Language Processing. We'd like to build a system that automatically reads scientific texts from literature databases, analyses them and together with our research data generates hypotheses that can be evaluated by DZD scientists. Also conceivable: predictive models that can prescribe the course of the disease to a certain degree of probability.

This is all coming, and we are certain that our data management and analysis approach will take us to the next level in precision medicine, prevention and treatment of diabetes. In general, technology and data absolutely have a central role to meeting the Grand Challenge the UK wants to take on.

Dr Alexander Jarasch is head of data and knowledge management at Munich’s head-office of the German Centre for Diabetes Research, the DZD.

Professor Martin Hrabe de Angelis is speaker and member of the DZD board.
Research misconduct

Mark Newton, a consultant from Heartland QA, gives his take on the scale of research misconduct taking place in laboratories

Research misconduct – and other integrity issues – are poised to become a major issue for research publications, eventually forcing major changes in the way laboratory data is summarised, reviewed, and retained. Greater transparency and data lifecycle management will not be a best practice, but a mandate for researchers who want to continue to publish original research.

Allow me a personal moment. By background, I am a scientist whose career has been largely spent in quality control laboratories and lab informatics teams, so I have experienced the lab from both scientific and data management views. For the past several years, I have been deeply involved in data integrity (misconduct) identification and remediation for pharmaceutical labs and manufacturing, so I have a view of this topic from a GMP (good manufacturing practice) perspective. As I read articles about misconduct in research, I see parallels to the numerous data integrity citations given to pharmaceutical and clinical research firms by regulators around the world. These parallels form the basis of my opinion here.

The issue is far larger than believed

There are several reasons to believe this issue will be big: (1) the issue of misconduct was more prevalent in GMP labs and manufacturing than anyone would have believed; (2) there is already indirect evidence that misconduct is widespread in research labs; (3) systematic controls to prevent misconduct are less rigorous in research than in GMP-regulated QC labs; (4) research labs are less likely to be inspected than their GMP counterparts; and (5) research labs have similar motives to GMP-regulated QC labs and manufacturers.

Data integrity among GMP manufacturers

In the past three years, about 130 US FDA Warning Letters have been given to pharmaceutical manufacturers or clinical research organisations for data integrity-related infractions. These are serious infractions that cost companies millions (or even billions) of dollars, and typically two years or more to remediate. So, what caused the spike in these infractions? Regulators learned to do data forensic auditing. Once they understood how to look for data discrepancies, and the ways that data could be manipulated to create a desired outcome, they were able to increasingly find bad data practices that were missed in the past.

They stopped looking at procedures in a

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.

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

- 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
conference room and started looking at data in the company's systems. And they found data manipulation: re-running samples to get a desired (passing) test result, keeping 'official' and 'unofficial' batch records for manufactured pharmaceutical products, deleting unfavourable data or storing it in other places to hide it from inspectors – these are but a few examples.

The shift in focus to inspecting original data directly in the electronic system exposed an industry-wide issue that will require several more years of efforts to improve.

Indirect evidence in research

Gupta [1] lists two relevant statistics in discussing the matter: nearly 40 per cent of researchers were aware of misconduct and did not report it, and 17 per cent of surveyed authors of clinical drug trials reported that they personally knew of fabrication in research occurring over the previous 10 years.

Systematic controls

By the term 'systematic controls', we are describing processes or procedures that are routinely used to assure that reported data values are complete and accurate. GMP regulations require companies to use equipment that is calibrated periodically, use test reference standards, formally train personnel, retain all testing data (even if not used to report a result), have all testing data reviewed by another scientist prior to using the data to make any product decisions.

Computer systems must not allow people to share accounts, and must limit key activities (enhanced access, such as administrator) to a limited set of people who have no conflict of interest in the work they perform.

Manufacturers typically use standard reports to look for product issues or to potential data integrity issues. They are required to record and track any unplanned events, and determine the root cause of the event, so its recurrence is unlikely.

In contrast, research labs may calibrate instruments, provide basic science training for personnel and will use reference standards and controls for tests. But they have no requirements for individual user accounts. Sometimes accounts are shared to save license fees, and many labs allow everyone in the lab to be a system administrator and make system changes as needed. Unplanned events have no requirement for investigation.

