Journal of Business Chemistry

Commentary

The Digital Revolution is coming to chemical laboratories

Michael Ulbrich* and Vikas Aggarwal*

* Accenture Strategy Chemicals & Natural Resources, Campus Kronberg 1, 61476 Kronberg im Taunus, Germany, michael.ulbrich@accenture.com

DOI: 10.17879/76199421140 ; URN: urn:nbn:de:hbz:6-76199423853

1 Introduction

When discussing the potential of digital technologies in research and development (R&D), people typically draw the line at so-called creative tasks. They see a clear division of labor, where creativity is the domain of humans and more mundane tasks are handled by machines. But that assumption comes into question when one considers both the current state of digital technologies and the nature of creativity itself.

Creativity is essentially the ability to develop or generate something novel. It can be something intangible like a magazine article, something physical like a new tool, or something artistic like a painting. But those things are usually the end result of a fairly long chain of events- the outcomes of a creative process. That is especially true when it comes to R&D. When a new material, for example, enters the development phase, that step has typically been preceded by years of research, trial-and-error testing and experience that inform development. In most cases, that creative development step is a logical continuation of historic findings, rather than a flash of insight from out of the blue. R&D creativity, like genius, is usually "1 percent inspiration and 99 percent perspiration", as the saying goes.

That hard "pre-work" is exactly where digital technologies are strong- and in many cases much stronger than humans. For example, machines (Microsoft Corporation, 2018):

- Are capable of accessing huge amounts of data within seconds, and are therefore fast when it comes to handling data.
- Have a near-unlimited storage capacity, and thus never "forget."
- Analyze data without prejudice and bias, and thus can draw objective conclusions based on probability and hard data.

With these capabilities, technology can be highly useful in many parts of the creative R&D process. And now, artificial intelligence (AI) is enhancing technology's capabilities even further, and making it suitable for tasks previously considered impossible for machines, such as predicting global market development and customer behaviors- or the behavior of new materials. As a consequence, digital and AI-enabled R&D is already state-of-the-art in industries such as life-sciences and healthcare, where it has proven to be especially useful in activities such as analyzing pharmaceutical activity profiles or identifying targets for personalized therapy within human DNA.

The chemical industry, however, has been slower to embrace this technology revolution in the lab. Often, researchers and developers in chemical laboratories are still engaged in manual, routine work, setting up experiments and monitoring or evaluating results with only rudimentary software support. They operate in disconnected information-technology islands, as well. Even if data is stored in an electronic laboratory note book (ELN) rather than on paper, a lack of technology integration between the business units and problems with validating data mean that many employees in the organization will struggle to quickly access that information. But evolving technology is now creating new opportunities for chemical companies to address those problems- and transform R&D to make it more efficient and effective in helping the company compete.

2 A new era of digitalization and automation

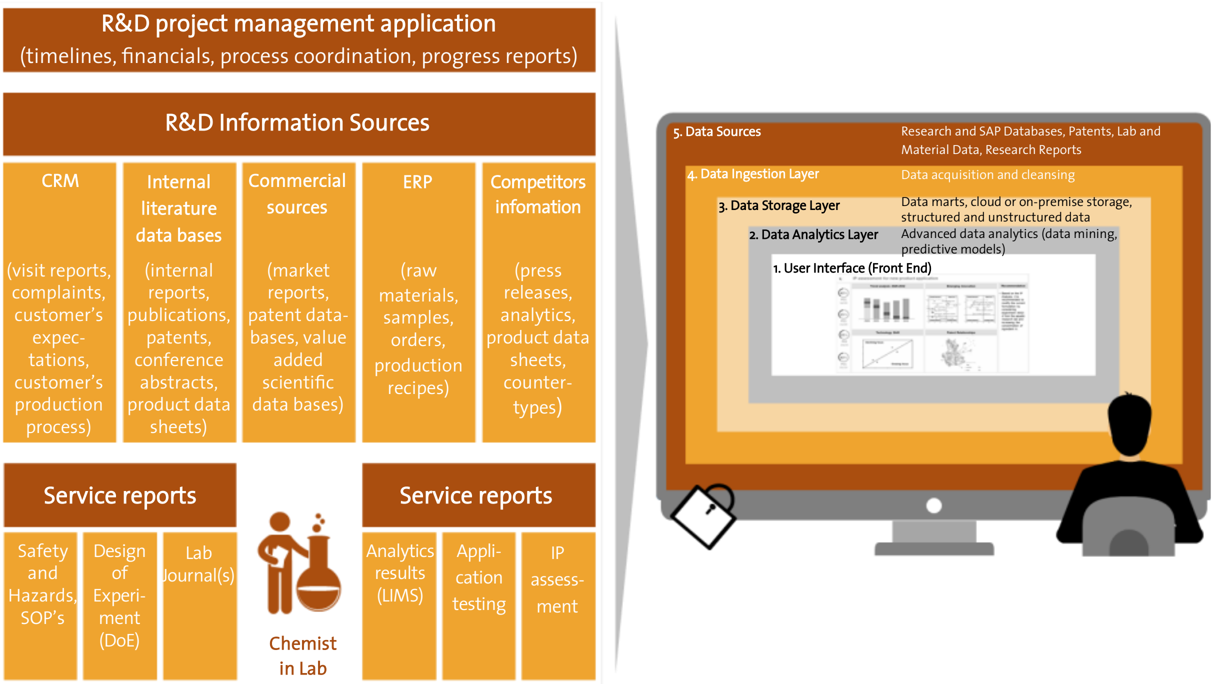
Why have chemical companies not embraced digital R&D, while other similar industries have? The answer lies in the nature of the data chemical companies work with, and their traditional approach to testing. For example, in the pharmaceutical industry, regulations essentially require drugs to be tested using defined protocols on large sets of subjects, with the effects being closely monitored over time-sometimes for years. This process produces a comprehensive set of data around the substance being tested.

