

# Leveraging research informatics to accelerate drug discovery

Research informatics has a vital array of tools which can advance the speed and success rate of drug discovery. This article discusses that if pharma companies want to benefit from these tools and are serious about overhauling their research capabilities, then they must realise that this is a major transformation effort encompassing complex inter-dependencies between data and visualisation, as well as collaboration and workflow.

The life science industry is constantly seeking new ways of advancing its research techniques. Research innovation in this sector is driven by patent expiries, pricing pressures, evolving therapeutic needs and the advent of biologics for drug development. However, recently these pressures have been exacerbated by the global scale and spread of research, new research methods, overflow of scientific research data, and the need for collaborative research practices. In spite of this, having a wide range of technologies at their disposal such as Genomics, Proteomics, Marker Based Assessment and Microarray Technology, organisations are still finding it challenging to realise the full potential of their research. In this article we examine the various research challenges facing the industry and discuss the ways in which these can increase drug research through Research Informatics tools.

In recent years the transformation of drug discovery research has been propelled by the need to replenish a dwindling product pipeline. With \$60 billion worth of products going off patent by 2011, life science companies have to identify novel and innovative methods to compensate for falling research productivity<sup>1</sup>.

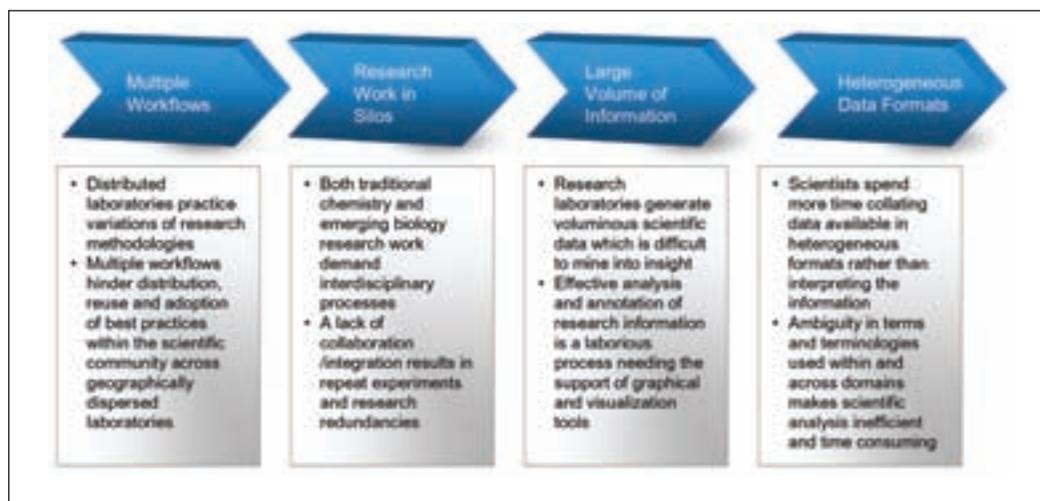
To complicate matters further, research labs are generating data faster than can be fully integrated. This is because companies are increasingly adopting omics-based scientific methods to gain information and knowledge. While the pharmaceutical industry has been adept at optimising the drug development process, it has rarely implemented different structures<sup>2</sup> to make the discovery process more efficient<sup>3</sup>. Therefore, scientists must use their creativity to constantly innovate, align informatics and data management needs to meet an integrated cross-disciplinary discovery process.

## Challenges in discovery research

The drug discovery process is complex and interdisciplinary in nature; however, there are portfolios of support tools and applications, collectively named Research Informatics, which can help immensely. The life science industry faces several key challenges which can hinder drug discovery; in particular, coping with data that is generated by registering biological or chemical entities and testing their biological, physical or chemical character or their pharmacological action. Information related to the registration and assay workflows, form the basis of all scientific innovation in any disease

