Life Science Leader

HOW AI IS TRANSFORMING THE LIFE SCIENCE SECTOR - PART 1

As you consider your next, or first, venture into AI, how can you ensure that this technology effectively accelerates and streamlines your work?

TABLE OF CONTENTS

ARE YOU UP TO SPEED ON AI TERMINOLOGY?

ROB WRIGHT, CHIEF EDITOR
LIFE SCIENCE LEADER

When sitting down to interview Ameet Nathwani, M.D., EVP of Sanofi's medical function for an upcoming feature article (i.e., *Readying Sanofi For The Era Of Digital Health*) in our August 2018 issue, he was frequently using some artificial intelligence (AI) terms with which I was only vaguely familiar. So I asked the good doctor if he could "define this, and explain that" to help me better understand. For example, many people use the terms AI and digital assistants (e.g., Siri) interchangeably. Though digital assistants utilize AI, AI doesn't have to take the form of a digital assistant. And machine learning (ML) is not AI, though it is necessary for the development of AI, because machines, like humans, must be taught.

WHAT'S A DATA LAKE?

"A data lake is basically a platform that has data from every conceivable source and is stored in a way that can be segregated, aggregated, and analyzed in multiple dimensions," explains Nathwani. "You can do this even if the data has come from completely different sources or has different levels of specificity." According to Nathwani, such data could be structured or unstructured, and it could include doctors' notes and laboratory data.

"A data lake puts all this information in one place and enables a user to find ways to link the data together to make sense of it," he concludes.

WHAT IS DIGITAL HEALTH?

According to Nathwani, the term digital health means many things to many people. "I define digital health as a convergence of advanced biological [e.g., the human genome] and engineering science with computer-assisted computational analytics."

Rob Wright
Chief Editor
Life Science Leader

The amount of data is so big that to unravel and make sense of it you need advanced computational technologies. "But digital health also includes the attachment of physical science and materials, because you now can create new devices capable of delivering drugs tailored by algorithms. But the end goal of digital health must be to improve the precision and personalization of healthcare delivery," he contends.

ADDITIONAL RESOURCES

There are many more AI terms (e.g., genetic algorithm, heuristics, and recurrent neural network) being thrown about beyond those defined by Nathwani above. One resource I found instructive was put together by Richaldo Elias, a computer science engineering student and AI enthusiast. His 2017 article is about a 5-minute read and a good investment of your time. For as Albert Einstein said, "Any fool can know. The point is to understand."

Another resource you might enjoy is this 2016 TED Talk, *Can We Build AI Without Losing Control Over It?* by Sam Harris, an American author, philosopher, and neuroscientist. This talk is instructive, for it goes beyond just facilitating your understanding of what AI is, but what it portends. ▪

HOW AI CAN ACCELERATE LIFE SCIENCES

ROY NICHOLSON & NATE REGIMBAL
GRANT THORNTON LLP

There's no doubt that artificial intelligence holds great promise for the life sciences industry. A recent report on AI trends through 2023 by Research and Markets states "the tremendous demand for AI in life sciences applications, such as drug discovery and patient monitoring, is opening up new opportunities," but it added that organizations may be hesitant about solutions that reduce jobs or come with high initial costs.

Indeed, many life sciences organizations are wary about the escalating buzz around AI. So, as you consider your next, or first, venture into AI, how can you ensure that this technology effectively accelerates and streamlines your work?

START WITH THE PROCESS

It can be hard to clarify exactly where AI has succeeded and where it's likely to succeed for your organization. One common thread among companies that have successfully implemented AI is that they began by looking for optimization. Rather than asking, "How can we use this new technology?" they asked, "What processes within our organization provide the biggest opportunity for optimization?"

To help identify these opportunities, some organizations created teams within their finance divisions. These teams look holistically across the organization to find processes where optimization could provide business benefits.

