ARTIFICIAL INTELLIGENCE and biopharma R&D IT

As noted in the Ernst & Young report *Beyond borders Biotechnology report 2017*[1], "…R&D productivity remains an ongoing concern. Artificial Intelligence and the accompanying analytics are now so advanced that these tools promise to improve the traditional drug target selection and R&D process."

By The PRISME Forum

The PRISME Forum is the biopharmaceutical industry R&D IT leadership group that meets twice a year. It addresses common industry challenges, shares use cases and catalyses more rapid creation, adoption and application of solutions to increase the efficiency and effectiveness of biopharmaceutical R&D.

As such, there is a contribution that the PRISME Forum should make to the development and implementation of AI to reduce the time and cost of bringing new medicines to market to treat unmet patient needs. With that in mind, the PRISME Forum focused its Fall 2017 Technical Meeting on the potential of Artificial Intelligence (AI) to improve biopharma R&D and healthcare.

Definition

What is AI? A google search reveals a practical definition, ie "the theory and development of computer systems able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making and translation between languages"[2]. However, IBM Research provides a more pragmatic definition of AI: "By AI we mean anything that makes machines act more intelligently"[3] and this is the definition that will be adopted in this article.

There are many technical definitions and taxonomies that can be used to stratify the various computer science tools that live under the umbrella term AI. Indeed, AI experts can get excited about the nuances between terms such as machine intelligence and human intelligence or deep learning methods and pattern matching.

For further background, Stanford University's initiative on the One Hundred Year Study on Artificial Intelligence or AI100[4] provides a perspective on the history and future of AI. This article identifies the start of the 'AI100 timeline' with Alan Turing's 1950 paper on Computing Machinery and Intelligence[5].

Technologists working in Life Science R&D should be interested in the practical applications within Biopharma R&D. IBM's pragmatic definition of AI is helpful and helps identify the opportunities in automation that are not necessarily in alignment with precise, computer-science definitions of AI.

AI at a tipping point

AI and Machine Learning have reached a tipping point. While many of the AI-based, computer science techniques have been available for some time (eg neural networks) the increase in data availability, vastly more powerful and easily accessed computer power (eg GPUs in the cloud), increased interest in data science and analytics, and improvements in algorithms have transformed the AI landscape.

A recent Harvard Business Review Article[6] states that "In the sphere of business, AI is poised to have a transformational impact, on the scale of earlier general-purpose technologies… Although AI is already in use in thousands of companies around the world, most big opportunities have not yet been tapped."

In addition, a recent Forbes article[7] cites "The necessity to start embracing AI technologies and revamping human resource strategies to create data science-driven interdisciplinary teams has become a matter of the future business sustainability for biopharma organisations."

Biopharma opportunities in AI

A crowd-sourced, multi-author paper entitled 'Opportunities and Obstacles for Deep Learning in Biology and Medicine'[8] was created by a broad list of contributors (many academic) from around the world. The paper highlights a number of opportunities and challenges in applying deep learning to biology and medicine. It examines applications of deep learning to a variety of biomedical problems in particular: (i) patient classification, (ii) fundamental biological processes and (iii) treatment of patients. The conclusion was that the future would see "deep learning powering changes at the bench and bedside with the potential to transform several areas of biology and medicine".

Discussion at the meeting revealed that many technology-based companies were highlighting their AI capabilities as the market quickly responded to the enthusiasm for the potential of AI, and in particular to the use of AI in life science R&D and healthcare. More than 30 relevant examples (see **Table 1**) were quickly identified, but it was widely recognised that any list of organisations active in the evolving R&D IT/healthcare AI landscape would be rapidly evolving.

Table 1 illustrates that there are many opportunities for the application of AI across the life science R&D/healthcare pipeline. However, the timescales in which they might create benefit varies. Importantly, there are today many near-term opportunities for AI that are not as yet adopted broadly, but that have clear benefits for biopharmaceutical R&D. Examples might include:

- Image analysis and phenotypic screening of drug candidates.
- Drug repositioning and competitive intelligence though data integration.
- Clinical trial cost and timeline improvement (eg protocol authoring, patient recruiting, site monitoring and risk assessment have already been implemented in commercial products and services from CROs).
- Cost savings in pharmacovigilance and regulatory reporting (eg through Robotic Process Automation [RPA]).