While original data exists in computer systems, that is usually not the data reviewed by another scientist; rather, the summarised test data is reviewed before determining the test result as acceptable for use. Data is retained for future use, but it is the summary data, rather than the complete set of original data values collected at the time of testing. Articles submitted for publication are peer reviewed, but the host lab provides the summary data to a peer reviewer. And there is no requirement to save all data – even the results that were not summarised for publications.

The lack of original data review, and the lack of a requirement to retain all data – even data not included in a summary – collectively permit a scientist to test and retest until a desired result is obtained, then to ignore or delete all other data values and report the desired one.

Inspection

GMP manufacturers can be inspected by regulators at any time, with no notice (some notice is required for foreign inspections). Once in the facility, they can look at anything and interview anyone involved in manufacturing. Forensic data inspections can be conducted, and regulatory bodies (such as FDA, MHRA, EMA) have experts they bring for these inspections.

The FDA attempts to inspect firms every two years, although requests to market a new drug nearly always result in an inspection prior to an approval to market. It is not uncommon for large pharmaceuticals to be inspected a dozen or more times by different global regulators within a year.

In contrast, research labs might be inspected for safety by university personnel, by a local committee, or perhaps by a certifying authority (if certified at all), but they have no inspection authority looking at their operation in detail, unless a ‘for cause’ is initiated by a local committee. This lack of direct, independent, detailed oversight provides an environment where misconduct can continue for extended periods without discovery.

Similar motives

Research labs and QC labs that test pharmaceuticals have similar motives, which means that they will have similar reasons to manipulate results to be more favourable to them.

QC labs perform mostly routine tests, using written procedures and analytical equipment often configured to efficiently do one job. By contrast, research labs reconfigure their lab equipment for each experiment, and seldom use written procedures. So how could they be the same? Motive is the answer, and the problem.

People will manipulate data when they are rewarded (or, not punished) for doing it.

A lab scientist can be pushed to misconduct when a deadline looms and the 'right' result is needed, or someone's pay/position will suffer. If the 'wrong' result is reported, the material must be discarded (manufacturing), or the experimental thesis abandoned and/or revised (research). For both manufacturer or researcher, money and time are lost. The manufacturer must have new batches of medicines to sell for profit, while the researcher needs data to stay ahead – the 'publish or perish' challenge of research. Since both manufacturer and researcher share similar motivations to achieve desirable data on a schedule, data integrity issues in one (manufacturing) makes it likely that similar issues exist in the other on an equal basis (research).

Driving the engine of change

Given the high percentage of indirect evidence of research misconduct, the lack of data forensic inspections and independent oversight of research labs, the lack of requirements for strict security and access controls to data management systems, research labs appear to be more at risk than GMP manufacturers for misconduct, manipulation and hiding of data.

So what will cause the issue to be exposed, and how will the improvements in data integrity be pushed into research? Will it come from governments, NIH, publishers or the universities themselves? Will it be voluntary, tied to standards or external accreditations, or codified as law? Misconduct also raises questions about all those studies that are contradictory; coffee is bad for you, it is good for you, etc. Is the problem a lack of statistical power in the data, or was it due to data selection, trying to support an unsupportable hypothesis?  

Reference

The science of appliance
Discover Scientific Computing World Online

- Read news as it is published
- See industry press releases
- Refer to archived feature content
- View webcasts
- Study white papers
- Find relevant suppliers for your business
- Subscribe to the magazine, find our Twitter feed, or connect to us on LinkedIn

www.scientific-computing.com
Laboratory news round-up

Robert Roe presents some of the top stories throughout the year

Optibrium and Intellegens collaboration applies deep learning to drug discovery

In September Optibrium, a developer of software for drug discovery, announced it signed a collaboration agreement with Intellegens, a company developing novel artificial intelligence (AI) software.

The collaboration will provide Optibrium clients with access to advanced, proprietary deep learning methods that will extend and improve predictive models to guide more efficient design and selection of high-quality drug candidates.