Roy Nicholson
Principal, Advisory Services
Grant Thornton LLP

Even if an emerging technology solution is sold as not requiring process changes, a process evaluation and an organizational technology evaluation can help create a strong business case for a technology implementation. Some examples of how automation, machine learning, and other AI capabilities can be used in life sciences include:

Nate Regimbal
Sr. Manager
Grant Thornton LLP

- Automated lab testing in R&D that can drive cost reductions or the reallocation of resources for other development initiatives.
- Predictive solutions that identify, for instance, which chronic disease patients a physician is likely to see soon in order to help field reps present any relevant treatment alternatives.
- Call-center transcription analysis that helps identify trending key words and questions for training call center and medical staff.
- Sales forecasting, especially for chronic disease treatments, analyzing a combination of data from patients, prescriptions, and physicians.

When an organization has zeroed in on some opportunities for optimization, the next question is, “How?” With a host of AI vendors and consultant agencies, life sciences organizations don’t need to go it alone. But they do need to balance costs and develop key in-house AI expertise.

BUILD EXPERTISE WHERE IT MATTERS

Given the dynamic and complex nature of AI, most organizations can't afford to wait while their in-house expertise matures. But organizations miss out on important value and create a dependency if they use only external AI agencies.

With the AI models available now, organizations can design use cases and formulate solutions on their own. So, what types of expertise matter for in-house teams? How can organizations establish the right balance of external expertise with internal insight on proprietary information?

- **IN-HOUSE LEADERSHIP:** AI efforts should be led by internal staff who can capture and integrate the lessons learned and proprietary information that needs to be passed on to future projects.
- **EXTERNAL EXPERTISE:** External consultants, teams, and tools can help with implementing new solutions, ERP systems, and an enterprise platform or data warehouse across the organization.

It's important for internal resources to understand how the organization feeds data to its new solutions, how the master data is managed, and other structural aspects for a data platform. A central enterprise data platform helps provide a standard for the tools and solutions that ultimately deliver business value.

As an organization considers how much AI and machine- learning expertise it should develop internally, it can consider three general models:

- **PROPRIETARY-PLATFORM MODEL:** An organization can build its own team of technologists and data scientists who are entirely responsible to build solutions on the organization's unique data in order to achieve unique AI goals.
- **DATA-SCIENCE MODEL:** An organization can acquire an AI platform, hiring a supporting team that includes one or more technologists and data scientists who help the organization use the platform to achieve its goals.
- **MANAGED-SERVICE MODEL:** An organization can call upon a cloud ma-

chine-learning capability or other service-based AI capabilities, providing data for external data scientists to process and report upon. This is common in pilot projects.

Once an organization aligns itself with an AI approach, it needs to choose the tools that will best integrate with users and processes.

FIND A PLATFORM AND FLEXIBLE TOOLS

Behind any successful AI solution is a source (or sources) of data and a set of tools that deliver the key capabilities. To deliver those capabilities, you need the tools to perform three types of work:

- **MINING:** The data mining, or the plumbing, extracts data from various sources and puts it into an environment ready for analytic processing.
- **ANALYTICS:** The analysis applies the logic, calculations, hierarchies, and algorithms that are specific for analytics (as opposed to transaction processing).
- **PRESENTATION:** The presentation tells a story or presents a visualization that informs business decision making, transformative action, or even compliance reporting.

Most organizations already have access to a range of mining, analytics, and presentation solutions. For life sciences organizations, some tools that have proven particularly valuable include:

- **BIRST:** business intelligence software that can help mine volumes of financial, prescription, and clinical data
- **BLACKLINE:** a cloud-based accounting tool that can automate reconciliation processes
- **MICROSOFT EXCEL:** most AI platforms offer connection to Excel
- **NETSUITE:** cloud-based ERP solutions useful for processing financial data

- **POWER BI:** creates interactive dashboards to analyze data for decision making
- **QLIK:** helps analyze physician data and tracks in-market performance and more
- **SALESFORCE:** helps track sales opportunities and forecast business or inventory
- **TABLEAU:** financial and marketing analytics for identifying when to combine different data sets
- **WER.AI:** provides use case-specific AI solutions that are supported via an AI-as-a-Service subscription agreement

It's critical to remember that your processes for data generation are the foundation of your AI solution's quality. It's important to have tight controls and process discipline to help your solution provide accurate real-time data to the field.