What are the implications of AI for the biopharmaceutical industry?

The 60 or so participants at the meeting were organised into five different discussion groups to consider the actions that biopharmaceutical companies would need to take to be in a position to derive advantage of this emerging AI technology. Each group considered one of the following five perspectives viz: Skills, Data, Organisation, Infrastructure and Metrics.

Skills

Leveraging AI in biopharma required the IT function, and the R&D IT groups, to have a strong foundation of the traditional skills and the willingness, flexibility and capability to acquire new skills. As illustrated in **Figure 1** these skills span science, mathematics and technology.

There are also foundational capabilities required that are cultural and cross-disciplinary and include a process to innovate rapidly and to assess value from 'placed bets' in the quickly-changing AI landscape. The MIT Sloan article entitled 'How Innovative is Your Company's Culture'[9] provides practical guidance for assessing a corporate culture and highlights that an innovative culture rests on a foundation of six building blocks, viz: resources, processes, values, behaviour, climate and success.

The adoption of AI follows a similar trajectory as with other technology innovations (reference the Gartner Hype Cycle[10]). There are many examples and articles on innovation management processes which generally focus on:

- Knowing what problem it is one needs to solve.
- Establishing success criteria.
- Experimenting rapidly – either to fail quickly or to demonstrate value.
- Scaling up successful experiments.

The Global Innovation Management Institute describes one 'Rapid Iterative Experimentation Process' (RIEP – pronounced 'reap') in the article 'Rapid, Iterative Experimentation Process – a Lean Startup-style Approach to Innovation'[11]. Sanjoy Ray at Merck has been an influential voice on this topic in pharma, describing a hypothesis-driven

Table 1

Arterys	Medical Imaging Cloud AI – The AI assistant for radiologists
Asimov	Programmes living cells with genetic circuits
Atomwise	Artificial Intelligence for drug discovery
BenevolentAI	Artificial Intelligence for scientific innovation
Berg	Therapeutic discovery using its unique AI-based Interrogative Biology® platform
Biovia	Collaborative, knowledge-driven innovation and predictive analytics
BioXel	AI and Big Data technologies to power the next wave of medicines
Cyclica Inc	Harnesses biophysics, bioinformatics and artificial intelligence (AI) to help pharmaceutical companies navigate the drug discovery pipeline by assessing the safety and efficacy of drugs
Deep 6 AI	Artificial Intelligence and natural language processing to medical records to find more patients for clinical trials
Deep Genomics	Geneticists, molecular biologists and chemists are supported by its biologically-accurate artificial intelligence technology
Digital Reasoning Syntheses	Reads and combines data from all sources, including human language to build a comprehensive picture of individual patients, revealing insights that aid clinicians' care decisions
Exscientia	AI-driven systems to automate drug design
Genpact Cora	An artificial-intelligence (AI)-based platform for digital transformation
Google DeepMind	Delivers tools that clinicians can use to make sense of the huge inflows of information which overwhelms them
Healx	To transform the lives of rare disease patients by intelligently matching drug treatments
IBM Watson	Machine learning capabilities to find unique connections between recorded symptoms and other clinical data, such as timing and dosing of medicine
Infosys Mana	A platform that brings machine learning together with the deep knowledge of an organisation, to drive automation and innovation
Insilico Medicine	Artificial Intelligence for drug discovery and ageing research
InveniAI	To extend the human experience in the use of artificial intelligence over the discovery process
Knowledgent	Enabling advanced and agile analytics, the digital enterprise and robotics
Microsoft Hanover	AI for precision medicine
Nference	AI to tackle the challenge of synthesising the world's biomedical knowledge
nQ Medical	Supports diagnosis of neurodegenerative disease years earlier than current gold standard tools
Palantir Technologies	Analysing real-world data for differentiated biomedical insight requires dealing with tremendous scale and complexity, as well as potential privacy concerns
Phenomic AI	State-of-the-art deep learning-based algorithms for analysing cell and tissue phenotypes in microscopy data
Precision Digital Health	Accelerates the adoption of digital health for researchers by providing a real-world evidence solution
QuantumBlack	In its raw form, data can be stubbornly unyielding. We use it to seek out the incisive insights and clear-cut response
Recursion Pharmaceuticals	Discovering transformative new treatments by leveraging the speed of automation with the intelligence of computers
Salesforce Einstein	A layer within the Salesforce platform that infuses Artificial Intelligence features and capabilities across all Salesforce Clouds. Einstein takes care of the data prep and modelling
Sema4	To revolutionise clinical diagnostics by combining comprehensive screening and diagnostic testing, predictive modelling, cutting-edge technologies and open-access data
Syntel SyntBots	Transform your business with intelligent automation
Transformative AI	Using cutting-edge artificial intelligence and novel analysis techniques to transform the treatment of serious medical conditions by collecting and translating clinical data into real-time, predictive assessments that guide the actions of patients and healthcare providers
twoXAR	Improving health on a global scale by building a profit-generating business utilising Artificial Intelligence
Viyasa	Enterprise scale deep learning platform, with consideration for Big Data scale data handling and provision to a range of deep learning approaches
WCG	Delivers transformational solutions that stimulate growth, foster compliance and maximise efficiency for those who perform clinical trials
Wellth	Application of behavioural economics through scalable technology to achieve better adherence, engagement and health
Wipro HOLMES	AI and automation
Zebra Medical Vision	Provides radiologists with the tools they need to make the next leap in patient care
...and many more!	