In contrast, innovation in the chemical industry is usually focused on various, highly customized products for specific applications. After successful testing within the R&D lab, the innovation process gradually shifts via pilot plant to technical service departments. There, the product is fine-tuned to specific customer production lines through a few trial-and-error iterations. With the new product complete, the process is essentially repeated for the next product and customer. With this approach, data is generated at different points along the internal and external value chain, and various customized products are run through the process separately. Thus, the data is usually scattered among many stakeholders, and is typically not kept in a standardized format. Historically, dealing with this disparate data has been more than rules-based software can handle. However, that has changed, with technologies such as big data analytics, AI, data-lake architectures and lab automation making it possible to leverage that disparate data to accelerate R&D.

Through the targeted use of technology, R&D departments in the chemical industry can achieve greater efficiency and contribute significantly to competitive advantage. Potential benefits include a reduction of up to 5 percent in lead time from idea generation to market launch of new products; a 5 percent to 10 percent increase in R&D throughput in the form of new commercial products; a portfolio shift to higher value products; and a 25 percent to 35 percent increase in employee efficiency (Accenture, 2018).

Looking further, the technology can enable greater benefits over time as it transforms

Figure 1 Legacy IT systems a lab researcher has to deal with, contrasted with a fully data-integrated Laboratory 4.0 layered architecture (source: own representation).



Journal of Business Chemistry

R&D. Chemical companies can create "Laboratory 4.0" operations by weaving lab technologies into a cohesive whole, with data from various pieces of lab equipment being automatically captured and routed to where it is needed- helping to both reduce costs and enhance creativity (see Figure 1).

3 Productivity boost through AI and simulations

Today, so-called strong AI, an intelligence indistinguishable from human intelligence, is yet to be realized. But so-called weak AI, which focuses on performing specific tasks, is here and having a clear impact. Indeed, AI is opening the door to a quantum leap in innovation and productivity. Using historic and current data, AI -capable supercomputers can now identify hidden relationships between molecular structure or formulation composition and the observed overall performance of a material. This makes it possible to quickly find new structure-effect relationships- potential innovations- that probably would not be discovered without the use of AI. Over time, machine learning, a type of AI, can create a database of such relationships, which can be used to develop new, currently unknown products, or to precisely reverse-engineer already-formulated products. This may sound futuristic, but in some areas it is now a reality. Take for example detergents or surface coatings: Based on physical measurements of the observed macroscopic properties of applied coatings and hundreds of calibration formulations, AI-based algorithms can propose the exact ingredients required to reproduce a new coating which exhibits the identical visible color after application on the substrate. Similarly, chemical compounds for tire production can be calculated, or enzyme formulations can be optimized for ethanol production. AI allows R&D to achieve results far beyond what was possible with previous rule-based approaches.

AI can also be used in conjunction with analytics to automate the analysis and interpretation of test series. Those technologies can also be used to conduct intelligent research in literature and patent databases or ELN to support the search for leads for new applications.

4 Tapping into the power of data

For many chemical companies, there may be no clear business case for an AI-equipped supercomputer. However, there are other options that may be less leading-edge, but are nevertheless effective. For example, the integration of laboratory technologies can significantly improve data collection. An IT infrastructure that integrates laboratory information management systems (LIMS) and ELNs can enable companies to gather R&D data into data lakes- centralized collections of large amounts of data that can be readily accessed by employees. In addition, a variety of tools make it easier to use data. Semantic search and chemical entity extraction can help identify relevant documents containing chemical structural information in text form, such as patents, or develop insights from unstructured presentations. And analytics software can be used to explore data and derive concrete suggestions for structured problem-solving, or to develop inspirations for new products.