ENSURE QUALITY WITH GOVERNANCE

An enterprise AI strategy can help teams ensure that they work together to build value in a central data standard that will sustain business value for present and future solutions. The enterprise strategy needs to clearly articulate the processes and procedures that govern data . And it falls squarely within the realm of an organization's internal expertise. Process optimization, data management, and governance may not grab many headlines, but they are the foundation of AI solutions that accelerate the search for tomorrow's critical treatments and cures. ▪

AI, DATA INTEGRITY, & THE LIFE SCIENCES: LET'S NOT WAIT UNTIL SOMEONE DIES

KIP WOLF, PRINCIPAL
TUNNELL CONSULTING

The idea for machines that can think became the topic of science fiction in the early parts of the 20th century and made for interesting reads. Science caught up and the term "artificial intelligence" (AI) was coined by John McCarthy at the Dartmouth Summer Research Project on Artificial Intelligence (DSRPAI) in 1956, where the first AI program, the *Logic Theorist*, was presented by Allen Newell, Cliff Shaw, and Herbert Simon.[1]

AI research flourished in the early years until it was slowed by limits in computational power, but it was reinvigorated in the 1980s by both computational tools and investment when John Hopfield and David Rumelhart popularized deep learning techniques that allowed computers to learn using experience.[1] The next limitation to AI advancement was in computer storage, which by the late 1990s was no longer a problem, as storage advancements produced cheap and ubiquitous solutions. In our modern world, we carry devices in our daily lives that dwarf the storage capability of supercomputers of only a few decades ago. AI has now gone mainstream, leaving the labs and coming into our living rooms with intelligent assistants (i.e., Alexa and Siri) and smart TVs. AI is on the news and on our tongues, as scarcely a week goes by without a television commercial or someone in our social

circles mentioning AI. But what is AI and how might it impact our lives when applied to life sciences?

GARBAGE IN, GARBAGE OUT

For the sake of this discussion, we can agree that AI may result in predictions, classifications, and decisions from computational analysis of large data sets that is based on machine learning (ML) from representative data sources and further informed by said data and related results. In this context, AI may, for example, present significant potential for improving efficiency of research and development activities such as discerning viable drug targets for further investigation. Or, AI may offer greater capacity for manufacturing by reducing the potential for defects and accelerating product review, release, and disposition for shipping through the supply chain. However, there remains great risk with this potential for great reward, as mistakes or losses caused by poor AI results could cause negative impacts on public health.

Kip Wolf, Principal
Tunnell Consulting

One AI challenge today is in developing and managing ML and AI to handle the great volume of disparate and non-standardized data that is available. AI is being rapidly adopted in our lives, made obvious in examples of consumer electronics. Karan Bedi, COO of Blaupunkt Televisions India, reports that “consumer goods companies are leaving no stone unturned to empower their products with digital and AI technologies” and “many household appliance manufacturers integrate the Internet of Things (IoT) and AI in the household products.”[2] A common example is the smart TV, the global unit share of which rose to over 70 percent of TVs sold in 2018, up from 55 percent in 2015.[3]

Consider how bad data could affect the AI experience. Data integrity and data quality play key roles in AI results. Poor quality input may produce unexpected or erroneous AI output. Take, for example, the use of a smart TV where Netflix data was

carelessly entered (e.g., randomly selected programs of interest) or a Hulu account login was entered by a houseguest. When an algorithm used for targeted advertising or suggested programming in either of these services is applied to the data set, the results might have no relevance whatsoever to the current viewer. While this may be irritating or unhelpful, it is not life-threatening.

However, careless data entry or incorrect data sets related to life sciences applications could have consequences that include mortality. “Machine learning algorithms are very dependent on accurate, clean, and well-labeled training data to learn from so that they can produce accurate results,” says Ron Schmelzer.[4] During ML, biased or erroneous inputs can cause inaccurate or anomalous outputs that have no relevance to the patient at hand. While mistakes are unlikely, the acceptability of error drops dramatically when considering any negative impact to patient health and public safety. Viewing Netflix programming advertisements that are of no interest is one thing – being dosed with the incorrect medicine is something else entirely.