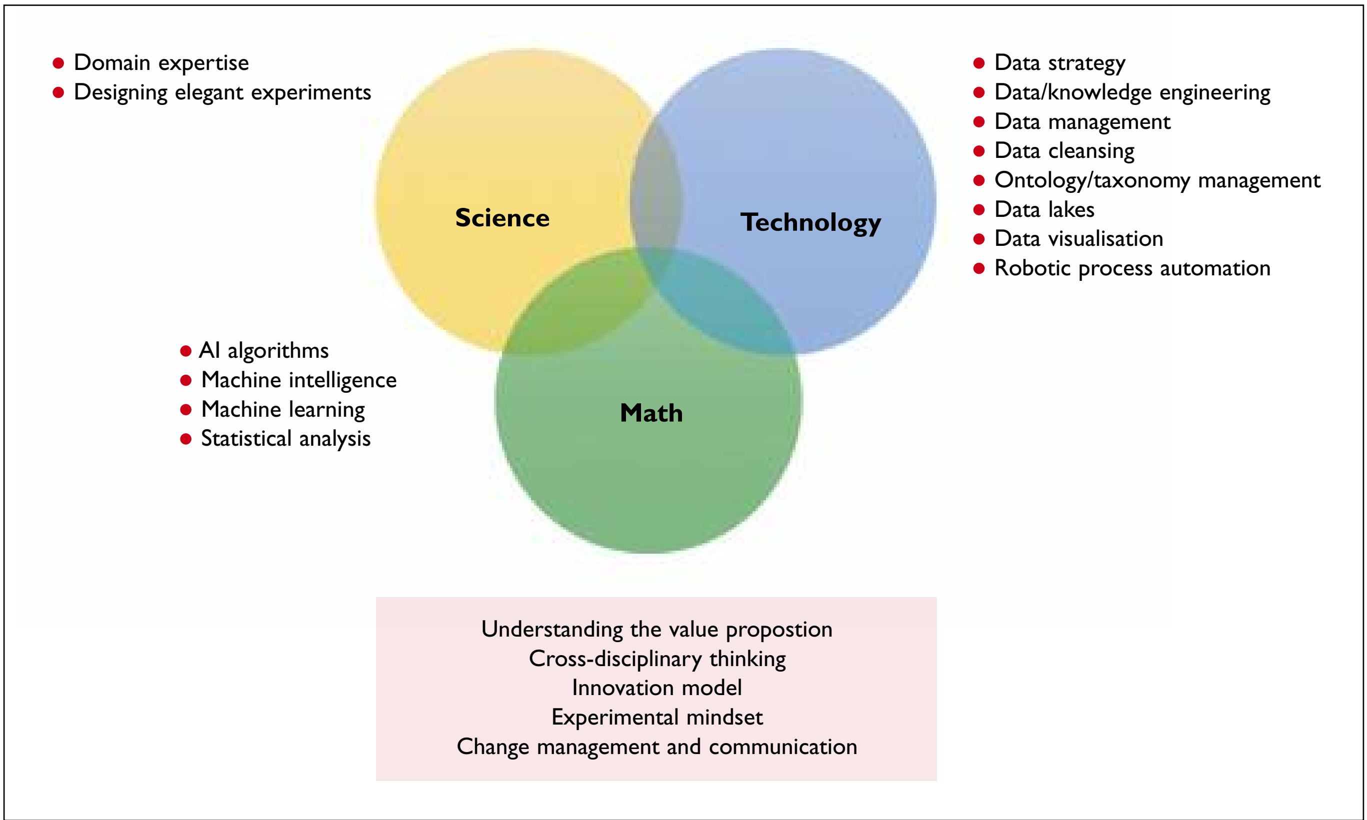

Figure 1
Skillsets and foundational capabilities required for AI adoption

approach and "rapid short experiments where 'good failures' are celebrated"[12].

Data

The recent boom in Big Data, and the development of the management principles to exploit Big Data, is one of the driving forces for the re-emergence of AI. Without Big Data (preferably of high quality and lots of it), AI would be starved of the raw material upon which it depends. Four key areas of focus for biopharma successfully to adopt AI capabilities were highlighted, viz: (i) Data Strategy, (ii) Data Governance, (iii) Knowledge Representation and (iv) Data Stewardship.

(i) Data Strategy focuses on establishing the overarching strategy for how a company wants to create, manage and use its data assets. A Data Strategy provides a roadmap for a company to advance its data capabilities, and includes addressing key questions such as:

- What is the data that will generate competitive advantage, that we should maintain for the enterprise over time?
- What data do we need that is external to our company?
- What data do we need to generate to be competitive?

(ii) Data Governance is comprised of the overall management of the FAIR data principles (findability, availability, integrity, reusability) along with the security and confidentiality of the data used in an enterprise. A good Data Governance approach defines:

- Who is the responsible owner for data.
- What quality standards should be applied to data.
- How data will be managed to those standards going forward.

Data privacy regulations (eg the European Union GDPR) are an important topic. Organisations need to ensure that they were knowledgeable about the regulations and were in a position to be in compliance when the GDPR came into force in May 2018.

(iii) Knowledge representation focuses on:

- How we store and represent data.
- How we structure data for usability.

(iv) Data Stewardship focuses on the skilled resources required to put a Data Strategy and Data Governance into practice. It addresses the challenge of how to resource and operate on a daily basis the curation and maintenance of an organisation's data and the continuous improvement of data quality and value.