**By Dr Anirban Ghosh  
and Siddharth  
Sawhney**

## Informatics

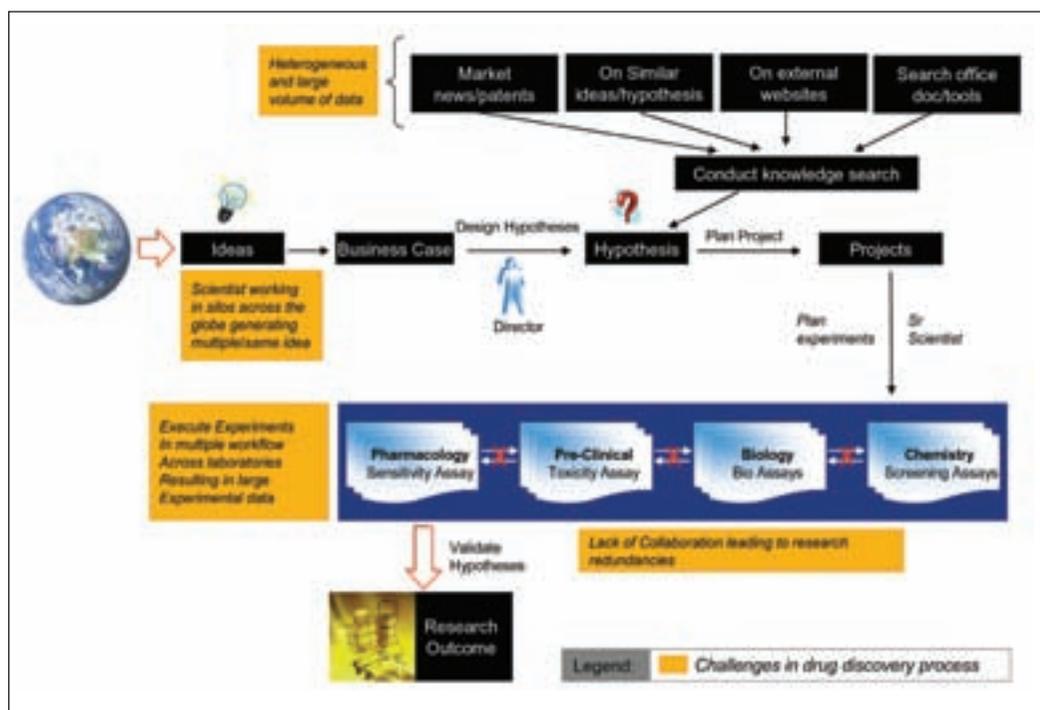


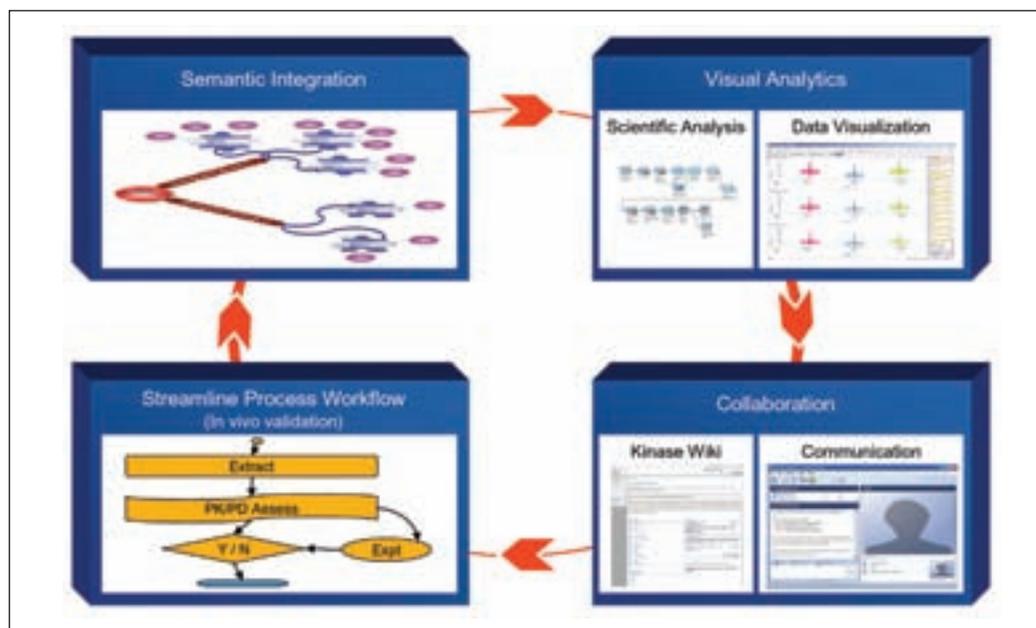
programme. In particular, researchers are faced with several challenges including process complexity, data indecipherability and questionable technology efficacy:

- Multiple workflows** – As the life science sector has expanded, multinational pharmaceutical companies have established global laboratories all engaged in related or similar research activities. However, there has been a deficiency in establishing suitable methods of sharing findings. In addition to this, the industry has seen a rise in mergers and acquisitions, which has led to redundancy and

lack of harmony between laboratory workflows. These multiple workflows often hinder the distribution, reuse and adoption of best practices within the global scientific community.

- Research work in silos** – Traditional chemistry and emerging biology research teams often work alone and fail to make the most of their results due to a lack of inter-disciplinary activity. An inability to collaborate between biology and chemistry processes leads to repeat and redundant work and inconsistent results. In order to make the most of their research, each team must work together by utilising each other's results. For





example, by integrating chemistry data within the context of biochemical processes or integrating genetic data with biological pathway, information can facilitate better understanding of disease.

- **Heterogeneous data formats** – The primary focuses in drug discovery research are diseases, pathways, proteins (along with their interactions) and genes. These are the foundation stones on which new molecule research is built. The biggest obstacle to the integration of research information is that data is usually only available in heterogeneous formats and stored in silos, hence cannot be shared easily. In addition, frequent duplication of information or ambiguity in terms adds to the difficulty of making timely informed decisions. Scientists often first spend time generating raw data sets and then some more time interpreting pieces of data to create knowledge assets. This lack of integration across research entities makes scientific analysis inefficient and time consuming.