But software and data have the potential to do more than inspire chemists. Recently, a research group developed an algorithm that uses deep neural networks to develop retrosynthetic routes to organic compounds. This approach handled nearly twice as many molecules, and worked 30 times faster, than non-AI computer-aided methods. Furthermore, those routes were of such quality that students were not able to tell whether they had been derived by a human or an artificial chemist (Segler, et al., 2018): . The question then is, are we seeing "creative" work being performed by a machine? In time, such algorithms are expected to be able to solve more complex chemical problems and significantly improve today's computer-aided synthesis planning.

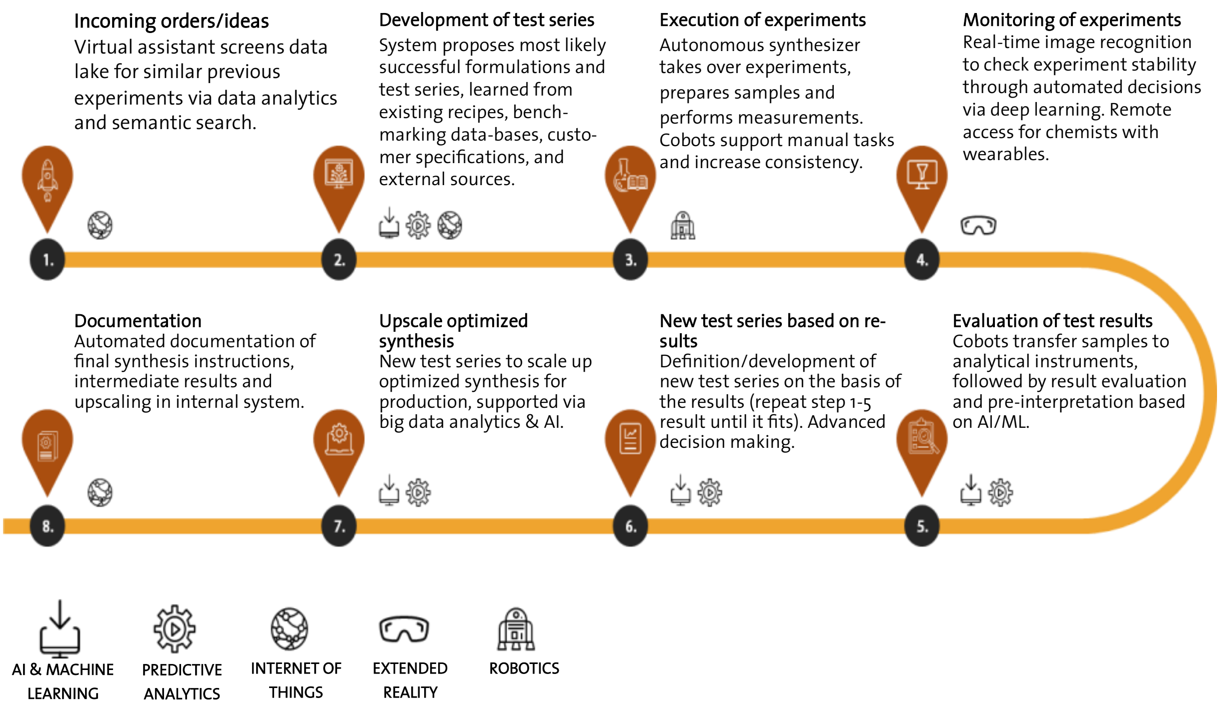
The availability of standardized information on reactions from various scientific databases is fundamental to this type of achievement. Companies can also take steps to leverage internal data more effectively through the use of data lakes- but it is vitally important that only validated data be allowed to enter the lake. Today, the process of data validation can be easily automated at the point of ingestion with the help of AI. In addition, the integration of knowledge from external sources, such as literature or patent databases, can augment the use of internal data in R&D. Overall, these efforts can make it possible to work with uniform data in laboratories at multiple geographically dispersed locations. This can make R&D faster and more cost-efficient, and shorten the innovation process by up to 15 percent, increase project throughput and considerably improve time-to-market (Accenture, 2018).

5 Taking lab automation to the next level

New automation technology has the potential to further raise productivity, standardization and reproducibility in the lab. For example, collaborative robots, called cobots, can supplement the work of humans, physically interacting with them in a shared workspace. Cobots are typically used for easy, monotonous and repetitive tasks. Unlike larger industrial robots, they can work in a safe and interactive manner without being "caged" and segregated from employees. In addition, cobots can adapt to real -world variability and are agile enough to be modified for different applications quickly. They perform their tasks without the errors associated with human work, and with greater speed and precision. They can, for example, reduce standard deviation when conducting specific experiments to one third of that typically seen in human work, while making globally determined test results comparable and more reproducible. Also, cobots can be manually set up and trained by non-engineers within minutes. Cobots are already freeing up R&D employees from repetitive tasks in some chemical company labs, and they are likely to be more and more common in the near future. Indeed, the overall industrial cobot market in 2020 is expected to be 10 times the size that it was in 2015, making it one of the fastest growing markets in the robotics space (Twentyman, 2017).