Intellectual property and data pose new challenges in the context of AI. As an industry, the following questions arise: (i) When are we able to share pre-competitively? (ii) When are we willing to share pre-competitively? (iii) When does sharing put intellectual property at risk? The biopharmaceutical industry has, in general, considered knowledge of software and technology to be pre-competitive and sharable, while data and information about specific compounds, and the specific processes by which they are developed into drugs, is considered proprietary and not sharable. The question is: with an algorithm trained on proprietary data sets, would biopharma companies be willing to share the algorithm independently of the data? For many companies, the answer may well be 'no'. If the data cannot be shared, then the algorithm trained on that data cannot be shared. However, the untrained algorithm could potentially be considered sharable.

AI thrives on training data sets. The richer and more diverse the training data set, the richer and more diverse can be the interpolation from the AI machine. Cross-industry opportunities to share data can create richer training data sets and allow AI algorithms to function better. The IMI e-TOX Project[13] provides one example of such a pre-competitive, collaborative, data-sharing initiative.

Organisation

With the adoption of cloud-based technologies, the role of the IT organisation in the enterprise is evolving rapidly. The boundary between IT and other functions is changing as IT becomes a facilitator and more deeply embedded within business functions. This changing landscape comes even more sharply into focus as new skills and so-called 'double deep' resources (ie personnel knowledgeable in both business functions and the technology) are required to leverage data and the underlying technologies of AI.

There is no simple answer to the question of whether an enterprise should adopt a centralised or a decentralised approach, eg a single Chief Data Officer and Centers of Excellence versus a decentralised approach with data scientists embedded in the functions. Success depends instead on many factors ranging from company size and culture, to process, technology and data maturity.