Optibrium provides software solutions for small molecule design, optimisation and data analysis. By leveraging Intellegens’ Alchemite technology, the partnership will create a predictive modelling platform capable of delivering more accurate predictions and enabling better decision-making when it comes to the optimisation of compounds.

Dr Matthew Segall, Optibrium’s CEO, said: ‘We are delighted to be working with Intellegens. Their cutting-edge, deep learning technology, already proven in the field of materials design, has shown remarkable results when applied to challenging drug discovery data in our proof-of-concept studies. These methods will deliver a step-change for our users’ ability to make confident decisions regarding their research strategy.’

The Intellegens deep learning tool, Alchemite, can model multiple endpoints simultaneously, gaining more information from available data than traditional quantitative structure-activity relationship (SAR) models. Intellegens’ methods are uniquely capable of building accurate models based on the sparse and uncertain data typically available in drug discovery.

This latest collaboration with Intellegens supports Optibrium’s approach to working with the leading and most innovative providers of technology and services for drug discovery and other chemistry fields.

Dr Gareth Conduit, Chief Technical Officer and Intellegens co-founder, said: ‘We are excited to accelerate the adoption and the impact of our neural network engine, Alchemite, by taking it to market in the field of drug discovery with Optibrium. The world-class predictive capabilities of Alchemite, delivered lucidly by Optibrium, will be game-changing for the drug discovery market, helping to simplify elements of the compound selection and design process.’

The Intellegens deep learning tool, Alchemite, can model multiple endpoints simultaneously, gaining more information from available data than traditional quantitative structure-activity relationship models

IDBS continues expansion with key appointments

During November IDBS, a provider of enterprise scientific informatics platforms, 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.

The appointments strengthen IDBS’ ability to drive the creation of labs of the future, as well as their preclinical footprint in the scientific informatics market, furthering its position as a provider of data management solutions and services to preclinical and biopharmaceutical organisations.

Christian Marcazzo, IDBS general manager commented: ‘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 preclinical pharma and discovery informatics, holding senior roles at Accenture’s LabAnswer, Dotmatics, Vanderbilt University Medical Center and Harvard University.

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, working at Waters Corporation, Medidata Solutions and Oracle.

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 ranging from sample prep to anti-drug antibody.

April Pisek, solutions consultant at IDBS, said: ‘When it comes to supporting drug development, IDBS’ technology and its commitment to delivering a high-quality service that is 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 that delivers this innovation, ensuring IDBS continues its growth trajectory.’
In August a partnership was established between the University of Essex and Provide to improve the UK’s National Health Service (NHS) by implementing an AI-powered decision-making engine to determine the service a patient needs.

The Innovate UK-funded Knowledge Transfer Partnership (KTP) will look at how artificial intelligence (AI) can help the NHS cope with the increasing demands of the 21st century.

John Niland, Chief Executive of Provide, said: ‘Our analysis shows that accurate classification and grouping of people enables more effective interventions, this can be achieved via non-face-to-face appointments, such as Skype, and online self-assessment and other self-assessment approaches. We’re excited to have this opportunity to use AI to determine the right approach for patients, deliver high levels of service and provide value for money.’

The initial trial will be based on Provide’s musculoskeletal services – for people with conditions affecting joints, muscles and other soft tissues – and, if successful, will be deployed further via a web-based platform across community and primary healthcare services, improving their efficiency and effectiveness, and consistency in decision making – a huge benefit of machine learning and AI.

Professor Maria Fasli, Director of the university’s Institute of Analytics and Data Science and UNESCO Chair in Analytics and Data Science, is the lead academic on this project and will drive the development of the AI-powered decision-making engine for Provide.

Professor Fasli comments: ‘I’m very excited to have this unique opportunity to use AI to improve the way health services are delivered in the area where I live and work.’ The wider impact on the NHS of this project could be huge, as the technological solutions we’ll be developing in partnership with Provide could eventually be rolled out to many other services and areas.

‘I am very pleased that a central part of this project will focus on helping staff to understand the technology and truly embed it into their everyday working lives. Getting everyone behind our innovative approaches will ensure we get the most from this project and really improve the patient experience’ Fasli added.