BEGIN WITH THE END IN MIND

A market has emerged for data preparation solutions (including ClearStory Data, Datameer, Datawatch, Melissa Data, Oracle, Paxata, SAP, SAS, TIBCO Software, Trifacta, and Unifi Software) that perform data wrangling, data cleaning, and data preparation to enable ML and AI. In fact, “the vast majority of machine learning project time” is taken up by these activities.[4] However, at the rate that data is being created, it is likely that the ability to prepare data will be outstripped by the backlog of data to be prepared.

Data preparation continues to require some level of human interaction. At a minimum, a human must configure the specification for data transformation during ETL (extract, transform, and loading) when gathering data from multiple sources or migrating data to a central data store. Yet data wrangling represents potentially greater human involvement, as context may be necessary to perform advanced processing for transformation of robust data.

Whether data preparation solutions can keep up or not, the argument remains for

improving the integrity and quality of data as it is created, rather than attempting to clean it up later as it is being prepared for ML and AI. This is accomplished in large part through data management and information governance where data integrity and data quality are central tenets. It is incumbent upon those creating AI for life sciences solutions to apply the greatest attention to data integrity and data quality to mitigate the risk of negative impact on patient health and public safety. AI for life sciences solutions are held to a higher standard than in other industry sectors.

DATA INTEGRITY A CRITICAL SUCCESS FACTOR

Data integrity and data quality are critical success factors for AI solutions in life sciences. Standards for and verification of data integrity and data quality must be elevated for data sets where ML/AI is intended to be applied. Simply performing computer system validation (CSV) or managing computer systems under CGMP conditions is not enough to ensure data integrity and data quality.

Data integrity and data quality must be common themes in a mature quality management system and proactively integrated into data management and information governance as a core business activity. We find that when firms understand that data integrity and data quality are critical success factors, the result is a competitive advantage. Fewer human errors may occur, and investigations may be completed more rapidly and successfully when they do. M&A activities are made more effective and efficient as due diligence is more easily facilitated and valuation is more clear with defensible data and with human resources who understand and can explain it. These success factors lead to better AI results with improved ability to provide products to patients and increased value to owners and shareholders.

Now that AI has become more common, we are presented very directly with practical and ethical questions as we “allow AI to steadily improve and run amok in society.”[1] When will an AI result based on bad data end in a consequence of human injury or even death? At what point will malicious intent be involved to influence those AI outcomes to result in injury or death? Remember the “Tylenol murders of 1981” and how the resulting regulatory and industry actions changed forever

how we package medicines.[5] Will we behave reactively, waiting for a "Tylenol-level event" to force us to govern ourselves and the bad data we are pumping through AI? Or will we behave proactively to fervently manage our data to prevent negative impacts on patient health and human life? ▪

REFERENCES

1. Anyoha, R. (2017, August 28). The History of Artificial Intelligence. Retrieved June 8, 2019, from Science in the News website: http://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/
2. Bedi, K. (2019, February 21). How Artificial Intelligence is Reinventing Consumer Electronics Segment. Retrieved June 7, 2019, from Entrepreneur website: https://www.entrepreneur.com/article/328400
3. Global smart TV market share 2015-2018. (2018, July). Retrieved June 7, 2019, from Statistica website: https://www.statista.com/statistics/889000/worldwide-smart-tv-market-share/
4. Schmelzer, R. (2019, March 7). The Achilles' Heel Of AI [Forbes]. Retrieved June 7, 2019, from https://www.forbes.com/sites/cognitiveworld/2019/03/07/the-achilles-heel-of-ai/#d08829c7be7e
5. Markel, H. (2014, September 29). How the Tylenol murders of 1982 changed the way we consume medication. Retrieved June 8, 2019, from PBS NewsHour website: https://www.pbs.org/newshour/health/tylenol-murders-1982