Companies vary in the way they support technology innovation such as AI. Regardless of model, centralised or decentralised, successful AI technology implementation requires a focus on cross-functional partnership and teams. While the lines of responsibilities may be blurred, there are primary responsibilities that are clear and a hybrid model with platforms and centralised skillsets serving embedded domain experts is illustrated in **Figure 2**.

Platforms and data standards should be maintained at an enterprise level by IT organisations

Figure 2
Centralised versus decentralised skills in a hybrid model

Responsibility	IT centralised	R&D embedded
Technology Platforms	**✓**	✓
Data Stewardship	✓	**✓**
Data Standards Information Architecture	**✓**	✓
Data Analysis Data Reporting	✓	**✓**

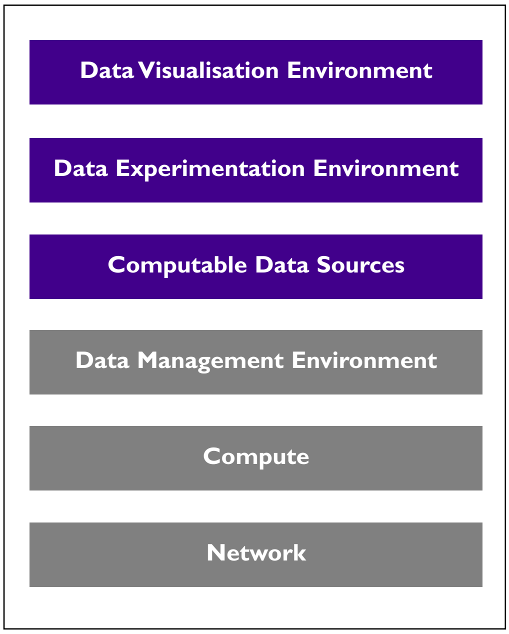

Figure 3
IT infrastructure – both foundational and emerging capabilities are required to be successful in AI

and leverage IT skillsets such as information architecture, systems architecture and knowledge representation. Data reporting and analysis should be the focus of embedded experts in the sub-disciplines applying AI. It can be argued that both centralised and decentralised approaches might increase the pace of adoption. For example, centralisation encourages the recruitment and retention of scarce skillsets, creates standards, shares use cases across functions, increases organisational commitment and the faster reuse of technology. However, decentralisation increases discipline agility and expertise, and discipline-level innovation keeps focus on value and with a lower governance hurdle, enabling faster deployment of resources.

Many capabilities founded on AI evolve into something else as they become successful. Examples might include help desk automation, image analysis, imaging biomarkers, supply chain analytics, genomic data analysis, computational chemistry, field force and promotion analytics, clinical protocol authoring. By the time the system is working well, responsibility for using the technology is more amenable to decentralisation.

An IT organisation should strive to create the environment for domain experts to self-serve their reports and analytics and to be a partner for advice and implementation of new platforms and technologies or to manage new types of data. This ensures less siloing of data and redundancy of platforms and builds enterprise level data assets.

Infrastructure
AI infrastructure requirements fall into two categories: (i) The foundational capabilities that most companies have today which might need enhanced capability or greater agility or more mature processes to be effective for AI management (shown in grey in **Figure 3**), and (ii) the emerging capabilities that are not yet as mature in most companies today and will require new investment to create (shown in blue in **Figure 3**).

Much of the infrastructure cannot be achieved if the enterprise does not first have a clear data strategy and data inventory as part of the data management environment.

Metrics
What metrics will be most effective in measuring AI maturity and the success of AI implementation efforts? There are at least three broad categories for measuring the success of AI efforts:
Experiments: A number of AI pilots, and within these pilots there should be cycle time reduction, cost savings or quality metrics for rapid assessment of value. There should be the rapid conclusion of pilots (both successes and failures) and then for production implementation of the successful pilots.
Automation: Where we should see a high percentage of the process becoming automated and human responsibilities becoming concentrated on the higher-level, decision making, organisation of personnel and scientific-interpretation tasks.
Decision quality: Metrics that show that AI-enabled decision-making increases decision quality.

Setting accountabilities and goals around these areas can provide the organisational incentives successfully to advance the AI agenda.