Clinicians currently assess people (known as clinical triage) over the phone for many services, such as adult therapy, stop-smoking support and general practice in the NHS, with administrative staff manually booking appointments. The new partnership is set to improve the efficiency of this process by creating a decision-making engine, powered by AI, which will identify the type of service people need, then signpost them to the bespoke support they require.

The project will also enable psychologists at the University of Essex to gain a deeper understanding of potential barriers to the use of new technology, so clinicians will trust and embrace it, where safe and appropriate to do so, as an effective tool to improve how services are delivered to people.

By the end of the three-year project, clinicians will be able to target their interventions more accurately, as patients will be prioritised more effectively. Those who do not need on-going clinical support can be discharged and patients can self-manage remotely – promoting independence and tackling a significant challenge in the NHS.

Professor Riccardo Russo, from the Department of Psychology at the University of Essex, said: ‘I am excited about this project, where scientists from different fields will work side by side to improve the quality of health services.’
The Pistoia Alliance has released results of a survey that found that 72 per cent of science professionals believe their sector is lagging behind other industries in AI development.

To accelerate the successful use of AI, The Pistoia Alliance has launched its Centre of Excellence for AI in Life Sciences, aiming to encourage greater collaboration between stakeholders to bridge the gap between technology and science. The aim of the centre is to bring together best practice, adoption strategy, events and hackathons covering a range of challenges.

The survey, carried out in June, found that adoption of AI is high, with 69 per cent of companies using AI, machine learning, deep learning, and chatbots; an increase from when the same question was asked in September 2017, where 44 per cent of respondents were using or experimenting with AI.

“This survey shows interest in AI remains strong, but there is still a challenge with moving past the hype to a reality where AI is delivering insights with the power to truly augment researchers’ work,” commented Dr Steve Arlington, president of The Pistoia Alliance.

‘It is significant that a majority of people in our own industry believe we are trailing other sectors in the use of AI, and we must address this issue by working closely with each other and with stakeholders in other sectors.

‘AI is poised to have a radical impact on life sciences and healthcare, but the industry must give researchers the best chance of success.’

A further 19 per cent of respondents signalled that they plan to use AI in the next 12 months, with just 12 per cent of life science professionals not using AI at all.

While it is clear that life science companies are using or planning to use AI, there was an indication by survey respondents that the technology was not yet providing value.

Of those currently using AI in their organisations, approximately 21 per cent felt that their projects were not yet providing meaningful outcomes and 21 per cent ‘didn’t know’ if projects were delivering meaningful outcomes.

The Pistoia Alliance believes collaboration between interested parties will be essential in ensuring AI projects lead to results that positively impact R&D.

In May Agilent announced its plans to acquire Genohm, a developer of highly differentiated, on-premise and cloud-based software solutions for laboratory management. This acquisition will expand Agilent’s laboratory informatics software portfolio, adding LIMS and workflow management, while expanding ELN capability.

‘We were impressed with the team and the technology,’ said John Sadler, vice president and general manager of Agilent’s Software and Informatics Division. ‘The modern architecture of SLIMS is perfectly aligned with the values of Agilent’s OpenLab products. By integrating this technology with our instrument portfolio, we are in a position to support and enhance the operations of modern laboratories.’

Genohm’s main laboratory software automation suite, SLIMS, is a digital platform that provides laboratories with a rapidly deployable laboratory information management system (LIMS) and electronic lab notebook (ELN) environment that is used in biobanks, research labs and next-generation sequencing facilities.

The platform tracks data and samples, tests and users, results and workflows, from the original material shipment to the result from lab instruments and in-silico analysis pipelines. Genohm also has an application marketplace with preconfigured workflows to enable rapid system implementation across a broad range of industries and scientific workflows.

‘We are very excited to join the Agilent team and believe that together, we can accelerate development of the digital lab to help our customers advance science and discovery, while ensuring compliance and traceability,’ said Frederik Decouttere, founder and CEO of Genohm. ‘Our laboratory management platform is highly configurable, easily deployable and leverageable across many different workflows, which makes our technology a perfect fit for Agilent.’