The use of cobots will be just the beginning, and lab automation will soon go far beyond simple tasks such as standardized injection of samples into measuring instruments. Take, for example, recent research results published by the University of Glasgow: A robotic laboratory platform equipped with a chemical reaction protocol compiler, based on the translation of commonly reported organic reactions, was able to produce three pharmaceutical compounds without any manual interaction. The platform covers all synthetic laboratory processes-from reaction to workup, isolation and purification of the product- and the yields and purities of products and intermediates are comparable

Figure 2 Potential automated process (source: own representation).



Journal of Business Chemistry

or even better than those achieved in traditional manual syntheses. The ability to track the work digitally greatly enhances reproducibility, as well (Steiner, et al. 2019).

Looking ahead, that type of system could be combined with cobots to automate the transfer of product samples to analytical equipment and testing processes (see Figure 2). Using AI to interpret results (which is already being done in some cases today) and using sensors for real-time measurements could enable a completely autonomous synthetic laboratory that could iterate and supervise reaction procedures until a verified route to a desired compound has been developed. Chemists could run and oversee experiments via remote access and augmented-reality tools from anywhere in the world, while robots prepare and run experiments 24/7. What's more, shorter product development cycles could be achieved through integration of data analytics and machine learning and by making optimized processes and chemical analytics easily available in form of a digital code. The future is clear, and close: With advancing digital technologies, a fully autonomous lab is not a distant vision but something that is very likely to become reality soon.

6 Further development of employees is crucial

There is no doubt that digitalization and automation will revolutionize R&D work. However, chemical companies moving to Laboratory 4.0 need to look beyond the technology itself. It is vital to incorporate employees in the transformation process and implement cloud and analytics solutions that are centered on the user- in this case the researchers/ lab technicians. Only then will users embrace new technologies as an opportunity to drive effectiveness and success, rather than regard it as a threat to their jobs. In addition, companies will need "translators"- specialists who are able to understand the functionality desired by chemists and translate it into technical requirements that can be understood and processed by the IT staff. At the same time, companies will have to expand their already-emerging teams of data scientists in order to make the best use of data, AI and analytics. Ultimately, the job descriptions for all R&D staff will continue to evolve, as technology and data knowledge become increasingly interwoven with traditional lab work.

7 A call for action: From lighthouse projects to transformation

Like other enterprise functions, R&D can now benefit from the technological revolution. Therefore, it is crucial that chemical companies provide appropriate IT structures and platforms to enable digitalization, AI and automation. But companies that want to bring laboratory digitalization to a higher level need to move beyond lighthouse projects. Often, companies remain in the stage of digital experimentation for too long with efforts limited to isolated proof-of-concept studies or pilot projects. Instead of such incremental change, they should pursue systematic transformation based on a tailored strategy for the standardization and industrialization of digital activities. The associated scaling up of digitalized processes will require a close collaboration between the IT and R&D departments.

Those companies that remain at the level of individual lighthouse projects are facing the risk of creating unwanted complexity, duplicating effort and resources and leaving potential value on the table. Conversely, if investments are made wisely, with the focus on transformation, companies can position themselves to use digital and AI technologies to drive creativity- and revolutionize chemicals R&D.

References

Accenture (2018): Digital disruption in the lab: The case for R&D digitization in chemicals, available at https://www.accenture.com/_acnmedia/PDF-86/Accenture-Chemicals-Blog-Digital-Disruption-Lab.pdf, accessed 19 March 2019.

Segler, M.H., Preuss, M. and Waller, M.P. (2018): Planning chemical syntheses with deep neural networks and symbolic AI, *Nature*, **555** (7698), p.604.

Microsoft Corporation (2018): *The Future Computed- Artificial Intelligence and its role in society*, available at https:/news.microsoft.com uploads/2018/01/The-Future-Computed.pdf, p. 38, accessed 19 March 2019.

Steiner, S., Wolf, J., Glatzel, S., Andreou, A., Granda, JM, Keenan, G., Hinkley, T., Aragon-Camarasa, G., Kitson, PJ, Angelone, D., Cronin, L. (2019): Organic synthesis in a modular robotic system driven by a chemical programming language, *Science*, **363** (6423).

Twentyman, J. (2017): *IIoT and the rise of the cobots,* available at https://internetofbusiness.com/iiot-rise-cobots/, accessed 09 April 2019.