BIG DATA AND AI IN PHARMACEUTICAL DEVELOPMENT & MANUFACTURING — AN INSIDE LOOK

JERRY MARTIN

NEW TECHNOLOGIES CAN LOWER COSTS, SPEED UP PRODUCTION, AND IMPROVE RESEARCH AND DEVELOPMENT

With the explosion of health-related data in recent years, pharmaceutical companies are looking to Big Data to reduce costs in research and development and manufacturing. The market for artificial intelligence (AI) in drug development, valued at $200 million in 2015, ballooned to $700 million in 2018 and is predicted to value more than $5 billion in 2024, according to a report by Big Data Analytics.[1]

One of the reasons for this rapid growth is improved capabilities in generating and harnessing data. Many pharmaceutical manufacturing systems already embed data in process development and optimization, connecting to the Internet of Things (IoT) on the factory floor. However, having integrated platforms for smart devices can enhance the accuracy and implementation of data throughout the life cycle — from research to manufacturing to the point of purchase and beyond (including patient monitoring).

AI and Big Data have the potential to lower the cost and time of drug trials, to better determine patient outcomes with established drugs, and to better design

new drugs. Computer software and algorithms can provide better analytics before and during the manufacturing processes and stimulate insights to fuel better decisions in the pharmaceutical industry.

Jerold Martin
Independent Consultant

MONITORING MANUFACTURING

Over the last few years, the FDA has encouraged companies toward continuous manufacturing and away from the step-by-step approach of batch processing. By continuously monitoring manufacturing processes and collecting feedback, pharmaceutical companies can monitor factors like temperature and moisture change, therefore reducing variability in the final product and patient outcomes, through a method called process analytical technology (PAT).

Continuously testing drugs through the production process using PAT is more efficient and cost-effective than batch testing, particularly when synthesizing biological proteins at millions of dollars per batch. By better monitoring pharmaceutical manufacturing process operations, we can potentially predict performance and reduce variability in clinical outcomes. Applying AI to performance and feedback response data generated with PAT can be used not only to control manufacturing processes, but also to improve them through automated machine learning.

In the case of biopharmaceutical production, companies can capitalize on Big Data by employing logistics in operations. Manufacturing biopharmaceuticals require the use of living genetically engineered cells. Production teams monitor for more than 200 variables to ensure consistency and purity. Two batches of a particular substance produced using an identical process can exhibit a variable yield of between 50 and 100 percent, creating issues with capacity, product quality and regulation adherence.

A McKinsey report described how a top-five specialty chemicals manufacturer used Big Data to significantly increase vaccine production yield. The company gathered

data from its production process in groups of related activities and entered it into a central database. The team then applied statistical analysis to determine interdependencies and sensitivities that had gone unnoticed. The manufacturer then made targeted process changes and was able to increase its vaccine yield by more than 50 percent — saving between $5 million and $10 million per year on production of a single substance.[2]

TRACK AND TRACE

The industry-wide serialization mandate also complements data integration efforts. FDA-required barcodes necessitate the pharmaceutical industry consider data and utilize software for coding product, creating a system wherein packaging operations rely on data for distribution. Serialization mandates that pharmaceutical companies modify their internal operations to be more data compliant, so an outside organization may also read and recognize their data. Data transparency could open pathways to new trends and best practices but presents many challenges due to the proprietary nature of this information.

Additionally, data gathered from serialization can have applications in drug development and patient adherence. Non-adherence poses risks to patient safety and financial burdens to the healthcare system. Lack of adherence is estimated to cause nearly 125,000 deaths and 10 percent of hospitalizations annually and costs the healthcare system between $100 billion and $289 billion per year.[3]

In response to this challenge, some companies level up their track and trace applications by serializing drugs down to the pill. In 2017, the FDA approved the first drug with a digital ingestion tracking system. Abilify MyCite, a medication for patients with schizophrenia, has an ingestible sensor embedded in the pill to record medication consumption. It works by sending a message from the pill's sensor to a wearable patch. The patch then transmits information to a mobile application so that patients can track the ingestion of the medication on their smartphone. Patients may allow their caregivers and physician to access the information through a web-based portal.[4] This practice could be a huge step towards solving the national medication adherence crisis.