Financial terms of the deal were not disclosed.

By integrating this technology with our instrument portfolio, we are in a position to support and enhance the operations of modern laboratories.
Empower Data-Driven Decisions with Digitized, Live Analytical Data

Standardize your analytical data, no matter the analytical technique or instrument vendor

Automate routine processing and analysis workflows

Simplify and improve data exchange and collaboration

ACD/Spectrus Platform

Assure Data Integrity
Enhance Regulatory Compliance
Drive Innovation

www.acdlabs.com/Spectrus

Paperless Lab Academy® 2019
Lake Maggiore, Italy | April 9-10
Grand Hotel Dino - Baveno

FREE REGISTRATION NOW
www.paperlesslabacademy.com

#eDATALIFECYCLE@WORK
TURN YOUR LABORATORY INTO A DATA-DRIVEN KNOWLEDGE CENTER

THOUGHT LEADERS KEYNOTES – END USERS CASE STUDIES
INTERACTIVE WORKSHOPS – LIVE DEMONSTRATIONS
STIMULATING NETWORKING

• eCONNECT: Data capture embracing Internet of Lab Things (IoLT)
• eMANAGE: Generate useful information from trustable data
• eDECIDE: Enable the power of Industry4.0 at the laboratory
• eARCHIVE: Secure valuable data archiving

PREMIUM SPONSORS

ORGANIZED BY
www.nl42.com
Directory of suppliers

A list leading of suppliers, consultants and integrators

Abbott Informatics – Your total LIMS solution provider
Abbott Informatics Europe Ltd.
Crossgate House, Cross Street
Sale, Manchester M33 7FT, UK
Tel: +44 (0) 161 711 0340
Fax: +44 (0) 161 711 0341
AISalesUK@abbott.com
www.informatics.abbott

From R&D to manufacturing to quality compliance, Abbott Informatics STARLIMS solutions support your entire product lifecycle. Invest in a LIMS solution with a flexible, modular design that will transform the way your lab works.

Advanced Chemistry Development, Inc. (ACD/Labs)
8 King Street East, Suite 107
Toronto, ON M5C 1B5
Canada
Tel: +1 (416) 368-3435
info@acdlabs.com  l  www.acdlabs.com

ACD/Labs is a leading provider of scientific software to R&D organizations. We help our customers assemble analytical, structural, and molecular information for effective decision-making, problem-solving, and product lifecycle control. Our enterprise technologies enable automation of molecular characterization, product development, and facilitate chemically intelligent knowledge management and collaboration—within organizations and between partners chemically intelligent knowledge management and collaboration—within organizations and between partners.

AGILAB
157 Boulevard MacDonald
75019 PARIS – France
Tel: +33 (0) 1 83 64 42 48
contact@agilab.com

AgiLab provides a lab digital platform for scientific data management and collaboration across R&D, combining Electronic Lab Notebook (ELN), Lab Execution System (LES) and Laboratory Information System (LIMS) functionality into a innovative and unified system. Built using the latest future-proof web technology powered by Oracle®, AgiLab leverages technologies to secure, enrich and offer the latest ergonomic design concepts and digital connectivity. AgiLab Science Platform sets a new digital lab benchmark and delivers out of the box functionality to enforce user experience, data integrity and regulatory compliance.

Autoscribe Informatics
1 Venus House, Calleva Park
Aldermaston, Reading
RG7 8DA, UK
Tel: +44 (0) 118 984 0610
info@autoscribeinformatics.com
www.autoscribeinformatics.com

Autoscribe Informatics offers industry leading configurable future-proof LIMS software. With 30 years’ experience our solutions are configured to match user requirements and can easily adapt as needs evolve. Matrix Gemini LIMS (LES, ELN & QMS) allows you to track and manage testing activities within a controlled environment with ease.