### Ways to overcome research challenges

The use of Research Informatics tools such as data semantics, visual analytics, collaboration and workflow streamlining in drug research can increase research effectiveness, improve predictability and foster teamwork among scientists. When applied effectively, Research Informatics can help to address the following issues:

- **Streamline process workflows** – Pharmaceutical companies continuously seek to optimise their

processes and workflows in existing fields of research and adopt new ones to drive innovation in discovery. Procuring new products before studying their alignment with research processes only adds to license fees and maintenance costs, without adding value to knowledge capabilities. One way to optimise costs is to establish a tight linkage between processes and applications. By streamlining workflows, researchers can select processes that will improve competitiveness, prioritise business activities and enable IT solutions.

- **Collaborative research** – Research scientists from different disciplines need to actively understand and address various facets of the disease problem together. For example, results can be amplified when findings for cell line-based screening assays against a class of inhibitor compounds, are jointly interpreted by a biologist and a pharmacologist. Currently, there is only a moderate level of collaboration between research disciplines. Scientists are not known to freely share and exchange concepts or findings from their experiments or computations. Most often, interchange of ideas and information sharing happens via handwritten notes, whiteboard or electronic mail and lacks any formal structure. For research collaboration to be successful, it is imperative that researchers use all the tools available to yield meaningful benefits.

- **Visual analytics for large data sets** – Scientists evaluate a hypothesis by gathering large volumes of multi-dimensional data for inspection. While raw alphanumeric data can be cumbersome to

## Informatics

### References

- 1 Ruffalo, RR. Why has R&D productivity declined in the pharmaceutical industry? *Expert Opin Drug Discovery*, 1, 99-102, 2006.
- 2 Sams-Dodd, Frank. Optimizing the discovery organization for innovation. *Drug Discovery Today*, 10(15), August 2005.
- 3 R&D Efficiency, Tuft Center for the Study of Drug Development. Tufts Univ., Outlook 2009.

handle, a pictorial rendition can facilitate analysis. Even as scientists slice and dice through mountains of 2-D graphical data, they often need other types of graphical presentation for drill down analysis. Pharmaceutical companies have two options, either to develop bespoke applications for molecule and data visualisation or buy third party applications. Researchers can use intuitive Research Informatics tools to guide them towards standard data views for inspection, which offer a fresh perspective on how to tackle multi-dimensional data views.

● **Semantics for data interoperability** – Aggregation of information across the discovery value chain is pivotal to creating an integrated discovery engine and is considered one of the leading tools for pharmaceutical companies worldwide. However, still more important to successful drug discovery is a company's ability to carry out cross-functional search. Effective integration infrastructure enhances the ability to undertake cross-functional searches on biological and chemical information categories, greatly reducing time-to-market for new drugs.

Building data service methods around key domain entities is a good way to broker information across multiple points of access. These services can fetch data from diverse scientific silos in the context of the research investigation. Currently, life science companies are wading through biological and chemical semantics to create a web of standard ontology. This helps scientists link a compound to a product, relate clinical protocols with an indication, associate a protocol with an experiment, determine synonym company identification with a generic name of a compound, or connect pharmacovigilance signals to genes in a pathway. Presently, product companies are building ontology-driven search capabilities and creating a connected graph of terms and concepts. Using semantic technologies, researchers and programme directors can discover relationships that enable them to make better and faster decisions about disease targets and drug compounds.

### Conclusion

Drug discovery holds huge potential benefits for both the pharmaceutical industry and its customers. Research Informatics is a vital array of tools which can help advance the speed and success rate of drug research. However, companies looking to overhaul their research capabilities must recognise that this is a major transformation effort and that there are complex inter-dependen-

cies between data and visualisation, as well as collaboration and workflow.

In the face of major challenges and pressure to replenish the revenue pipeline, research organisations have an opportunity to re-emerge as the growth engines of industry. This will need significant commitment from the leadership team and a strong vision for the future supported by the ability to acquire and deliver value in a phased manner. Pharmaceutical organisations, patients, payers and governments alike will welcome faster and cost-effective discovery of innovative therapies for present day medical challenges. DDW

---

*Dr Anirban Ghosh is a Principal at Infosys Consulting. He has made significant contributions in scientific innovation and research informatics through international publications, patents and solution implementations at top pharmaceutical and applied life sciences companies. He can be contacted at [ghosh\\_anirban@infosys.com](mailto:ghosh_anirban@infosys.com).*

*Siddharth Sawhney is a Senior Project Manager with the Life Sciences practice at Infosys Technologies. He leads the effort for design and development of informatics solutions for discovery research. He has experience in software engineering, process consulting, and programme management. He can be reached at [siddharths@infosys.com](mailto:siddharths@infosys.com).*