RESEARCH AND DEVELOPMENT

Big Data and AI show promise in reducing the time and cost of pharmaceutical R&D as well. Billions of dollars are invested each year into the discovery and development of new medicines, but it usually takes 10 to 15 years for a medicine or vaccine to evolve from discovery to commercialization. Manufacturers invest a huge amount of capital long before seeing any return, which means they must allocate a high percentage of budget to R&D. Big Data Analytics reports that the fail rate for clinical trials is 92 percent and drug companies spend an average of $2.6 billion developing a single drug.[1]

When a pharmaceutical company invests in R&D, it screens 5,000 to 10,000 chemical or biological compounds to find one that exhibits potential for treating new or existing conditions, according to the white paper "Five Factors Spur Growth in Pharmaceutical Packaging and Processing Equipment," produced by PMMI, The Association for Packaging and Processing Technologies.[5] There is also risk associated with investing in R&D for a drug if it fails to meet regulatory approval.

Using Big Data to better design clinical trials and predict outcomes makes it commercially feasible to develop drugs for smaller patient populations. Pharmaceutical manufacturers collect data on patient responses to drugs during the clinical trial process. Combining that data with AI will enable them to determine which aspects of a patient subpopulation are associated with responses to drugs.

LOOKING AHEAD

Despite the many benefits of Big Data, incorporating it into manufacturing processes poses some challenges. Companies need to be prepared to handle the volume of data, the speed at which it accumulates, and organizing it best to create a more detailed future.

For manufacturers seeking to improve production with Big Data, the first step is to determine how much useful data the company already has at its disposal. Com-

panies that collect process data for tracking purposes, not as a basis for improving operations, may invest in systems or third-party analysts to organize and then use that existing information to improve operations. Other companies have too little data available to draw statistical conclusions. The challenge then becomes exploring practices to collect more data.

Researchers recognize that it would be valuable for companies to share data and learning with others who can apply it to improve the safety and reliability of all manufacturing. But the competitive nature of the pharmaceutical industry means that all information is propriety, and developing an open source system can be complicated, as companies use different technology platforms. A standardization process would have to be developed. Such a solution may be possible, however, as organizations like the Allotrope Foundation, a worldwide coalition of scientists with a mission to improve data access and integrity, are seeking to revolutionize the way we acquire, share, and gain insights from scientific data.[5]

CONCLUSION

While the industry is in the early stages of applying electronic data to manufacturing processes, harnessing new capabilities can lead to a better understanding of drug structure and function that ultimately impact patient outcomes. Big data, AI, and machine learning have the potential to change the way we develop, monitor, manufacture, and apply drugs. ▪

REFERENCES

1. http://analytics.dkv.global/data/pdf/AI-for-DD-Q2/Available-for-Purchase/Infographic-Summary-Advanced-R&D-and-Drug-Discovery-2018-Q2-Sample.pdf
2. https://www.mckinsey.com/business-functions/operations/our-insights/how-big-data-can-improve-manufacturing
3. https://www.nacds.org/news/the-cost-of-medication-non-adherence/ https://www.fda.gov/newsevents/newsroom/pressannouncements/ucm584933.htm
4. https://pmmi.org/report/2018-pharmaceutical-packaging-processing-white-paper
5. https://www.allotrope.org/

ABOUT US

Life Science Leader magazine is an essential business journal for life science executives who work for everything from emerging biotechs to Big Pharmas. Our content is designed to inform readers of industry best practices and motivate them to implement those practices in their businesses.

Each month, *Life Science Leader* features interviews with executives from the leading pharma and biotech companies, including Pfizer, Merck, Daiichi Sankyo, Genentech, Genzyme, Eli Lilly, and Janssen, as well as the FDA and more. This exclusive thought leadership provides an in-depth look into industry trends and hot topics that will impact the growth of the pharma and biopharma markets.

As the market is continually growing and innovating, these insights provide readers with actionable information and strategies to implement when navigating the evolving clinical, manufacturing, entrepreneurial, and regulatory landscapes.

Life Science Leader

LifeScienceLeader.com
info@LifeScienceLeader.com
814.897.9000
5340 Fryling Rd., Ste 300
Erie, PA 16510