Collaborative Drug Discovery, Inc.
20 Station Road
Cambridge CB1 2JD
United Kingdom
Tel: +44 (0) 1223 803830
info@collaborativedrug.com  l  www.collaborativedrug.com

Collaborative Drug Discovery, Inc. (www.collaborativedrug.com) flagship product, CDD Vault, is a secure hosted platform for management and sharing of biological and chemical information. It enables researchers to intuitively organize chemical structures and biological study data, and collaborate with partners through a straightforward web-interface. CDD Vault functionality includes Activity, Registration, Visualization, Modelling, Inventory and ELN.

e-Science Solutions Ltd
33 Chapel Lane,
Whittington
Lichfield, WS14 9JT, U.K.
Tel: +44 (0)1543 547607
info@e-sciencesolutions.com
www.e-sciencesolutions.com

We provide independent expertise to help Life Sciences organizations transform their businesses through better use of technology. We support every stage of your journey to a digital lab, from business case development and vendor selection to deployment planning and oversight.
iVention Lab Automation

Since its establishment in 2009, iVention has shaken up the market for Laboratory Information Management Systems. iVention’s iLES platform offers laboratories a flexible software system designed around the laboratory process. It enables laboratories to streamline their workflow and manage and computerize their work, even in highly regulated environments. iLES is a powerful management tool that allows significant efficiency and quality improvements.

IDBS

IDBS is a leading global provider of enterprise-wide scientific knowledge management software to pharma, biotechs and others performing biological & chemical research for drug discovery and development, enabling them to securely capture, manage, share and gain insight from their structured and unstructured data, so they can spend more time on what they do best: science.

LabWare

LabWare provide a full-featured Enterprise Laboratory Platform incorporating LabWare LIMS and ELN. LabWare LIMS is supplied with an extensive range of application modules and instrument data acquisition delivering business advantage. LabWare focuses on customer success, providing global support and services backed by some of the worlds most experienced LIMS professionals.
MODA™ is a comprehensive platform providing solutions for QC and Production, with a focus on automation and integration of data for non-automated manual processes. It eliminates paper-based monitoring and testing that can be expensive, error-prone, time and labour-intensive, therefore reducing timelines enhancing data integrity and potentially saving clients resources and costs.

SoftN Labs is an IT services company specialized in Lab Informatics (LIMS, ELN, LES, etc.). We can also develop specific software solutions and mobile applications able to work independently or to be integrated to your information system. We work globally.

Thermo Fisher Scientific has been providing state-of-the-art Integrated Informatics solutions for more than 30 years. Our portfolio includes the largest global organizations across every industry, from pharma, food & beverage, environmental and oil and gas, representing the broadest range of complex customer requirements related to integrated, enterprise-level informatics solutions.

Waters Corporation creates business advantages for laboratory-dependent organizations by delivering practical and sustainable scientific innovation. Waters helps customers make profound discoveries, optimize laboratory operations, deliver product performance, and ensure regulatory compliance by providing a connected portfolio of separations and analytical science, laboratory informatics, mass spectrometry, as well as thermal analysis.

PerkinElmer delivers a comprehensive suite of scientific informatics and software solutions from instrument generated data, to enterprise solutions, providing scientists with the necessary tools to aggregate, search, mine, analyze and visualize critical data to help turn data into actionable insights in an automated, predictive and scalable way.

Scientists from ADME/DMPK, in vivo/in vitro Pharmacology, Inventory and Formulations rely on the BioRails platform to improve productivity, protect IP and manage data. Our unique combination of software, content and services delivers a highly effective collaborative environment for project teams, assay laboratories and contract research organizations to improve efficiency, transparency and data quality.
FROM FIRE TO MOLECULAR NANOBOOTS, TIME HAS HAD A POSITIVE IMPACT ON THE FUTURE OF SCIENTIFIC DISCOVERY. SO HAS OUR SOFTWARE.

Take a look at our timeline to see some of the most important moments in scientific history www.idbs.com/timeline

IDBS
info@idbs.com
www.idbs.com

UK (HQ)
Tel: +44 1483 595 000
2 Occam Court,
Surrey Research Park
Guildford, Surrey, GU2 7QB

USA
Tel: +1 781 272 3355
285 Summer Street
Fifth Floor, Boston,
MA 02210
LABWARE 7

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

Offices worldwide supporting customers in more than 100 countries

www.labware.com