Summary
AI is a potentially transformative technology for the biopharmaceutical and healthcare industries and has many applications for which a rapidly-evolving set of technology vendors and services is emerging.

The PRISME Forum Technical Meeting in November 2017 brought together R&D IT experts from across 30 top biopharma companies to map a path to successful adoption and to prioritise the actions that biopharma R&D IT organisations should take to be competitive in the adoption of AI. There are implications for all facets of R&D IT, viz: Skills, Data, Organisation, Infrastructure and Metrics.

Of these, staffing with the right set of skills, redefining the role of IT and measuring success are

Additional materials
'The State of AI', by Andrew Ng, Deeplearning.ai; Stanford University[14]. Andrew Ng has authored or co-authored more than 100 research papers in machine learning, robotics, and related fields.
Related: Coursera online course 'Neural Networks and Deep Learning'[15] If you want to break into cutting-edge AI, this course will help you do so.
McAfee, A, Brynjolfsson, E (2017). Machine, Platform, Crowd: Harnessing Our Digital Future[16]. Andrew McAfee (@amcafee), a principal research scientist at MIT, studies how digital technologies are changing business, the economy and society. Erik Brynjolfsson is the director of the MIT Center for Digital Business and one of the most cited scholars in information systems and economics.
Davenport, T, Kirby, J (2016). Only Humans Need Apply: Winners and Losers in the Age of Smart Machines[17]. Thomas H. Davenport and Julia Kirby reframe the conversation about automation, arguing that the future of increased productivity and business success isn't either human or machine. It is both. The key is augmentation, utilising technology to help humans work better, smarter, and faster.

the most important factors in equipping a company to obtain competitive advantage with AI. New capabilities must be built on a solid R&D IT foundation of people, process, culture and technology. Strong partnership and cross-functional teams that bridge organisational constructs are essential to successful innovation. DDW

The PRISME Forum (http://www.prismeforum.org) is the de facto *R&D IT leadership group of the global biopharmaceutical industry. Currently it has more than 40 members drawn from 30 different biopharmaceutical companies representing nine of the top 10 companies, and 25 of the top 30 companies by R&D expenditure. The PRISME Forum's mission is "to enhance the efficiency, effectiveness and impact of global" R&D IT in biopharma. Each PRISME Forum hosts a Technical Meeting which brings together relevant opinion leaders and technical experts to increase the awareness of the PRISME Forum members of the opportunities for biopharmaceutical R&D IT to advance patient-centred, drug discovery and development by optimising R&D IT.*

*For further details contact **John Wise**, the PRISME Forum Programme Co-ordinator, at jcmwise@prismenforum.org*

References

1 Giovannetti, Glen T et al (2017). E&Y – www.ey.com/Publication/vwLUAssets/ey-beyond-borders-biotech-report-2017/$FILE/ey-beyond-borders-biotech-report-2017.pdf.
2 https://en.oxforddictionaries.com/definition/artificial_intelligence.
3 http://researcher.watson.ibm.com/researcher/view_group.php?id=842.
4 https://ai100.stanford.edu/about.
5 www.csee.umbc.edu/courses/471/papers/turing.pdf.
6 https://hbr.org/cover-story/2017/07/the-business-of-artificial-intelligence.
7 www.forbes.com/sites/forbestechcouncil/2017/08/03/artificial-intelligence-in-drug-discovery-a-bubble-or-a-revolutionary-transformation/#584ed4264494.
8 www.biorxiv.org/content/early/2017/05/28/142760.
9 https://sloanreview.mit.edu/article/how-innovative-is-your-companys-culture/.
10 www.gartner.com/technology/research/methodologies/hype-cycle.jsp.
11 www.giminstitute.org/lean-startup-style-innovation.
12 www.pharmafocusasia.com/information-technology/innovative-it-enabling-pharmaceutical-randd.
13 www.imi.europa.eu/projects-results/project-factsheets/etox.
14 https://events.technologyreview.com/video/watch/andrew-ng-stanford-state-of-ai.
15 www.coursera.org/learn/neural-networks-deep-learning.
16 www.amazon.com/Machine-Platform-Crowd-Harnessing-Digital/dp/0393254291.
17 www.amazon.com/Only-Humans-Need-Apply-Machines/dp/006243861l/.